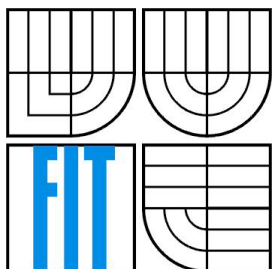


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INTELLIGENT SYSTEMS

INTELLIGENCE SKUPINY

SWARM INTELLIGENCE

DISERTAČNÍ PRÁCE

PhD THESIS

AUTOR PRÁCE

AUTHOR

Ing. ZDENKA WINKLEROVÁ

VEDOUCÍ PRÁCE

SUPERVISOR

Doc. Ing. FRANTIŠEK V. ZBOŘIL, CSc.

BRNO 2015

Abstrakt

Záměrem disertační práce je *aplikovaný výzkum skupinové (kolektivní) inteligence*. K prokázání použitelnosti inteligence skupiny je zkoumán algoritmus na bázi roje částic (*Particle Swarm Optimization PSO*), v němž je problém *inteligence skupiny* převeden na matematickou optimalizaci, kdy roj částic (*particle swarm*) hledá globální optimum ve vymezeném prostoru problému a prohledávání je řízeno podle předem nadefinované účelové funkce (*objective function*), která zastupuje řešený problém. Byla navržena a experimentálně ověřena strategie prohledávání, v níž částice průběžně přizpůsobují své chování charakteristikám prostoru řešeného problému, a bylo experimentálně zjištěno, jak se vliv řídicí účelové funkce zastupující řešený problém projevuje v chování částic. Výsledky experimentování s navrženou strategií prohledávání byly porovnány s výsledky experimentů s referenční verzí algoritmu *PSO*. Experimenty ukázaly, že klasické prohledávání, kde jedinou podmínkou je stabilní trajektorie, po níž se částice pohybuje v prostoru řešeného problému, a kde je ve výsledku eliminován vliv řídicí účelové funkce, může selhat a že dynamická stabilita trajektorií částic sama o sobě není ukazatelem prohledávacích schopností algoritmu ani konvergence algoritmu ke správnému, globálnímu řešení. Byl navržen způsob prohledávání prostoru řešeného problému, v němž algoritmus *PSO* reguluje stabilitu algoritmu průběžným přizpůsobováním chování částic charakteristikám prostoru problému. Navržený algoritmus usměrňoval vývoj prohledávání prostoru problému tak, že vzrostla pravděpodobnost úspěšnosti řešení.

Klíčová slova

Kolektivní inteligence, inteligence skupiny, rojová inteligence, optimalizace rojem částic

Citace

WINKLEROVÁ, Z.: *Inteligence skupiny*. Brno: Vysoké učení technické v Brně. Fakulta informačních technologií. Ústav inteligentních systémů, 2015. 154 stran, 28 stran dodatků, 7 stran příloh. Vedoucí doktorské práce doc. Ing. František V. Zbořil, CSc.

Abstract

The intention of the dissertation is the applied research of the *collective (group) (swarm) intelligence*. To demonstrate the applicability of the collective intelligence, the *Particle Swarm Optimization (PSO) algorithm* has been studied in which the problem of the collective intelligence is transferred to mathematical optimization in which the particle swarm searches for a global optimum within the defined problem space, and the searching is controlled according to the pre-defined *objective function* which represents the solved problem. A new search strategy has been designed and experimentally tested in which the particles continuously adjust their behaviour according to the characteristics of the problem space, and it has been experimentally discovered how the impact of the objective function representing a solved problem manifests itself in the behaviour of the particles. The results of the experiments with the proposed search strategy have been compared to the results of the experiments with the reference version of the *PSO* algorithm. Experiments have shown that the classical reference solution, where the only condition is a stable trajectory along which the particle moves in the problem space, and where the influence of a control objective function is ultimately eliminated, may fail, and that the dynamic stability of the trajectory of the particle itself is not an indicator of the searching ability nor the convergence of the algorithm to the true global solution of the solved problem. A search strategy solution has been proposed in which the *PSO* algorithm regulates its stability by continuous adjustment of the particles behaviour to the characteristics of the problem space. The proposed algorithm influenced the evolution of the searching of the problem space, so that the probability of the successful problem solution increased.

Keywords

Collective intelligence, group intelligence, swarm intelligence, particle swarm optimization.

Intelligence skupiny

Prohlášení

Prohlašuji, že jsem tuto disertační práci vypracovala samostatně pod vedením doc. Ing. Františka V. Zbořila, CSc. Další informace mi poskytl doc. Ing. Josef Schwarz, CSc. Uvedla jsem všechny literární prameny a publikace, ze kterých jsem čerpala.

.....
Zdenka Winklerová
15. května 2015

Poděkování

Ráda bych poděkovala vedoucímu disertační práce docentu Ing. Františku V. Zbořilovi, CSc. za návrh zpracování tématu intelligence skupiny a za možnost pracovat na tomto tématu, za přínosné kritické připomínky a doporučení ke struktuře a obsahové stránce disertační práce, za vedení, konzultace, podporu a čas, který mi věnoval.

Děkuji rovněž docentu Ing. Josefu Schwarzovi, CSc. za projevený zájem a uvedení do problému rojové intelligence.

Děkuji i profesoru RNDr. Milanu Čěškovi, CSc. za podněty a připomínky při průběžných prezentacích výsledků práce.

© Zdenka Winklerová, 2015.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

Seznam obrázků	7
Kapitola 1 Úvod do problematiky, charakteristika řešeného problému a záměr disertační práce	9
1.1 Model skupinové inteligence roje částic.....	9
1.2 Charakteristika problému skupinové inteligence roje částic.....	10
1.2.1 Určení počtu částic.....	10
1.2.2 Stanovení vah pro názorové posuny	10
1.2.3 Problematika návrhu reálné řídicí účelové funkce.....	11
1.3 Záměr disertační práce.....	14
Kapitola 2 Metodologie řešení problému a struktura předkládané práce	15
2.1 Metodologie COBP.....	15
2.1.1 Formulace problému a strategie řešení.....	16
2.1.2 Metriky, sociální faktory a scénáře	17
2.1.3 Metody, nástroje a data.....	19
2.1.4 Posouzení rizik studie	19
2.1.5 Výstupy	20
2.2 Struktura předkládané práce.....	20
Kapitola 3 Návrh metriky pro měření skupinové inteligence roje, formulace problému a cíle disertační práce	21
3.1 Charakteristika algoritmu PSO.....	21
3.2 Zralost rozhodovacího prostředí jako metrika pro skupinovou inteligenci roje	23
3.3 Zhodnocení algoritmu PSO vzhledem k atributům spolupracujícího chování.....	25
3.3.1 Reprezentace	26
3.3.2 Existence všeobecného povědomí.....	26
3.3.3 Zpravodajství	29
3.3.4 Interakce	29
3.3.5 Integrace	30
3.3.6 Koordinace	31
3.4 Posouzení zralosti operačního prostředí roje	32
3.5 Specifikace problému k řešení	33
3.6 Formulace problému	34
3.7 Cíle disertační práce.....	35
Kapitola 4 Meta-analýza a kritické zhodnocení současného stavu poznání v oblasti analytického stanovení hodnot koeficientů algoritmu PSO.....	37
4.1 Fyzikální pohled na pohyb částic v roji	38
4.1.1 Pohybové rovnice tělesa v konzervativním silovém poli	38

4.1.2	Pohybové rovnice částice v konzervativním silovém poli	40
4.1.3	Částice PSO v konzervativním silovém poli.....	41
4.1.4	Fyzikální význam ladícího parametru ω	43
4.1.5	Částice v nekonzervativním silovém poli	44
4.1.6	Pohybové rovnice částice v nekonzervativním visko-elastickém silovém poli	46
4.1.7	Pohybové rovnice nabité částice v nekonzervativním visko-elastickém silovém poli	47
4.1.8	Otázky k zodpovězení.....	49
4.2	Částice roje PSO jako dynamický systém.....	50
4.2.1	Částice PSO jako diskretní případ tlumeného harmonického oscilátoru.....	51
4.2.2	Částice PSO jako deterministický dynamický systém.....	53
4.2.3	Částice PSO jako dynamický systém popsán vektorem rychlosti.....	55
4.2.4	Částice PSO jako lineární dynamický systém	55
4.2.5	Otázky k zodpovězení.....	57
4.3	Vztahy mezi koeficienty plynoucí z deterministické analýzy stability chování částice PSO	58
4.3.1	Podmínky vnitřní systémové stability částice PSO.....	58
4.3.2	Analýza konvergence a volba koeficientů algoritmu PSO.....	62
4.3.3	Exploze, stabilita a konvergence částice PSO	65
4.3.4	Otázky k zodpovězení.....	70
4.4	Vztahy mezi koeficienty plynoucí ze stochastické analýzy stagnace částice PSO	71
4.4.1	Částice PSO jako stochastický systém popsán vektorem rychlosti	72
4.4.2	Analýza stagnace nejlepší částice	73
4.4.3	Porovnání s výsledky plynoucími z deterministické analýzy.....	74
4.4.4	Převzaté závěry experimentálních ověření výsledků stochastické analýzy stagnace nejlepší částice.....	75
4.4.5	Otázky k zodpovězení.....	76
4.5	Vztahy mezi koeficienty plynoucí ze stochastické analýzy standardní odchylky distribuce vzorků pozic stagnujících částice PSO	78
4.5.1	Trajektorie částice PSO jako stochastický systém popsán polohovým vektorem a dynamika momentů prvního řádu	78
4.5.2	Dynamika momentů druhého řádu	80
4.5.3	Stacionární body soustavy diferenčních rovnic pro momenty druhého řádu	81
4.5.4	Otázky k zodpovězení.....	84
4.6	Vztahy mezi koeficienty plynoucí ze stochastické analýzy stability trajektorií částice zobecněného algoritmu PSO	85
4.6.1	Zobecněný model algoritmu PSO (<i>Generalized PSO</i> , <i>GPSO</i>).....	86
4.6.2	Algoritmus <i>GPSO</i> jako analogie tlumeného pružinového systému.....	87
4.6.3	Vztahy mezi trajektoriemi v modelu algoritmu <i>GPSO</i>	88
4.6.4	Převzaté závěry experimentálních ověření výsledků stochastické analýzy stability trajektorie částice algoritmu <i>GPSO</i>	92

4.7	<i>Kritické zhodnocení současného stavu poznání v oblasti analytického stanovení hodnot koeficientů</i>	94
4.7.1	Stabilita a konvergence verze algoritmu PSO s konstantními koeficienty ...	95
4.7.2	Stabilita a konvergence verze algoritmu PSO s proměnnými koeficienty	97
4.7.3	Interpretace dynamického modelu systému částice PSO.....	98
Kapitola 5	Návrh řešení	99
5.1	<i>Zajištění dynamické stability systému částice PSO</i>	99
5.1.1	Odvrácení katastrofického scénáře.....	100
5.1.2	Vztahy mezi koeficienty	101
5.2	<i>Způsob detekce výlučnosti hodnoty účelové funkce</i>	102
5.2.1	Pravděpodobnostní model roje částic PSO	103
5.2.2	Testování výlučnosti globální nejlepší hodnoty účelové funkce	104
5.3	<i>Sledování chování částice v závislosti na účelové funkci a úspěšnost řešení</i>	107
5.4	<i>Výsledný algoritmus PSO</i>	108
Kapitola 6	Experimentální ověření navrhovaného řešení a dosažené výsledky	113
6.1	<i>Experimenty s originální verzí algoritmu PSO s konstantními koeficienty</i>	113
6.1.1	Simulace chování částic dle katastrofického scénáře.....	113
6.1.2	Ověření konvergence ke globálnímu optimu	118
6.1.3	Vliv inicializace na výsledek prohledávání	121
6.1.4	Vliv účelové funkce na průběh prohledávání.....	123
6.1.5	Vyhodnocení výsledků experimentování s originální verzí algoritmu PSO s konstantními koeficienty	125
6.2	<i>Experimenty s navrhovanou verzí algoritmu PSO</i>	126
6.2.1	Vliv inicializace na výsledek prohledávání, sledování rovnoměrnosti a detekce výlučnosti globální nejlepší <i>fitness</i>	126
6.2.2	Simulace chování částic dle katastrofického scénáře.....	131
6.2.3	Ověření konvergence ke globálnímu optimu	134
6.2.4	Vliv účelové funkce a inicializace na hodnoty koeficientů	136
6.2.5	Vyhodnocení výsledků experimentování s navrhovanou verzí algoritmu PSO	139
Kapitola 7	Posouzení rizik, návazné problémy a doporučení dalšího směru výzkumu	141
7.1	<i>Posouzení rizik navrhovaného řešení a návazné problémy</i>	141
7.1.1	Zhodnocení výsledků experimentů vzhledem k dosažení globálního cíle disertační práce.....	141
7.1.2	Zhodnocení navrhovaného řešení vzhledem ke splnění dílčích cílů disertační práce	142
7.1.3	Návazné problémy k řešení.....	144
7.2	<i>Doporučení dalšího směru výzkumu</i>	145
7.2.1	Roj jako komplexní systém	145
7.2.2	Model roje	147

Závěr		149
Odkazy		151
Dodatky		155
D1	<i>Odvození Eulerovy-Lagrangeovy rovnice (9)</i>	155
D2	<i>Odvození pohybové rovnice nabitě částice v nekonzervativním visko-elastickém silovém poli (50) (Reakční síla radiačního vyzařování a Abrahamova-Lorentzova formule)</i>	156
D3	<i>Odvození modifikované Eulerovy-Lagrangeovy rovnice (45)</i>	159
D4	<i>Odvození rovnice (52) charakterizující částici PSO jako dynamický systém popsáný polohovým vektorem</i>	161
D5	<i>Odvození rovnice (61a) charakterizující částici PSO jako dynamický systém popsáný vektorem rychlosti</i>	161
D6	<i>Odvození podmínky (73) vnitřní stability systému částice</i>	162
D7	<i>Úprava diskriminantu ve výrazu (85)</i>	163
D8	<i>Odvození rovnice (97) pro stochastickou analýzu stagnace nejlepší částice</i>	164
D9	<i>Charakteristiky náhodných veličin v rovnici (97) a jejich vliv na rychlost částice</i>	166
D10	<i>Odvození vzorce (103) pro maximální hodnotu koeficientu c_d plynoucího z deterministické analýzy</i>	170
D11	<i>Odvození soustavy diferenčních rovnic (109)</i>	171
D12	<i>Odvození vztahu (119) pro standardní odchylku pozic trajektorie částice od jejich očekávaných hodnot</i>	173
D13	<i>Ověření podmínky stability 2. řádu systému částice pro dvojici koeficientů $c, \omega \cong 1,48, 0,72$</i>	176
D14	<i>Vymezení oblasti stability momentů prvního řádu</i>	178
D15	<i>Odvození mezní křivky pro dvojice a, ω a parametr α</i>	179
D16	<i>Distribuční funkce a kritická hodnota pro Dixonův test odlehlých hodnot</i>	181
Příloha A		183
Příloha B		187

Seznam obrázků

- Obr. 1:** Porovnání strategií prohledávání.
- Obr. 2:** Diagram ilustrující hlavní fáze řešení komplexního problému podle metodologie *COBP*.
- Obr. 3:** Ontologické konceptuální schéma problému “*Intelligence skupiny*”.
- Obr. 4:** Model zralosti operačního prostředí (*Maturity Model*).
- Obr. 5:** Změna hodnoty $f(x)$ při posunu biologických částic po vodorovné ose x .
- Obr. 6:** Částice *PSO* jako tlumený harmonický oscilátor.
- Obr. 7:** Domény dynamického chování systému částice v prostoru koeficientů (ω, φ) .
- Obr. 8:** Příklady dynamického chování částice *PSO* pro různé volby koeficientů (ω, φ) .
- Obr. 9:** Horní hranice stability 2. řádu pro dvojice $[c, \omega]$.
- Obr. 10:** Změna deterministické oblasti stability zobecněného algoritmu *GPSO* v prostoru koeficientů (ω, φ) v závislosti na Δt .
- Obr. 11:** Rozšíření vyhledávacího prostoru $(-2x_{max} < x < 2x_{max})$ pro účelovou funkci $f(x)$ *Griewank* (*Příloha A*).
- Obr. 12(a):** Doslovný přepis původního algoritmu *PSO* [2].
- Obr. 12(b):** Sekvenční algoritmus *PSO* použitý v návrhu standardu [4] s koeficienty dle (87) a (88).
- Obr. 12(c):** *Pseudo-paralelní* přepis původního algoritmu [2] převzatý z [26:p.24].
- Obr. 13:** Částice *PSO* jako buzený oscilátor (*forced oscillator*).
- Obr. 14(a):** Návrh inicializační fáze výsledného algoritmu *PSO*.
- Obr. 14(b):** Návrh fáze prohledávání vymezeného vyhledávacího prostoru výsledného algoritmu *PSO*.
- Obr. 15(a):** Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k počtům úniků jednorozměrných částic mimo vymezený prostor $[-5.12, +5.12]$ (osa x) a odpovídající tendenční křivky pro originální algoritmus při působení funkce *Sphere*.
- Obr. 15(b):** Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k počtům úniků jednorozměrných částic mimo vymezený prostor $[-5.12, +5.12]$ (osa x) a odpovídající tendenční křivky pro originální algoritmus při působení funkce *Rastrigin*.
- Obr. 16:** Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k počtům úniků jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) a odpovídající tendenční křivky pro originální algoritmus při působení funkce *Rastrigin*.
- Obr. 17:** Histogram absolutních četností vyjadřující počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) a odpovídající tendenční křivka pro originální algoritmus při působení funkce *Schweffel*.

- Obr. 18:** Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k rychlosti konvergence (osa x) a odpovídající tendenční křivky pro originální algoritmus, při inicializaci a vymezení vyhledávacího prostoru $[-512, +512]$, při působení funkcí *Sphere*, *Rastrigin* a *Schwefel*.
- Obr. 19:** Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k rychlosti konvergence (osa x) a odpovídající tendenční křivky pro originální algoritmus, při inicializaci a vymezení vyhledávacího prostoru $[-512, +512]$ a $[-5.12, +5.12]$, při působení funkcí *Sphere* a *Rastrigin*.
- Obr. 20:** Jednodimenzionální profil funkce *Schwefel* při inicializaci vyhledávacího prostoru $[-512, +512]$ a rozmístění částic ve vyhledávacím prostoru po inicializaci pro simulace s vybranými, předem vytvořenými soubory *pseudo*-náhodných čísel.
- Obr. 21:** Pozice částic (osa y) v závislosti na pořadí iterace (osa x).
Porovnání odpovídajících trajektorií částic při působení funkcí *Rastrigin* a *Sphere* pro vybranou částici a vybraný soubor s *pseudo*-náhodnými čísly, při inicializaci vyhledávacího prostoru $[-512, +512]$ a neomezeném pohybu částic.
- Obr. 22:** Histogram absolutních četností vyjadřující počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) a odpovídající tendenční křivka pro navrhovaný algoritmus při působení funkce *Schwefel*.
- Obr. 23:** Porovnání histogramů absolutních četností vyjadřujících počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) navrhovaného algoritmu a originálního algoritmu při působení funkce *Schwefel*.
- Obr. 24:** Porovnání absolutních četností vyjadřujících počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) navrhovaného algoritmu a originálního algoritmu při působení funkce *Rastrigin*.
- Obr. 25:** Porovnání absolutních četností vyjadřujících počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) navrhovaného algoritmu a originálního algoritmu při působení funkce *Sphere*.
- Obr. 26:** Hodnoty koeficientů (osa y) v závislosti na pořadí iterace (osa x).
Porovnání vývoje hodnot koeficientů ω , a_l , a_g pro vybraný soubor s *pseudo*-náhodnými čísly při omezení pohybu částic do vyhledávacího prostoru $[-512, +512]$
(a) při působení funkce *Schwefel*,
(b) při působení funkcí *Rastrigin* a *Sphere*.
- Obr. 27:** Hodnoty koeficientů (osa y) v závislosti na pořadí iterace (osa x).
(a) koeficient setrvačnosti ω ,
(b) koeficient lokálního zrychlení a_l ,
(c) koeficient globálního zrychlení a_g .
Vzájemné porovnání průběhů hodnot koeficientů pro vybraný soubor s *pseudo*-náhodnými čísly při omezení pohybu částic do vyhledávacího prostoru $[-512, +512]$ při působení funkcí *Schwefel*, *Rastrigin* a *Sphere*.
- Obr. 28:** Vazby mezi třídami objektů navrhovaného algoritmu *PSO*.

Kapitola 1

Úvod do problematiky, charakteristika řešeného problému a záměr disertační práce

Komplexní systémy (*complex systems*) jako jsou hejna ryb, hejna ptáků, roje včel, či kolonie mravenců, můžeme považovat za skupinové entity zpracovávající informace (*information processing systems*) s kolektivním povědomím o prostředí, v němž se tyto entity nacházejí. Pochopení schopnosti prohledávání a poznávání prostředí vykazované skupinovými entitami, jehož výsledkem je nalezení určitých žádoucích pozic v prohledávaném prostoru, je podnětné pro vývoj aplikací distribuovaných technologií, jako jsou senzorické robotické roje (*olfactory robot swarms*), ale i pro řešení problémů abstraktní povahy, pokud pro daný problém dokážeme nadefinovat analogii prohledávaného prostoru a účelovou funkci (*objective function*), podle níž má být prohledávání řízeno. Daný problém lze potom převést na matematickou optimalizaci, kdy roj částic (*particle swarm*) hledá globální optimum ve vymezeném prostoru. Pod pojmem *intelligence skupiny* bude v předkládané disertační práci chápána právě takto definovaná schopnost prohledávání.

1.1 Model skupinové intelligence roje částic

Ke zkoumání vlastností roje částic z pohledu intelligence skupiny byl v roce 1995 navržen optimalizační algoritmus na bázi roje částic (*Particle Swarm Optimization PSO*) [1] [2]. Algoritmus *PSO*, jenž je inspirován biologickými komplexními systémy, je typický tím, že slabé částice v roji nezanikají, ale vzájemně spolupracují tak, že využívají informace od ostatních částic ke svému zlepšení. Model spolupracujícího chování roje na úrovni jednotlivých částic přitom nemá přímou oporu a vysvětlení v Darwinově teorii. Altruismus částic v roji je totiž v přímém protikladu s Darwinovou teorií, že individua jsou podněcována k co nejvyšší reprodukci a že selekce se projevuje změnou četnosti výskytu genů ve skupině.

Mechanismus selekce spočívající v nahrazení jednoho individua jiným je způsob informační výměny typický pro *genetické algoritmy*. Model roje *PSO* naproti tomu představuje *superorganismus* [3], v němž namísto selekce genů probíhá skupinová sociální adaptace formou optimalizace, v níž je maximalizována nikoli schopnost přežití jednotlivce, ale schopnost přežití roje jako celku (*group fitness maximization*). Ve shodě s formální teorií skupinové adaptace, uveřejněné v časopise *Journal of Evolutionary Biology* v roce 2009 [3], je sociální chování částic v roji vysvětlováno tím, že všechny částice mají stejný genotyp, a tudíž nejsou motivovány k soupeření za účelem reprodukce.

Tvůrci algoritmu *PSO*, sociolog *James Kennedy* a elektroinženýr *Russell Eberhart*, usoudili, že podobná pravidla jsou rovněž základem sociálního chování skupin lidí, avšak důležitým rozdílem je právě abstraktnost. Podobně jako živočichové vzájemně přizpůsobují svůj fyzický pohyb proto, aby se vyhnuli predátorům, našli zdroj potravy, či vyhledali optimální parametry prostředí, jednotlivé osoby mají při řešení problémů tendenci přizpůsobovat své názory na řešení problému ostatním ve skupině. Individuální názorové posuny lze potom interpretovat jako pohyby v prostoru řešeného problému.

Z psychologického hlediska má dle autorů [1] [2] každé individuum ve skupině (i) *tendenci neměnit svůj stávající názor* na řešení problému, má (ii) svou *autobiografickou paměť*, která ho přiklání k názoru, který se mu v minulosti nejlépe osvědčil, a zohledňuje (iii) *všeobecně přijatý názor*. Každý dílčí názorový posun má svoji váhu, která určuje jeho vliv na výsledný názor. Jednotlivé názorové složky jsou v algoritmu *PSO* modelovány jako vektory rychlosti posunu, které se ve výsledku sčítají a určují výslednou pozici v bodě prostoru problému v dané iteraci. Relevance názoru v daném bodě prostoru je měřena hodnotou účelové funkce (*objective function*), která je nadefinována tak, aby zastupovala řešený problém. Výsledný sociální simulátor je modelován jako matematický optimalizátor pomocí diferenčních rovnic vyjadřujících pohyb ve vymezeném prostoru problému.

1.2 Charakteristika problému skupinové inteligence roje částic

Algoritmus *PSO* je napsán v pouhých několika řádcích kódu, avšak pro své fungování vyžaduje (i) určení počtu částic, (ii) stanovení hodnot vah pro jednotlivé názorové složky a (iii) reprezentaci problému řešeného rojem částic v podobě účelové funkce.

1.2.1 Určení počtu částic

Počet částic v aplikacích algoritmu *PSO* je volen empiricky, obvykle 20 nebo 30 částic. Počet částic má být co nejnižší, neboť účinnosti algoritmu má být dosaženo aktivitou částic spočívající v důkladném prohledávání prostoru problému, a to zejména v počátcích. Zvýšení účinnosti algoritmu je dosahováno též uplatněním náhodných prvků ovlivňujících hodnoty vah.

1.2.2 Stanovení vah pro názorové posuny

Hodnoty vah pro názorové posuny byly v počátečních aplikacích algoritmu *PSO* zjišťovány experimentálně jako ladící parametry. Až do roku 2005 byl přijímán referenční model algoritmu *PSO* s proměnnými parametry, v němž váha setrvačné složky v průběhu simulace lineárně klesá, rovněž tak klesá váha složky reprezentující příklon k vlastnímu názoru. Váha příklonu ke všeobecnému názoru je naopak na počátku nejnižší, v průběhu simulace lineárně stoupá a nejvyšší je v závěru prohledávání.

Hodnoty vah v průběhu simulace zásadně ovlivňují úspěšnost aplikace, a proto se logicky objevují snahy o stanovení vah analyticky. Skutečnost, že názorové posuny jsou modelovány pomocí diferenčních rovnic vyjadřujících pohyb částice v prostoru, vedly řadu analytiků k interpretacím systému částice jako lineárního dynamického systému

(kmitajícího fyzikálního tělesa) a ke stanovení hodnot vah plynoucích z analýzy stability systému částice. V dané souvislosti byl v roce 2007 rovněž navržen nový referenční model pro algoritmus *PSO* [4], v němž váhy nabývají hodnot stanovených analyticky. Tyto hodnoty jsou neměnné po celou dobu simulace a váha příklonu k vlastnímu názoru a váha příklonu ke všeobecnému názoru jsou shodné.

V analytickém přístupu převládá dle názoru autorky předkládané disertační práce snaha o jednotné řešení společné pro široké spektrum účelových funkcí. Hledá se řešení, v němž figurují pouze váhy a jejich vzájemné vztahy, a jedinou podmínkou je stabilita trajektorie částice tak, aby částice konvergovala ke společnému globálnímu řešení. Vliv účelové funkce na trajektorii částice je ve výsledku eliminován. Výsledky příslušných analytických studií jsou experimentálně ověřovány na uměle vytvořených účelových funkcích navržených pro testování optimalizátorů, přičemž jsou voleny funkce symetrické, mající podobné tendenční křivky připomínající parabolu, kdy hledané globální optimum se nachází v oblasti vrcholu. Experimenty analytickým výsledkům nenasvědčují a řešení je nestabilní [5], ač podmínky stability mají být splněny. Jsou tedy zkoumány podmínky stability vyšších řádů. Do algoritmu *PSO* musí být včleněn dodatečný mechanismus omezování rychlosti, jenž je všeobecně vnímán jako pomocný, řešící pouze ojedinělé případy, a jehož skutečná funkce není zmapována.

Výše uvedený problém může být dle autorky předkládané disertační práce způsoben odklonem od sociálního modelu, přesněji od teorie [1], podle níž byl vytvořen algoritmus *PSO* [2], k ryze technickému pojetí modelování. Teorie [1] předpokládá názorové posuny, které mají mít určitou váhu. *Lze však hovořit o inteligenci skupiny, v níž všechny individuální názory na řešení problému mají od počátku do konce tutéž váhu, a problém, který má skupina řešit, není těmito váhami zohledněn?* Opomenutí vlivu účelové funkce na trajektorie částic totiž v důsledcích znamená ignorování problému, který řídící účelová funkce reprezentuje.

Odpovědi na uvedenou otázku však přesto nebude jednoznačné “ne”, jak by čtenář mohl logicky usoudit. Odpověď bude totiž záviset na charakteru účelové funkce. Pro jednoduché, hladké funkce typu *Sphere*,¹ které mají pouze jedno globální optimum, mohou být váhy konstantní, stanovené analyticky, a částice vždy k tomuto optimu správně konvergují. U složitějších problémů, reprezentovaných účelovou funkcí s více lokálními optimy, či *multi-objektivních* problémů, reprezentovaných dvěma i více účelovými funkcemi, jednoznačnou odpověď dle autorky předkládané práce není možno dát a je nutno vyslovit hypotézu, že hodnoty vah vedoucí ke konvergenci jsou na účelové funkci závislé.

1.2.3 Problematika návrhu reálné řídící účelové funkce

Za jedny z nejkompexnějších problémů, jejichž řešení se opírá o inteligenci skupiny, jsou pokládány expertní odhady (*expert estimation*). Skupinou, jejíž inteligence má být využita, je v daném případě expertní tým, jehož členové se pohybují v prostoru řešeného problému. Ne každý problém určený pro řešení týmem expertů však bude řešitelný prostředky umělé inteligence a možnost převedení problému na optimalizační úlohu bude rovněž záležitostí individuální. Stejně tak se může jevit, že úvaha o spojení sofistikované činnosti expertního

¹ Funkce *Sphere* má průběh $f(x) = \sum_{j=1}^n x_j^2$, kde n je dimenze problému.

týmu s činností relativně primitivních částic roje *PSO* je bezpředmětná. Autorka předkládané práce však přesto soudí, že sociální model navržený v roce 1995 [1], který se stal předlohou pro algoritmus *PSO*, vykazuje základní rysy shodné s modelem inteligence skupiny, který byl vytvořen v 50. letech 20. století pro řešení expertních odhadů a který je znám jako metoda *Delphi* [6].

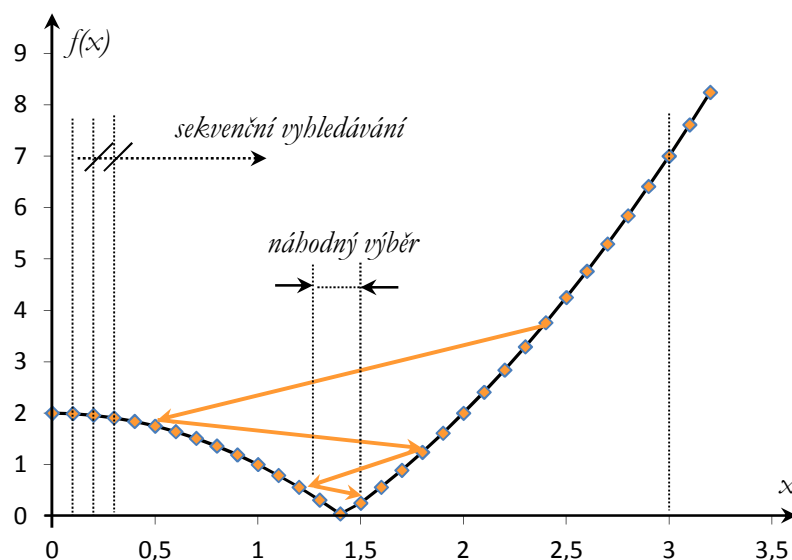
Zatímco skupinová inteligence roje částic je označována jako *rojová inteligence* (*swarm intelligence*), se skupinovou inteligencí týmu expertů je spjat pojem *kolektivní inteligence* (*collective intelligence*). K posouzení míry blízkosti obou modelů bude nyní naznačen postup řešení komplexního problému využitím metody *Delphi* [6] a možný způsob převedení takto formulovaného problému na optimalizační problém.

Nejdříve bude na jednoduchém příkladě vysvětlena podstata optimalizační metody implementující skupinovou inteligenci roje částic. Postup metody *Delphi* bude přiblížen na příkladu její aplikace z roku 1964, v níž byl řešen problém odhadu počtu lidí na zeměkouli v roce 2000 [6].

1.2.3.1 Jednoduchý problém pro matematický optimalizátor

Příklad

Řešme rovnici $x^2 = 2$, a to za předpokladu, že neznáme její analytické řešení. Potřebujeme tedy zjistit, pro kterou hodnotu x nabývá funkce $f(x) = x^2 - 2$ nulové hodnoty. Zaujímá nás pouze kladný výsledek, o němž tušíme, že by se měl vyskytovat v intervalu $(0,3)$. Hledáme tedy minimální hodnotu funkce $f(x) = |x^2 - 2|$ tak, že prohledáváme interval $(0,3)$. K prohledávání mohou být využity dvě základní strategie: (i) sekvenční vyhledávání nebo (ii) náhodný výběr hodnot $x \in (0,3)$. Porovnání obou strategií znázorňuje obrázek Obr. 1.



Obr. 1: Porovnání strategií prohledávání: Hledáme minimální hodnotu účelové funkce $f(x) = |x^2 - 2|$. Funkce $f(x)$ má jedno globální optimum v bodě $x = 1,414$. Vymezený prostor problému je interval $(0,3)$.

Při sekvenčním prohledávání postupujeme přes zvolený interval pomocí zvoleného kroku Δx a v každém kroku spočteme (či změříme) hodnotu funkce $f(x)$. Výsledkem je hodnota proměnné x , pro niž byla hodnota funkce $f(x)$ minimální (optimální). Je zřejmé, že přesnost popsané metody závisí na hodnotě kroku Δx a že výsledek nezískáme s absolutní přesností. Výpočetní náročnost sekvenční metody při požadované přesnosti je značná, uvážíme-li navíc, že definiční obor funkce $f(x)$ může být obecně n -dimenzionální.

Při náhodném výběru hodnoty $x \in (0,3)$ není nutné prověřovat postupně všechny hodnoty z vymezeného prostoru, avšak očekáváme, že algoritmus si při náhodném výběru sám zvolí velikost kroku Δx inteligentním způsobem tak, aby se krok postupně zmenšoval s tím, jak se pozice x přibližuje k hledanému řešení. Dosáhne-li hodnota kroku Δx určité hodnoty ε , kterou si předem zvolíme s ohledem na požadovanou přesnost, algoritmus může vyhledávání ukončit. (*Konec příkladu.*)

Představme si nyní částici pohybující ve vymezeném prostoru problému, která postupně zaujímá náhodné pozice a která má schopnost zapamatovat si svou dosud nejlepší dosaženou pozici. S využitím uvedené informace si částice může svou náhodnou pozici korigovat a dosahovat tak požadované konvergence. Pokud bude částicových průzkumníků více, každý z nich může obdržet informaci o celkově nejlepší dosažené pozici a prohledávání prostoru bude efektivnější. Popsanou strategii prohledávání používá i optimalizační metoda PSO [1] [2].

1.2.3.2 Komplexní problém pro tým expertů řešený metodou Delphi

Organizování diskusních skupin u kulatých stolů, kdy tým odborníků je vyzván k podání stanoviska k řešení komplexního problému nebo poskytnutí expertního odhadu, není počinem neznámým ani novým. Tradiční diskuse typu *brainstorming* se však často potýkají s psychologickými faktory, jako je přítomnost dominantní, přesvědčivé osobnosti, tendence jednotlivce přijmout řešení doporučované skupinou, sklon přičítat při rozhodování vyšší důležitost informacím, které jsou novější, a neochota jednotlivce či skupiny změnit rozhodnutí již jednou veřejně deklarovaná.

K eliminaci uvedených faktorů vedoucích k jevu označovanému idiomem “*hop on the bandwagon*”, jenž v daném kontextu značí nekritické akceptování většinového názoru, byla vyvinuta již zmíněná metoda *Delphi* [6]. Skupina expertů, kteří pracují anonymně, má odpovědět na otázku, která je formulována tak, aby odpověď byla numerickým údajem. Odpovědi jednotlivých respondentů jsou zpracovávány statisticky, aby bylo zaručeno jejich objektivní vyhodnocení. Výsledná odpověď je upřesňována v několika kolech a je zpětně poskytována jednotlivým respondentům. Každý respondent má rovněž umět svou vlastní odpověď zdůvodnit.

Příklad

Řešení problému odhadu počtu lidí na zeměkouli v roce 2000 [6] bylo jednou z prvních aplikací metody *Delphi* a probíhalo celkem ve čtyřech kolech dotazování. V prvním kole dotazování byli respondenti vyzváni, aby uvedli své odhady počtu lidí na zeměkouli v roce 2000. Odpovědi byly statisticky zpracovány a výsledný údaj o počtu lidí byl poskytnut jako zpětná vazba. Ve druhém kole měli respondenti zrevidovat svůj odhad a případně jej zpřesnit. Měli též své odpovědi zdůvodnit a uvést faktory, ze kterých jejich

odhady vycházely. Mezi faktory, které snižovaly odhad, byly *vzrůstající ekonomická prosperita, vzrůst životní úrovně a vzdělání, ztráty důsledkem válek, přírodních pohrom a epidemií*. Mezi důvody pro vysoké odhady byly *pokrok v medicíně, pokroky v zemědělství a globalizace* zajišťující efektivní distribuci potravin. Respondenti uváděli, že do svých odhadů promítli *odhady změn porodnosti a úmrtnosti*. Odpovědi byly statisticky zpracovány a výsledný údaj byl opět poskytnut jako zpětná vazba spolu se shrnutím důvodů vysokých a nízkých odhadů.

Ve třetím kole dotazování se měli respondenti navíc vyjádřit k argumentům, které byly nashromážděny ve druhém kole, a měli uvést, které z nich jsou dle jejich názoru nepřesvědčivé a proč. Do čtvrtého kola byl opět poskytnut výsledný odhad.

Ve čtvrtém kole byly shrnuty protiargumenty důvodů vysokých a nízkých odhadů. Byly dány k dispozici majoritní a minoritní odhady porodnosti a úmrtnosti a respondenti byli znovu vyzváni, aby zrevidovali své odhady. Medián odhadů jednotlivých respondentů byl potom prezentován jako výsledný odhad.

Na závěr byly rovněž sumarizovány faktory, které napomohly odhadu a které byly typu: *Jaká je světová populace v současné době? Jaká byla míra růstu populace během posledních padesáti let? Jaká je očekávaná délka života nyní narozených dětí? Kolik procent populace je starší 65 let?* Krom výsledného odhadu byl tedy k dispozici i pečlivě sestavený dotazník spolu s návody, jak dané faktory promítnout do výsledného odhadu. Výsledné číselné odhady učiněné aplikací metody *Delphi* publikace [6] z roku 1964 neuvádí.² (*Konec příkladu.*)

Z uvedeného příkladu je patrné, že metoda *Delphi* na rozdíl od metody *PSO* problém nejen řeší, ale iterativním způsobem jej i průběžně formuluje. Problém, který je expertním týmem formulován a jehož faktory jsou známy, by mohl být převeden na optimalizační úlohu v případě, že stěžejní otázka nebude definována samoučelně. Při odhadu populace bylo jistě účelem porovnání reálného odhadu populace s optimálním počtem vzhledem k faktorům, jakými jsou kupříkladu (omezené) životní zdroje.

Kolektivní inteligence expertního týmu *Delphi* a skupinová inteligence roje částic *PSO* vykazují při pohybu v prostoru řešeného problému základní společné charakteristiky, jimiž jsou (i) *anonymita respondentů*, (ii) *strukturování informací*, (iii) *zpětná informační vazba* a (iv) *role koordinátora*. Role koordinátora však není v algoritmu *PSO* [2] ani v teorii [1] deklarována tak, jako je deklarována přítomnost částic, ač dle autorky předkládané práce je i v algoritmu *PSO* koordinátor přítomen.

1.3 Záměr disertační práce

Záměrem předkládané disertační práce (*research plan*) je aplikovaný výzkum orientovaný na zjištění, jak hodnoty jednotlivých názorových vah ovlivňují inteligenci roje *PSO* v závislosti na problému, který roj řeší. Přesněji, jaký vliv na skupinovou inteligenci roje má posun algoritmu *PSO* od sociálního simulátoru [1], který se přibližuje metodě *Delphi*, k modelu [4], který má pevně stanoveny parametry vah. K realizaci záměru bude zvolena vhodná strategie řešení, jejíž součástí bude rovněž experimentální zjištění, jak se vliv řídicí účelové funkce zastupující řešený problém projevuje na chování částic.

² Metoda byla vyvinuta v letech 1950-1959 americkou společností *RAND* (*American Research and Development Corporation*) k odhadu budoucích technologií, které by mohly být využity ve vojenství, a metoda samotná i její výsledky byly na počátku utajovány.

Kapitola 2

Metodologie řešení problému a struktura předkládané práce

Z charakteristiky problému *skupinové (kolektivní) inteligence roje částic* lze vysledovat, že řešení problému *vlivu hodnot vah na skupinovou inteligenci roje částic PSO* dominují informační, kognitivní, sociální a behaviorální aspekty. Tyto aspekty naznačují, že analytický prostor řešeného problému bude složitý. Vzhledem k předpokládané závislosti hodnot vah na účelové funkci nebude možno provádět analýzy související se skupinovou inteligencí roje částic odděleně od účelové funkce a je třeba předpokládat, že při analýze komplexního problému, jakým je vliv hodnot vah na skupinovou inteligenci roje částic PSO, bude třeba více opakovaných průchodů a že průběžně získané poznatky budou ovlivňovat další jednotlivé kroky řešení.

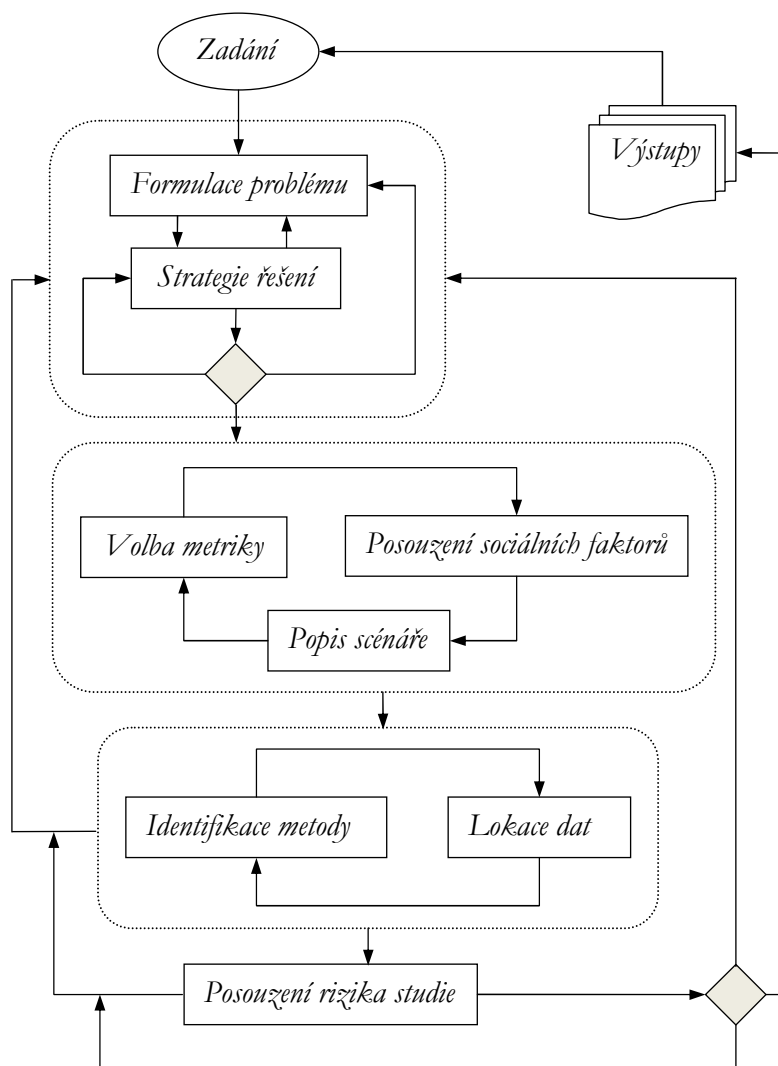
Osvědčený postup (*best practice*) při řešení problémů komplexní povahy poskytuje metodologie s názvem “*Code of Best Practice for Command and Control Assessments*” (COBP) [7], jež byla pro řešení problému *vlivu hodnot vah na skupinovou inteligenci roje částic PSO* zvolena. Metodologie COBP [7] je k dispozici ve formě manuálu, který byl vytvořen pro hodnocení účinnosti velení a řízení vzhledem k povaze plánované mise, avšak postupy zde uvedené jsou natolik obecné, že při určitém stupni abstrakce mohou sloužit i jako vodítko pro řešení problémů s obdobnou charakteristikou. Diagram ilustrující hlavní fáze řešení komplexního problému podle metodologie COBP je graficky znázorněn na obrázku Obr. 2.

Ve zbývajících částech kapitoly jsou popsány jednotlivé prvky diagramu, jimiž jsou (i) *formulace problému*, (ii) *strategie řešení*, (iii) *metriky*, (iv) *vliv lidského činitele*, (v) *scénáře*, (vi) *modely, nástroje a data*, (vii) *posouzení rizik* a (viii) *výstupy*, včetně vzájemných vztahů mezi jednotlivými prvky. Průběžně je diskutován způsob, jakým je metodika použita k řešení problému *vlivu hodnot vah na skupinovou inteligenci roje částic PSO*. V závěrečné části kapitoly je uvedena výsledná struktura předkládané disertační práce.

2.1 Metodologie COBP

Postup řešení podle metodologie COBP je iterativní proces. V kontextu metodologie COBP má řešený problém vždy mít svého garanta (*sponsor problem*), jenž schvaluje zadání problému a s nímž je řešení průběžně konzultováno.

Postup řešení předkládané disertační práce metodologii COBP odpovídal, neboť práce vznikla jako školní dílo v rámci výzkumného záměru MSM0021630528. Zadání práce obsahovalo (i) *návrh programového systému pro demonstraci použitelnosti skupinové inteligence*, (ii) *implementaci navrženého systému*, (iii) *provedení experimentů*, (iv) *porovnání s klasickými přístupy a výsledky* a (v) *doporučení směru dalšího výzkumu*.



Obr. 2: Diagram ilustrující hlavní fáze řešení komplexního problému podle metodologie COBP. Rozhodovací bloky značí konzultace se zadavatelem.

2.1.1 Formulace problému a strategie řešení

Formulace problému podle metodologie COBP je iterativní proces, který vychází z počáteční formulace problému (*initial problem formulation*) a je ukončen specifikací všech aspektů formulace problému, jimiž jsou (i) *otázky, které je třeba zodpovědět* (*issues to be addressed*), (ii) *předpoklady* (*assumptions*), (iii) *metriky* (*measures of merit*), (iv) *nezávislé proměnné* (*independent variables*) a (v) *rozsah hodnot nezávislých proměnných* (*constraints on the values of the independent variables*).

Proměnné jsou charakteristiky problému, jejichž hodnota se může změnit. Je třeba určit, které proměnné jsou výstupy (tj. závislé proměnné), a které proměnné jsou vstupy (tj. nezávislé proměnné). Význam metrik bude patrný později.

Formulace problému může být znovu upřesněna po stanovení strategie řešení a posouzení rizika studie. Platí, že problém může být přeformulován kdykoli v průběhu studie, pokud se objeví nové poznatky. Z charakteristiky tématu disertační práce s názvem “*Intelligence skupiny*” vyplynulo, že problém se vztahuje k *inteligenci skupiny relativně primitivních*

jedinců a že počáteční formulací problému ve smyslu metodologie COBP je *aplikace inteligence skupiny pomocí nástrojů umělé inteligence*.

Strategie řešení vychází z počáteční formulace problému a definuje jednotlivé analytické kroky vedoucí k vyřešení problému. Stanovení strategie řešení je rovněž iterativní proces. Metodologie COBP doporučuje, aby byl na počátku vytvořen počáteční *konceptuální model*, který bude obsahovat pouze *metriky a vstupní proměnné*. Konceptuální model může být upřesněn v dalších iteracích, po (i) *identifikaci sociálních (lidských a organizačních) faktorů*, které mají vliv na proměnné a jejich vztahy, a (ii) *definici scénářů*, které poskytují kontext pro zkoumání těchto proměnných a jejich vztahů.

2.1.2 Metriky, sociální faktory a scénáře

Metriky (*Measures of Merit MoM*) podle metodologie COBP tvoří množinu proměnných, které měří či jinak kvalitativně odrážejí *vliv velení a řízení na úspěšnost splnění mise*. Metriky jako takové jsou tedy závislými (výstupními) proměnnými a mohou tvořit hierarchii. Na počátku studie mají být definovány metriky na nejvyšší hierarchické úrovni. Autorka předkládané disertační práce definovala analogickým způsobem počáteční metriku, která hodnotí *vliv inteligence skupiny na úspěšnost vyřešení problému (společného úkolu)*, kde *inteligence skupiny* je vstupní proměnná a *míra úspěšnosti vyřešení problému* je metrika.

Sociální faktory zahrnují lidské činitele a organizační faktory. Lidské činitele jsou proměnné charakterizující pojmy jako je víra (*beliefs*), kulturní norma (*cultural norms*), stres (*stress*), inteligence (*intelligence*) a úroveň zkušeností (*level of experience*). V analýzách souvisejících s velením a řízením jsou to obvykle nezávislé (vstupní) veličiny. Metodologie COBP člení lidské činitele do tří hlavních kategorií:

- (i) *lidské chování spojené s degradací výkonu*, jako je stres;
- (ii) *chování spojené s rozhodovací pravomocí (decision making behaviour)* včetně problematiky složitosti rozhodování a schopností osob s rozhodovací pravomocí;
- (iii) *styl velení (command style)* charakterizující (ne)ochotu delegovat rozhodovací pravomoci na podřízené složky.

Organizační faktory jsou proměnné charakterizující vztahy mezi skupinami jednotlivců a mezi jednotlivci ve skupině, jako je soudržnost (*cohesion*), rozhodovací (velitelská) struktura (*command structure*), formální a neformální vztahy (*explicit and tacit relationships*) a informační toky (*information flows*). V analýzách souvisejících s velením a řízením jsou to rovněž typicky nezávislé (vstupní) veličiny. Navození situací, v nichž se mohou projevit proměnné a jejich vztahy, je účelem scénářů (*scenarios*). Scénář obsahuje následující prvky a jejich popisy:

- (i) *kontext (context)*;
- (ii) *účastníky (jejich záměry, schopnosti) (participants (their intentions, capabilities))*;
- (iii) *prostředí (environment)*;
- (iv) *vývoj prvků a událostí v čase (evolution of events in time)*.

Při posuzování sociálních faktorů vycházela autorka předkládané disertační práce z počátečního předpokladu, že jednotlivci ve skupině budou entity realizované prostředky umělé inteligence, a tedy že negativní lidský činitel, jako je stres, se nebude uplatňovat.

Aby byly eliminovány další negativní lidské faktory, jako je *bandwagon effect* zmíněný v úvodu práce, budou názory jednotlivých entit (řešitelů) na řešení zadaného problému hodnoceny podle relevance, nikoli podle toho, který řešitel daný názor vyslovil, a budou vyhodnocovány centrálně, entitou označenou jako koordinátor (moderátor).

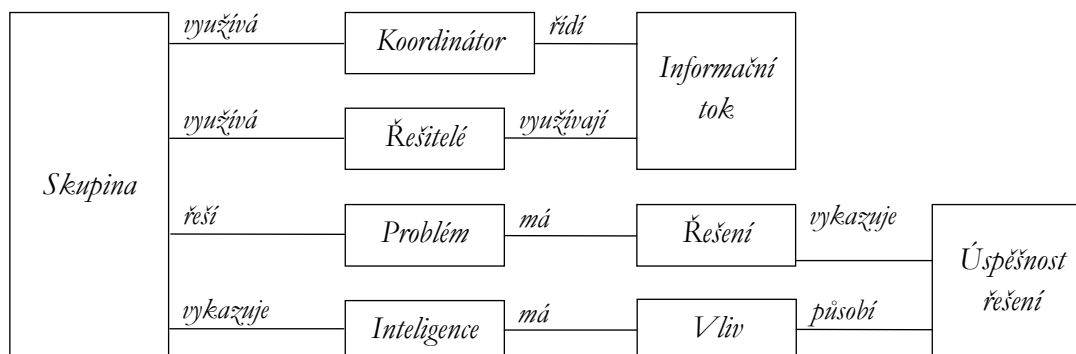
Přítomností koordinátora nebude dotčena samostatnost rozhodování jednotlivých řešitelů. Všechny entity ve skupině budou plně autonomní, nebude zde centrální rozhodovací prvek. Úlohou koordinátora bude pouze vyhodnocení informací a zpětné poskytování vyhodnocené informace řešitelům, aby řešitelé mohli na jejím základě opakovaně zpřesňovat své názory. Organizačním schématem nebude tedy tradiční “kulatý stůl”, ale spíše panelová diskuse, kdy řešitelé spolu nekomunikují přímo, ale prostřednictvím koordinátora tak, že každý řešitel bude mít svůj identifikátor. Koordinátor nebude znát identitu jednotlivých řešitelů, ani řešitelé nebudou znát vzájemně svou identitu.

Pro popis situačního scénáře bylo nejdříve nutno specifikovat pojem “relativně primitivní jedinec”, který dle zadání charakterizuje řešitele coby entitu realizovanou prostředky umělé inteligence, to jest softwarového agenta nebo inteligentního robota. U řešitele se tudíž *nepředpokládá* schopnost *kontextového* porozumění problému, který má skupina společně řešit. Skupina jako celek tedy nebude schopna naformulovat řešený problém, a proto musí být každý řešitel předem seznámen s jednotlivými parametry problému (otázkami, na které je nutno odpovědět) a možným rozsahem odpovědí. Dále musí být seznámen s algoritmem, jak z odpovědí na jednotlivé otázky vyhodnotit závěr.

Závěry jednotlivých řešitelů budou kvalitativně vyhodnoceny koordinátorem. Ani on však, stejně jako každý jednotlivý řešitel, nedokáže posoudit, zda výsledné “nejlepší” řešení je skutečně nejlepší možnou odpovědí. Koordinátor je tedy ve vztahu ke skupině řešitelů v pozici *vnitřního pozorovatele*.

Je téměř jisté, že skupina nevyřeší problém v prvním průchodu. Řešení bude tedy probíhat iterativně a skončí okamžikem, kdy se odpovědi jednotlivých řešitelů nebudou příliš lišit a/nebo po určitém počtu iterací. Jednotlivé dimenze problému budou ohodnoceny numericky, rovněž výsledkem řešení problému bude numerická hodnota.

Počáteční konceptuální schéma problému “*Intelligence skupiny*” je na obrázku Obr. 3. Pro vytvoření schématu zvolila autorka ontologický přístup, v němž vlastnosti entit jsou rovněž entitami.



Obr. 3: Ontologické konceptuální schéma problému “*Intelligence skupiny*”.

2.1.3 Metody, nástroje a data

Základní *metodou* používanou k ověřování výsledků analýz v oblasti velení a řízení je konstrukce modelů přesně simulujících reálné scénáře v aspektech, které odpovídají situacím, k nimž byly scénáře navrženy. K implementaci modelu slouží *nástroje* (zpravidla softwarové), kde *data* jsou údaje spojené s proměnnými simulačního modelu.

Důvodem k modelování v oblasti velení a řízení je skutečnost, že experimenty v této oblasti nelze provádět přímo.³ Simulační modely jsou zde nástrojem k experimentálnímu ověření výsledků analýz. V oblasti umělé inteligence může být vytváření modelů oproti tomu smyslem a výsledkem snažení. Modely nemusí přesně odrážet realitu, mohou být realitou pouze inspirovány, nebo může být vytvořen koncept zcela nový.⁴

Jelikož součástí zadání předkládané disertační práce byl návrh modelu (programového systému) pro demonstraci použitelnosti inteligence skupiny, autorka práce zkoumala existující modely označované jako kolektivní, spolupracující, sociální a/nebo inteligentní, a to jak v oblasti humánní, tak v oblasti umělé inteligence. Inspirativním modelem inteligence skupiny se pro autorku stala metoda *Delphi* stručně charakterizovaná v úvodní kapitole 1.2.3.2, která byla jako metoda kolektivní inteligence jako první takto označena.

Pro možnost implementace metody zkoumala autorka multiagentní modely sociálního chování, a sice model, který byl použit v základní struktuře projektu *ARCHON* [8], dále v systému pro jeho verifikaci *Coverage* [8] a v systému *Pleiades* [8], zejména však *twin*-bázový model (*Twin-Base Model*) [9] a zdokonalený *tri*-bázový model (*Tri-Base Model*, *3bA*) publikovaný a podrobněji zformalizovaný např. v [10].

Sociální agent je obecně navržen pro kolektivní řešení složitých úloh, které jsou nad rámec jeho individuálních schopností. Agent si sám s využitím vlastní znalosti schopností a aktuálního vytížení ostatních agentů vyjedná spolupráci s ostatními agenty a stane se koordinátorem řešené úlohy. Řešení úlohy neurčuje skupina, ale agent koordinátor, který je navržen k maximalizaci individuálního profitu a není tudíž vůči skupině v pozici nestranného vnitřního pozorovatele. Autorka usoudila, že tato skupina sociálních agentů nenaplňuje koncept kolektivní inteligence a obrátila pozornost k biologicky inspirovaným systémům a metodě *PSO*, stručně charakterizované v úvodní kapitole 1.2. V uvedeném smyslu byla rovněž upřesněna počáteční formulace problému k řešení jako *aplikace inteligence skupiny pomocí roje částic PSO*.

2.1.4 Posouzení rizik studie

Analýzy související s komplexními problémy bývají složité, často řeší nejasně ohraničené problémy, které nemusí být zcela zevrubně pochopeny. Tyto faktory zvyšují úroveň rizika spojeného s úspěšným vyřešením problému. Riziko dle metodologie *COBP* je spojeno zejména s formulací problému a identifikací sociálních faktorů. Postup řešení podle metodiky *COBP* pomůže rizika minimalizovat, avšak nikoli zcela eliminovat. Pro úspěšné zvládnutí rizik je doporučeno vytvářet seznam rozpoznaných rizik a tato rizika potom podrobněji analyzovat.

³ Nelze uskutečnit misi pouze proto, aby byl experimentálně vyzkoušen určitý způsob velení.

⁴ Příkladem, který má autorka na mysli, může být Neumannova koncepce počítačové architektury z roku 1945. Tento model nemá předlohu v žádném reálném systému, přesto provádí aritmetické operace neporovnatelně rychleji než lidský mozek.

2.1.5 Výstupy

Mají-li být výsledky studie prezentovány jasným a vyčerpávajícím způsobem, měly by výstupy studie podle metodologie COBP zahrnovat:

- (i) *deklaraci problému (statement of the problem)* zahrnující *formulaci problému (problem formulation)*, *cíle (objectives)*, *rozsah řešení a vstupní předpoklady (scope and assumptions)*;
- (ii) *strategii řešení (solution strategy)* obsahující jednotlivé analytické kroky, odrážející přístup k provádění studie, a jejich návaznosti;
- (iii) *dosažené výsledky, poznatky a závěry s diskusí (findings/conclusions with caveats)*;
- (iv) *doporučení (recommendations)*;
- (v) *návazné problémy k řešení (future challenges)*;
- (vi) *dodatky (appendices)*.

2.2 Struktura předkládané práce

Způsob řešení i prezentace výsledků předkládané disertační práce na téma “*Intelligence skupiny*” odpovídají metodologii COBP.

První, úvodní kapitola prezentující záměr předkládané disertační práce je výsledkem prvního průchodu metodologií podle diagramu na obrázku Obr. 2. Popis metodologie je obsahem druhé, stávající kapitoly, jejíž obsah odpovídá druhému průchodu. Struktura zbývajících částí předkládané disertační práce je výsledkem třetího průchodu metodologií a odpovídá způsobu, jakým mají být prezentovány výstupy podle kapitoly 2.1.5.

Formulace problému jako součást *deklarace problému* je včleněna do třetí kapitoly a opírá se o volbu metricky pro měření *skupinové (kolektivní) inteligence roje částic* a posouzení vlivu jednotlivých parametrů algoritmu PSO na tuto inteligenci. *Cíle práce, rozsah řešení a vstupní předpoklady* jako další součást *deklarace problému* vycházejí z *formulace problému k řešení* a jsou obsaženy v závěru třetí kapitoly.

Strategii řešení nelze stanovit bez předchozí důkladné znalosti stavu poznání (*State of Art*) ve sledovaných oblastech. Dalším analytickým krokem, který je obsahem čtvrté, nejrozsáhlejší kapitoly, proto byla podrobná *meta-analýza⁵ a kritické zhodnocení současného stavu poznání v oblasti analytického stanovení hodnot parametrů algoritmu PSO*.

Následuje *návrh řešení problému* v kapitole páté a *experimentální ověření navrhovaného řešení a prezentace dosažených výsledků* v kapitole šesté.

Posouzení rizik, návazné problémy a doporučení dalšího směru výzkumu je obsahem kapitoly sedmé. Disertační práce je autorkou celkově zhodnocena *závěrečným shrnutím*.

Předkládaná disertační práce se opírá o obsáhlý matematický aparát. Veškeré matematické vztahy uváděné v literatuře, o něž se autorka opírá, jsou též autorkou odvozeny, přičemž detaily jsou vyčleněny v dodatcích.

⁵ *Meta-analýzou* míní autorka syntézu současných poznatků ve sledované oblasti, které již byly analyzovány v literatuře, doplněnou o analýzu nových poznatků touto syntézou vzniklých.

Kapitola 3

Návrh metriky pro měření skupinové inteligence roje, formulace problému a cíle disertační práce

Aby bylo možno formulovat problém k řešení a stanovit cíle disertační práce, je třeba nejdříve vymezit pojem “inteligence skupiny”, přičemž předpokladem je implementace skupiny jako roje částic podle [1] [2]. První stěžejní otázkou, na niž hledala autorka předkládané disertační práce odpověď, je proto otázka *skupinové (kolektivní) inteligence roje částic*. Protože inteligenci nelze obecně definovat, ale pouze měřit, a kolektivní inteligence expertního týmu a skupinová inteligence roje částic jsou nesouměřitelné, bylo třeba nejdříve zvolit *metriku pro měření skupinové (kolektivní) inteligence roje částic*. Prvním autorčinným analytickým krokem proto bylo studium existujících variant algoritmu *PSO* s proměnnými parametry a posouzení vlivu jednotlivých parametrů na inteligenci roje podle zvolené metriky. Takto vymezená inteligence skupiny je označena jako “zralost chování”.

Předmětem kapitoly 3.1 je charakteristika roje částic implementovaného algoritmem *PSO*. Otázka metriky je diskutována v kapitole 3.2. Následuje zhodnocení algoritmu *PSO* vzhledem k atributům spolupracujícího chování v kapitole 3.3 a posouzení zralosti chování roje v kapitole 3.4, jež je východiskem pro bližší specifikaci problému v kapitole 3.5. Formulace problému k řešení je obsahem kapitoly 3.6. Cíle disertační práce jsou stanoveny v kapitole 3.7.

3.1 Charakteristika algoritmu *PSO*

Optimalizátor využívající roje částic (*Particle Swarm Optimizer PSO*) [2] je stochastický algoritmus aplikovatelný na jakýkoli problém, který lze charakterizovat účelovou funkcí (*objective function*) tak, že má být nalezen její globální extrém. Algoritmus přitom nevyžaduje žádnou další informaci, jako např. gradient účelové funkce, neboť vyhledávání globálního extrému je řízeno stochastickými komponentami vektoru rychlosti částice při jejím průletu prohledávaným n -dimenzionálním prostorem (hyperprostorem), kdy každá částice reprezentuje možné řešení problému. Stručný popis činnosti algoritmu je následující:

Na počátku prohledávání je určitá částice označena jako nejlepší, a to na základě hodnoty účelové funkce v pozici hyperprostoru, v níž se částice právě nachází. Ostatní částice jsou potom směřovány k této nejlepší částici, ale jsou též vychylovány směrem ke své vlastní nejlepší nalezené pozici. Každá částice má příležitost při svém průletu objevit lepší globální řešení a ostatní částice potom mění směr a zamíří k této nové nejlepší částici.

Pozice $\mathbf{x}$ a rychlost $\mathbf{v}$ každé částice jsou definovány jako vektory

$$\mathbf{x} = (x_1, x_2, \dots, x_N, t)$$

$$\mathbf{v} = (v_1, v_2, \dots, v_N, t)$$

jež jsou v originálním algoritmu *PSO* [2] aktualizovány podle následujících rovnic:

$$v_n^i(t + \Delta t) = \omega v_n^i(t) + a_l u_{1n}^i [l_n^i(t) - x_n^i(t)] \Delta t + a_g u_{2n}^i [g_n(t) - x_n^i(t)] \Delta t \quad (1)$$

$$\mathbf{x}^i(t + \Delta t) = \mathbf{x}^i(t) + \mathbf{v}^i(t + \Delta t) \Delta t \quad (2)$$

kde $n \in \mathbb{N}$ je počet stupňů volnosti optimalizačního problému, $\mathbf{l}^i$ je nejlepší dosud dosažená pozice i -té částice, $\mathbf{g}$ je nejlepší globální pozice dosud nalezená celým rojem, $\omega \in \mathbb{R}$ je reálné číslo zvané koeficient setrvačnosti, $u_{1n}^i, u_{2n}^i \in U(0,1)$ jsou náhodná čísla rovnoměrně rozložená v intervalu $(0,1)$, jimiž jsou váženy koeficienty zrychlení $a_l, a_g \in \mathbb{R}$ (jednotka *metr.sekunda⁻² m.s⁻²*), a Δt je diskrétní časový krok (jednotka *sekunda s*), jehož implicitní hodnota je rovna jedné. Složky vektoru pozice $x_1, x_2, \dots, x_N$ (jednotka *metr m*) i složky vektoru rychlosti $v_1, v_2, \dots, v_N$ (jednotka *metr.sekunda⁻¹ m.s⁻¹*) jsou reálná čísla.⁶

Rovnice (1) je interpretována tak, že *vektor aktualizace rychlosti částice* se skládá ze tří složek zvaných *složka setrvačnosti*, *kognitivní složka* a *sociální složka*. Kognitivní složka i sociální složka vektoru aktualizace rychlosti částice obsahují koeficienty zrychlení (a_l, a_g) v roli učících faktorů, které ovlivňují rychlost a pozici částice tak, že v kognitivní složce se odráží vlastní zkušenost částice, zatímco sociální složka vyjadřuje interakci mezi částicemi. Složka setrvačnosti vyjadřuje tendenci částice udržovat svou předchozí rychlost.

Je evidentní, že roj vykazuje známky inteligentního chování, neboť (i) *každá částice je schopna si zapamatovat svoje nejlepší řešení*, (ii) *mezi částicemi je informační tok, pomocí něhož částice komunikují s nejlepší částicí v závislosti na topologické struktuře roje*, (iii) *roj vykazuje kooperativní chování bez známek soupeření*.

Algoritmus *PSO* je primárně označován jako stochastický algoritmus proto, že částice při svém průletu hyperprostorem využívá náhodných prvků. Částice se však primárně orientuje pomocí své vlastní informace v kognitivní složce vektoru rychlosti a pomocí informace získané ostatními částicemi v sociální složce vektoru rychlosti. Tyto primární informace jsou až následně modifikovány, jednak pomocí koeficientů zrychlení, jejichž hodnota je získána empiricky, a jednak náhodně. Teprve souhrn získané a modifikované informace udává částici nový směr a modul rychlosti. *Chování částice je tedy stochastické*, neboť částice využívá náhodné prvky při svém rozhodování, a rovněž *deterministické*, neboť se učí ze své vlastní zkušenosti a/nebo ze zkušenosti ostatních částic.

Problém řešení pomocí *PSO* je reprezentován účelovou funkcí. Experimenty ukazují, že v závislosti na konkrétní účelové funkci (na konkrétním problému) je v některých případech převládající stochastická povaha chování částic pro řešení výhodnější, neboť

⁶ Každá částice je popsána N -ticí reálných čísel představujících souřadnice v kartézském souřadnicovém systému, kde každá osa může mít jiné jednotky, a sice podle povahy řešeného problému. Předpokládá se normované zobrazení sledovaných veličin do příslušné osy tak, že jednotkou je normovaný vektor, jehož délka je rovna jedné délkové jednotce.

lépe vede k nalezení optima [11:p.243]. Experimentálním měřítkem úspěšnosti je přitom rychlost konvergence algoritmu a kvalita výsledného řešení.

Je tedy možno pokládat stochastické chování rovněž za inteligentní? Na počátku prohledávání totiž nemá částice prakticky žádnou informaci o prohledávaném prostoru a jedinou její možností je vydat se náhodným směrem a tuto informaci posléze získat. Částice se na počátku prohledávání stejně tak nemůže spoléhat na zkušenosti ostatních částic, protože jejich informace jsou na počátku stejně chabé jako její vlastní informace.

Je tedy žádoucí posílit kognitivní schopnosti částice? Experimenty ukázaly, že relativně vysoká hodnota kognitivní složky má v raných fázích prohledávání za následek rozsáhlé, a tedy žádoucí putování částic ve vyhledávacím prostoru [11]. Pokud však bude částice využívat výhradně svou kognitivní složku i v pozdějších fázích prohledávání, bude výkonnost algoritmu výrazně menší [12]. Jedním z důvodů je skutečnost, že při absenci sociální složky vektoru rychlosti není mezi částicemi žádná interakce, a tedy možnost využití i zkušeností ostatních částic. Relativně vysoká hodnota sociální složky v raných fázích optimalizace může však vést k předčasné konvergenci algoritmu k lokálnímu optimu, neboť částice by měly nejdříve získat svou vlastní zkušenost a teprve potom ji předávat ostatním částicím. Avšak pro některé aplikace [13] je naopak výkonnost algoritmu, který používá výhradně sociální složku, výrazně vyšší než výkonnost originálního algoritmu.

Existence kognitivního orgánu je předpokladem inteligence individua. Uvážíme-li, že kognitivní schopnost je schopnost rozpoznávat, pamatovat si, učit se a přizpůsobovat se, kritickou vlastností částice bude patrně schopnost přizpůsobit svoje chování charakteru prostředí řešeného problému a rovněž schopnost učit se.

Otázka tedy zní: *Jak měřit skupinovou inteligenci roje jako celku*, víme-li, že částice se mohou jednotlivě chovat stochasticky nebo zcela potlačit svou vlastní zkušenost, a přesto algoritmus může konvergovat rychle ke správnému řešení, a tedy roj jako celek se zachová inteligentně?

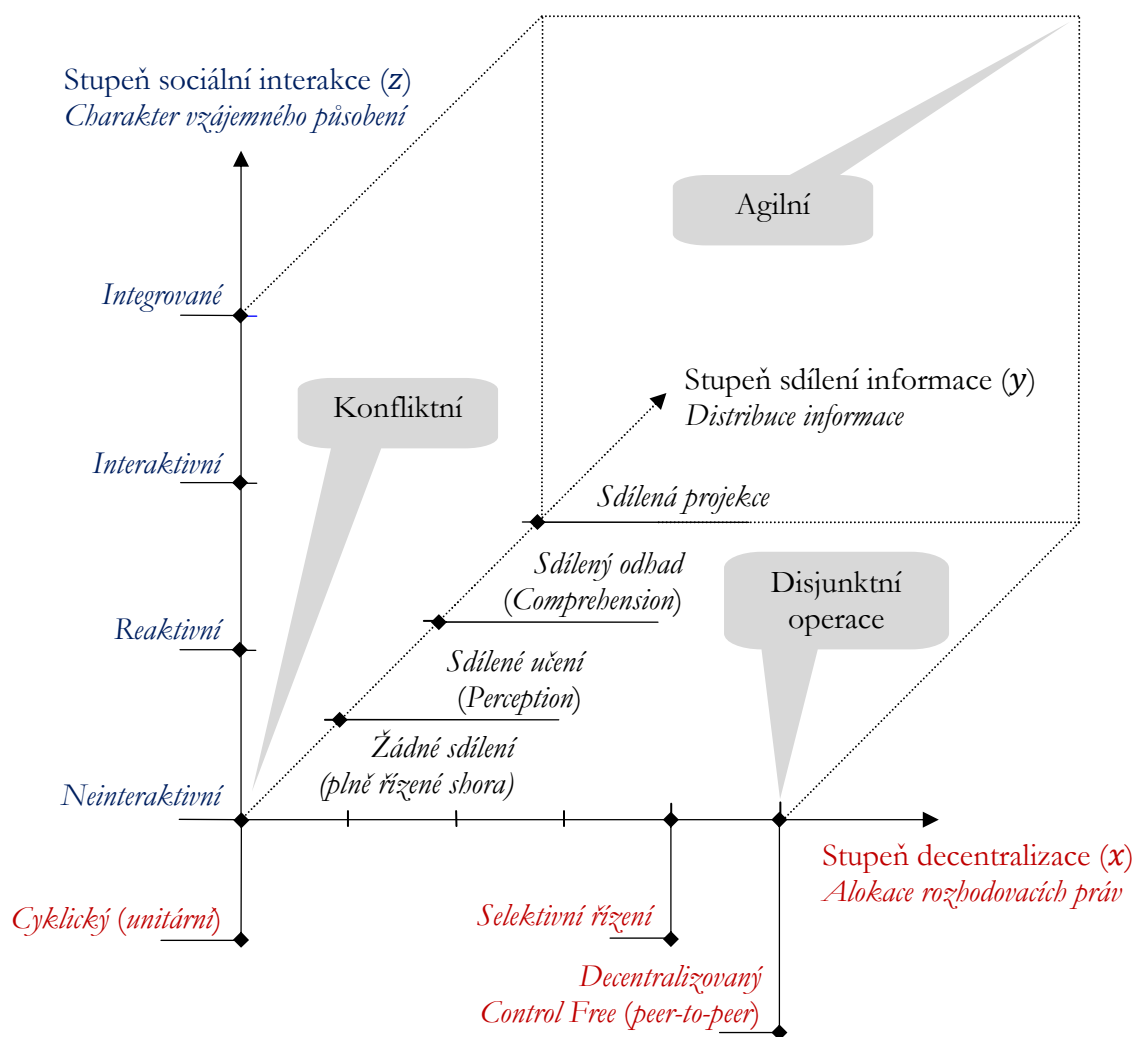
3.2 Zralost rozhodovacího prostředí jako metrika pro skupinovou inteligenci roje

Moffat a Albers navrhli v roce 2006 model zralosti operačního prostředí (*Maturity Model*) [14], jenž byl následně rozpracován v roce 2010 [15]. Model je navržen pro distribuované operační prostředí, které je charakterizováno tím, že obsahuje více autonomních, zpravidla geograficky rozptýlených entit tvořících systém, komunikujících na sémantické úrovni prostřednictvím výměny informace za účelem řešení společné úlohy. Zralost operačního prostředí má být posuzována podle tří základních ukazatelů, a sice podle (i) stupně alokace rozhodovacích práv (*Allocation of Decision Rights*), (ii) způsobu distribuce informace (*Distribution of Information*) a (iii) charakteru vzájemného působení (*Patterns of Interaction*) mezi entitami operujícími v daném prostoru.

Myšlenku znázorňuje diagram na obrázku Obr. 4. V původním diagramu [14], [15] nejsou explicitně ohodnoceny jednotlivé osy. Osy na Obr. 4 jsou proto ohodnoceny autorkou disertační práce dodatečně, a sice způsobem, jakým tento diagram autorka na základě svých vlastních znalostí charakteristik operačního prostředí interpretuje pro účely posuzování skupinové inteligence roje.

Osa x v modelu zralosti reprezentuje stupeň alokace rozhodovacích práv skupině entit. Osu x ohodnotila autorka ve shodě s klasifikací stupňů alokace rozhodovacích práv [16:pp.24-25] od unitárního až po decentralizovaný stupeň rozhodování. Decentralizované rozhodování odpovídá plně autonomním entitám. Roj by mohl být v tomto kontextu pokládán za decentralizovaný systém autonomních částic, pokud by se každá částice o svém chování rozhodovala samostatně, pouze na základě dostupných informací.

Osa y reprezentuje stupeň sdílení informace mezi částicemi. Při ohodnocení osy y autorka aplikovala všeobecně akceptované schéma JDL⁷ [17]. Sdílení informace od učení přes odhad situace až po projekci sdílené informace mohou částice dosáhnout tak, že je zabezpečeno šíření informace v roji pomocí zvolené topologické struktury.



Obr. 4: Model zralosti operačního prostředí (Maturity Model) podle [14], [15]

Dostupná informace poskytnutá ke sdílení by však neměla žádný význam, pokud by částice na tuto informaci nereagovala. Stupeň sociální interakce mezi částicemi proto

⁷ Pracovní skupina JDL (Joint Directors of Laboratories Data Fusion Working Group) byla vytvořena v roce 1996 ke koordinaci činností US DoD (United States Department of Defence).

udává osa z . Při ohodnocení osy z vycházela autorka z obecné definice sociální interakce jako posloupnosti sociální akce a následné sociální reakce dosažené pomocí výměny informace. Smyslem sociální akce je vybudit reakci, jejímž následkem je změna chování. Z hlediska výměny informace mezi entitami lze rozlišit několik úrovní: (i) *neinteraktivní chování*, kde není žádná reakce, (ii) *reaktivní chování*, kde reakce je vztažena pouze k jediné akci a (iii) *interaktivní chování*, kdy reakce je vztažena k více akcím a jejich oboustranným následkům.

Výslednou zralost operačního prostředí určuje trojrozměrný vektor, jehož souřadnice (x, y, z) jsou vyjádřeny jako stupně (*decentralizace*, *sdílení informace*, *sociální interakce*). Podle hodnoty souřadnic (x, y, z) může potom výsledná zralost operačního prostředí analogicky s [14], [15:p.80] nabývat hodnot (*konfliktní*, *nekonfliktní*, *koordinovaná*, *spolupracující*, *agilní*).

Pro názornost lze uvést následující příklady určení zralosti podle diagramu na obrázku Obr. 4: Při *decentralizovaném řízení*, *sdílené projekci* a *integrováných* entitách bude prostředí *agilní*. Při *unitárním řízení*, *nesdílené informaci* a *neinteraktivních* entitách se jedná o *konfliktní* prostředí.

Roj, v němž budou částice využívat pouze svou kognitivní složku, a kde tedy nebude žádná sdílená informace ani interakce mezi částicemi, bude patrně *nekonfliktní*, avšak jistě nebude *spolupracující*. Zhoršené chování kognitivní verze algoritmu *PSO* bylo experimentálně prokázáno již v roce 1997 [12] a tento poznatek je rovněž v souladu s filosofií modelu zralosti [14], [15]. Má-li být tudíž dosaženo efektivnějšího chování roje, je třeba se zaměřit na zbývající ukazatele, na *distribuci informace* a *interakci* mezi částicemi.

Pokud bychom stanovovali hodnoty zralosti pro všechny možné kombinace hodnot (x, y, z) , zjistíme, že ne všechny kombinace dávají smysl. Je-li kupříkladu ($z = \text{reaktivní}$), nemůže být ($y = \text{žádné sdílení}$), protože má-li částice reagovat na informaci, musí se k ní tato informace nejdříve dostat. Při hodnocení zralosti rozhodovacího prostředí roje je tedy nutno se zaměřit na zjištění, jakých hodnot by měly nabývat souřadnice (y, z) při konstantní hodnotě ($x = \text{decentralizovaný}$), aby roj mohl být označen jako *koordinovaný* a případně *spolupracující*.

3.3 Zhodnocení algoritmu *PSO* vzhledem k atributům spolupracujícího chování

Daniel D. Corkill stanovil v roce 2003 [18] základní atributy spolupracujícího chování mezi softwarovými moduly tvořícími systém (*Collaborating Software*), jež jsou všeobecně akceptovány. Těmito atributy jsou: (i) *vhodná reprezentace informace* (*representation*), (ii) *existence všeobecného povědomí* (*awareness*), (iii) *zpravodajství* (*investigation*), (iv) *interakce* (*interaction*), (v) *integrace* (*integration*) a (vi) *koordinace* (*coordination*).

Pojmy *interakce*, *integrace* a *koordinace* figurují i v navrhovaném modelu zralosti operačního prostředí. Budeme-li pokládat částice roje za softwarové moduly a algoritmus *PSO* za systém, dostaneme možnost propojit oba modely a posoudit roj částic podle charakteristik jednotlivých atributů spolupracujícího chování.

3.3.1 Reprezentace

*Vhodná reprezentace informací*⁸ [18] znamená, že částice chápou význam *sdílené* informace. Částice při své kooperativní činnosti nekomunikují přímo, ale vyměňují si informace o nejlepších globálních pozicích přes sdílené proměnné (vektory) v prostředí algoritmu PSO. Při komunikaci mezi roji (algoritmy) [26:pp.134-137] je informace sdílena přes kontextový vektor.

Pro úplnost je v dané souvislosti vhodné poznamenat, že částice při své činnosti používá další informace, které nevznikly učením od ostatních částic, a nejedná se tedy o sdílenou informaci. Jsou to jednak náhodné prvky u_1 , u_2 , jejichž hodnotu si částice sama vygeneruje a jež zapojuje do svého rozhodování, a jednak koeficient setrvačnosti ω a koeficienty zrychlení a_l , a_g (1), o jejichž hodnotě samy částice ani případný pozorovatel uvnitř systému nerozhodují, neboť tyto hodnoty jsou u stávajících variant algoritmu stanovovány experimentálně a/nebo analyticky a jsou jako parametry dodávány zvnějšku.

3.3.2 Existence všeobecného povědomí

*Existence všeobecného povědomí (awareness)*⁹ je vlastnost systému taková, že se částice dozví, pokud se v systému přihodí cokoli významného z hlediska její činnosti [18]. Obecně tato vlastnost znamená, že systém musí být schopen reagovat na jakoukoli strukturální nebo funkční změnu, jakou by bylo přidání nebo ztráta entity, snížení její spolehlivosti nebo její zahlcení, a musí být schopen pokračovat ve své operaci vzhledem k danému cíli.

V algoritmu PSO a jeho variantách nehrozí ztráta nebo zahlcení částice, částice pouze nesmí opustit vyhledávací prostor. Stávající varianty algoritmu PSO rozpoznají, že částice opustila vyhledávací prostor, a vrátí částici zpět, takže algoritmus PSO se jako systém zotaví a pokračuje ve svém běhu.

Z pohledu spolehlivosti částice nebo ztráty její schopnosti pokračovat v prohledávání má algoritmus PSO jednu nežádoucí vlastnost: implicitní vlastností roje částic je, že částice komunikuje pouze s omezeným počtem svých sousedů, což implikuje *lokální spíše než globální prohledávací schopnost* roje. Lokální schopnost prohledávání představuje úskalí pro multimodální účelové funkce (funkce s více lokálními optimy), kdy algoritmus může předčasně konvergovat k některému lokálnímu optimu.

Z hlediska vzájemné komunikace částic v roji existují dva základní modely algoritmu PSO zvané model *gbest* a model *lbest*. Rozdíl mezi oběma modely je dán množinou částic, které se budou navzájem ovlivňovat. Model *gbest* udržuje pouze jednu částici jako nejlepší řešení v celém roji. Tato částice působí na ostatní částice jako atraktant tak, že nakonec budou všechny částice konvergovat k tomuto nejlepšímu řešení. Nebude-li tudíž globální nejlepší částice aktualizována správně, může dojít k předčasné konvergenci algoritmu.

Model *lbest* se snaží zabránit předčasné konvergenci algoritmu tím, že udržuje více atraktantů. Pro každou částici je definována podmnožina částic jako její okolí, ze kterého je potom vybírána nejlepší lokální částice. Podmnožiny nejsou disjunktní, tatáž částice může patřit do několika okolí a jednotlivá okolí jsou tímto způsobem propojena. Pro

⁸ *Representation* – getting software modules to understand one another (Corkill, 2003 [18])

⁹ *Awareness* – making modules aware when something relevant to them occurs (Corkill, 2003 [18])

model *lbest* jsou konstruovány alternativní topologie [19, 20, 21], neboť bylo vypořádováno, že změnou topologie lze měnit rychlost informačního toku v modelu, a sice ve smyslu zpomalení. Pomalejší šíření informace v modelu umožní částicím prozkoumat více oblastí v prohledávaném prostoru, což sníží pravděpodobnost předčasné konvergence při aplikaci na *multi*-modální funkce.

Model *gbest* je speciálním případem modelu *lbest*, kdy okolí částice představuje přímo množinu všech částic v roji. Model *gbest* tak popisuje plně propojený roj. Topologie *gbest* produkuje lepší řešení na *uni*-modálních funkcích, a to díky rychlejšímu šíření informace v plně propojeném modelu [21].

Další okolností, jež má vliv na konvergenci algoritmu, je *tendence částice zachovávat svou původní rychlost pohybu*, z níž je částice kognitivní a sociální složkou rovnice aktualizace rychlosti (1) následně vychylována. Do originální verze algoritmu byl proto zaveden součinitel setrvačnosti částice ω , aby bylo vyváжено globální a lokální prohledávání během optimalizačního procesu [22]. Myšlenka je taková, že v prvních iteracích algoritmu je zapotřebí podstatně vyšší setrvačnosti částic, aby byl využit celý rozsah vyhledávacího prostoru, zatímco během pozdějších iterací, kdy algoritmus konverguje ke globálnímu optimu, je třeba zpomalení, a tedy menší setrvačnosti částic. Postupného zpomalování částic je dosaženo tak, že součinitel setrvačnosti lineárně klesá pro každou iteraci. Koncept je označován jako *TVIW* (*Time Variable Inertia Weight*) [23], [11]. Posléze bylo zjištěno, že tento lineární koncept není efektivní pro dynamické systémy (měnící se v čase), a byl zaveden koncept náhodného součinitele setrvačnosti *RANDIW* (*Random Inertia Weight*) [24], [11], v němž je součinitel pro každou iteraci vygenerován pomocí rovnoměrného náhodného rozložení.

Ačkoli algoritmus *PSO-TVIW* lokalizuje optimální řešení mnohem rychleji než originální verze algoritmu *PSO*, jeho schopnost přesného doladění optimálního řešení je stále nedostatečná. Důvod byl přičítán *nedostatečnému rozptylu (diversity) částic v závěrečné fázi* prohledávání [25], [11]. Proto byl zkoumán další koncept, jak rozšířit globální prohledávání v raných fázích a jak přimět částice ke konvergenci ke globálnímu optimu v závěru prohledávání, a sice koncept časově proměnných koeficientů zrychlení, které určují vliv kognitivní a sociální složky na aktualizaci rychlosti a pozice částice [11]. Koncept je označován *TVAC* (*Time Varying Acceleration Coefficients*). Experimenty ukázaly, že relativně vysoká hodnota kognitivní složky má v raných fázích za následek rozsáhlé putování částic prohledávaným prostorem. Naproti tomu relativně vysoká hodnota sociální složky v raných fázích optimalizace může částice předčasně dovést k lokálnímu optimu. Z tohoto důvodu je doporučováno přizpůsobit hodnoty obou koeficientů tak, aby oba druhy prohledávání byly vyvážené.

Další koncept pro zvýšení rozptylu částic je znám jako *CPSO* (*Cooperative Particle Swarm Optimizer*) [26], v němž je vektor pozice částice (2) rozložen na jednotlivé složky a každá složka je optimalizována zvlášť samostatným rojem (*sub-swarm*). K doložení, že varianta *CPSO* zvyšuje rozptyl (*diversity*) částic, byla navržena a aplikována měřítka rozptylu [27].

Přestože modifikace *TVIW*, *RANDIW* a *TVAC* zlepšují konvergenci originálního algoritmu *PSO*, nejsou schopny zabránit *předčasné konvergenci algoritmu PSO z důvodu stagnace částic*. Pokud bude aktuální pozice částice shodná s nejlepší globální pozicí, tj. bude-li

$x^i = l^i = g$ v rovnici (1), částice se bude pohybovat pouze ve směru své rychlosti. Pokud se předchozí rychlosti částic budou blížit nule, potom všechny částice zastaví svůj pohyb, jakmile narazí na tuto nejlepší globální částici. Přitom není zaručeno ani to, že algoritmus skončí v některém lokálním optimu – skončí v nejlepší globální pozici, která byla do té doby rojem rozpoznána. Proto byly vyvinuty strategie, jak zabránit předčasné konvergenci algoritmu *PSO* z důvodu stagnace částic, a sice (i) definicí výběrového prostoru pro nejlepší globální částici (algoritmus *GCPSO*), (ii) zavedením mutací a dodatečného učení (algoritmus *HPSO*) a (iii) rozšířením učení částic (algoritmus *CLPSO*).

Algoritmus s garantovanou konvergencí *GCPSO* (*Guaranteed Convergence Particle PSO*) [26:p.100-103] je založen na sledování počtu úspěchů nebo neúspěchů v předchozích iteračních krocích, přičemž neúspěch je detekován, jestliže se hodnota fitness nejlepší globální částice během poslední iterace nezmění. Algoritmus provádí náhodný výběr vzorku v oblasti obklopující nejlepší globální částici a modifikuje vektor rychlosti nejlepší globální částice tímto vzorkem.

Samo-organizující hierarchický algoritmus *HPSO* (*Self-organizing Hierarchical PSO*) [11] řeší stagnaci částic v důsledku ztráty jejich rychlosti. Pokud je detekována ztráta rychlosti, je částici dodán nový impuls tak, že modul vektoru rychlosti částice je nastaven na novou počáteční hodnotu náhodně vygenerovanou rychlostí. Jelikož bylo pozorováno, že výkonnost algoritmu je nízká při konstantních hodnotách koeficientů zrychlení, byly prováděny experimenty s kombinací metody *HPSO* metodami *TVAC*, *TVIW* a *RANDIW*, přičemž výrazně lepší na většině testovacích funkcí se jevila kombinace *HPSO-TVAC*, avšak s jedinou výjimkou¹⁰ [11:p.249-250].

Algoritmus s úplným učením *CLPSO* (*Comprehensive Learning Particle Swarm Optimizer*) [28] vyhledává částice, které byly nasměrovány do některého z lokálních optim (*weak particles*) a urychluje tyto částice ve směru lepší částice v roji. Proces urychlování je spojen s úplným učením tak, že částice se učí pro jednotlivé dimenze a v různém čase. Rozhodnutí, od které částice se daná částice v dané dimenzi bude učit, závisí na učící pravděpodobnosti, která může nabývat různých hodnot pro různé částice a je zjištěna experimentálně. Tato nová učící strategie dává částici schopnost uletět z oblasti lokálního optima, a to díky kooperativnímu chování celého roje. Pro další zlepšení efektivity využívá algoritmus *CLPSO* metodu *TVIW* [28:p.286].

Ve všech uvedených strategiích, jak zabránit předčasné konvergenci z důvodu stagnace částic, se částice sama snaží vyřešit si svůj problém stagnace, aby neposkytovala špatný učící příklad ostatním částicím. Ostatní částice nejsou o jejím problému nijak informovány – informace o stagnaci není sdílena. Vlastností algoritmu *PSO* také je, že částice může v dané iteraci přeletět globální optimální pozici, aniž tuto skutečnost nějak zjistí. Jednou z příčin může být, že si částice nerozhoduje o svých koeficientech setrvačnosti a zrychlení, které jsou jí v každé iteraci předem dány. Příkladem, že řídicí

¹⁰ Pro porovnání se strategií *HPSO* byly prováděny experimenty s modifikací *MPSO-TVAC* (*Particle Swarm Optimizer With Mutation and Time-Varying Acceleration Coefficients*). Pokud se globální optimum nezlepšilo po určitém počtu iterací, byla náhodně zvolena částice a byla změněna rychlost této částice v náhodně zvolené dimenzi a tato částice byla použita jako globální nejlepší částice. Výkonnost tohoto algoritmu byla srovnatelná s modifikací *PSO-RANDIV*, avšak s jedinou výjimkou pro účelovou funkci *Schaffer*, kde metoda *MPSO-TVAC* překonala všechny ostatní testované metody.

parametry mohou být dynamicky přizpůsobovány během prohledávání, je varianta *SACLPSO* (*Self-adaptive Comprehensive Learning Particle Swarm Optimizer*) [29].

3.3.3 Zpravodajství

Zpravodajství (*investigation*)¹¹ [18] slouží k rychlému nalezení informací vztahujících se k právě prováděným činnostem. Zpravodajská informace je v systému zpravidla vyhodnocována průběžně a postupována entitě, která ji potřebuje nebo která o ni požádá.

Lze konstatovat, že prakticky veškerá sdílená informace mezi částicemi má charakter zpravodajské informace, neboť vzniká vyhodnocováním informace po celou dobu operace roje. Způsob sdílení informace je dán topologickou strukturou *gbest* nebo *lbest* a systém zpravodajství je dán zvolenou učicí strategií.

Učicí strategie *lbest*, která vychází z konceptu “inteligence”, či “sociální inteligence”, vedla rovněž k myšlence “učinit částice plně informovanými” [30]. Ve variantách plně informovaného roje (*Fully Informed Particle Swarm FIPS*) se částice učí z více příkladů ve svém okolí, kdy je experimentováno s počtem učících příkladů a velikostí okolí částice. Originální algoritmus *PSO* lze tedy v pojetí *FIPS* pokládat za speciální případ, kdy se částice učí pouze od jedné, té nejlepší částice ve svém okolí a ze své dosud nejlepší dosažené pozice, a tedy ze dvou příkladů. Varianty *FIPS* vykazují lepší výsledky než originální algoritmus, avšak se zvětšujícím se okolím částice výkonnost jednotlivých variant *FIPS* klesá [30:p.210]. Pro topologie, kde jsou částice opravdu plně informovány, to jest, kde každá částice získává informaci od všech ostatních částic v roji, jsou výsledky experimentů velmi špatné, jak bylo ostatně předvídáno samotnými autory tohoto konceptu, a zvyšování schopnosti sdílení informace až do tohoto extrému se tudíž nejeví jako řešení, jak posílit částici v jejím rozhodování.

Ve variantě *CLPSO* [27], [18:p.135], která je typu *gbest*, je kognitivní a sociální složka nahrazena novou, jedinou komponentou, v níž je nejlepší pozice utvořena tak, že každá dimenze částice se učí od odpovídající dimenze jiné částice zvané *vzor* (*exemplar*). Částici je umožněno učení, pokud nezlepší svou pozici po určitém počtu iterací v závislosti na učicí pravděpodobnosti získané experimentálně. Je vygenerováno náhodné číslo, a pokud je toto číslo větší než pravděpodobnost učení, potom se částice učí od své nejlepší pozice, jinak se učí od jiné částice, která je vybrána jako lepší ze dvou náhodně zvolených částic. Každá částice se tedy může stát učícím příkladem a určovat směr pohybu ostatních částic, navíc každá dimenze částice se může učit od odpovídající dimenze jiné částice.

3.3.4 Interakce

*Interakce*¹² [18] je druh vzájemného působení v systému, kdy dvě nebo více entit vykazuje vzájemné ovlivňování činnosti. Ovlivňování činnosti se projevuje tak, že entita změní svoje chování jako reakci na určitou informaci. Působení může být jednostranné – reaktivní, avšak pro interakci je podstatné právě oboustranné působení.

¹¹ *Investigation* – helping modules to quickly find information related to their current activities (Corkill, 2003 [18])

¹² *Interaction* – creating modules that are able to use the concurrent work of others while working on a shared task (Corkill, 2003 [18])

Částice v roji jsou ve vzájemné interakci, neboť (i) informace pro částice zabezpečuje zpravodajský systém algoritmu *PSO* a (ii) sdílení informace mezi částicemi je potom umožněno topologickou strukturou částic (*lbest* nebo *gbest*), kterou jsou částice propojeny a (iii) částice si aktualizuje svou rychlost a pozici na základě své vlastní informace a sdílené informace ve shodě se zvolenou učící strategií.

Jelikož vektor aktualizace rychlosti dané částice (1) směřuje tuto částici k dosud nejlepší částici v jejím topologickém okolí, daná částice se posune k dosud nejlepší částici, ale v dalších iteracích se může situace vyměnit a daná částice se může sama stát nejlepší částicí, k níž ostatní částice směřují. Protože se do rychlosti částice v dalších iteracích promítají jak její vlastní zkušenosti, tak zkušenosti ostatních částic, je působení částic vzájemné, interaktivní.

Je důležité si uvědomit, že i při interaktivním působení částic může dojít k předčasné konvergenci částic z důvodu jejich stagnace v prohledávaném prostoru. Existující strategie, jak vyvést částice ze stagnace, mohou používat jiný způsob aktualizace vektoru rychlosti. Obecně tedy bude platit, že chování částic bude interaktivní, pokud částice budou využívat sociální složku vektoru aktualizace své rychlosti.

3.3.5 Integrace

*Integrace*¹³ [18] spočívá ve schopnosti částice slučovat vlastní dosažené výsledky s výsledky ostatních částic. Integraci lze obecně popsat jako stav systému, kdy entity jsou schopny souběžné činnosti a průběžné reakce na mezivýsledky při plnění společného cíle.

Souběžná činnost částic a průběžná reakce na mezivýsledky je v algoritmu *PSO* úspěšně simulována *pseudo*-paralelním způsobem. Činnost částic probíhá sekvenčně ve dvou po sobě následujících cyklech přes všechny částice, kdy v jednom cyklu se provede aktualizace nejlepších pozic a ve druhém cyklu následuje provedení změn v prohledávaném prostoru, to jest aktualizace rychlostí a pozic jednotlivých částic. V originální verzi algoritmu je nejdříve aktualizován celý n -dimenzionální vektor rychlosti částice podle rovnice (1) a teprve potom je aktualizována pozice i -té částice podle rovnice (2). Aktualizace probíhají v pořadí, které bylo částicím přiděleno při inicializaci.

Pro varianty algoritmu *HPSO* [11:p.244] a *CLPSO* [28:p.286] probíhá aktualizace pro jednotlivé dimenze částice. Dva cykly běžící přes n jednotlivých dimenzí částice jsou vnořeny do hlavního cyklu přes všechny částice. V prvním cyklu se provede aktualizace nejlepších pozic a ve druhém cyklu následuje provedení změn ve vyhledávacím prostoru. Souběžná činnost a průběžná reakce na mezivýsledky je tak simulována na úrovni dimenzí částice.

Ve variantě algoritmu *CPSO-S* [26:pp.134-137], která zavádí n rojů, z nichž každý optimalizuje jednu dimenzi částic původního n -dimenzionálního roje sekvenčním způsobem, si jednotlivé roje předávají informace přes kontextový vektor. Aktualizace probíhají analogicky ve dvou cyklech jako ve standardní verzi *PSO*, avšak přes jednotlivé roje. Souběžná činnost a průběžná reakce na mezivýsledky je tak simulována jednak na úrovni rojů a jednak uvnitř rojů na úrovni dimenzí částic.

¹³ *Integration* – combining results produced by other modules (Corkill, 2003 [18])

Integrací výsledků činnosti entit při vzájemné interakci má vzniknout nový typ řídicí činnosti a organizace na vyšší úrovni. Vzniká *metasystém*, který je složitější, inteligentnější a pružnější ve své činnosti než původní systém nezávislých entit. Ve shodě s obecnou teorií systémů bývá v pokročilé *meta*-přeměně systému vzájemná integrace částí typicky následována jejich specializací, kdy různé části přebírají různé role v metasystému a podstupují příslušné modifikace. Specializované části zůstávají samostatnými autonomními entitami, avšak se speciálními funkcemi. Proces se nazývá diferenciací částí.

Roj částic v originálním algoritmu *PSO* lze v uvedeném smyslu pokládat za *meta*-systém, v němž však nepozorujeme diferenciaci částí. Za specializované části by mohly být pokládány roje *CPSO-S* [26:p.135], neboť každý roj může být specializován na optimalizaci jedné dimenze. Vyšší stupeň integrace lze pozorovat u varianty *HCPSO-S_K* [26:p.144], kde je *pseudo*-paralelní činnost a průběžná reakce na mezivýsledky dovedena až na úroveň dvou algoritmů.

3.3.6 Koordinace

*Koordinace neboli součinnost*¹⁴ je úsilí spojující a usměrňující entity tak, aby zaměřily své činnosti na potřebné věci ve správný čas [18]. Podstatou součinnosti (synergie) je relativně autonomní činnost jednotlivých entit řešících dílčí úlohy a komunikujících za účelem usměrňování své činnosti. Koordinace spolupracujících entit se má uskutečňovat na principu samo-organizace při absenci centrálního prvku.

Dílčí úlohou částice v roji je posunout se v každé iteraci blíže ke globálnímu optimu a společnou úlohou roje částic je globální optimum najít. Koordinace činností částic v jejich společném úsilí je dosahováno jednak (i) regulací rychlosti informačního toku s využitím topologické struktury částic modelu *lbest* a jednak (ii) regulací rychlosti částic využitím součinitele setrvačnosti a koeficientů zrychlení.

Jak již bylo naznačeno v podkapitole 3.3.2, různé topologie pro model *lbest* jsou konstruovány proto, že umožňují zpomalovat rychlost informačního toku mezi částicemi [19]. Pro každou částici je nadefinováno okolí jako podmnožina částic, přičemž jedna částice je součástí okolí dvou a více částic. Všechny částice v roji jsou tímto způsobem propojené a mohou si vyměňovat informace nepřímo. Relativně dlouhá přenosová cesta mezi vzdálenými částicemi z různých okolí potom zpomaluje informační výměnu informace mezi vzdálenými částicemi, což částicím umožňuje zkoumání různých oblastí ve vyhledávacím prostoru současně při zachování schopnosti sdílení informací. Částice jsou vybírány do okolí podle indexu částice bez ohledu na pozici částice ve vyhledávacím prostoru. Topologie je nadefinována při inicializaci algoritmu a stává se charakteristikou prostředí, které se částice přirozeně přizpůsobí. Topologie tedy neomezuje autonomní činnost jednotlivých entit.

Regulace rychlosti částic modifikací součinitele setrvačnosti a koeficientů zrychlení, jak je popsána v kapitole 3.3.2, je způsob koordinace částic, který vyvažuje nedostatek všeobecného povědomí v systému. Částice si však setrvačnost ani zrychlení při svém průletu hyperprostorem neurčuje sama, na základě postupného poznávání charakteru

¹⁴ *Coordination* – getting modules to focus their activities on the right things at the right time (Corkill, 2003 [18])

prohledávaného prostoru. Částice totiž nemá informace, na základě kterých by se mohla rozhodnout, jakou setrvačností a jakým zrychlením by měla aktualizovat svou rychlost. Rychlost částice při průletu hyperprostorem je totiž uměle regulována z vnějšku systému, avšak tento způsob koordinace je v protikladu s principem samo-organizace.

3.4 Posouzení zralosti operačního prostředí roje

Předmětem diskuse při hodnocení zralosti rozhodovacího prostředí roje *PSO* vzhledem k atributům spolupracujícího chování bylo posouzení, jakých hodnot by měly nabývat souřadnice (y, z) v modelu zralosti na obrázku Obr. 4 při konstantní hodnotě $(x = \text{decentralizovaný})$, aby roj mohl být označen jako *koordinovaný* a případně *spolupracující*.

Nejvyšší dosažený stupeň sdílení informace (osa y) podle diagramu na obrázku Obr. 4 posoudila autorka předkládané disertační práce jako *sdílenou projekci*, neboť každá částice má při své činnosti možnost vyhodnotit dostupnou informaci o své pozici a o globální nejlepší pozici (učí se), aktualizuje informaci o globální nejlepší pozici, pokud sama objeví lepší pozici (provede odhad situace) a posune se očekávaným směrem (promítne tuto informaci do své činnosti). Částice vzájemně řídí svou činnost podle dosud nejlepších objevených pozic, a tedy mají schopnost sdílené projekce.

Z hlediska schopnosti částice slučovat vlastní dosažené výsledky s výsledky ostatních částic (osa z) pokládá autorka roj za *integrováný systém*, neboť částice v roji jsou standardně schopny souběžné činnosti a průběžné reakce na mezivýsledky při plnění společného cíle.

Z hlediska alokace rozhodovacích práv (osa x) autorka práce usoudila, že *částice nejsou zcela autonomní*, neboť rychlost částice při jejím průletu hyperprostorem je uměle regulována z vnějšku systému, a sice pomocí (*proměnného*) součinitele setrvačnosti ω a (*proměnných*) koeficientů zrychlení a_l, a_g (1) stanovených experimentálně. Oproti původnímu výchozímu předpokladu $(x = \text{decentralizovaný})$ by stupeň decentralizace měl být tedy označen spíše jako $(x = \text{selektivní řízení})$ [16:pp.24-25].

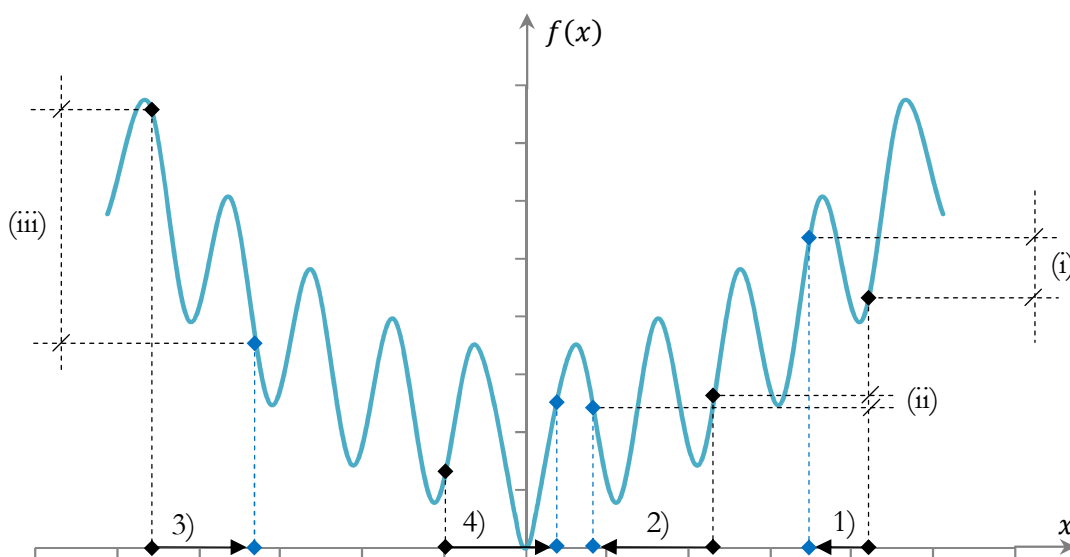
Posuzuje-li autorka *zralost* operačního prostředí roje podle metriky, kterou si stanovila a jejíž diagram je na Obr. 4, dochází k závěru, že chování roje jako systému kooperujících částic není zcela “*vyzrálé*”. Aby roj mohl být označen jako *koordinovaný* a případně *spolupracující*, bylo by nutno *posílit autonomní činnost jednotlivých částic*. Částice nemá schopnost získat veškerou informaci potřebnou pro aktualizaci své rychlosti učením z vlastní zkušenosti a/nebo od ostatních částic, případně od nějakého pozorovatele uvnitř systému, a pokud má částice problém, ostatní částice se o něm nedozvědí. *Nedostatečná existence všeobecného povědomí* v systému kooperujících částic je dle názoru autorky předkládané práce hlavní příčinou přetrvávající nespolehlivosti algoritmu *PSO* a jeho variant při optimalizaci *multi*-modálních funkcí s mnoha lokálními optimy. Zpravodajský podsystém roje by měl být proto rozšířen o *koncept vnitřního pozorovatele*, který by přispíval k lepší *koordinaci* částic.

Funkce podsystémů *interakce* a *integrace* částic u originálního algoritmu [2] a jeho variant se autorce jeví v daném kontextu inteligence skupiny jako dostatečná.

3.5 Specifikace problému k řešení

Problém *měření skupinové inteligence roje jako celku* lze po posouzení zralosti operačního prostředí roje upřesnit na problém *zvýšení skupinové inteligence roje posílením autonomní činnosti jednotlivých částic*. Předmětem zkoumání nyní bude, zda lze posílit autonomní činnost jednotlivých částic v roji takovým způsobem, aby částice používala k regulaci své rychlosti podle rovnice (1) koeficienty (ω , a_l , a_g), jejichž hodnota by byla stanovena samotnou částicí a/nebo pozorovatelem uvnitř systému, který sleduje a vyhodnocuje činnost jednotlivých částic a poskytuje jim příslušnou informaci, a to bez nutnosti vnějšího řízení, při zachování stochastického charakteru operace roje.

Řešením problému by byla adjustace rychlosti částice jako funkce její důvěry ve správnost (i) směru a (ii) velikosti (modulu) její aktuální rychlosti v dané pozici. Podle biologicky inspirovaného modelu publikovaného v prosinci 2009 pod názvem *Context – dependent interaction leads to emergent search behaviour in social aggregates* [31:p.22056], jenž vychází ze studia kolektivní kognice sociálních živočichů [32], by se chování individua mělo řídit intenzitou “signálu” detekovaného v dané pozici tak, že (i) při prudkém zhoršení detekovaného signálu má individuum zamířit do středu roje, (ii) při nevýrazné změně kvality signálu se má řídit podle svého okolí a (iii) při prudkém zlepšení signálu má individuum ignorovat své okolí a má se vydat dosavadním směrem.



Obr. 5: Změna hodnoty $f(x)$ při posunu biologických částic po vodorovné ose x .

Nejlépe hodnota $f(x) = 0$ je v pozici $x = 0$, (i) zvýšení hodnoty u částice 1) a 4),

(ii) nevýrazná změna hodnoty u částice 2), (iii) snížení hodnoty u částice 3).

Ve všech případech 1), 2), 3), 4) se pozice částice zlepšila.

Pokud by autorka předkládané disertační práce ztotožnila hodnotu signálu s hodnotou účelové funkce v dané pozici hyperprostoru a aplikovala by myšlenku adjustace rychlosti podle uvedené, biologicky inspirované strategie na roj částic *PSO*, pak by zlepšení či zhoršení signálu nemohlo být měřeno pouhou detekcí změn hodnoty účelové funkce. Intenzita signálu by musela být vyjádřena hodnotou *vylučnosti hodnoty účelové funkce* dané

částice v dané pozici. Hodnota *vylučnosti hodnoty účelové funkce* by nebyla závislá pouze na hodnotě účelové funkce v pozici, v níž se tato částice právě nachází, ale *kvantitativně stanovená hodnota* účelové funkce by musela být *kvalitativně ohodnocena* ve vztahu k ostatním hodnotám. Samotná částice PSO však ze svého subjektivního pohledu takového kvalitativního ohodnocení není schopna.

Situaci ilustruje obrázek Obr. 5., který znázorňuje změnu hodnoty účelové funkce $f(x)$ při posunu jednorozměrných biologických částic po vodorovné ose x . U částice označené 1) se hodnota účelové funkce v nové pozici zhoršila (zvýšila), přesto částice postupuje správným směrem. U částice 2) došlo k nepatrnému zlepšení (snížení) a částice by měla volit “průměrné” chování, přestože se blíží k optimu.

Ve skutečnosti došlo ke zlepšení pozice u všech částic. Autorka předkládané práce usuzuje, že žádná z hodnot účelové funkce na obrázku Obr. 5 patrně není *výlučná* (hodná zřetele) a patrně žádná změna hodnoty účelové funkce není výrazná. Jedná se tedy o případy, kdy částice se budou řídit podle svého okolí a koeficienty (ω, a_l, a_g) budou mít vyváženou konfiguraci.

3.6 Formulace problému

Jak je uvedeno v kapitole 2.1.1, problém dle metodologie COBP [7] je formulován, jsou-li specifikovány (i) *otázky, které je třeba zodpovědět*, (ii) *počáteční předpoklady*, (iii) *metriky*, (iv) *nezávislé proměnné* (to jest *vstupní proměnné*) a (v) *hodnoty*, jichž mohou nezávislé proměnné nabývat.

Při řešení problému *inteligence skupiny* vychází autorka z předpokladu, že skupina je tvořena rojem částic a odpovídajícím scénářem je *sociální simulátor* podle [1], [2], charakterizovaný v kapitole 3.1, a problémem je *zvýšení skupinové inteligence roje posílením autonomní činnosti jednotlivých částic* takovým způsobem, aby částice používala k regulaci své rychlosti podle rovnice (1) koeficienty (ω, a_l, a_g) , jejichž hodnota by byla stanovena samotnou částicí a/nebo pozorovatelem uvnitř systému.

Z takto specifikovaného problému vyplývají dvě základní otázky, které je třeba zodpovědět, a to (i) *Jak stanovit hodnoty koeficientů (ω, a_l, a_g) v závislosti na hodnotě účelové funkce a vylučnosti této hodnoty?* (ii) *Jak definovat a detekovat vylučnost hodnoty účelové funkce vzhledem k ostatním hodnotám?*

Metrikou na nejvyšší úrovni je *úspěšnost řešení* dle konceptuálního schématu na obrázku Obr. 3. Dílčí řešení bude pokládáno za úspěšné, dospěje-li algoritmus do globálního optima, a to za předpokladu, že počet iterací nevzroste dramaticky ve srovnání s počtem iterací, které vykáže testovaná konfigurace algoritmu při aplikaci na vstupní účelovou funkcí *Sphere*,¹⁵ přičemž rozptyl prozkoumaných pozic v dynamické historii roje bude dostatečný. Je třeba vzít v úvahu, že požadavek na dostatečný rozptyl prozkoumaných pozic je protichůdný k požadavku na co nejmenší počet iterací.

¹⁵ Funkce *Sphere* má průběh $f(x) = \sum_{j=1}^n x_j^2$, kde n je dimenze problému. Jedná se o jednoduchou, hladkou funkci, která má pouze jedno globální optimum, do něhož částice vždy správně konverguje, a proto je tato funkce používána jako referenční funkce při experimentování, kdy počet evaluací funkce *Sphere* ve srovnání s počtem evaluací jiné testované funkce je zvolen jako kritérium úspěšnosti testované varianty algoritmu PSO. Informativní grafické znázornění funkce *Sphere* nalezne čtenář v sekci *Příloha A*.

Relativní úspěšnost řešení bude dána poměrem dílčích úspěšných řešení k celkovému počtu pokusů (běhů testované konfigurace algoritmu).

Vstupní proměnnou podle schématu na obrázku Obr. 3 je *inteligence skupiny*. Dle závěrů v kapitole 3.5 lze usoudit, že inteligence skupiny je funkcí proměnných (ω , a_l , a_g) a hodnoty těchto proměnných závisí na účelové funkci reprezentující řešený problém. Další vstupní proměnnou je tedy testovaná účelová funkce (*objective function*). Autorka zvolila za stěžejní testovanou funkci účelovou funkci *Schwefel*, jejíž tendenční křivkou¹⁶ je přímka, a za srovnávací funkci účelovou funkci *Sphere*. Počet testovaných funkcí může být rozšířen v experimentální fázi řešení disertační práce.¹⁷

Je třeba ještě specifikovat vztah mezi účelovou funkcí a funkcí *fitness* ve vztahu k *vylučnosti*. Při řešení *maximalizačního* problému je hodnota účelové funkce v dané pozici totožná s hodnotou *fitness*. Při řešení *minimalizačního* problému, jehož případ znázorňuje obrázek Obr. 5, nabývá maximální hodnoty *fitness* pozice s nejnižší hodnotou účelové funkce.

Kvantitativně vyjádřená hodnota *fitness* je tedy mírou kvality. Autorka předkládané práce zavádí pojem *vylučnosti hodnoty fitness* v dané pozici jako další metriku. Výlučnost hodnoty *fitness* v dané pozici slouží jako *kvalitativní ukazatel* (číslo) vyjadřující relativní poměr k ostatním hodnotám *fitness* v ostatních pozicích.

3.7 Cíle disertační práce

Cílem disertační práce je *zlepšení výkonnosti algoritmu, jehož modelem je originální algoritmus PSO* [1], [2], a to *posílením autonomní činnosti částic roje*, kde rychlosti $\mathbf{v}$ a pozice $\mathbf{x}$ budou i nadále aktualizovány podle rovnic (1) a (2). Tento globální cíl je autorkou předkládané disertační práce strukturován následovně:

- (i) Originální algoritmus *PSO* bude rozšířen o *koncept vnitřního pozorovatele* v roli *koordinátora*, tak jak byl tento koncept nastíněn v kapitole 2.1.2. Vnitřní pozorovatel bude sledovat a vyhodnocovat činnost částic a bude poskytovat informaci o tom, jakou sadu koeficientů (ω , a_l , a_g) by si měla částice v dané pozici zvolit k regulaci své rychlosti podle rovnice (1). Tuto informaci bude vnitřní pozorovatel poskytovat bez nutnosti vnějšího řízení. Vnitřní pozorovatel bude při volbě sady parametrů zodpovídat za dynamickou stabilitu roje při zachování stochastického charakteru operace roje, a to ve smyslu kapitoly 1.2.2.
- (ii) Při stanovení sady parametrů bude pozorovatel vycházet z hodnoty *fitness* dané částice v dané pozici a *vylučnosti hodnoty fitness* ve vztahu k hodnotám *fitness* v ostatních pozicích, tak jak byla výlučnost specifikována v kapitolách 3.5 a 3.6. Bude navržen způsob, jak stanovit hodnotu výlučnosti.
- (iii) Součástí řešení disertační práce bude rovněž *výzkum chování částice v závislosti na účelové funkci a definice výstupní proměnné*, která by mohla indikovat dosažení globálního

¹⁶ *Tendenční křivkou* míní autorka tzv. *spojnici trendu*, která je pro danou funkci spočtena polynomicky podle implicitního vzorce obsaženého v programu MS Excel.

¹⁷ Informativní grafické znázornění všech účelových funkcí, na něž je v předkládané disertační práci odkazováno, nalezne čtenář v sekci *Příloha A*.

optima, jako další metriky. Touto metrikou je *úspěšnost řešení* dle konceptuálního schématu na obrázku Obr. 3. Vedlejší úlohou vnitřního pozorovatele tedy bude sledování, zda pozice prozkoumané částicemi v dynamické historii roje dostatečně pokryly celý vymezený (hyper)prostor v okamžiku, kdy částice konvergují k nějakému optimu. K tomu si pozorovatel bude vytvářet vhodnou mapu vymezeného prostoru. Pokud nebyl dosud dostatečně prozkoumán celý vymezený prostor (na mapě prostoru jsou bílá místa), nelze mít jistotu, že se jedná o globální optimum.

Návrh (strategii) řešení problému skupinové inteligence roje, tak jak byl problém formulován v kapitole 3.6, musí vycházet z důkladné znalosti stavu poznání (*State of Art*) ve sledovaných oblastech. Sledované oblasti, které je třeba prozkoumat, si autorka předkládané disertační práce stanovila ve shodě se stanovenými cíli jako (i) *oblast analytického stanovení koeficientů* (ω, a_l, a_g), (ii) *koncept vnitřního pozorovatele*, (iii) *kvalitativní analýza hodnot fitness* v dané iteraci a (iv) *chování částice v závislosti na účelové funkci*.

Oblast analytického stanovení koeficientů ω, a_l, a_g je obsáhle analyzována v literatuře, a proto je této oblasti věnována samostatná, následující kapitola 4. Zbývající oblasti jsou jednotlivě pojednány a zpracovány v návrhu řešení v kapitole 5.

Koncept vnitřního pozorovatele není v literatuře explicitně zmiňován, i když souvisí s (*pseudo*)*paralelním způsobem zpracování* dosažených výsledků, jak bylo popsáno v kapitole 3.3.5. Koncept vnitřního pozorovatele je tedy novým přínosem předkládané disertační práce, stejně jako kvalitativní analýza hodnot *fitness* a výběr vhodné metody pro tuto analýzu. Chování částice v závislosti na účelové funkci je v literatuře zmiňováno jako návazný problém k řešení (*future challenge*).

Nutnost kvalitativní analýzy hodnot *fitness* již byla objasněna v kapitole 3.5. Narušení *pseudo*-paralelního způsobu zpracování a eliminace vlivu účelové funkce na trajektorii částice jsou předmětem kritického zhodnocení současného stavu v oblasti analytického stanovení koeficientů (ω, a_l, a_g) v závěru kapitoly 4.

Kapitola 4

Meta-analýza a kritické zhodnocení současného stavu poznání v oblasti analytického stanovení hodnot koeficientů algoritmu *PSO*

Účelem *zkoumání vlivu jednotlivých koeficientů algoritmu PSO na inteligenci roje* podle zvolené metriky, jež bylo předmětem kapitoly 3.3 předkládané disertační práce a jež vyústilo v podrobnou specifikaci řešeného problému, bylo posouzení, do jaké míry algoritmus *PSO* jako abstraktní matematický model vystihuje dynamiku chování kooperující sociální skupiny. Dalším analytickým krokem bude *meta-analýza současného stavu poznání v oblasti analytického stanovení koeficientů algoritmu PSO*, cíleně zaměřená na pochopení smyslu koeficientů a jejich vztahu k účelové funkci. *Meta-analýzou* míní autorka syntézu současných poznatků ve sledované oblasti, které již byly analyzovány v literatuře, doplněnou o analýzu nových poznatků touto syntézou vzniklých.

Model operace roje, který velmi přispívá k intuitivnímu pochopení významu koeficientů (ω , a_l , a_g), vychází z fyzikálního pohledu na pohyb částic v roji, jenž byl publikován v roce 2008. Úloha koeficientů je však zdůrazňována již od roku 1995, kdy byl algoritmus *PSO* představen poprvé. Přístup k algoritmu *PSO* a zejména pak k jeho aplikacím byl zpočátku převážně evoluční a koeficienty (ω , a_l , a_g) byly stanovovány experimentálně jako ladící parametry vektoru aktualizace rychlosti částice (1). Analytický význam koeficientů začíná být patrný od roku 1999, při vyšetřování jejich vlivu na stabilitu algoritmu, neboť úplná absence koeficientů v rovnici (1) způsobovala nekontrolovaný růst rychlostí částic a tzv. “explozi” roje. Stabilita algoritmu je zkoumána metodami deterministické a stochastické analýzy chování systému částice, jejímž cílem je najít vzájemné vztahy mezi jednotlivými koeficienty a stanovit omezující podmínky pro jejich hodnoty tak, aby částice konvergovala k hledanému optimu. Algoritmus *PSO* je v kontextu stability vyšetřován jako dynamický systém.

Fyzikální pohled na pohyb částic v roji je obsahem kapitoly 4.1. Částice roje *PSO* jako dynamický systém je pojednána v kapitole 4.2. Vztahy mezi koeficienty (i) plynoucí z deterministické analýzy stability jsou odvozeny v kapitole 4.3, (ii) plynoucí ze stochastické analýzy stagnace nejlepší částice v kapitole 4.4, (iii) plynoucí ze stochastické analýzy standardní odchylky distribuce vzorků pozic stagnující částice *PSO* v kapitole 4.5 a (iv) plynoucí ze stability trajektorií částic zobecněného algoritmu *GPSO* v kapitole 4.6.

Poznátky jsou řazeny nikoli chronologicky, ale podle logické návaznosti, kdy předcházející poznátky vyvolávají otázky, na něž autorka pomocí dalších poznatků postupně hledá odpověď. Získané poznátky jsou podrobeny kritickému zhodnocení v kapitole 4.7.

Matematické vztahy, jimiž jsou jednotlivé poznatky prezentovány, jsou odvozeny v dodatcích D1 až D14. Důvodem detailního přístupu je skutečnost, že autoři studií příslušná odvození ve většině případů neuvádějí a stejně tak explicitně neuvádějí ani podmínky, za nichž lze původní matematický popis modelu systému částice nahradit popisem novým. Notace zápisu matematických vztahů je sjednocena, aby odpovídala jednotnému způsobu zápisu použitému v celém dokumentu.

4.1 Fyzikální pohled na pohyb částic v roji

Fyzikální pohled na pohyb částic v roji *PSO* [33:pp.13-38] umožňuje způsob modelování operace roje tak, že rovnice pro pohyb částice *PSO* (1), (2) jsou porovnávány s diskrétním tvarem rovnic pro pohyb částice v silovém poli. Jelikož hledání optima většiny známých fyzikálních procesů vede k nalezení pohybových rovnic, hlavní myšlenkou tohoto fyzikálního modelu z roku 2008, jehož autory jsou *Said M. Mikek* a *Abmed A. Kishk*, je zakomponovat stochastické prvky algoritmu *PSO* a inteligenci roje do Newtonova zákona síly, kdy částice na sebe budou vzájemně působit tak, že si budou vyměňovat informaci přes klasické silové pole.

Stručný úvod do problematiky fyzikálního tělesa pohybujícího se v konzervativním silovém poli přináší kapitola 4.1.1. Vzájemnému porovnání pohybových rovnic částice v konzervativním silovém poli s pohybovými rovnicemi částice *PSO* a fyzikálnímu významu koeficientu ω v rovnici (1) jsou věnovány kapitoly 4.1.2 až 4.1.4.

Částice v nekonzervativním silovém poli je obsahem kapitoly 4.1.5. Porovnání pohybových rovnic částice *PSO* s pohybovými rovnicemi částice ve *visko-elastickém* silovém poli je uvedeno v kapitole 4.1.6. Úplný ztrátový model *nabité* částice ve *visko-elastickém* silovém poli je pojednán v kapitole 4.1.7 a pro doplnění odvozen autorkou předkládané práce v dodatku D2. Další doplňující informace vztahující se ke kapitole 4.1 jsou obsaženy v dodatcích D1 a D2.

4.1.1 Pohybové rovnice tělesa v konzervativním silovém poli

Uvažujme v prostoru vektorové silové pole, tedy takové pole, kdy v každém místě prostoru působí na hmotný bod síla F (jednotka *Newton*, N , $kg.m.s^{-2}$) jednoznačně určená velikostí a směrem. Pro přesunutí hmotného bodu z místa A do místa B po dráze s_1 je třeba vykonat určitou práci W_1 (jednotka *Joule*, $kg.m^2.s^{-2}$). Předpokládejme nyní, že se hmotný bod přesune zpět z místa B do místa A po jiné dráze s_2 , přičemž se vykoná práce W_2 . Hmotný bod tak vykoná pohyb po uzavřené dráze. Pokud platí $W_1 = -W_2$, je celková práce po uzavřené dráze nulová. Body A a B byly zvoleny libovolně, a tedy práce nezávisí na dráze, kterou musí hmotný bod projít, ale pouze na počáteční a konečné poloze hmotného bodu. Taková silová pole se nazývají *konzervativní*, též *potenciálová*, nebo též *pole potenciálových sil*.

Fyzikální těleso nacházející se v konzervativním silovém poli má určitou potenciální energii, která je dána jeho pozicí v tomto poli. Na těleso působí síla, která má tendenci přemístit těleso na pozici s nejnižší hladinou potenciální energie. Potenciální energie tělesa se může změnit na kinetickou energii a obráceně. Kinetická energie tělesa je potom dána rozdílem potenciální energie mezi oběma pozicemi.

Těleso ve fyzice je modelováno jako hmotný bod nebo jako soustava vzájemně vázaných hmotných bodů. Hmotný bod je bezrozměrný útvar, kterému je přiřazena určitá hmotnost. V mechanice mluvíme o posuvném pohybu hmotného bodu podél určité dráhy. Jestliže těleso nepodléhá účinkům okolí, pak součet kinetické a potenciální energie hmotných bodů, z nichž se soustava skládá, zůstává stálý.

K popisu polohy hmotného bodu v prostoru slouží polohový vektor (3), což je spojnice počátku soustavy souřadnic a hmotného bodu s orientací k hmotnému bodu:

$$\mathbf{r} = (q_1, q_2, \dots, q_N, t) \quad (3)$$

kde q jsou zobecněné souřadnice a $n = 1 \dots N$ je počet stupňů volnosti systému. Pohyb hmotného bodu lze popsat změnou polohového vektoru v čase, dráha hmotného bodu představuje trajektorii pohybu.

Infinitesimální změna souřadnic hmotného bodu (3) v čase je vyjádřena jako:

$$\frac{d\mathbf{r}}{dt} = \frac{\partial \mathbf{r}}{\partial t} + \sum_{j=1}^N \frac{\partial \mathbf{r}}{\partial q_j} \frac{dq_j}{dt} \quad (4)$$

Infinitesimální změna souřadnic hmotného bodu, při níž nedochází ke změně času, se označuje jako virtuální posunutí a značí se $\delta \mathbf{r}$:

$$\delta \mathbf{r} = \left(\frac{d\mathbf{r}}{dt} \right)_{t=0} = \left(\frac{\partial \mathbf{r}}{\partial t} + \sum_{j=1}^N \frac{\partial \mathbf{r}}{\partial q_j} \frac{dq_j}{dt} \right)_{t=0} = 0 + \sum_{j=1}^N \frac{\partial \mathbf{r}}{\partial q_j} \delta q_j \quad (5)$$

Posunutí $\delta \mathbf{r}$ je označeno jako virtuální proto, že žádné skutečné posunutí nemůže beze změny času nastat. Protože (virtuální) čas je neměnný, je $\partial \mathbf{r} / \partial t = 0$, analogicky $dq_j / dt = \delta q_j$, a tedy:

$$\delta \mathbf{r} = \sum_{j=1}^N \frac{\partial \mathbf{r}}{\partial q_j} \delta q_j$$

Pohyb tělesa pod vlivem síly v závislosti na čase představuje fyzikální systém, jehož vývoj v čase je popsán fyzikální akcí (*action*) A . Akce (účinek) ve fyzice je měřitelná veličina. Je definována jako integrál *Lagrangeovy* funkce L mezi časy t_0 a t_1 , které odpovídají počátečnímu a konečnému stavu systému. *Lagrangeova* funkce vyjadřuje dynamický zákon, podle něhož se těleso pohybuje.

Jednotkou akce A je *Joule.second* ($\text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2} \cdot \text{s} \approx \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$). Jednotkou *Lagrangeovy* funkce L je *Joule* ($\text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2}$).

Akce je vyjádřena jako matematický funkcionál, jenž zpravidla zahrnuje určité integrály a obsahuje neznámé funkce a jejich derivace a jehož argumentem je trajektorie systému a výsledkem reálné číslo. Je zřejmé, že akce systému bude nabývat různých hodnot pro různé trajektorie. Klasická mechanika postulují, že trajektorie fyzikálního systému je ta, při níž je akce systému minimalizována, neboli ta z trajektorií, při níž je akce stacionární. Diferencovatelný funkcionál je stacionární ve svém lokálním maximu, minimu, nebo flexibilním bodě. Hledáme tedy externí funkce L , pro něž je daný funkcionál A stacionární, a sice s využitím variačního počtu, kdy extrémy mohou být

lokalizovány nalezením bodů, v nichž je derivace funkcionálu rovna nule. Aplikací uvedeného postupu, označovaného též jako *princip stacionární akce*, můžeme nalézt pohybové rovnice fyzikálního systému.¹⁸

Zvláštní význam ve fyzice mají funkcionály ve tvaru:

$$A = \int_{t_0}^{t_1} L(t, f(t, \mathbf{r}), f'(t, \mathbf{r}')) dt \quad (6)$$

Pro uvedený typ funkcionálu lze problém hledání nulové variace převést na řešení *Eulerovy-Lagrangeovy rovnice* podle postupu, který je podrobně uveden v dodatku D1, kde jsou obsaženy i rovnice (7) a (8) zde neuvedené:

$$\frac{\partial L}{\partial \mathbf{r}} - \frac{d}{dt} \frac{\partial L}{\partial \mathbf{r}'} = 0 \quad (9)$$

Vztah (9) se nazývá *Eulerova-Lagrangeova rovnice*, v níž L je *Lagrangeova* funkce mající konstantní hodnotu v čase, a tedy splňující podmínku $\delta L = 0$. Pokud známe *Lagrangeovu* funkci, můžeme aplikací principu stacionární akce získat pohybové rovnice fyzikálního systému. Jinými slovy, pokud známe dynamický zákon, podle něhož se těleso pohybuje (dynamiku pohybu), můžeme získat vztah mezi silami působícími na těleso. V konzervativním silovém poli je to zákon zachování mechanické energie, který stanoví, že rozdíl kinetické a potenciální energie tělesa zůstává konstantní.

4.1.2 Pohybové rovnice částice v konzervativním silovém poli

Uvažujeme nyní ve shodě s autory [33] částici pohybující se v konzervativním silovém poli a tuto částici pokládejme za hmotný bod. Pohybové rovnice částice v konzervativním silovém poli odvodili autoři [33] z podmínky, že akce systému A (jednotka *Joule.second*, $\text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2} \cdot \text{s} \approx \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$) musí být stacionární:

$$\delta A = \int_{t_0}^{t_1} \delta L(t, \mathbf{r}(t), \mathbf{v}(t)) dt = 0 \quad (10)$$

Zde t_0 je konstanta značící počáteční čas vývoje systému a t_1 je konstanta značící koncový čas vývoje systému. *Lagrangeovou funkci* pro daný případ i -té částice pohybující se v konzervativním silovém poli je rozdíl kinetické energie částice a potenciální energie konzervativní síly působící na tuto částici (jednotka *Joule* $\approx \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2}$). Takto definovaná *Lagrangeova funkce* má konstantní hodnotu, a tedy $\delta L = 0$:

$$L(t, \mathbf{r}^i, \mathbf{v}^i) = \frac{1}{2} m \sum_{n=1}^{n=N} \mathbf{v}_n^i \cdot \mathbf{v}_n^i - U(\mathbf{r}^i) \quad (11)$$

kde m je hmotnost částice, $n = 1 \dots N$ jsou souřadnice vektorového hyperprostoru $\mathbb{R}^N$, $\mathbf{r}^i(t) = (r_1^i, r_2^i, \dots, r_N^i, t)$ je polohový vektor i -té částice a $\mathbf{v}^i(t) = (v_1^i, v_2^i, \dots, v_N^i, t)$ je vektor rychlosti i -té částice a platí

¹⁸ Pohybovou rovnicí rozumíme vztah mezi silami, které působí na těleso. Pro dané těleso je celkem N pohybových rovnic, kde N je počet stupňů volnosti systému.

$$v_n^i = \frac{dr_n^i}{dt} \quad (12)$$

Pohybové rovnice částice v konzervativním silovém poli určili autoři [33] tak, že spočetli příslušné derivace *Lagrangeovy* funkce (11) a dosadili je do *Eulerovy-Lagrangeovy* rovnice (9), a to za podmínky, že hmotnost částice m je konstantní a nenulová:

$$\frac{\partial L}{\partial r_n^i} = \frac{\partial}{\partial r_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) \right) = - \frac{\partial U_n^i}{\partial r_n^i} \quad (13)$$

$$\frac{d}{dt} \frac{\partial L}{\partial v_n^i} = \frac{d}{dt} \frac{\partial L}{\partial v_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) \right) = \frac{d}{dt} m v_n^i = m \frac{d^2}{dt^2} r_n^i = m a_n^i \quad (14)$$

Dosazením (13) a (14) do (9) odvodili autoři [33] pohybovou rovnici i -té částice v n -té dimenzi v konzervativním silovém poli:

$$- \frac{\partial U_n^i}{\partial r_n^i} - m a_n^i = 0 \quad (15)$$

kde $F_n^i = m a_n^i = (-\partial U^i / \partial r_n^i)$ je síla (jednotka *Newton*, N, $kg \cdot m \cdot s^{-2}$) působící na i -tou částici v n -té dimenzi udělující výsledné zrychlení a_n^i (jednotka $m \cdot s^{-2}$).

4.1.3 Částice *PSO* v konzervativním silovém poli

Autoři [33] následně upravili rovnice (1) a (2) pro pohyb částice v algoritmu *PSO* tak, aby bylo možno porovnat s diskrétním tvarem rovnic pro pohyb částice v konzervativním silovém poli. Autorka předkládané disertační práce upřesňuje, že roj *PSO* bude obsahovat $i = 1 \dots M$ částic, kdy i -tá částice bude popsána polohovým vektorem:

$$\mathbf{r}^i(t) = (r_1^i, r_2^i, \dots, r_N^i, t) \quad (16)$$

a vektorem rychlosti:

$$\mathbf{v}^i(t) = (v_1^i, v_2^i, \dots, v_N^i, t) \quad (17)$$

Protože se částice pohybují ve vektorovém (hyper)prostoru $\mathbb{R}^N$ ($n = 1 \dots N$), každá částice bude popsána N -ticí reálných čísel představujících souřadnice v kartézském souřadnicovém systému, kde každá osa může mít jiné jednotky (tj. $kg, s, m, \dots$), a sice podle povahy řešeného problému. Autorka proto předpokládá normované zobrazení sledovaných veličin do příslušné osy tak, že jednotkou je normovaný vektor, jehož délka je rovna jedné délkové jednotce. Souřadnice vektoru pozice a vektoru rychlosti částice pokládá autorka za zobecněné souřadnice¹⁹ a vektor pozice částice $\mathbf{x}_n^i$ v rovnicích (1) a (2) symbolicky nahrazuje zobecněným polohovým vektorem $\mathbf{r}_n^i$.

¹⁹ Pojem “zobecněné souřadnice” (*generalized coordinates*) odlišuje tyto souřadnice od tradičně používaného pojmu “souřadnice”, jímž jsou míněny kartézské souřadnice. Souřadnice ve fyzice jsou totiž zpravidla voleny tak, aby vyjádření pohybových rovnic bylo matematicky co nejjednodušší. Kartézské souřadnice tedy též spadají pod pojem “zobecněné souřadnice”.

Autoři [33] dále zavedli náhodné proměnné $\varphi_{1n}^i = a_l u_{1n}^i$ a $\varphi_{2n}^i = a_g u_{2n}^i$, dosadili tyto proměnné do rovnice pro rychlost částice (1) a obdrželi upravenou rovnici pro rychlost i -té částice v n -té dimenzi:

$$v_n^i(t + \Delta t) = \omega v_n^i(t) + \varphi_{1n}^i [l_n^i(t) - r_n^i(t)] \Delta t + \varphi_{2n}^i [g_n(t) - r_n^i(t)] \Delta t \quad (18)$$

Dále zavedli proměnnou $o_n^i(t)$ jako vážený průměr proměnných $l_n^i(t)$ a $g_n(t)$

$$o_n^i(t) = \frac{\varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)}{\varphi_{1n}^i + \varphi_{2n}^i} \quad (19)$$

a náhodnou proměnnou φ_n^i jako součet proměnných φ_{1n}^i a φ_{2n}^i

$$\varphi_n^i = \varphi_{1n}^i + \varphi_{2n}^i \quad (20)$$

Proměnnými $o_n^i(t)$ a φ_n^i nahradili autoři [33] odpovídající proměnné v rovnici (18) a odvodili nový tvar rovnice pro rychlost částice v algoritmu *PSO*:²⁰

$$v_n^i(t + \Delta t) = \omega v_n^i(t) + \varphi_n^i [o_n^i(t) - r_n^i(t)] \Delta t \quad (21)$$

Autoři [33] nyní uvažovali, že hodnota koeficientu ω v rovnici (21) je rovna jedné a rovnici jednoduše upravili:

$$\frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} = \varphi_n^i [o_n^i(t) - r_n^i(t)] \quad (22)$$

Výraz na levé straně rovnice (22) vyjadřuje zrychlení a_n^i i -té částice v n -té dimenzi. Časový krok Δt zde má význam integračního časového kroku:

$$\frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} \approx a_n^i(t) \quad (23)$$

Připustíme-li ve shodě s autory [33], že částice *PSO* je hmotným bodem, a tedy $m \neq 0$, dostáváme pohybovou rovnici:

$$m \cdot a_n^i(t) = m \cdot \varphi_n^i [o_n^i(t) - r_n^i(t)] \quad (24)$$

Neboli:

$$F_n^i = m \cdot a_n^i(t) = m \cdot \varphi_n^i [o_n^i(t) - r_n^i(t)]$$

Porovnáním pohybové rovnice částice v konzervativním silovém poli (15) s rovnici pro pohyb částice *PSO* (21) získali autoři [33] pohybovou rovnici i -té částice *PSO* v n -té dimenzi za podmínky $\omega = 1$:

$$-\frac{\partial U_n^i}{\partial r_n^i} - m \cdot \varphi_n^i [o_n^i(t) - r_n^i(t)] = 0 \quad (25)$$

Protože se síly působící na částici v jejích jednotlivých dimenzích sčítají, vyjádřili autoři [33] výslednici sil jako:

²⁰ O ekvivalenci rovnic (21) a (18) se čtenář může nejnázne přesvědčit zpětným dosazením (19) a (20) do (18).

$$F^i = m \cdot a^i(t) = m \cdot \phi^i[o^i(t) - r^i(t)]$$

kde ϕ^i je diagonální matice s nenulovými prvky představující jednotlivé náhodné proměnné ϕ_n^i , a získali pohybovou rovnici i -té částice PSO:

$$-\frac{\partial U^i}{\partial r^i} - m \cdot \phi^i[o^i(t) - r^i(t)] = 0 \quad (26)$$

Částice PSO se autorům [33] jeví tak, jako by byla přitahována silou F^i , která je přímo úměrná posunutí částice z její aktuální pozice vzhledem k nějaké střední pozici o^i (19), která je dána součtem vážených průměrů všech pozic $l_n^i(t)$ s pozicí $g_n(t)$. Tato výsledná pozice o^i (19) je nelineární funkcí interakcí mezi částicemi. Proměnná U^i , jejíž hodnotu neznáme, představuje potenciální energii, mající charakter konzervativní elastické (deformační) potenciální energie podle *Hookova* zákona.²¹ Síla F^i je náhodná, a tedy výsledná trajektorie částice se jeví jako náhodná každému pozorovateli vně systému, avšak (Newtonův) dynamický zákon pohybu částice je deterministický.

4.1.4 Fyzikální význam ladícího parametru ω

Aby autoři [33] zjistili význam ladícího parametru ω v rovnicích (1), (18), (21), o němž je z výsledků experimentů známo, že nabývá hodnot menších než jedna a že je funkcí času,²² provedli úpravu rovnice (21) pro rychlost částice v algoritmu PSO a zahrnuli do úprav i koeficient $\omega(t)$:

$$\begin{aligned} v_n^i(t + \Delta t) &= \omega(t) \cdot v_n^i(t) + \phi_n^i[o_n^i(t) - r_n^i(t)]\Delta t \\ v_n^i(t + \Delta t) - v_n^i(t) &= \omega(t) \cdot v_n^i(t) + \phi_n^i[o_n^i(t) - r_n^i(t)]\Delta t - v_n^i(t) \\ \frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} &= \frac{\omega(t) - 1}{\Delta t} v_n^i(t) + \phi_n^i[o_n^i(t) - r_n^i(t)] \end{aligned}$$

Výraz na levé straně upravené rovnice vyjadřuje zrychlení a_n^i n -té částice PSO v i -té dimenzi:

$$\frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} \approx a_n^i(t)$$

Zavedeme-li po vzoru autorů [33] koeficient β

$$\beta = \frac{\omega(t) - 1}{\Delta t} \quad (27)$$

²¹ Elastická potenciální energie U (jednotka *Joule*) je potenciální energie tělesa, které je namáháno (*stressed*). Vzniká jako důsledek reakční síly F (jednotka *Newton*, N), která se snaží vrátit objekt do původního tvaru. Často je to elektromagnetická síla mezi molekulami (částicemi), z nichž je těleso tvořeno. Jestliže je napětí uvolněno, elastická potenciální energie je transformována na kinetickou energii. *Hookův* zákon definuje reakční elastickou sílu $F = (-k\Delta r)$ nebo působící sílu $F^i(r) = (k\Delta r)$. V případě částice PSO se jedná o působící sílu, kde vektor $\Delta r = o^i(t) - r^i(t)$, $k = m\phi$; k se nazývá konstanta tuhosti (*spring constant*).

²² Čas se při simulaci algoritmu PSO neměří v sekundách, ale v počtu iterací, kdy jedna iterace odpovídá Δt .

a připustíme-li, že částice *PSO* je hmotným bodem, a tedy $m \neq 0$, dostáváme výslednou pohybovou rovnici i -té částice v n -té dimenzi, kdy $\omega \neq 1$:

$$m \cdot \mathbf{a}_n^i(t) = m \cdot \beta \mathbf{v}_n^i(t) + m \cdot \varphi_n^i [\mathbf{o}_n^i(t) - \mathbf{r}_n^i(t)] \quad (28)$$

Protože se síly působící na částici v jejích jednotlivých dimenzích sčítají, autoři [33] vyjádřili výslednici sil jako:

$$\mathbf{F}^i = m \cdot \mathbf{a}^i(t) = m \cdot \beta \mathbf{v}^i(t) + m \cdot \boldsymbol{\phi}^i [\mathbf{o}^i(t) - \mathbf{r}^i(t)]$$

kde $\boldsymbol{\phi}^i$ je diagonální matice s nenulovými prvky představující jednotlivé náhodné proměnné φ_n^i , a získali pohybovou rovnici i -té částice *PSO*:

$$\begin{aligned} \mathbf{F}^i = m \cdot \mathbf{a}^i(t) &= -\frac{\partial U^i}{\partial \mathbf{r}^i} = m \cdot \beta \mathbf{v}^i(t) + m \cdot \boldsymbol{\phi}^i [\mathbf{o}^i(t) - \mathbf{r}^i(t)] \\ -\frac{\partial U^i}{\partial \mathbf{r}^i} - m \cdot \beta \mathbf{v}^i(t) - m \cdot \boldsymbol{\phi}^i [\mathbf{o}^i(t) - \mathbf{r}^i(t)] &= 0 \end{aligned} \quad (29)$$

Autoři [33] poukazují, že rychlost částice *PSO* je omezena koeficientem β (27). Experimenty s variantami algoritmu *PSO* ukazují, že koeficient ω jako ladící parametr má v průběhu času nabývat klesajících hodnot 0,9 až 0,2. Uvážíme-li, že v experimentech $\Delta t = 1$, a tedy $\beta = \omega(t) - 1$, znamená to, že koeficient β je v praxi záporný a má proto tendenci zpomalovat rychlost částice. Koeficient $\omega(t)$ jako ladící parametr v rovnicích (1), (18), (21) se tudíž podle autorů [33] spíše než k setrvačnosti vztahuje k (vnitřnímu) tření či viskozitě²³ uvnitř roje tak, že koeficient $\beta = \omega(t) - 1$ nabývající hodnot -0,1 až -0,8 je tzv. tlumící faktor (*damping factor*), což značí, že prostředí *PSO* je ve skutečnosti nekonzervativní, disipativní,²⁴ a nelze jej popsat *Lagrangeovou* funkcí pro konzervativní systém ve tvaru (11).

4.1.5 Částice v nekonzervativním silovém poli

Eulerova-Lagrangeova rovnice (9) v podobě, jak byla odvozena v dodatku D1, platí pro systémy, v nichž se energie nepřeměňuje nevratně. Autorka předkládané práce si ve snaze o sestavení úplné pohybové rovnice částice v nekonzervativním silovém poli a porovnání této rovnice s rovnicí (29) položila otázku, zda lze formulovat *princip stacionární akce* též pro systémy, v nichž dochází k energetickým ztrátám. Tuto otázku autoři [33] neřeší.

Problém spočívá v působení nekonzervativní (disipativní) síly F_d (jednotka *Newton*, *N*) odpovídající mechanické práci, která přemění kinetickou energii částice na jiný druh (ztrátové) energie. Tato síla by měla být včleněna do *Eulerovy-Lagrangeovy* pohybové rovnice (9):

²³ Viskozita neboli vnitřní tření charakterizuje odpor, který klade medium vlastnímu pohybu (toku). Odpovídá síle potřebné v médiu k posunu jednoho kilogramu o jeden metr za jednu sekundu (jednotka *Pascal.sekunda* $\text{Pa.s} = \text{kg.m}^{-1}.\text{s}^{-1}$). Jedná se o zjednodušený model, kdy síla je přímo úměrná rychlosti pohybu. Se zvyšováním teploty viskozita klesá, a to mnohem rychleji než hustota.

²⁴ Nekonzervativní síly jsou síly, jejichž práce na uzavřené křivce je nenulová. Při jejich působení dochází k rozptýlení (disipaci) energie. Příkladem nekonzervativní síly je třecí síla – na uzavřené křivce vykoná nenulovou práci. Příkladem konzervativní síly je gravitační síla – na uzavřené křivce vykoná nulovou práci.

$$\frac{\partial L}{\partial \mathbf{r}} - \frac{d}{dt} \frac{\partial L}{\partial \mathbf{r}'} + F_d = 0 \quad (30)$$

Pro řešení problému působení nekonzervativní síly volila autorka postup analogický se studií [34] uveřejněnou v digitální knihovně pro fyziku a astronomii v lednu 2012 pod názvem “*Is it possible to formulate Least Action Principle for Dissipative Systems?*”, jejímiž autory jsou *Qinping A. Wang* a *Ru Wang*.

V případě pohybové rovnice (30) je ve shodě s autory [34] nezbytné vzít v úvahu dva systémy, a sice (i) pohybující se částici jako první systém a (ii) prostředí, v němž se částice pohybuje, jako druhý systém, propojené navzájem disipativní silou F_d . Oba dva systémy je třeba pokládat za jeden konzervativní celek, pro nějž již bude možno *princip stacionární akce* uplatnit. Autorka po vzoru studie [34] předpokládá, že energetický transfer z prvního systému do druhého systému probíhá výhradně přes disipativní sílu F_d . Celková energie systému H (jednotka *Joule*) potom bude součtem kinetické energie K částice, potenciální energie U konzervativní síly působící na částici a energie H_2 systému okolí částice.

Autorka rovněž tak předpokládá, že okolí částice se pohybuje po téže trajektorii $s = \mathbf{r}(t)$ jako částice (tj. částice se pohybuje spolu se svým okolím). Částice se pohybuje z bodu $\mathbf{r}(t_0)$ do bodu $\mathbf{r}(t_1)$ během časového intervalu $(t_1 - t_0)$. Množství energie E_d rozptýlené z prvního do druhého systému v časovém okamžiku t , $(t_0 \leq t \leq t_1)$ je dáno negativní prací disipativní síly:

$$E_d(\mathbf{r}(t_0, t)) = - \int_{\mathbf{r}(t_0)}^{\mathbf{r}(t)} F_d \cdot d\mathbf{s} = - \int_{\mathbf{r}(t_0)}^{\mathbf{r}(t)} F_d \cdot d\mathbf{r} = - \int_{t_0}^t F_d \cdot \mathbf{r}'(\tau) d\tau \quad (31)$$

kde t je libovolný časový okamžik mezi t_0 a t_1 a $d\mathbf{s} = d\mathbf{r} = \mathbf{r}' d\tau$ je malé posunutí podél trajektorie $s = \mathbf{r}(\tau)$. Zápis $E_d(\mathbf{r}(t_0, t))$ vyjadřuje, že rozptýlená energie závisí na celé trajektorii od počátku $\mathbf{r}(t_0)$ až po okamžitou polohu $\mathbf{r}(t)$. Okamžitý přírůstek energie E_d ve druhém systému je kompenzován souběžným úbytkem $K + U$ v prvním systému, což zaručuje konstantní hodnotu energie celkového systému $H = K + U - E_d + H_2 + E_d$.

Rozptýlená energie E_d způsobí změnu kinetické a potenciální energie všech ostatních částic v prostředí druhého systému, avšak vzhledem k platnosti zákona zachování energie může být suma těchto změn energií částic vyjádřena prací disipativní síly F_d .

Disipativní sílu lze vyvodit z (31) jako: $F_d = -\partial E_d / \partial \mathbf{r}$, a tedy disipativní síla se odvozuje z rozptýlené energie, podobně jako se konzervativní síla odvozuje z potenciální energie: $F_c = -\partial U / \partial \mathbf{r}$ (13). Rozptýlená energie není ovšem potenciální energií, neboť nezávisí pouze na $\mathbf{r}(t)$, ale také na předchozí trajektorii, podél níž byl proveden integrál (31). Je navíc nemožné obnovit rozptýlenou energii E_d , která narůstá s časem, tak, jako je možné obnovit potenciální energii částice pouhým návratem zpět do výchozí pozice.

Přesto má rozptýlená energie E_d dle autorů [34] společné znaky s potenciální energií: její infinitesimální navýšení dE_d v čase t je rovno negativní práci vykonané disipativní silou F_d na infinitesimálním posunutí $d\mathbf{r}$. Tato skutečnost je dle autorů [34] dostačující pro to, abychom mohli napsat *Lagrangeovu* funkci pro disipativní systém ve tvaru $L = K - U - E_d$. Fyzikální akce celkového systému na dráze mezi body $\mathbf{r}(t_0)$ a $\mathbf{r}(t_1)$ je potom:

$$A = \int_{t_0}^{t_1} L dt = \int_{t_0}^{t_1} (K - U - E_d) dt = \int_{t_0}^{t_1} \left(K - U - \int_{r(t_0)}^{r(t)} F_d \cdot d\mathbf{r} \right) dt \quad (32)$$

Potud aplikace studie [34]. Autorka předkládané práce má nyní aparát k odvození pohybové rovnice částice v nekonzervativním *visko-elastickém silovém poli* z podmínky, že akce A (32) (jednotka *Joule.second*, $\text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2} \cdot \text{s} \approx \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$), musí být stacionární, to jest z podmínky, že $\delta A = 0$. Pohybové rovnice částice určí autorka analogicky jako pohybové rovnice částice v konzervativním silovém poli, ale použije *Lagrangeovu* funkci zahrnující energetické ztráty s tím, že pro určení rozptýlené energie E_d potřebuje znát vztah pro disipativní sílu F_d . Autorka vychází ze znalosti, že vztahy pro disipativní síly se ve fyzice určují empiricky. Disipativní síla bývá úměrná velikosti rychlosti částice a má opačný směr než pohyb částice.

4.1.6 Pohybové rovnice částice v nekonzervativním *visko-elastickém silovém poli*

Jelikož je prostředí roje modelováno jako visko-elastické, a tedy nekonzervativní, pro získání pohybových rovnic částice hledá autorka předkládané práce modifikovanou *Lagrangeovu* funkci L obsaženou v (32), zahrnující energetické ztráty.

Autorka uvažuje ztrátový model v non-relativistickém pojetí²⁵ a předjímá, že celková síla F^i působící na částici i o hmotnosti m udělující částici výsledné zrychlení $\mathbf{a}^i$ je složena ze síly vnitřního tření (viskozity) F_s^i a konzervativní elastické síly F_c^i :

$$F^i = m\mathbf{a}^i = F_s^i + F_c^i$$

kde $F_c^i = -\partial U^i / \partial \mathbf{r}^i$ je konzervativní složka síly podle (13) a F_s^i je ztrátová složka síly vnitřního tření. Viskozita neboli vnitřní tření charakterizuje odpor, který klade medium vlastnímu pohybu (toku). Odpovídá síle potřebné v médiu k posunu jednoho kilogramu hmoty o jeden metr za jednu sekundu (jednotka *Pascal.sekunda*, $\text{Pa} \cdot \text{s} = \text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-1}$). Jedná se o zjednodušený model, kdy síla je přímo úměrná rychlosti pohybu. Síla vnitřního tření F_s^i se určí podle *Stokesova* zákona jako síla na překonání odporu části prostředí o hmotnosti $\mathcal{M}$, které bylo částicí uvedeno do pohybu:²⁶

$$F_s^i = -\mu \mathcal{M} \mathbf{v}^i \quad (33)$$

Autorka nejdříve spočetla rozptýlenou energii E_s^i podle (31), kam dosadila (33):

$$E_s^i = - \int_{r(t_0)}^{r(t)} F_s^i \cdot d\mathbf{r}^i = \mu \mathcal{M} \int_{r(t_0)}^{r(t)} \mathbf{v}^i \cdot d\mathbf{r}^i \quad (34)$$

Eulerovu-Lagrangeovu pohybovou rovnici (30) nyní autorka formálně odvodí tak, že spočte příslušné derivace *Lagrangeovy* funkce obsažené v (32), a to za podmínky, že

²⁵ Pojetí před teorií relativity a kvantovou fyzikou, pro rychlosti mnohem menší než je rychlost světla.

²⁶ Záleží rovněž na tvaru tělesa. *Stokesův zákon* je odvozen pro těleso ve tvaru koule. Částice ovšem žádný tvar nemá, nemá ani objem, má pouze hmotnost, neboť je pokládána za hmotný bod.

hmotnost m částice je konstantní a nenulová, a dosadí tyto derivace do *Eulerovy-Lagrangeovy* rovnice (9):

$$\begin{aligned}\frac{\partial L}{\partial r_n^i} &= \frac{\partial}{\partial r_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) - E_{sn}^i(r_n^i) \right) = -\frac{\partial U_n^i}{\partial r_n^i} - \frac{\partial E_s^i}{\partial r_n^i} \\ \frac{\partial L}{\partial r_n^i} &= -\frac{\partial U_n^i}{\partial r_n^i} - \mu \mathcal{M} v_n^i\end{aligned}\quad (35)$$

$$\begin{aligned}\frac{d}{dt} \frac{\partial L}{\partial v_n^i} &= \frac{d}{dt} \frac{\partial}{\partial v_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) - E_{sn}^i(r_n^i) \right) \\ \frac{d}{dt} \frac{\partial L}{\partial v_n^i} &= \frac{d}{dt} m v_n^i = m \frac{d^2}{dt^2} r_n^i = m a_n^i\end{aligned}\quad (36)$$

Dosazením (35) a (36) do (9) odvodila autorka pohybové rovnice částice v nekonzervativním *visko-elastickém* silovém poli jako vztah mezi silami působícími na i -tou částici v n -té dimenzi (jednotka *Newton*, N, $\text{kg} \cdot \text{m} \cdot \text{s}^{-2}$):

$$-\frac{\partial U_n^i}{\partial r_n^i} - \mu \mathcal{M} v_n^i - m a_n^i = 0 \quad (37)$$

kde $F_{cn}^i = -\partial U_n^i / \partial r_n^i$ je konzervativní síla pole potenciálních sil, $F_{sn}^i = -\mu \mathcal{M} v_n^i$ je síla vnitřního tření a $F_n^i = m a_n^i$ je síla udělující částici zrychlení a_n^i (jednotka $\text{m} \cdot \text{s}^{-2}$).

Porovnáním pohybové rovnice (37) s pohybovou rovnicí i -té částice *PSO* (29), kterou zde autorka pro přehlednost znovu uvádí:

$$-\frac{\partial U^i}{\partial \mathbf{r}^i} - m \cdot \beta \mathbf{v}^i(t) - m \cdot \phi^i[\mathbf{o}^i(t) - \mathbf{r}^i(t)] = 0$$

dospívá autorka k závěru, že v modelu *PSO* skutečně dochází ke ztrátám, které by bylo možno ve shodě s autory [33] interpretovat jako ztráty třením podle *Stokesova zákona*, kdy síla tření je úměrná velikosti rychlosti s koeficientem vnitřního tření $m\beta = \mu\mathcal{M}$. Autorka však současně upozorňuje, že *Stokesův zákon* je odvozen pro těleso ve tvaru koule. Částice ovšem žádný tvar nemá, nemá ani objem, má pouze hmotnost, neboť je pokládána za hmotný bod, kde hmotnost částice můžeme dle autorů [33] zvolit libovolně.

Další okolností, kterou studie [33] neřeší, je závislost koeficientu β (27) na časovém okamžiku Δt .

4.1.7 Pohybové rovnice nabitě částice v nekonzervativním *visko-elastickém* silovém poli

K energetickým ztrátám v nekonzervativním silovém poli může docházet rovněž radiací. Ztráty radiací mohou být buď (i) v podobě infračerveného záření, kdy množství radiace závisí na teplotě tělesa (roje), anebo (ii) pokud by částice byly elektricky nabitě, každé nenulové zrychlení částice by způsobilo vyzařování elektromagnetických vln, které by způsobily ztrátu části mechanické energie částice.

Autoři [33] proto porovnávají pohybovou rovnici částice *PSO* v nekonzervativním silovém poli (29) s pohybovou rovnicí úplného ztrátového modelu částice a konstatují, že částice v roji elektromagneticky nevyzařují. Autoři [33] navrhuji, že by model částice *PSO* mohl být o ztráty radiací rozšířen. Autorka předkládané práce se proto pokusila o sestavení úplné pohybové rovnice částice v nekonzervativním silovém poli zahrnující i ztráty radiací, kterou autoři [33] uvádějí, ale jejíž odvození chybí.

K odvození úplného ztrátového modelu autorka předpokládala, že periodicky pohybující se částice nese elektrický náboj, a visko-elastický model sil působících na částici rozšířila o tzv. reakční sílu radiačního vyzařování (*radiation reaction force*) F_r^i

$$F^i = m\mathbf{a}^i = F_r^i + F_s^i + F_c^i$$

Celý postup odvození úplného ztrátového modelu je uveden v dodatku D2.

Jako první krok zde autorka odvodila obecný vzorec pro sílu F_r^i odpovídající ztrátám vyzařováním, který je znám jako *Abrahamova-Lorentzova formule* pro reakční sílu radiačního vyzařování:

$$F_r = \frac{2q^2}{3c^3} \frac{d^2\mathbf{v}(t)}{dt^2} = \frac{2q^2}{3c^3} \frac{d\mathbf{a}(t)}{dt} = \gamma \frac{d\mathbf{a}(t)}{dt}$$

kde c je rychlost světla ve vakuu (jednotka $m.s^{-1}$), q je náboj nesený částicí (jednotka *Coulomb*), $\mathbf{a}(t)$ je zrychlení částice (jednotka $m.s^{-2}$).

Koeficient $2q^2/3c^3$ označila autorka jako γ . Veličina $J = d\mathbf{a}(t)/dt$ se nazývá *jerk* (jednotka $m.s^{-3}$) a je rovna časové změně zrychlení. K časové změně zrychlení může dojít pouze při změně velikosti síly F způsobující pohyb náboje. Je-li tato síla konstantní, konstantní bude rovněž zrychlení náboje, a tedy $J = 0$.

Odvození vzorce pro sílu radiačního vyzařování v D2 zahrnuje rovnice (38)-(41).

Jelikož reakční síla radiačního vyzařování F_r je funkcí gradientu zrychlení částice $d\mathbf{a}/dt$, bylo třeba nejdříve odvodit modifikovanou *Eulerovu-Lagrangeovu* rovnici pro příslušnou *Lagrangeovu* funkci L_r , kdy L_r je nejen funkcí času, polohy a rychlosti, ale i zrychlení:

$$\frac{\partial L_r}{\partial \mathbf{r}} - \frac{d}{dt} \frac{\partial L_r}{\partial \mathbf{r}'} + \frac{d^2}{dt^2} \frac{\partial L_r}{\partial \mathbf{r}''} = 0$$

Modifikovaná *Eulerova-Lagrangeova* rovnice je odvozena v dodatku D3. Odvození zahrnuje rovnice (42)-(45).

Následně autorka spočetla jednotlivé členy modifikované *Eulerovy-Lagrangeovy* rovnice a sestavila pohybovou rovnici i -té částice v n -té dimenzi nesoucí náboj ve visko-elastickém nekonzervativním poli:

$$-\frac{\partial U_n^i}{\partial r_n^i} - \mu \mathcal{M} v_n^i - m a_n^i + \gamma \frac{d}{dt} a_n^i = 0$$

kde $F_r = \gamma * d\mathbf{a}_n^i/dt$ je reakční síla radiačního vyzařování, kterou autorka předkládané disertační práce ztotožňuje s *Abrahamovou-Lorentzovou* formulí. Úplný postup sestavení je obsažen v dodatku D2. Sestavení zahrnuje rovnice (46)-(49).

Protože se síly působící na částici v jejich jednotlivých dimenzích sčítají, vyjádřila autorka výslednici sil jako

$$-\frac{\partial U^i}{\partial \mathbf{r}^i} - \mu \mathcal{M} \mathbf{v}^i - m \mathbf{a}^i + \gamma \frac{d}{dt} \mathbf{a}^i = 0 \quad (50)$$

a výslednou pohybovou rovnicí i -té částice (50) následně porovnála s pohybovou rovnicí i -té částice *PSO* (29). Poslední člen rovnice (50) v rovnici (29) evidentně chybí, a tedy koeficient ztrát radiací γ v modelu *PSO* je roven nule. Znamená to, že v roji dochází k tepelným ztrátám a roj se postupně “ochlazuje” tím, jak částice postupně ztrácejí rychlost vlivem vnitřního tření, avšak že částice v roji *PSO* elektromagneticky nevyzařují.

Bude-li koeficient ($\beta = (\omega(t) - 1)/\Delta t$) (27) v rovnici (29) roven nule, a sice pro $\omega(t) = 1$, dostáváme pohybovou rovnici částice v konzervativním silovém poli (25).

Celý postup odvození úplného ztrátového modelu je uveden v dodatku D2, odvození modifikované *Eulerovy-Lagrangeovy* rovnice je obsaženo v dodatku D3.

4.1.8 Otázky k zodpovězení

Autorka předkládané disertační práce se nepřiklání k návrhu modelu částice *PSO* rozšířeného o ztráty radiací a po provedené analýze upozorňuje na následující úskalí spojená s tímto návrhem:

(i) Mechanismus reakční síly radiačního vyzařování není ve fyzice dosud jednoznačně objasněn. Úplný ztrátový model (50) předpokládá periodický (buzený) pohyb nabitě částice, která tudíž spolu se svým okolím musí tvořit konzervativní celek. Reakční síla radiačního vyzařování $F_r = \gamma (d\mathbf{a}^i/dt)$ zde tedy musí působit jako budící síla (*drive force*) proti odporové disipativní síle $F_d = -m \cdot \beta \mathbf{v}^i(t)$. Pro rovnoměrně zrychlenou částici bude reakční síla radiačního vyzařování rovna nule, a tedy rovnoměrně zrychlená částice nebude vykazovat žádnou radiační reakci, což je v rozporu s tvrzením, že nabitá částice vždy vyzařuje elektromagnetické vlnění. Vysvětlení se dotýká fyzikálního významu zrychlení částice a není dosud ve fyzice zcela jednoznačné (tj. existuje několik možných vysvětlení).

(ii) Předpoklad periodického pohybu částice je splněn, neboť částici *PSO* lze analyzovat jako diskrétní případ tlumeného harmonického oscilátoru. Lze nahlédnout, že pohybová rovnice částice v nekonzervativním visko-elastickém silovém poli (29) je zápisem pohybové rovnice kmitající soustavy, a sice tlumeného (*damped*) harmonického oscilátoru:

$$\frac{\partial U}{\partial \mathbf{r}} = m \mathbf{a} = -b \frac{d\mathbf{r}}{dt} - k \Delta \mathbf{r}$$

kde disipativní síla tření způsobující ztráty je úměrná velikosti rychlosti s tlumícím faktorem $b = m\beta$. Posunutí z rovnovážné polohy $\Delta \mathbf{r} = [\mathbf{o}^i(t) - \mathbf{r}^i(t)]$, kde $k = m \cdot \phi^i$ je konstanta tuhosti (*spring constant*). Z rovnice (50) je patrné, že reakční síla radiačního vyzařování má působit *proti* ostatním silám. Bylo by třeba specifikovat, jakou veličinu ve vztahu k *sociálnímu simulátoru* [1] by měla tato síla modelovat.

(iii) Pohybová rovnice tlumeného oscilátoru je lineární, avšak proměnná $\mathbf{o}^i(t)$ (19) je nelineární funkcí, což může náhled na předpokládané chování systému částice zcela měnit. Zatímco pohybová rovnice částice (29) představuje diferenční rovnici druhého řádu, rovnice (50) je diferenční rovnicí třetího řádu, a bylo by tudíž třeba pečlivě posoudit, zda by takový komplikovaný krok vedl ke *zvýšení inteligence skupiny*.

4.2 Částice roje *PSO* jako dynamický systém

Dynamický systém je obecně definován pomocí *dynamických podmínek*, které popisují změnu chování tohoto systému v čase. *Stav systému* v libovolném časovém okamžiku je potom popsán vektorem ve stavovém prostoru (stavovým vektorem). Dynamické podmínky jsou zadány soustavou diferenciálních rovnic (spojitý případ) nebo diferenčních rovnic (diskrétní případ), které popisují změnu stavového vektoru v čase. Změna stavu dynamického systému se děje provedením těchto rovnic a nahrazením stávající hodnoty stavového vektoru hodnotou novou.

Fyzikální model obsažený v kapitole 4.1 interpretuje systém částice *PSO* jako dynamický systém, jenž je popsán polohovým vektorem $r_n^i(t)$ (16), vektorem rychlosti $v_n^i(t)$ (17) a vektorem zrychlení $a_n^i(t)$ (23) a jehož dynamické podmínky jsou zadány soustavou diferenciálních rovnic (28), které lze převést na diskrétní tvar.

Dynamický systém může být deterministický (chaotický) nebo náhodný, stochastický. Částice *PSO* je stochastický diskrétní dynamický systém, neboť rovnice (28) popisující změnu stavového vektoru v čase obsahují náhodné proměnné. V počátečních studiích [35], [36] je však částice analyzována primárně jako deterministický systém, a to za předpokladu, že náhodné proměnné u_{1n}^i , u_{2n}^i v rovnicích (1), (18) a (19) nabývají očekávaných hodnot $E[u_{1n}^i] = E[u_{2n}^i] = 0,5$. Ze stochastických rovnic (28) tak vzniknou obyčejné diferenční rovnice popisující chování částice, které umožňují matematickou analýzu dynamického chování částice v závislosti na hodnotě koeficientů ω, φ_n^i (a případně na hodnotě časového kroku Δt). Dynamické chování algoritmu *PSO* je analyzováno využitím teorie lineárních diskrétních dynamických systémů a teorie regulace s cílem dosažení systémové stability roje tím, že (stochastické) chování částic bude (částečně) usměrněno.

Pro deterministický typ matematické analýzy je využívána řada různých popisů systému částice *PSO*. Syntézou poznatků vydedukovaných z deterministických analýz [35], [36] sestavila autorka předkládané disertační práce ucelený přehled těchto popisů s cílem ukázat, že všechny tyto popisy lze odvodit ze soustavy diferenčních rovnic (28), a dále, za jakých podmínek jsou všechny tyto popisy vzájemně ekvivalentní. Autorka rovněž na základě vlastní dedukce usoudila, že analytikové pokládají systém částice *PSO* za *deterministický lineární systém* (tlumený harmonický oscilátor).

Interpretace rovnic (28) pro deterministický typ matematické analýzy je obsahem kapitoly 4.2.1. a slouží jako úvod pro navazující kapitolu 4.2.2, v níž autorka charakterizuje částici *PSO* jako deterministický *dynamický* systém. Částice *PSO* jako dynamický systém popsán vektorem rychlosti uvádí kapitola 4.2.3. Popis částice *PSO* jako jednoduchého lineárního dynamického systému je v kapitole 4.2.4.

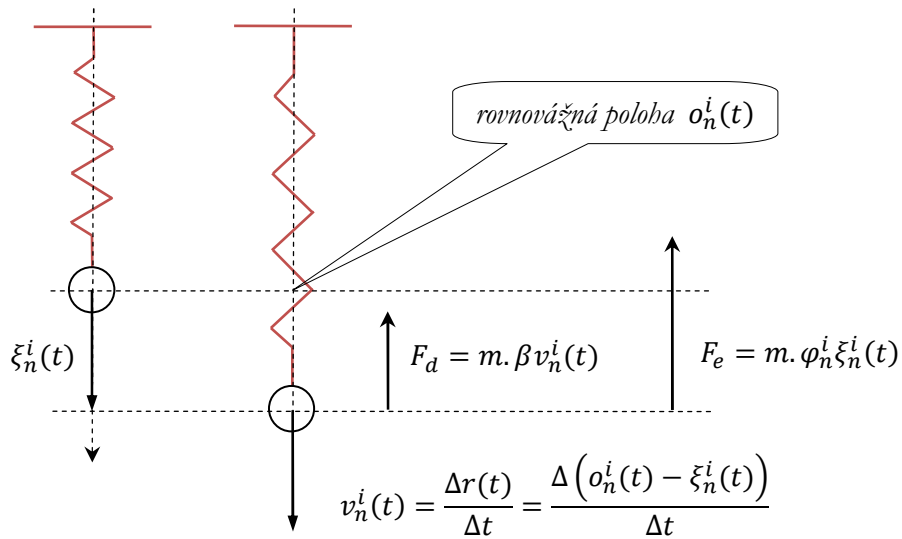
Autorka se shoduje s autory [35], [36] v názoru, že každá složka (rozměr n) vektoru rychlosti částice je aktualizována nezávisle na ostatních složkách, jak je zřejmé z rovnice (1), a že jedinou spojnici mezi jednotlivými složkami jsou hodnoty globální nejlepší pozice a lokálních nejlepších pozic odvozené z hodnot účelové funkce, reprezentované hodnotou proměnných $o_n^i(t)$ v rovnicích (28). Autorka stejně tak usuzuje, že pro účely matematické analýzy nebude na úkor obecnosti, pokud bude analyzována pouze jedna rovnice (28) pro jeden libovolný rozměr n libovolné částice i .

4.2.1 Částice *PSO* jako diskrétní případ tlumeného harmonického oscilátoru

Skutečnost, že fyzikální model algoritmu *PSO* popsany v kapitole 4.1 interpretuje částici jako těleso, na něž působí elastická síla a odporující síla, vedla analytiku k interpretaci částice *PSO* jako (diskrétního) případu tlumeného harmonického oscilátoru.

Tlumený harmonický oscilátor definuje autorka klasicky jako hmotný bod připojený k elastickému objektu zanedbatelné hmotnosti, který je na opačném konci fixován. Pohyb hmotného bodu je tak vázán na přímku, podél které se může hmotný bod pohybovat. Položíme-li do této přímky souřadnicovou osu y s počátkem v místě, k němuž směřuje elastická síla F_e a odporující síla F_d , je tento hmotný bod vázán na rovnovážnou polohu, od které se při svém pohybu vzdaluje do konečné vzdálenosti $\xi(t)$.

Analogicky s uvedenou definicí tlumeného harmonického oscilátoru působí na i -tou částici *PSO* v její n -té dimenzi odporující síla $F_d = m \cdot \beta v_n^i(t)$ a elastická síla $F_e = m \cdot \varphi_n^i[o_n^i(t) - r_n^i(t)]$ ve shodě s pohybovou rovnicí (28). Rovnovážnou polohou je pozice $o_n^i(t)$ podle (19) a vzdálenost $\xi_n^i(t) = o_n^i(t) - r_n^i(t)$, jak ukazuje obrázek Obr. 6:



Obr. 6: Částice *PSO* jako tlumený harmonický oscilátor.

Hmotný bod připojený k elastickému objektu zanedbatelné hmotnosti.

Na i -tou částici *PSO* v její n -té dimenzi působí odporující síla F_d a elastická síla F_e .

Rychlost $v_n^i(t)$ i zrychlení $a_n^i(t)$ lze vyjádřit jako funkce výchylky z $o_n^i(t)$.

Rovnici (28), v níž autorka vykrátí nenulovou hmotnost částice m , lze formálně přepsat do tvaru diferenční rovnice druhého řádu, kde $t = k\Delta t$ je čas v k -té iteraci, $k = 1, 2, \dots, K$, K je počet iterací:

$$a_n^i(k\Delta t) - \beta v_n^i(k\Delta t) - \varphi_n^i[o_n^i(k\Delta t) - r_n^i(k\Delta t)] = 0 \quad (51)$$

Do rovnice (51) dosadila autorka zpětně za $\beta = (\omega - 1)/\Delta t$ podle (27), zrychlení $a_n^i(k\Delta t)$ vyjádřila jako přírůstek rychlosti $v_n^i(k\Delta t)$ podle (23):

$$\frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} - \frac{\omega - 1}{\Delta t} v_n^i(t) - \varphi_n^i o_n^i(t) + \varphi_n^i r_n^i(t) = 0$$

a odvodila rovnici charakterizující částici *PSO* jako dynamický systém popsany polohovým vektorem $r_n^i(t)$ (16):

$$r_n^i(t + \Delta t) + (\varphi_n^i \Delta^2 t - \omega - 1) r_n^i(t) + \omega r_n^i(t - \Delta t) = \varphi_n^i o_n^i(t) \Delta^2 t \quad (52)$$

Detailní postup odvození rovnice (52) je obsažen v dodatku D4.

Za předpokladu, že časový krok Δt je implicitně roven jedné, přejde rovnice (52) k obvyklému zápisu diferenční rovnice druhého řádu

$$r_n^i(t + 1) + (\varphi_n^i - (\omega + 1)) r_n^i(t) + \omega r_n^i(t - 1) = \varphi_n^i o_n^i(t) \quad (53)$$

Lze nahlédnout, že koeficienty diferenční rovnice (53) nejsou vzájemně nezávislé. Nenulová pravá strana činí z rovnice (53) rovnici nehomogenní, v níž $o_n^i(t)$ představuje vstupní veličinu.

Řešení homogenní rovnice příslušné k rovnici (53) je ve tvaru $r_n^i(t) = c\lambda^t$, kde $t = k\Delta t$.²⁷ Časový krok Δt je implicitně roven jedné, a tedy $t = 1, 2, \dots, K$, kde K je počet iterací. Dosazením za $r_n^i(t + 1) = c\lambda^{t+1}$, $r_n^i(t) = c\lambda^t$ a $r_n^i(t - 1) = c\lambda^{t-1}$ do homogenní rovnice příslušné k rovnici (53) a následnou úpravou dostáváme:

$$c\lambda^{t+1} + (\varphi_n^i - (\omega + 1)) c\lambda^t + \omega c\lambda^{t-1} = 0$$

Vytkneme $c\lambda^{t-1}$:

$$c\lambda^{t-1} (\lambda^2 + (\varphi_n^i - (\omega + 1)) \lambda + \omega) = 0$$

a tedy:

$$(\lambda^2 + (\varphi_n^i - (\omega + 1)) \lambda + \omega) = 0 \quad (54)$$

Rovnice (54) je klasická kvadratická rovnice. Označme její kořeny, které mohou být reálné i komplexní, jako λ_1 a λ_2 a můžeme psát obecné řešení příslušné homogenní rovnice k rovnici (53):

$$r_n^i(t) = c_1 \lambda_1^t + c_2 \lambda_2^t$$

kde c_1 , c_2 jsou konstanty, které by bylo třeba určit ze znalosti počátečních podmínek v časech $(t - 1) = 0$ a $t = 1$. Obecné řešení nehomogenní rovnice by potom bylo

$$r_n^i(t) = c_1 \lambda_1^t + c_2 \lambda_2^t + f(t) \quad (55)$$

Člen $f(t)$ v rovnici (55) bychom určili řešením rovnice (53), kam bychom dosadili (55). Jednotlivé pozice $r_n^i(t)$ však pro stanovení podmínek stability systému částice není nutno počítat.

²⁷ Řešení vyplývá analogicky z řešení rovnice prvního řádu ve tvaru $r(t) = \lambda r(t - 1)$, které by bylo $\lambda^t r(0)$.

4.2.2 Částice *PSO* jako deterministický dynamický systém

Částice jako tlumený harmonický oscilátor dle obrázku Obr. 6 je *dynamický systém*, to jest systém měnící své chování v čase. Dynamické systémy (ve fyzice) vznikají z disipativních systémů: kdyby na částici nepůsobila odporující síla F_d (jednotka *Newton N*), bylo by chování částice harmonické, periodicky opakující se, neměnné.

Charakteristickým rysem (deterministického) dynamického chování je, že systém bez ohledu na rozdílné počáteční podmínky (počáteční vychýlení) dospěje vždy do určitého konečného stavu, v našem případě do požadovaného optima $o_n^i(t)$. Takovou vlastnost totiž u čistě náhodného chování nepozorujeme.

Zvažme nyní, že situace je složitější, než na obrázku Obr. 6, neboť pozice polohy $o_n^i(t)$ v rovnici (53) nezůstává fixní, ale mění se po každé iteraci. Tato změna se musí částici z pozice $r_n^i(t)$ nutně jevit jako náhodná, a to i v případě fixace náhodných koeficientů u_{1n}^i, u_{2n}^i v rovnici (1) na jejich předpokládané hodnoty.²⁸

Připomeňme si rovněž, že smyslem matematické analýzy chování částice není počítání jednotlivých pozic $r_n^i(t)$ v každé iteraci podle rovnice (53) s předpokládanými hodnotami koeficientů u_{1n}^i, u_{2n}^i , nýbrž pochopení dynamického chování částice v závislosti na hodnotách koeficientů ω, φ_n^i .

Analyzujeme proto pohybovou rovnici částice (53) podrobněji s tím, že potřebujeme docílit takového chování částice, aby pozice $r_n^i(t)$, která je ovlivňována veličinou $\varphi_n^i o_n^i(t)$, konvergovala k nějaké stacionární hodnotě $r_n^{i*}(t)$ tak, aby v budoucím vývoji systému platilo:

$$\lim_{t \rightarrow \infty} r_n^i(t) = r_n^{i*}(t) \quad (56)$$

Neboli, potřebujeme dospět do stádia vývoje systému, kdy se vektor pozice již nemění:

$$r_n^i(t + \Delta t) = r_n^{i*}(t) = r_n^i(t - \Delta t) \quad (57)$$

Za podmínky (57) nám rovnice (53) umožní zjistit hodnotu stacionárního bodu $r_n^{i*}(t)$:

$$r_n^{i*}(t) + (\varphi_n^i - (\omega + 1)) r_n^{i*}(t) + \omega r_n^{i*}(t) = \varphi_n^i o_n^i(t)$$

$$r_n^{i*}(t)(1 + \varphi_n^i - (\omega + 1) + \omega) = \varphi_n^i o_n^i(t)$$

$$r_n^{i*}(t) = o_n^i(t)$$

Uvážíme-li, že vstupní veličina $o_n^i(t)$ v rovnici (53) se jeví jako náhodná, můžeme říci, že náhodně měníme podmínky chování systému na počátku každé iterace. Navíc požadujeme, aby se stavový vektor $r^i(t)$ každé částice po určitém dostatečně velkém počtu iterací ustálil ve společném bodě hyperprostoru, což má být hledané optimum $o^i(t) = o^*(t)$.

²⁸ Autorka předkládané práce podotýká, že změny se mohou jevit částici jako náhodné, i když ve skutečnosti z hlediska vnějšího pozorovatele náhodné být nemusí.

Ptejme se tedy, jak nastavit koeficienty ω , φ_n^i v rovnici (53), aby se chování částice při (náhodných) změnách vstupní veličiny $o_n^i(t)$ na počátku každé iterace co nejvíce blížilo chování deterministickému.

Naším cílem je zkoumat vzdálenost bodů $o_n^i(t)$ a $r_n^i(t)$, která se má s postupujícím časem zmenšovat. Zavedeme tedy novou proměnnou $\xi_n^i(t)$:

$$\xi_n^i(t) = o_n^i(t) - r_n^i(t) \quad (58)$$

a vyjádříme $r_n^i(t)$ pomocí této nově zavedené proměnné:

$$r_n^i(t) = o_n^i(t) - \xi_n^i(t)$$

Analogicky platí:

$$r_n^i(t + \Delta t) = o_n^i(t) - \xi_n^i(t + \Delta t)$$

$$r_n^i(t - \Delta t) = o_n^i(t) - \xi_n^i(t - \Delta t)$$

Nyní dosadíme za $r_n^i(t + \Delta t)$, $r_n^i(t)$ a $r_n^i(t - \Delta t)$ do rovnice (53), v níž položíme ($\Delta t = 1$), a tedy ($\varphi_n^i \Delta^2 t \approx \varphi_n^i$), a použijeme označení $(\varphi_n^i - (\omega + 1)) = b$:

$$r_n^i(t + 1) + br_n^i(t) + \omega r_n^i(t - 1) = \varphi_n^i o_n^i(t)$$

$$(o_n^i(t) - \xi_n^i(t + 1)) + b(o_n^i(t) - \xi_n^i(t)) + \omega[o_n^i(t) - \xi_n^i(t - 1)] = \varphi_n^i o_n^i(t)$$

Provedeme úpravy:

$$o_n^i(t) - \xi_n^i(t + 1) + bo_n^i(t) - b\xi_n^i(t) + \omega o_n^i(t) - \omega \xi_n^i(t - 1) = \varphi_n^i o_n^i(t)$$

$$o_n^i(t)(1 + b + \omega) - (\xi_n^i(t + 1) + b\xi_n^i(t) + \omega \xi_n^i(t - 1)) = \varphi_n^i o_n^i(t)$$

$$(\xi_n^i(t + 1) + b\xi_n^i(t) + \omega \xi_n^i(t - 1)) = o_n^i(t)(1 + b + \omega) - \varphi_n^i o_n^i(t)$$

$$\xi_n^i(t + 1) + b\xi_n^i(t) + \omega \xi_n^i(t - 1) = o_n^i(t)(1 + b + \omega - \varphi_n^i)$$

Dosadíme zpět za $b = (\varphi_n^i - (\omega + 1))$ a vyčíslíme pravou stranu rovnice:

$$\xi_n^i(t + 1) + (\varphi_n^i - (\omega + 1))\xi_n^i(t) + \omega \xi_n^i(t - 1) = o_n^i(t)(1 + \varphi_n^i - (\omega + 1) + \omega - \varphi_n^i)$$

$$\xi_n^i(t + 1) + (\varphi_n^i - (\omega + 1))\xi_n^i(t) + \omega \xi_n^i(t - 1) = 0 \quad (59)$$

Získali jsme rovnici (59), která je homogenní a která má tutéž charakteristickou rovnici (54) jako diferenční rovnice (53). Autorka však upozorňuje na důležitou okolnost, a sice že rovnici (59) mohou autoři analýz [35], [36] použít namísto rovnice (53) pouze za splnění limitní podmínky (56).

4.2.3 Částice *PSO* jako dynamický systém popsaný vektorem rychlosti

Z definice proměnné $\xi_n^i(t)$ (58) a z rovnice (21) je patrný vztah proměnné $\xi_n^i(t)$ k rychlosti:

$$\xi_n^i(t) = \frac{v_n^i(t + \Delta t) - \omega v_n^i(t)}{\varphi_n^i \Delta t}$$

Rovněž tak koeficienty ω , φ_n^i figurující v rovnicích (1) a (21) slouží jako omezovače rychlosti, což je využíváno v analýze stability [36]. Jako alternativu k rovnici (59) obsahující ξ_n^i tedy autorka ještě odvodila popis chování systému pomocí diferenční rovnice obsahující rychlost v_n^i .

Autorka vycházela opět z rovnice (51), kam dosadila zpětně za $\beta = (\omega - 1)/\Delta t$ podle (27), dále dosadila za $\xi_n^i(t) = o_n^i(t) - r_n^i(t)$ podle (58), vyjádřila zrychlení a_n^i jako přírůstek rychlosti v_n^i :

$$a_n^i(k\Delta t) - \beta v_n^i(k\Delta t) - \varphi_n^i[o_n^i(k\Delta t) - r_n^i(k\Delta t)] = 0$$

$$\frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} - \frac{\omega - 1}{\Delta t} v_n^i(t) - \varphi_n^i \xi_n^i(t) = 0$$

a odvodila rovnici charakterizující částici *PSO* jako dynamický systém popsaný vektorem rychlosti $v_n^i(t)$ (17):

$$v_n^i(t + \Delta t) + (\varphi_n^i \Delta^2 t - (\omega + 1)) v_n^i(t) + \omega v_n^i(t - \Delta t) = 0 \quad (61a)$$

Detailní postup odvození rovnice (61a) je obsažen v dodatku D5.

Za předpokladu, že časový krok Δt je implicitně roven jedné, a tedy $(\varphi_n^i \Delta^2 t \approx \varphi_n^i)$, přejde rovnice (61a) k obvyklému zápisu diferenční rovnice druhého řádu:

$$v_n^i(t + 1) + (\varphi_n^i - (\omega + 1)) v_n^i(t) + \omega v_n^i(t - 1) = 0 \quad (61)$$

Získaná rovnice (61) je homogenní a má tutéž charakteristickou rovnici (54) jako rovnice (59) a (53).

4.2.4 Částice *PSO* jako lineární dynamický systém

Aby bylo možno analyzovat chování algoritmu *PSO* v závislosti na hodnotě koeficientů ω , φ_n^i využitím teorie lineárních diskrétních dynamických systémů a současně uplatnit poznatky z teorie regulace, jak to činí analýzy [35], [36], je třeba charakterizovat částici *PSO* jako jednoduchý lineární dynamický systém.

Protože budeme zkoumat stabilitu proměnné $\xi_n^i(t)$ definovanou vztahem (58), vyjdeme z diferenční rovnice (59) s charakteristickou rovnicí (54).

Označme levou stranu charakteristické rovnice (54) jako polynom $p(\lambda)$, který lze zapsat pomocí kořenových činitelů λ_1 , λ_2 :

$$p(\lambda) = \lambda^2 + (\varphi_n^i - (\omega + 1))\lambda + \omega = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

kde kořeny λ_1, λ_2 polynomu $p(\lambda)$ jsou komplexní čísla, která nemusí být nutně různá a která jsou řešením charakteristické rovnice (54).

Polynom $p(\lambda)$ je charakteristickým polynomem jisté čtvercové matice M , jež popisuje systém tak, že platí

$$p(\lambda) = \det(\lambda I - M) = 0 \quad (62)$$

a tedy $(\lambda I = M)$, kde I je (diagonální) jednotková matice. Vektor $\lambda = (\lambda_1, \lambda_2)$ je vlastní vektor matice M a čísla λ_1, λ_2 jsou vlastní čísla matice M .

Abychom mohli charakterizovat částici *PSO* jako jednoduchý lineární dynamický systém pomocí matice M , vyjádříme nejdříve pozice v hyperprostoru určené proměnnou ξ_n^i v časech $t, t + \Delta t$:

$$\xi_n^i(t) = o_n^i(t) - r_n^i(t)$$

$$\xi_n^i(t + \Delta t) = o_n^i(t) - r_n^i(t + \Delta t)$$

Pozice vyjádřené proměnnou ξ dosadíme do rovnice (2) a upravíme:

$$r_n^i(t + \Delta t) = r_n^i(t) + v_n^i(t + \Delta t)\Delta t$$

$$-o_n^i(t) + r_n^i(t + \Delta t) = -o_n^i(t) + r_n^i(t) + v_n^i(t + \Delta t)\Delta t$$

$$\xi_n^i(t + \Delta t) = \xi_n^i(t) - v_n^i(t + \Delta t)\Delta t$$

Za $v_n^i(t + \Delta t)$ dosadíme podle rovnic (1), (21) a (58):

$$\xi_n^i(t + \Delta t) = \xi_n^i(t) - \omega v_n^i(t)\Delta t - \varphi_n^i \xi_n^i(t)\Delta^2 t = -\omega v_n^i(t)\Delta t + \xi_n^i(t)(1 - \varphi_n^i \Delta^2 t)$$

Chování částice *PSO* jako zjednodušeného lineárního dynamického systému je potom popsáno soustavou rovnic:

$$v_n^i(t + \Delta t) = \omega v_n^i(t) + \varphi_n^i \xi_n^i(t)\Delta t \quad (63)$$

$$\xi_n^i(t + \Delta t) = -\omega v_n^i(t)\Delta t + (1 - \varphi_n^i \Delta^2 t)\xi_n^i(t) \quad (64)$$

kde rovnice (63) je odvozena z rovnice (21) dosazením za $\xi_n^i(t)$ podle (58) a je analogií rovnice (1). Rovnice (64) je analogií rovnice (2).

Systém charakterizovaný rovnicemi (63) a (64) vyjádříme v maticovém tvaru:

$$y_n^i(t + \Delta t) = M y_n^i(t)$$

$$y_n^i(t + \Delta t) = \begin{bmatrix} v_n^i(t + \Delta t) \\ \xi_n^i(t + \Delta t) \end{bmatrix} \quad y_n^i(t) = \begin{bmatrix} v_n^i(t) \\ \xi_n^i(t) \end{bmatrix}$$

Systém je popsán maticí M :

$$M = \begin{bmatrix} \omega & \varphi_n^i \Delta t \\ -\omega \Delta t & 1 - \varphi_n^i \Delta^2 t \end{bmatrix} \quad (65)$$

Charakteristickou rovnici spočteme podle (62):

$$\begin{aligned}
\det(\lambda I - M) &= \det \left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} \omega & \varphi_n^i \Delta t \\ -\omega \Delta t & 1 - \varphi_n^i \Delta^2 t \end{bmatrix} \right) = \begin{vmatrix} \lambda - \omega & -\varphi_n^i \Delta t \\ \omega \Delta t & \lambda - 1 + \varphi_n^i \Delta^2 t \end{vmatrix} \\
&= \lambda^2 - \lambda\omega - \lambda + \omega + \lambda\varphi_n^i \Delta^2 t - \omega\varphi_n^i \Delta^2 t + \omega\varphi_n^i \Delta^2 t = \\
&= \lambda^2 + \left(\varphi_n^i \Delta^2 t - (\omega + 1) \right) \lambda + \omega = 0 \\
&\left(\lambda^2 + \left(\varphi_n^i \Delta^2 t - (\omega + 1) \right) \lambda + \omega \right) = 0
\end{aligned}$$

Lze nahlédnout, že popis systému pomocí soustavy diferenčních rovnic (63), (64) má tutéž charakteristickou rovnici jako homogenní systém popsany diferenčními rovnicemi (59) a (61) a tutéž charakteristickou rovnici jako nehomogenní systém popsany rovnicí (53).

4.2.5 Otázky k zodpovězení

Všechny uvedené popisy systému částice vycházejí z fyzikálního modelu tlumeného harmonického oscilátoru, jak je znázorněn na obrázku Obr. 6. Oscilátor periodicky kmitá na mezi stability. Jeho rovnovážná poloha $o_n^i(t)$ se v čase mění a bod $o_n^i(t)$ se tedy pohybuje po trajektorii. *Odpovídá však tato představa skutečnému chování částice PSO?*

Typické chování tlumeného oscilátoru je, že vždy dospěje do klidového, rovnovážného bodu, který je jediným atraktorem. *Je zde záruka, že bod v hyperprostoru, v němž se trajektorie $o_n^i(t)$ ustálí, je jediným atraktorem, a sice globálním optímem?*

Dle autorky předkládané práce takovou záruku dát nelze. Experimenty ukazují, že částice může divergovat (opustit vyhledávací prostor), konvergovat k lokálnímu optimu, nebo může skončit v libovolné pozici, která byla rojem rozpoznána jako nejlepší globální pozice, přičemž se nemusí jednat ani o lokální, ani globální optimum, jak bylo podrobněji zmíněno v kapitole 3.3.2. Lokální optimum má charakter *klidového bodu (steady state)*, z něhož může být částice vyvedena, zatímco globální optimum má charakter *ekvilibria (equilibrium)*. Výsledné chování částice je přitom nepředvídatelné a závisí na počátečních podmínkách.

Popsané chování systému částice je typické pro (*deterministický*) *chaos*, kdy klidové body jsou předem nepředvídatelné body v hyperprostoru.²⁹ *Má tedy být chování částice interpretováno spíše jako chaotické? A jaký vliv na výsledné chování částice mají náhodné koeficienty? Zaručuje splnění limitní podmínky (56), že nedojde k divergenci systému částice?*

Na uvedené otázky se autorka zaměří při meta-analýze stability systému částice.

²⁹ Pojem *chaos* má přesný matematický význam a popisuje chování, kdy deterministické (diferenční) rovnice modelující chování systému mohou simulovat náhodné chování tak, že i nepatrné změny počátečních podmínek vyvolají podstatné změny budoucího chování, a budoucí stav *chaotického deterministického dynamického systému* proto není v důsledku citlivosti na počáteční podmínky předpověditelný.

Dynamický systém může mít za stabilní stavy *pevné body*, či *periodické body* (příkladem periodického bodu je trajektorie družice vyvedené na oběžnou dráhu), nebo tzv. *podivné atraktory* (útvary, které nelze popsat složením jednoduchých geometrických obrazců).

Chaotické dynamické systémy mohou vykreslovat ve stavovém prostoru tyto podivné, tzv. *fraktální struktury*, ale pro chaotický systém to není podmínkou. *Chaotický atraktor* je atraktor, který *nelze předem předpovědět*, a sice vlivem citlivosti na počáteční podmínky. *Chaotičnost neznamená náhodnost*.

Chaos může nastat ve spojitém systému, který je minimálně trojrozměrný. Pro diskrétní systémy, což je případ částice, může existovat i jednorozměrný chaos. Jako příklad jednorozměrného diskrétního chaotického systému je často uváděna tzv. *logistická mapa* definovaná předpisem $f(x) = c(1 - x)$. Funkce $f(x)$ je pro $c > 4$ chaotická.

V případě částice by se jednalo o chaotický systém, jehož atraktory jsou nepředpověditelné pevné body.

4.3 Vztahy mezi koeficienty plynoucí z deterministické analýzy stability chování částice *PSO*

Podmínka stability pro dynamický systém (56), která byla stanovena tak, že pozice $r_n^i(t)$, která je ovlivňována veličinou φ_n^i , má konvergovat k nějaké stacionární hodnotě $r_n^{i*}(t)$, je z hlediska teorie regulace podmínkou tzv. *vnitřní stability systému*.

Podmínka vnitřní stability systému, jak následně vyplýne, je pro diskrétní dynamický systém definována tak, že kořeny charakteristické rovnice popisující systém, ať komplexní, či reálné, musí ležet uvnitř jednotkového kruhu v komplexní rovině.

Podmínky vnitřní stability systému částice *PSO* vycházejí z obecných podmínek stability systému popsaného diferenčními rovnicemi druhého řádu a z poznatků teorie regulace. Tyto podmínky budou uvedeny v kapitole 4.3.1 a autorka předkládané práce rovněž odvodí související vztahy, na něž se odvolávají studie [35] a [36]. Interpretace závěrů uvedených studií je potom obsahem kapitol 4.3.2 a 4.3.3.

Autorka v dané souvislosti ještě poznamenává, že jiný pojem je *stabilita přenosu typu BIBO* (*Bounded Input Bounded Output*). Přenos je *BIBO-stabilní*, právě když odezva systému na každý omezený vstupní signál je také omezená. Pokud nějaký omezený vstupní signál vyvolá na výstupu neomezenou odezvu, pak přenos ze vstupu na výstup stabilní není. Vnitřně stabilní systém má všechny přenosy ze vstupu na výstup stabilní.³⁰

4.3.1 Podmínky vnitřní systémové stability částice *PSO*

Vnitřní systémovou stabilitu částice *PSO* na základě podmínky (56) a vztahu (58) definuje autorka pomocí proměnné $\xi_n^i(t)$ jako:

$$\lim_{t \rightarrow \infty} \xi_n^i(t) = 0 \quad (66)$$

Podmínky vnitřní systémové stability částice *PSO* autorka odvodí z obecného řešení diferenční rovnice (59) tak, že spočte kořeny λ_1 a λ_2 charakteristické rovnice (54), a sice jako kořeny kvadratické rovnice, a řešení bude autorka následně analyzovat. Rovnici (54) zde autorka pro přehlednost uvádí opakovaně:

$$\lambda^2 + (\varphi_n^i - (\omega + 1))\lambda + \omega = 0$$

Diskriminant charakteristické rovnice (54) je roven:

$$D = (\varphi_n^i - (\omega + 1))^2 - 4\omega \quad (67)$$

Pro stanovení kořenů λ_1 a λ_2 charakteristické rovnice (54), které mohou být reálné i komplexní, nastávají klasicky tři případy, a sice (i) diskriminant $D < 0$ pro komplexní kořeny, (ii) $D = 0$ pro jeden dvojnásobný reálný kořen a (iii) $D > 0$ pro reálné kořeny.

Analyzujme nejdříve případ $D < 0$. Má-li být diskriminant kvadratické rovnice (54) daný výrazem (67) menší než nula, musí platit:

³⁰ Stabilita přenosu je vlastností přenosu ze vstupu na výstup, nikoli vlastností systému. Vnitřně stabilní systém má sice všechny přenosy ze vstupu na výstup *BIBO-stabilní*, avšak opačně to neplatí.

$$\left(\varphi_n^i - (\omega + 1)\right)^2 < 4\omega$$

Kořeny λ_1 a λ_2 rovnice (54) budou v daném případě komplexně sdružená čísla:

$$\lambda_{1,2} = -\frac{\varphi_n^i - (\omega + 1)}{2} \pm j \frac{\sqrt{4\omega - \left(\varphi_n^i - (\omega + 1)\right)^2}}{2} = -\frac{b}{2} \pm j \sqrt{\omega - \frac{b^2}{4}} \quad (68)$$

kde $b = \varphi_n^i - (\omega + 1)$, j je imaginární jednotka, pro niž platí $j^2 = (-1)$, a obecné řešení rovnice (59) bude ve tvaru

$$\xi_n^i(t) = c_1 \lambda_1^t + c_2 \lambda_2^t \quad (69)$$

kde konstanty c_1, c_2 mohou být rovněž komplexní čísla.

Abychom snáze určili podmínky stability pro případ komplexních kořenů, navrhuje autorka předkládané práce použití Eulerova vzorce pro přepis komplexních čísel:

$$\lambda_{1,2} = |r|(\cos \theta \pm j \sin \theta)$$

kde j je imaginární jednotka, θ je reálné číslo, e je Eulerovo číslo, $|r|$ je modul čísel $\lambda_{1,2}$ a $\cos \theta, \sin \theta$ určíme z (68):

$$\cos \theta = -\frac{\varphi_n^i - (\omega + 1)}{2|r|}$$

$$\sin \theta = \frac{\sqrt{4\omega - \left(\varphi_n^i - (\omega + 1)\right)^2}}{2|r|}$$

Spočteme modul $|r|$:

$$|r| = \sqrt{\left(-\frac{\varphi_n^i - (\omega + 1)}{2}\right)^2 + \left(\frac{\sqrt{4\omega - \left(\varphi_n^i - (\omega + 1)\right)^2}}{2}\right)^2} = \sqrt{\omega}$$

Aby platila limitní podmínka stability (66), musí být modul $|r| = \sqrt{\omega}$ komplexních kořenů $\lambda_{1,2}$ rovnice (54) menší než jedna, neboli

$$|r| < 1 \quad 0 < \omega < 1 \quad (70)$$

a vnitřní stabilita systému tudíž závisí na umístění kořenů charakteristické rovnice:

(i) $|\lambda_1| < |\lambda_2| < 1$

Oba kořeny leží uvnitř jednotkové kružnice v komplexní rovině, řešení (69) je asymptoticky stabilní a blíží se k nule pro $t \rightarrow \infty$.

(ii) $|\lambda_1| < |\lambda_2| = 1$

Kořen λ_2 leží na jednotkové kružnici v komplexní rovině, řešení (69) je na mezi stability.

(iii) $|\lambda_1| = |\lambda_2| = 1$

Oba kořeny leží na jednotkové kružnici v komplexní rovině, řešení (69) je nestabilní

(iv) $|\lambda_1| < 1 < |\lambda_2|$

Kořen λ_2 leží vně jednotkové kružnice v komplexní rovině, řešení (69) je nestabilní

Komplexní kořeny $\lambda_{1,2}$ (68) charakteristické rovnice (54) splynou v jediný dvojnásobný reálný kořen

$$\lambda = -\frac{\varphi_n^i - (\omega + 1)}{2}$$

bude-li platit

$$\left(\varphi_n^i - (\omega + 1)\right)^2 = 4\omega \quad (71)$$

a tedy diskriminant charakteristické rovnice (54) daný výrazem (67) bude roven nule. Obecné řešení rovnice (59) bude v daném případě reálné, ve tvaru

$$\xi_n^i(t) = (c_1 + c_2 t)\lambda^t$$

Bude-li diskriminant kvadratické rovnice (54) daný výrazem (67) větší než nula, a tedy bude-li platit:

$$\left(\varphi_n^i - (\omega + 1)\right)^2 > 4\omega$$

budou kořeny $\lambda_{1,2}$ charakteristické rovnice (54) reálná čísla:

$$\lambda_{1,2} = -\frac{\varphi_n^i - (\omega + 1)}{2} \pm \frac{\sqrt{\left(\varphi_n^i - (\omega + 1)\right)^2 - 4\omega}}{2} = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - \omega} \quad (72)$$

kde $b = \varphi_n^i - (\omega + 1)$. Obecné řešení rovnice (59) bude v tomto případě reálné a bude ve tvaru (68). Aby platila podmínka stability $\lim_{t \rightarrow \infty} \xi_n^i(t) = 0$ (65), musí být reálná čísla λ_1 a λ_2 stanovená podle (70) v absolutní hodnotě menší než jedna. Platí vztah $\lambda_1 \geq \lambda_2$, ale $|\lambda_1| \leq |\lambda_2|$, a tedy pro podmínku vnitřní stability postačí zkoumat podmínku $|\lambda_2| < 1$:

$$\left| -\frac{b}{2} - \sqrt{\frac{b^2}{4} - \omega} \right| < 1$$

Úpravami, které jsou podrobně uvedeny v dodatku D6, se dopravujeme k podmínce:

$$|b| < \omega + 1$$

Dosadíme zpět za $b = \varphi_n^i - (\omega + 1)$ a výsledná podmínka vnitřní stability systému pro reálné řešení $\xi_n^i(t)$ pro reálné kořeny $\lambda_{1,2}$ (72) charakteristické rovnice (54) je:

$$|\varphi_n^i - (\omega + 1)| < \omega + 1 \quad (73)$$

Rozhraní mezi reálnými a komplexními kořeny udává vztah (71), na jehož základě spočteme hodnoty φ_n^i pro $\omega_{max} = 1$:

$$\left(\varphi_n^i - (\omega_{max} + 1)\right)^2 = 4\omega_{max}$$

$$(\varphi_n^i - 2)^2 = 4$$

$$(\varphi_n^i)^2 - 4\varphi_n^i + 4 = 4$$

$$\varphi_n^i(\varphi_n^i - 4) = 0$$

Rovnice má klasicky dvě řešení, a sice $\varphi_{n_{min}}^i = 0$, $\varphi_{n_{max}}^i = 4$. Výsledná podmínka stability systému částice pro φ_n^i pro reálné řešení $\xi_n^i(t)$ a reálné kořeny $\lambda_{1,2}$ je

$$0 < \varphi_n^i < 4 \quad (74)$$

Dosadíme $\varphi_{n_{min}}^i = 0$, $\varphi_{n_{max}}^i = 4$ do (73) a spočteme ω_{min} , ω_{max} :

$$\varphi_{n_{min}}^i = 2(\omega_{min} + 1) \quad \omega_{min} = (-1)$$

$$\varphi_{n_{max}}^i = 2(\omega_{max} + 1) \quad \omega_{max} = 1$$

Výsledná podmínka stability systému částice pro ω pro reálné řešení $\xi_n^i(t)$ a reálné kořeny $\lambda_{1,2}$ je

$$-1 < \omega < 1 \quad |\omega| < 1 \quad (75)$$

Reálné řešení rovnice (59) pro $\xi_n^i(t)$ podle (69) lze získat i v případě komplexních kořenů $\lambda_{1,2}$ charakteristické rovnice (54). Zapišeme-li obecné řešení rovnice (59) s použitím Eulerova vzorce:

$$\begin{aligned} \xi_n^i(t) &= c_1 |r|^t (\cos \theta t + j \sin \theta t) + c_2 |r|^t (\cos \theta t - j \sin \theta t) = \\ &= |r|^t [(c_1 + c_2) \cos \theta t + j(c_1 - c_2) \sin \theta t] \end{aligned} \quad (76)$$

a budou-li konstanty c_1 , c_2 komplexně sdružená čísla:

$$(c_1 + c_2) = R_1 = A \cos \alpha \quad (c_1 - c_2) = R_2 = -A \sin \alpha$$

pak R_1 , R_2 budou reálná čísla a výraz

$$R_1 \cos \theta t + R_2 \sin \theta t = A \cos \alpha \cos \theta t - A \sin \alpha \sin \theta t = A \cos(\theta t + \alpha)$$

je výrazem pro harmonický kmit s frekvencí θ . Rovnici (76) lze potom zapsat ve tvaru

$$\xi_n^i(t) = A |r|^t \cos(\theta t + \alpha)$$

což je rovnice tlumeného kmitu, pokud platí podmínka $|r| < 1$ (70).

Podmínky vnitřní stability chování částice (70), (73), (74) a (75) budou uplatněny i v případě, že použijeme popis systému pomocí nehomogenní rovnice (53), popis pomocí homogenní rovnice (61), anebo popis pomocí soustavy diferenciálních rovnic (63), (64), neboť všechny popisy mají tutéž charakteristickou rovnici (54) jako homogenní systém popsany diferenciální rovnicí (59), z něhož jsme vycházeli.

4.3.2 Analýza konvergence a volba koeficientů algoritmu *PSO*

Relativně jednoduchou teoretickou analýzu algoritmu *PSO*, jež využívá výsledků získaných aplikací teorie dynamických systémů, poskytuje široce citovaná publikace [35] uveřejněná v časopisu *Information Processing Letters* v roce 2003, jejímž autorem je *Ioan Cristian Trelea*.

Východiskem pro tuto analýzu je diferenční rovnice (53) s charakteristickou rovnicí (54). Je uvažována deterministická verze algoritmu, kdy náhodné proměnné $u_{1n}^i(t)$, $u_{2n}^i(t)$ v rovnicích (1), (18) a (19) nabývají očekávaných hodnot $E[u_{1n}^i(t)] = E[u_{2n}^i(t)] = 0,5$. Tedy:

$$E[\varphi_n^i] = E[\varphi_{1n}^i] + E[\varphi_{2n}^i] = a_l E[u_{1n}^i(t)] + a_g E[u_{2n}^i(t)] = \frac{a_l + a_g}{2}$$

Výsledkem teoretické analýzy jsou graficky vyjádřené pokyny pro výběr koeficientů (ω, φ) , jak znázorňuje obrázek Obr. 7.

Obrázek Obr. 7(a) znázorňuje rozhraní mezi reálnými a komplexními koeficienty λ_1 , λ_2 charakteristické rovnice (54) určené křivkou (71).

Obrázek Obr. 7(b) vymezuje doménu, kdy částice konverguje k optimu, ohraničenou podmínkami $|\omega| < 1$ podle (70), (75), $\varphi < 4$ podle (74) a $\varphi < 2\omega + 2$ podle (71), (75).

Bude-li mít alespoň jeden kořen charakteristické rovnice (54), ať reálný (72), či komplexní (68), negativní reálnou část, bude částice *překmitávat* (*zigzag*)³¹ kolem optima. Doména “*zig-zag*” na obrázku Obr. 7(c) je tak určena podmínkami $\varphi > \omega + 1$, $\omega < 1$.

Pozice typických dvojic koeficientů (ω, φ) jsou na obrázku Obr. 7(d). Odpovídající konfigurace kořenů λ_1, λ_2 jsou následující:

- (i) $Re \lambda_1 > 0, Re \lambda_2 > 0$

Komplexní kořeny s pozitivní reálnou částí.

Značí konvergující harmonické oscilace.³² Dvojice koeficientů (ω, φ) , které jsou blíže k těžišti trojúhelníku, vyvolávají rychlou konvergenci k optimu, zatímco dvojice koeficientů blíže k hranám trojúhelníku vyžadují ke konvergenci více iterací

- (ii) $Re \lambda_1 < 0, Re \lambda_2 < 0$

Komplexní kořeny s negativní reálnou částí.

Značí harmonické oscilace doprovázené *překmitáváním*.

- (iii) $\lambda_1 > 0, \lambda_2 > 0$

Reálné pozitivní kořeny.

Značí tlumené, konvergující chování bez oscilací a bez *překmitávání*. Tato konfigurace není obecně doporučována, i když v některých situacích může být použitelná.

- (iv) $\lambda_1 < 0, \lambda_2 < 0$

Reálné negativní kořeny.

Značí symetrické konvergující *překmitávání* bez oscilací. Konvergence může být urychlena nebo zpomalena stejně jako v případě (i).

³¹ *Překmitávání* (*zigzag*) je *periodické* kmitání, jehož průběh není sinusový (oblý). Průběh *překmitávání* může být trojúhelníkový (*triangle wave*), čtvercový (*square wave*) nebo pilovitý (*saw-tooth wave*). Matematicky lze *překmitávání* modelovat skládáním sinusoid podle *Fourierova* rozvoje.

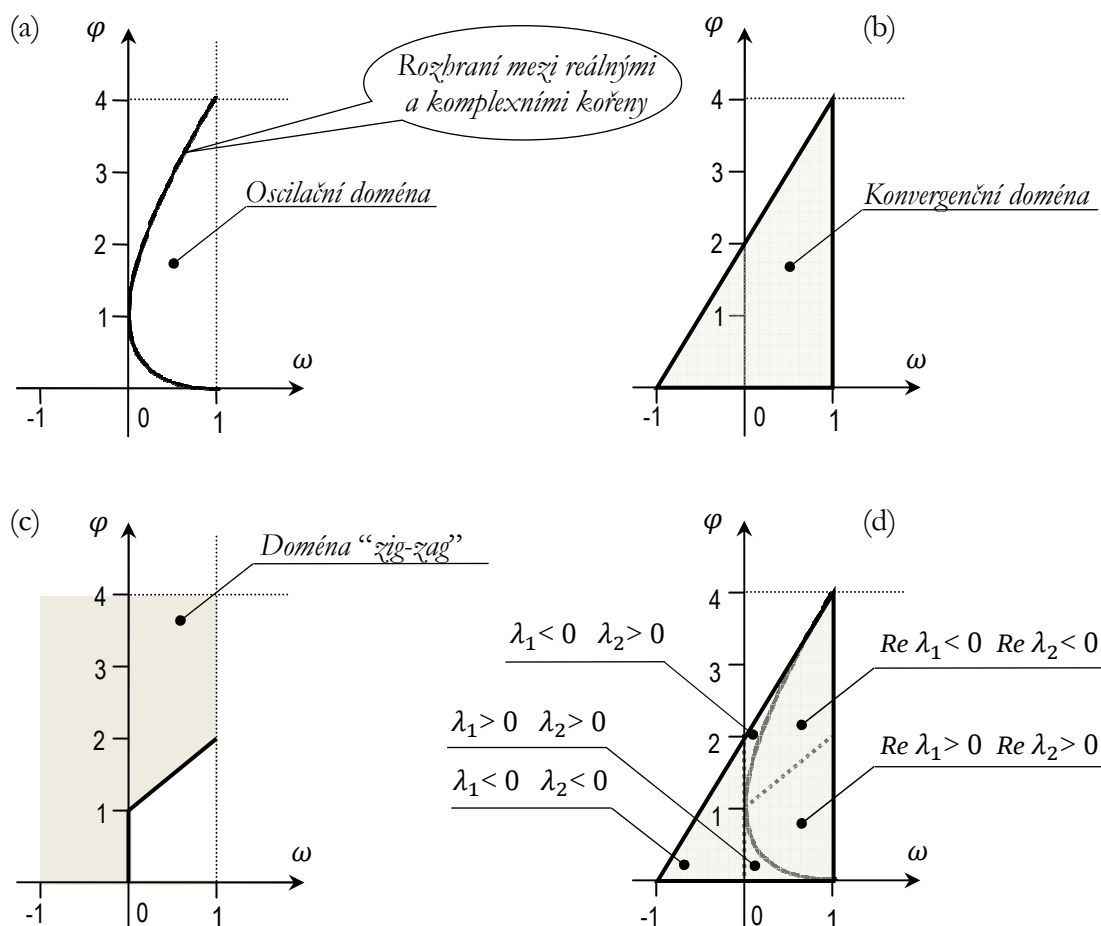
³² *Oscilace* je opakovaná změna (kmitání, typicky v čase) nějaké hodnoty kolem rovnovážné polohy nebo mezi dvěma různými stavy systému.

(v) $\lambda_1 < 0, \lambda_2 > 0$

Reálné kořeny s opačnými znaménky.

Značí konvergující asymetrické překmitávání.

Příklady dynamického chování odpovídající popisům (i)-(v) jsou na obrázku Obr. 8.



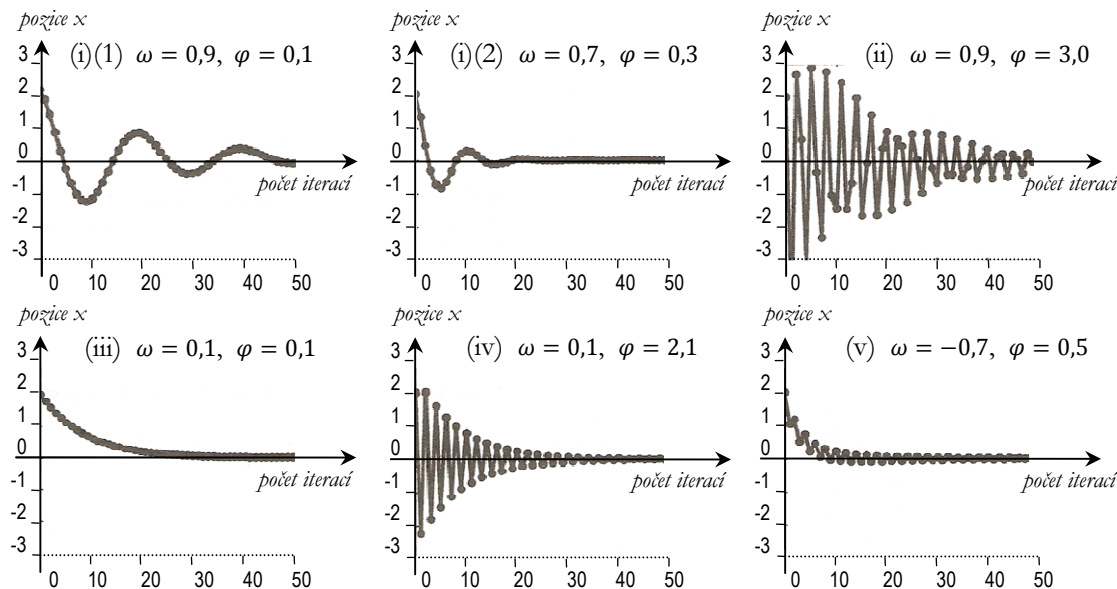
Obr. 7: Domény dynamického chování systému částice v prostoru koeficientů (ω, φ) .

(a) Oscilační doména. (b) Konvergenční doména. (c) Doména chování typu „zig-zag“.

(d) Pozice typických dvojic koeficientů (ω, φ) . [35] (Trelea 2003)

Využitím diagramu na obrázku Obr. 7(d) může být dle autora studie [35] dosaženo kvalifikovaného kompromisu při volbě koeficientů (ω, φ) , a sice mezi požadavkem na důkladné prohledávání prostoru řešení (*exploration*) a požadavkem na rychlou konvergenci (*exploitation*), to vše při naplnění podmínek vnitřní stability chování systému částice.

Autor studie [35] poznamenává, že konkrétní volba koeficientů (ω, φ) v daných rozmezích závisí na charakteru účelové funkce, zda převáží požadavek důkladného prohledávání, které však vede k většímu počtu evaluací účelové funkce, nebo požadavek na urychlení konvergence, což však může vést k neoptimálnímu řešení (k výslednému nalezení lokálního optima). Autor rovněž uvádí příklady dynamického chování částice pro různé volby koeficientů (ω, φ) , jež znázorňuje obrázek Obr. 8. Grafy jsou převzaty výsledky experimentů ze studie [35]:



Obr. 8: Příklady dynamického chování částice *PSO* pro různé volby koeficientů (ω, φ) podle obrázku Obr. 7(d): (i)(1) Oscilace s pomalejší konvergencí, (i)(2) Oscilace s rychlejší konvergencí, (ii) Oscilace s *překmitáváním*, (iii) Konvergence bez oscilací a *překmitávání*, (iv) Symetrické *překmitávání*, (v) Asymetrické *překmitávání* [35:p.322] (Trelea 2003)

Autor studie [35] dospěl po poměrně rozsáhlé sérii experimentů s účelovými funkcemi *Sphere*, *Rosenbrock*, *Rastrigin*, *Griewank* a *Schaffer's f6*³³ k návrhu koeficientů algoritmu *PSO* o hodnotách $\omega = 0,6$ a $\varphi = 1,7$. Algoritmus *PSO* je implementován jako simulace rovnic (21) a (1), s využitím vztahů (19) a (20).

Autor [35] v závěru své studie upozorňuje na nutnost dalšího výzkumu, který by objasnil vliv náhodných čísel u_{1n}^i, u_{2n}^i v algoritmu *PSO* (1), (18). Autor [35] zdůrazňuje, že úloha náhodných čísel v algoritmu *PSO* je neodmyslitelná, neboť přítomnost náhodných čísel zvyšuje kmitání, zpomaluje konvergenci a zabraňuje stagnaci, a tedy zlepšuje explorativní schopnost algoritmu.

Odvracení stagnace je důležité, i když nejlepší dosud dosažená pozice dané částice $l_n^i(t)$ je od nejlepší globální pozice $g_n(t)$ vzdálena. Náhodná čísla u_{1n}^i, u_{2n}^i totiž zaručí, že $o_n^i(t)$ (19) se změní i v případě, pokud by se ani jedna z pozic $l_n^i(t) \neq g_n(t)$ v následující iteraci nevylepší:

$$o_n^i(t) = \frac{\varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)}{\varphi_{1n}^i + \varphi_{2n}^i} = \frac{a_l u_{1n}^i(t)}{\varphi_{1n}^i + \varphi_{2n}^i} l_n^i(t) + \frac{a_g u_{2n}^i(t)}{\varphi_{1n}^i + \varphi_{2n}^i} g_n(t)$$

Přesněji řečeno, nastane-li situace:

$$l_n^i(t) \neq g_n(t), \text{ ale } l_n^i(t+1) = l_n^i(t) \text{ a } g_n^i(t+1) = g_n(t), \text{ potom bude také } o_n^i(t+1) \neq o_n^i(t), \text{ protože } u_{1n}^i(t+1) \neq u_{1n}^i(t) \text{ a } u_{2n}^i(t+1) \neq u_{2n}^i(t).$$

Přítomnost náhodných čísel zaručí, že nastane pohyb, který by jinak nenastal.

³³ Vzorce a informativní grafické znázornění zmiňovaných účelových funkcí nalezne čtenář v sekci *Příloha A*.

4.3.3 Exploze, stabilita a konvergence částice *PSO*

Publikace [35], o níž pojednává kapitola 4.3.2, poukazuje na předcházející obsáhlou teoretickou analýzu algoritmu *PSO* [36] uveřejněnou v časopise *IEEE Transactions on Evolutionary Computation* v roce 2002, jejímiž autory jsou *Maurice Clerc a James Kennedy*.

Analýza [36], podobně jako analýza [35], zkoumá vliv koeficientů algoritmu *PSO* na vnitřní stabilitu částice využitím výsledků získaných aplikací teorie dynamických systémů. Východiskem pro analýzu [36] je popis systému částice *PSO* jako zjednodušeného dynamického systému charakterizovaného rovnicemi (63), (64) s předpoklady $\omega = 1$ a $\Delta t = 1$, avšak pojetí koeficientu φ_n^i je poněkud odlišné od pojetí v originální verzi algoritmu *PSO* [2]. Pro další výklad proto autorka předkládané disertační práce přeznačila koeficient φ_n^i v rovnicích (63), (64) na ψ_n^i a takto pozměněné rovnice zde uvádí:

$$v_n^i(t+1) = v_n^i(t) + \psi_n^i \xi_n^i(t) \quad (77)$$

$$\xi_n^i(t+1) = -v_n^i(t) + (1 - \psi_n^i) \xi_n^i(t) \quad (78)$$

Odpovídající rovnici pro rychlost částice, z níž autoři [36] rovněž vycházejí, by autorka práce odvodila podobně, jako odvodila odpovídající diferenční rovnici (61), a to rovněž za předpokladu $\omega = 1$ a $\varphi_n^i \approx \psi_n^i$:

$$v_n^i(t+2) + (\psi_n^i - 2)v_n^i(t+1) + v_n^i(t) = 0 \quad (79)$$

Řešení rovnice (79) je ve tvaru

$$v_n^i(t) = c_1 \mu_1^t + c_2 \mu_2^t$$

kde μ_1 a μ_2 :

$$\mu_{1,2} = 1 - \frac{\psi_n^i}{2} \pm \frac{\sqrt{(\psi_n^i - 2)^2 - 4}}{2} = 1 - \frac{\psi_n^i}{2} \pm \frac{\sqrt{\psi_n^{i2} - 4\psi_n^i}}{2}$$

jsou kořeny charakteristické rovnice k rovnici (79). Kořeny odpovídají kořenům charakteristické rovnice (54) pro $\omega = 1$.

Jako další krok je ve studii [36] uvažován obecnější popis chování systému částice, a sice přidáním koeficientů $\alpha, \beta, \gamma, \delta, \eta$ do popisu (77), (78) takto:

$$v_n^i(t+1) = \alpha v_n^i(t) + \beta \psi_n^i \xi_n^i(t) \quad (80)$$

$$\xi_n^i(t+1) = -\gamma v_n^i(t) + (\delta - \eta \psi_n^i) \xi_n^i(t) \quad (81)$$

Studie [36] dále pokračuje detailním pojednáním, v němž autoři analyzují vzájemné vztahy a možné konfigurace koeficientů $\alpha, \beta, \gamma, \delta, \eta$ tak, aby systém částice splňoval podmínky vnitřní stability.

Z možných konfigurací koeficientů $\alpha, \beta, \gamma, \delta, \eta$ zde autorka předkládané disertační práce zmíní pouze tu variantu, která v zásadě nemění originální verzi algoritmu *PSO* a která byla autory [36] v roce 2007 navržena jako součást standardu³⁴ algoritmu *PSO* [4].

³⁴ Pojem “standard” zde není ve významu závazné normy, ale spíše *referenčního modelu* algoritmu.

Volbu koeficientů α , β , γ , δ , η , která odpovídá originální verzi algoritmu *PSO*, znázorní autorka nejlépe pomocí matice originálního systému (65) tak, že matici bude postupně modifikovat:

$$M = \begin{bmatrix} \omega & \varphi_n^i \Delta t \\ -\omega \Delta t & 1 - \varphi_n^i \Delta^2 t \end{bmatrix} \quad M = \begin{bmatrix} \alpha & \beta \psi_n^i \Delta t \\ -\gamma \Delta t & \delta - \eta \psi_n^i \Delta^2 t \end{bmatrix} \quad M = \begin{bmatrix} \chi & \chi \psi_n^i \Delta t \\ -\chi \Delta t & 1 - \chi \psi_n^i \Delta^2 t \end{bmatrix}$$

Výsledkem studie [36] je nalezení vzájemné závislosti koeficientu $\chi \approx \omega$ a koeficientu $\psi_n^i = \varphi_n^i / \chi$ stanovením podmínek vnitřní stability systému popsaného modifikovanou maticí M . V dalším pojednání se autorka soustředí na odvození tohoto výsledku vlastními postupy, přičemž se nechá postupy obsaženými v [36] pouze inspirovat.

Systém částice charakterizovaný modifikovanou maticí M je popsán soustavou rovnic:

$$v_n^i(t+1) = \chi v_n^i(t) + \chi \psi_n^i \xi_n^i(t) \quad (82)$$

$$\xi_n^i(t+1) = -\chi v_n^i(t) + (1 - \chi \psi_n^i) \xi_n^i(t) \quad (83)$$

Charakteristickou rovnici soustavy (82), (83) spočte autorka podle (62):

$$\begin{aligned} \det(\vartheta I - M) &= \det \left(\begin{bmatrix} \vartheta & 0 \\ 0 & \vartheta \end{bmatrix} - \begin{bmatrix} \chi & \chi \psi_n^i \\ -\chi & 1 - \chi \psi_n^i \end{bmatrix} \right) = \begin{vmatrix} \vartheta - \chi & -\chi \psi_n^i \\ \chi & \vartheta - 1 + \chi \psi_n^i \end{vmatrix} \\ &= \vartheta^2 - \vartheta \chi - \vartheta + \chi + \vartheta \chi \psi_n^i - \chi \chi \psi_n^i + \chi \chi \psi_n^i = \\ &= \vartheta^2 + (\chi \psi_n^i - (\chi + 1)) \vartheta + \chi = 0 \\ &(\vartheta^2 + (\chi \psi_n^i - (\chi + 1)) \vartheta + \chi) = 0 \end{aligned}$$

Kořeny $\vartheta_{1,2}$ charakteristické rovnice mohou být reálné

$$\vartheta_{1,2} = \frac{(\chi + 1) - \chi \psi_n^i}{2} \pm \frac{\sqrt{(\chi \psi_n^i - (\chi + 1))^2 - 4\chi}}{2}$$

anebo komplexní

$$\vartheta_{1,2} = \frac{(\chi + 1) - \chi \psi_n^i}{2} \pm j \frac{\sqrt{4\chi - (\chi \psi_n^i - (\chi + 1))^2}}{2}$$

Rozhraní mezi reálnými a komplexními kořeny určuje křivka analogická s křivkou (71) na obrázku Obr. 7(a), kde $\chi \approx \omega$ a $\varphi = \chi \psi_n^i$:

$$(\chi \psi_n^i - (\chi + 1))^2 = 4\chi$$

Soustavě rovnic (82), (83) odpovídá diferenční rovnice pro rychlost částice, která je analogií rovnice (61):

$$v_n^i(t+2) + (\chi \psi_n^i - (\chi + 1)) v_n^i(t+1) + \chi v_n^i(t) = 0$$

Autorka předkládané práce pro úplnost ještě uvádí, že řešení pro $v_n^i(t)$ je ve tvaru:

$$v_n^i(t) = d_1 \vartheta_1^t + d_2 \vartheta_2^t$$

Spíše než konkrétní řešení diferenční rovnice pro $v_n^i(t)$ však autorku zajímají podmínky vnitřní systémové stability obou soustav podle obecných vztahů, tak jak byly odvozeny v kapitole 4.3.1. Obě soustavy bude autorka nyní vzájemně porovnávat.

Autorka nejdříve porovná podmínky vnitřní stability pro komplexní kořeny $\mu_{1,2}$ soustavy (77), (78) a komplexní kořeny $\vartheta_{1,2}$ soustavy (82), (83).

Moduly kořenů $|\mu_1| = |\mu_2| = 1$, a tedy oba kořeny leží na jednotkové kružnici v komplexní rovině a systém (77), (78) je na mezi stability. Moduly $|\vartheta_1| = |\vartheta_2| = \sqrt{\chi}$. Pokud mají oba kořeny ϑ_1, ϑ_2 soustavy (82), (83) ležet uvnitř jednotkové kružnice, musí platit $0 < \chi < 1$. Platí tudíž vztahy:

$$|\vartheta_1| = \sqrt{\chi} |\mu_1| > \chi |\mu_1| \quad |\vartheta_2| = \sqrt{\chi} |\mu_2| > \chi |\mu_2|$$

Jelikož χ je reálné číslo menší než 1, a tedy $\sqrt{\chi} > \chi$, lze usoudit, že oba kořeny ϑ_1, ϑ_2 budou ležet uvnitř jednotkové kružnice i za podmínek

$$|\vartheta_1| = \chi |\mu_1| \quad |\vartheta_2| = \chi |\mu_2|$$

Nyní autorka porovná podmínky vnitřní stability pro reálné kořeny $\mu_{1,2}$ soustavy (77), (78) s podmínkami vnitřní stability soustavy (82), (83) pro reálné kořeny $\vartheta_{1,2}$.

Autorka postupuje tak,³⁵ že spočte součiny absolutních hodnot kořenů μ_1, μ_2 a koeficientu χ a výsledek porovná s absolutními hodnotami kořenů ϑ_1, ϑ_2 :

$$\chi |\mu_1| = \left| \frac{2\chi - \chi \psi_n^i}{2} + \frac{\sqrt{\chi^2 (\psi_n^i)^2 - 4\psi_n^i}}{2} \right| \quad (84)$$

Aby autorka mohla snáze porovnat výraz (84) s výrazem pro $|\vartheta_1|$, nejdříve si ve výrazu pro $|\vartheta_1|$ upraví výraz pro diskriminant:

$$\begin{aligned} |\vartheta_1| &= \left| \frac{(\chi + 1) - \chi \psi_n^i}{2} + \frac{\sqrt{(\chi \psi_n^i - (\chi + 1))^2 - 4\chi}}{2} \right| = \dots \\ &\dots = \left| \frac{(\chi + 1) - \chi \psi_n^i}{2} + \frac{\sqrt{\chi^2 \left(\left(\psi_n^i - \frac{1}{\chi} + 1 \right)^2 - 4 \frac{1}{\chi} \right)}}{2} \right| \quad (85) \end{aligned}$$

Podrobný postup úpravy výrazu pro $|\vartheta_1|$ na výraz (85) je obsažen v dodatku D7. Upravený výraz pro $|\vartheta_1|$ (85) bude autorka moci lépe porovnat s výrazem pro $\chi |\mu_1|$ (84). Hodnoty pro $\chi |\mu_2|$ a $|\vartheta_2|$ by autorka spočetla analogicky.

³⁵ Vzhledem k nesrovnalostem ve vzorci (4.14) ve studii [36] na straně 67 volí autorka vlastní, analogický postup jako [36] a dochází ke stejnému výslednému vzorci (5.3) uvedenému v [36] na straně 69.

Nyní autorka porovná výrazy (84) a (85). Uváží-li, že hodnota ψ_n^i musí být větší než čtyři, aby kořeny $\lambda_1, \lambda_2, \vartheta_1, \vartheta_2$ byly reálnými čísly, a dále že χ je kladné reálné číslo menší než jedna, dochází autorka k následujícímu závěru:

$$|\vartheta_1| > \chi|\mu_1| \quad |\vartheta_2| > \chi|\mu_2|$$

Pro reálné kořeny $\mu_1, \mu_2, \vartheta_1, \vartheta_2$ platí $|\mu_1| \leq |\mu_2|, |\vartheta_1| \leq |\vartheta_2|$ a musí rovněž platit $|\mu_1| \leq |\mu_2| \leq 1, |\vartheta_1| \leq |\vartheta_2| \leq 1$. Pro splnění podmínek vnitřní stability bude tedy dostačující, zvolíme-li

$$|\vartheta_2| \leq 1 \quad |\vartheta_2| > \chi|\mu_2| \quad \chi|\mu_2| \leq 1$$

$$\chi \leq \frac{1}{|\mu_2|}$$

Výsledné podmínky vnitřní stability systému popsaného rovnicemi (82), (83) budou tedy souhrnně formulovány takto:

$$0 < \chi < 1 \text{ pro } \psi_n^i \leq 4 \quad (86)$$

$$\chi = \frac{k}{|\mu_2|} = \frac{2k}{\left| 2 - \psi_n^i - \sqrt{\psi_n^{i2} - 4\psi_n^i} \right|} = \frac{2k}{\psi_n^i - 2 + \sqrt{\psi_n^{i2} - 4\psi_n^i}} \quad (87)$$

$$\text{pro } \psi_n^i > 4, 0 < k \leq 1$$

Originální algoritmus *PSO* charakterizovaný rovnicemi (1) a (2) zůstane zachován i po náhradě koeficientů $\omega = \chi, \varphi_n^i = (\varphi_{1n}^i + \varphi_{2n}^i) \approx \chi \psi_n^i = \chi(\psi_{1n}^i + \psi_{2n}^i)$. Pozmění se pouze popis algoritmu. Analogie rovnice (1) vznikne z rovnice (63) zpětným dosazením za φ_n^i :

$$v_n^i(t + \Delta t) = \omega v_n^i(t) + \varphi_n^i \xi_n^i(t) \Delta t$$

$$v_n^i(t + \Delta t) = \chi v_n^i(t) + \chi \psi_n^i \xi_n^i(t) \Delta t =$$

$$\chi v_n^i(t) + \chi(\psi_{1n}^i + \psi_{2n}^i)[o_n^i(t) - r_n^i(t)] \Delta t =$$

$$\chi v_n^i(t) + \chi(\psi_{1n}^i + \psi_{2n}^i) \left[\frac{\psi_{1n}^i l_n^i(t) + \psi_{2n}^i g_n(t) - r_n^i \psi_{1n}^i - r_n^i \psi_{2n}^i}{\psi_{1n}^i + \psi_{2n}^i} \right] \Delta t$$

Výsledná rovnice pro regulaci (omezení) rychlosti částice pomocí koeficientů χ, ψ_n^i definovaných vztahy (86), (87), a sice náhradou za koeficienty $\omega \approx \chi, \varphi_n^i \approx \chi \psi$ v rovnici (1), má tedy následující podobu:

$$v_n^i(t + \Delta t) = \chi v_n^i(t) + \chi \psi_{1n}^i [l_n^i(t) - r_n^i] \Delta t + \chi \psi_{2n}^i [g_n(t) - r_n^i] \Delta t \quad (88)$$

Potud studie [36], jejíž výklad autorka soustředila na variantu charakterizovanou soustavou rovnic (82), (83) označenou jako “*Constriction Type 1*”. Pouze tato varianta je rozšířením originální verze algoritmu *PSO*. K opomenutí ostatních variant, jimiž se zabývá

studie [36], však měla autorka podstatnější důvod, než je posouzení shody příslušné varianty s originální verzí algoritmu.

K výsledné soustavě rovnic (82), (83), již se autorka předkládané práce zabývala, se totiž můžeme dopracovat méně složitým způsobem než autoři analýz [36], a sice pokud vyjdeme z jedné z úvah autora studie [35] a uvědomíme si, že koeficienty γ , η , δ v rovnici (81) jsou ve skutečnosti nadbytečné a že mají být rovny jedné. Zmíněná úvaha je následující:

Mějme zjednodušený dynamický systém popsáný rovnicemi:

$$\begin{aligned}v(t+1) &= av(t) + b(o(t) - x(t)) \\ x(t+1) &= cx(t) + dv(t+1)\end{aligned}$$

Postupem obdobným, jako autorka odvodila rovnici (53), bychom odvodili diferenční rovnici odpovídající tomuto systému:

$$x(t+1) + (bd - a - c)x(t) + acx(t-1) = bdo(t)$$

Lze nahlédnout, že koeficienty bd se v této rovnici vyskytují pouze v součinu. Proto hodnotu jednoho z nich můžeme zvolit pevně a měnit pouze hodnotu zbývajících z nich. Zvolíme hodnotu koeficientu $d = 1$ a koeficient b necháme proměnný, ladící. Dále můžeme z podmínky (56), že $\lim_{t \rightarrow \infty} x(t) = o(t)$, vyvodit:

$$\begin{aligned}o(t) + (b - a - c)o(t) + aco(t) &= bo(t) \\ o(t)(1 - a - c + ac) &= 0 \\ (a - 1)(c - 1) &= 0\end{aligned}$$

Máme nyní volbu $a = 1$ nebo $c = 1$. Volíme tedy $c = 1$ a koeficient a ponecháme jako ladící. Koeficienty c a d jsou tudíž nadbytečné. Potud úvaha obsažená v [35]. Autorka se nyní pokusí o podobné úvahy ohledně koeficientů rovnic (80), (81):

Přepíšme nejdříve rovnici (81) do zpět do podoby analogické s rovnicí (2):

$$\begin{aligned}o_n^i(t) - r_n^i(t+1) &= -\gamma v_n^i(t) + (\delta - \eta \psi_n^i)(o_n^i(t) - r_n^i(t)) \\ o_n^i(t) - r_n^i(t+1) &= -\gamma v_n^i(t) - \eta \psi_n^i(o_n^i(t) - r_n^i(t)) + \delta(o_n^i(t) - r_n^i(t)) \\ + r_n^i(t+1) &= +\delta r_n^i(t) + \gamma v_n^i(t) + \eta \psi_n^i(o_n^i(t) - r_n^i(t)) - \delta o_n^i(t) + o_n^i(t) \\ r_n^i(t+1) &= \delta r_n^i(t) + \gamma v_n^i(t) + \eta \psi_n^i(o_n^i(t) - r_n^i(t)) + o_n^i(t)(1 - \delta)\end{aligned}$$

Ze studie [35] již víme, že vektor $r_n^i(t+1)$ máme spočítat složením vektorů $r_n^i(t)$ a $v_n^i(t+1)$ bez dalších redundantních koeficientů. Vektor $v_n^i(t+1)$ bychom tudíž spočetli podle rovnice (21) zcela zbytečně, kdybychom ho potom nahradili jiným vektorem takto:

$$\begin{aligned}v_n^i(t+1) &= \alpha v_n^i(t) + \beta \psi_n^i(o_n^i(t) - r_n^i(t)) \\ v_n^i(t+1) &= \gamma v_n^i(t) + \eta \psi_n^i(o_n^i(t) - r_n^i(t))\end{aligned}$$

Logicky tedy vyplývá jediná možnost, a sice $\alpha = \gamma$ a $\beta = \eta$. Zbývá koeficient δ , pomocí něhož bychom měli posunovat středem přitažlivosti $\phi_n^i(t)$ a měli bychom tak dodatečně ovlivnit vektor $r_n^i(t+1)$. Střed přitažlivosti však již modifikujeme pomocí $\beta\psi_n^i$ v rovnici pro $v_n^i(t+1)$. Nastavíme tedy $\delta = 1$ a objevuje se nám již známý popis chování částice *PSO* pomocí soustavy rovnic (63), (64).

Potud úvaha autorky předkládané práce inspirovaná úvahou uvedenou ve studii [35] a autorka se opět vrací ke studii [36] a k rovnici (88).

Ve srovnání s původní rovnicí (1) pro rychlost částice mohou koeficienty ψ_{1n}^i, ψ_{2n}^i v nové rovnici (88) nabývat v součtu hodnoty větší než čtyři, avšak po vynásobení koeficientem χ musí tato hodnota pod hodnotu čtyři opět klesnout tak, aby se hodnota koeficientu $\phi_n^i = (\phi_{1n}^i + \phi_{2n}^i) \approx \chi\psi_n^i = \chi(\psi_{1n}^i + \psi_{2n}^i)$ nacházela v pásmu stability na obrázku Obr. 7(b). Jak již bylo uvedeno, záměrem autorů [36] bylo nalezení závislosti mezi koeficienty ϕ a ω v rovnici (1).

4.3.4 Otázky k zodpovězení

Stále však zůstává otázkou, jaké hodnoty koeficientů $\chi, \psi_{1n}^i, \psi_{2n}^i$ máme konkrétně volit. Může se koeficient χ blížit k jedné? Mají být koeficienty a_l, a_g obsažené uvnitř ψ_{1n}^i, ψ_{2n}^i shodné? Mají nabývat svých maximálně možných hodnot nebo hodnot nižších? Jaký vliv na vnitřní stabilitu částice budou mít náhodná čísla skrytá uvnitř koeficientů ψ_{1n}^i, ψ_{2n}^i ?

Rovnice (88) je totiž výsledkem *deterministické analýzy* vnitřní stability chování systému částice. To znamená, že koeficienty $\phi \approx \psi$ byly až dosud pokládány pouze za reálná čísla. Tyto koeficienty jsou však ve skutečnosti realizacemi náhodných proměnných a měli bychom tedy přesněji, ve shodě s rovnicemi (1), (18), psát:

$$\begin{aligned} \phi_{1n}^i &\approx \phi_{1n}^i(t) = a_l u_{1n}^i(t) & \phi_{2n}^i &\approx \phi_{2n}^i(t) = a_g u_{2n}^i(t) \\ \chi\psi_{1n}^i &\approx \chi\psi_{1n}^i(t) = \chi a_l u_{1n}^i(t) & \chi\psi_{2n}^i &\approx \chi\psi_{2n}^i(t) = \chi a_g u_{2n}^i(t) \end{aligned}$$

Pro získání odpovědí na výše uvedené otázky ohledně hodnot koeficientů algoritmu *PSO* je proto třeba přistoupit k metodám *stochastické analýzy* vnitřní stability chování systému částice.

4.4 Vztahy mezi koeficienty plynoucí ze stochastické analýzy stagnace částice *PSO*

Deterministická analýza stability, jež byla obsahem kapitoly 4.3, se zaměřuje na vnitřní stabilitu chování systému částice a hledá podmínky, za nichž se systém bude z hlediska rychlosti částice pohybovat na mezi stability tak, aby rychlost částice blížící se mezní rychlosti ještě nezpůsobila divergenci částice a prohledávací schopnost částice přitom zůstala stále dobře zachována.

Opačným extrémem k divergenci částice je stagnace částice neboli úplná ztráta pohybu postihující nejlepší částici v roji. Pro nejlepší částici platí, že její lokální (subjektivní) nejlepší pozice je současně nejlepší pozicí dosaženou v celém roji, a tedy platí $l_n^i = g_n$ pro i -tou částici v rovnici (1), (18).

Vývoj rychlosti nejlepší částice je typicky takový, že rychlost částice rychle klesá k nule, neboť částice má tendenci se přibližovat k nejlepší částici v roji, což je v daném případě ona sama. Prakticky to znamená, že další evaluace účelové funkce jsou prováděny marně, neboť částice se po několika iteracích zcela zastaví. Jinými slovy, částice přestává prohledávat, stagnuje.

Pokud i -tá částice není nejlepší částicí v roji, to jest, pro i -tou částici neplatí $l_n^i = g_n$ v rovnici (1), (18), nemusí k úplné stagnaci dojít, i když se pozice l_n^i , g_n v následujících iteracích nezlepší, neboť vliv náhodných koeficientů vyvolává kmitání i -té částice mezi pozicemi l_n^i a g_n . Je vyzpozorováno, že modul rychlosti i -té částice má v tomto případě klesající statistické rozložení. To znamená, že malé oscilace se vyskytují často a velké oscilace se vyskytují zřídka. Avšak i malé oscilace přispívají k prohledávání a mohou částici vyvést ze stagnace.

Analýza stagnace nejlepší částice z hlediska působení náhodných veličin a jejich vlivu na vývoj rychlosti částice je obsahem výzkumné zprávy [37] vydané na *University of Essex* v roce 2006. Východiskem pro analýzu je rovnice analogická s rovnicí (61), avšak při jejím sestavování autor výzkumné zprávy *Maurice Clerc* sleduje i hodnoty realizací náhodných veličin a sestavení rovnice není tudíž zcela přímočaré. Tato rovnice spolu se základními charakteristikami náhodných veličin včetně jejich vlivu na rychlost částice bude obsahem kapitoly 4.4.1.

Autor studie [37] popisuje problém stagnace částice a účinku náhodných veličin podobným způsobem, jakým naznačil tento problém autor studie [35] *Ioan Cristian Trelea* již v roce 2003 a jak je též zmíněno v závěru kapitoly 4.3.2. Autor [37] navíc zmiňuje i případ, kdy $l_n^i = g_n$ v rovnici (1), (18). Analýza stagnace globální nejlepší částice bude předmětem kapitoly 4.4.2. Autorka předkládané práce pro úplnost ještě poznamenává, že studie [37] z roku 2006 na studii [35] z roku 2003 nijak neodkazuje.

Výsledky analýzy stagnace částice a souvislosti se vztahy mezi koeficienty plynoucími z deterministické analýzy konvergence systému částice jsou uvedeny v kapitole 4.4.3. Stručné závěry experimentálních ověření výsledků analýz provedené autorem studie [37] shrnuje kapitola 4.4.4.

4.4.1 Částice *PSO* jako stochastický systém popsáný vektorem rychlosti

Východiskem pro analýzu stagnace nejlepší částice v roji [37] je rovnice pro rychlost částice

$$v_n^i(t+1) = (\omega - s + q)v_n^i(t) - \omega q v_n^i(t-1) + d(l - g) \quad (97)$$

v níž veličiny s , q , d jsou realizacemi náhodných veličin S , Q , D . Je definována také náhodná veličina

$$Z = \omega - S + Q$$

Náhodná veličina D , jejíž realizace je obsažená v rovnici (97), je ve studii [37] označena jako W a je interpretována jako šum. Čím větší je rozdíl $(l - g)$, tím větší je vliv tohoto šumu na rychlost.

Hodnota realizace náhodné veličiny Q je vždy kladná, a tedy hodnota $(-\omega Q)$ v rovnici (97) je vždy záporná, a proto je veličina ωQ označována ve studii [37] jako odporující síla (*back force*). Její hodnota závisí na ω .

Hodnota realizace náhodné veličiny $Z = \omega - S + Q$ bude častěji kladná, ale může být i záporná. Veličina Z je ve studii [37] označována jako dopředná síla (*forth force*).

Odvození rovnice (97), jež zahrnuje rovnice (89) až (96) a jež je ve studii [37] pouze naznačeno, provedla autorka podrobně v dodatku D8. Odvození vychází z rovnic (1) a (2) a popisuje rychlost částice v časech $(t - \Delta t)$, t a $(t + \Delta t)$ ve dvou po sobě jdoucích iteracích tak, že jsou sledovány i realizace náhodných veličin v rovnici (1). Autor studie [37] předpokládá, že pro koeficient ω platí $0 < \omega < 1$. Jelikož autor [37] hledá maximální hodnotu koeficientů $a_l = a_g = a_{max}$, použil v rovnici (1) pro hodnotu a_{max} společné označení c .

Rozsah realizací náhodné veličiny S je $(0, 2c)$, střední hodnota $\mu(S) = c$. Rozsah realizací náhodné veličiny Q je $(0, \infty)$, střední hodnota $\mu(Q) = 2 \ln(2)$. Rozsah realizací náhodné veličiny Z je $(\omega - 2c + 1, \infty)$. Položíme-li $\omega - 2c + 1 = 0$, a to splněním podmínky

$$c = (\omega + 1)/2,$$

dosáhneme toho, že realizace náhodné veličiny Z budou v intervalu $(0, \infty)$. Střední hodnota veličiny Z je dána součtem obsahujícím střední hodnoty S , Q :

$$\mu(Z) = \omega - \mu(S) + \mu(Q) = \omega - c + 2 \ln(2)$$

Veličina D je symetrická se střední hodnotou $\mu(D) = 0$. Rozsah realizací veličiny D je $(-c, +c)$.

Odvození vztahů pro určení rozsahu realizací a středních hodnot náhodných veličin uvedených ve studii [37] provedla autorka v dodatku D9, jenž zahrnuje rovnice (98) až (100). Účelem odvozování bylo verifikovat ekvivalenci rovnic (61) a (97), dále pak ukázat, že tyto rovnice jsou ekvivalentní pouze za předpokladu platnosti podmínky stagnace, a rovněž ověřit vztahy pro určení rozsahu realizací a středních hodnot náhodných veličin.³⁶

³⁶ Výzkumná zpráva [37] je zatížena určitým počtem tiskových chyb, a tedy snahou autorky předkládané práce bylo odvrácení případných nejasností.

Při odvozování rovnice pro rychlost částice (96) v dodatku D8 byl sledován časový krok Δt . Ve studii [37] je však zkoumána diferenční rovnice pro rychlost za předpokladu $\Delta t = 1$. Časový krok Δt je ve studii [37] implicitně roven jedné od počátku, a tedy jeho vliv na rychlost uvažován není.

4.4.2 Analýza stagnace nejlepší částice

Při analýze stagnace částice z hlediska působení náhodných veličin vychází autor [37] z poznatku, že vývoj rychlosti částice závisí také na hodnotách koeficientů ω, a_l, a_g v rovnicích (1), (18), neboť v závislosti na hodnotách těchto koeficientů může částice divergovat a stejně tak její rychlost může klesnout k nule. Autor [37] usuzuje, že pravděpodobně existují hodnoty koeficientů, při nichž částice nebude divergovat ani stagnovat, a hledá tyto hodnoty tak, že analyzuje vliv realizací náhodných veličin na stabilitu chování systému částice vedoucí k (žádoucí) stagnaci nejlepší částice.

Autor [37] hledá podmínky stability stagnující nejlepší částice uplatněním poznatku z deterministické analýzy tak, že sestaví matici charakterizující systém chování stagnující nejlepší částice a spočte charakteristickou rovnici podle (62). Autorka předkládané práce zvolila ekvivalentní postup vycházející přímo z rovnice (97), jenž je dle jejího názoru přímočařejší a vede ke stejným výsledkům.

Pro nejlepší částici v roji platí rovnost $l = g$, a proto se vliv náhodné veličiny D v rovnici (97) neuplatní, neboť poslední člen v této rovnici bude roven nule:

$$v_n^i(t+1) - (\omega - s + q)v_n^i(t) + \omega q v_n^i(t-1) = 0$$

Ve shodě s autorem studie [37] budou hodnoty náhodných proměnných nahrazeny jejich středními hodnotami:

$$z = (\omega - s + q) \approx \mu(Z), \quad \omega q \approx \omega \mu(Q)$$

$$v_n^i(t+1) - \mu(Z)v_n^i(t) + \omega \mu(Q)v_n^i(t-1) = 0 \quad (101)$$

Autorka předkládané práce pro úplnost připomíná, že rovnice (97) byla odvozena za předpokladu stagnace lokální nejlepší pozice i -té částice i globální nejlepší pozice:

$$l_n^i(t) = l_n^i(t-1) \quad g_n = g_n(t-1)$$

Podmínky stability autorka nyní nalezne tak, že spočte kořeny charakteristické rovnice k rovnici (101), jejichž modul v případě komplexních kořenů, či absolutní hodnota v případě reálných kořenů, musí být menší než jedna:

$$\lambda_{1,2} = \frac{\mu(Z)}{2} \pm \frac{\sqrt{\mu(Z)^2 - 4\omega\mu(Q)}}{2}$$

Pro komplexní kořeny platí

$$\mu(Z)^2 - 4\omega\mu(Q) < 0, \text{ a tedy } |\mu(Z)| < 2\sqrt{\omega\mu(Q)}, \text{ neboli:}$$

$$|\omega - c + \mu(Q)| < 2\sqrt{\omega\mu(Q)}$$

Jelikož dle výchozího předpokladu je $\omega > 0$ a také střední hodnota náhodné veličiny $\mu(Q) > 0$, kořeny charakteristické rovnice k rovnici (101) jsou komplexní pro hodnoty koeficientu $c \approx a_{max}$ nacházejícího se v intervalu

$$\omega + \mu(Q) - 2\sqrt{\omega\mu(Q)} < c < \omega + \mu(Q) + 2\sqrt{\omega\mu(Q)} \quad (102)$$

Podmínku stability pro komplexní kořeny stanoví autorka z podmínky, že modul $|\lambda_{1,2}| < 1$

$$|\lambda_{1,2}| = \sqrt{\frac{\mu(Z)^2}{4} + \frac{4\omega\mu(Q) - \mu(Z)^2}{4}} = \sqrt{\omega\mu(Q)} < 1$$

$$\omega < \omega_{max} \quad \omega_{max} * \mu(Q) = 1 \quad \omega_{max} = \frac{1}{\mu(Q)} = \frac{1}{2 \ln(2)} = \frac{1}{2 * 0,693147} \cong 0,72$$

Realizace náhodné proměnné Z budou kladnými čísly při splnění podmínky (100):

$$c = \frac{\omega + 1}{2} = \frac{0,72 + 1}{2} = 0,86$$

Autorka ještě ověří, zda hodnota $c = 0,86$ vyhovuje také podmínce (102):

$$\omega_{max} + 2 \ln(2) - 2\sqrt{\omega_{max} * 2 \ln(2)} = \frac{1}{2 * 0,6931} + 2 * 0,6931 - 2\sqrt{1} \cong 0,10$$

$$\omega_{max} + 2 \ln(2) + 2\sqrt{\omega_{max} * 2 \ln(2)} = \frac{1}{2 * 0,6931} + 2 * 0,6931 + 2\sqrt{1} \cong 4,10$$

$$0,10 < 0,86 < 4,10$$

a tedy hodnota $c = 0,86$ podmínce (102) vyhovuje.

Pro reálné kořeny platí $|\mu(Z)| > 2\sqrt{\omega\mu(Q)}$, a tedy

$$|\omega - c + \mu(Q)| > 2\sqrt{\omega\mu(Q)}$$

Jelikož $\omega > 0$ a také střední hodnota náhodné veličiny $\mu(Q) > 0$, kořeny charakteristické rovnice k rovnici (101) jsou reálné pro hodnoty koeficientu $c \approx a_{max}$ nacházejícího se vně intervalu (102). Mají-li být současně splněny podmínky $c > 0$, $\mu(Q) = 2 \ln(2)$, $\omega < 0$, je možný rozsah koeficientu c pro reálné kořeny nepatrný, a proto autor studie [37] volbu reálných kořenů z hlediska stability neuvažuje.

4.4.3 Porovnání s výsledky plynoucími z deterministické analýzy

Autor [37] uvažuje hodnotu c jako maximální hodnotu pro koeficienty a_l, a_g v rovnicích (1), (18) a (19) plynoucí ze stochastické analýzy stagnace částice a porovnává tuto hodnotu $c \approx c_{max} \approx a_l \approx a_g$ s maximální hodnotou odpovídajícího koeficientu, která vyplývá z deterministické analýzy. Označme deterministicky získaný koeficient jako c_d . Pro výpočet hodnoty koeficientu c_d uvádí autor [37] vztah:

$$c_d = \frac{(1 + \chi_{max})^2}{2} \approx \frac{(1 + \omega_{max})^2}{2} \cong \frac{(1 + 0,72)^2}{2} \cong 1,48 \quad (103)$$

Vzorec (103) má být odvozen z následujících vztahů plynoucích z deterministické analýzy stability systému částice:

$$\omega_{max} \approx \chi_{max} = \frac{1}{|\mu_2|} = \frac{2}{\psi_n^i - 2 + \sqrt{\psi_n^{i^2} - 4\psi_n^i}} \approx \chi$$

$$\chi_{max} * \psi_n^i = 2c_d$$

Jelikož je takto stanovená hodnota koeficientu c_d navrhována jako standardní hodnota pro koeficienty $a_l \approx a_g$ [4] a odvození vzorce (103) je ve studii [37] opět pouze naznačeno, vzorec pro koeficient c_d autorka předkládané práce odvodila v dodatku D10.

Zpětným dosazením autorka pro úplnost ještě ověří, že hodnotě $c_d = 1,48$ spočtené podle vzorce (103) odpovídá hodnota $\psi_{n_{max}}^i = 4,1 = (2,05 + 2,05)$:

$$\begin{aligned} \omega_{max} \psi_{n_{max}}^i &= \omega_{max} (\psi_{1n_{max}}^i + \psi_{2n_{max}}^i) = \omega_{max} (\psi_{1n_{max}}^i + \psi_{2n_{max}}^i) = \\ &0,72 * (2,05 + 2,05) = 1,48 + 1,48 = c_d + c_d = 2,96 \cong 4,1 * 0,72 \end{aligned}$$

4.4.4 Převzaté závěry experimentálních ověření výsledků stochastické analýzy stagnace nejlepší částice

K prověření výsledků stochastické analýzy stagnace nejlepší částice navrhl autor [37] celkem pět variant algoritmu *PSO* a prováděl experimenty na účelových funkcích *Sphere* (*Parabola*), *Griewank*, *Rosenbrock*, *Rastrigin* a *Ackley*.³⁷ Jako kritérium úspěšnosti si autor určil počet evaluací dané funkce ve srovnání s počtem evaluací funkce *Sphere*, která je použita jako referenční.³⁸

První dvě varianty algoritmu *PSO* jsou implementovány jako simulace rovnic (1) a (2). Jako koeficienty v rovnici (1) jsou použity hodnoty získané stochastickou analýzou.

První testovaná varianta algoritmu *PSO* je implementována s koeficienty $\omega = 0,72$, $a_l = a_g = 0,86$:

$$\begin{aligned} v_n^i(t + \Delta t) &= 0,72 v_n^i(t) + 0,86 u_{1n}^i(t) [l_n^i(t) - x_n^i(t)] \Delta t \\ &\quad + 0,86 u_{2n}^i(t) [g_n(t) - x_n^i(t)] \Delta t \end{aligned}$$

Ve druhé testované variantě jsou použity koeficienty $a_l = a_g$, které odpovídají hodnotě c v situaci, kdy střední hodnota “dopředné síly” Z je rovna střední hodnotě “zpětné síly” Q :

$$\mu(Z) = \mu(Q)$$

$$\omega - c + 2 \ln(2) = \omega 2 \ln(2)$$

³⁷ Informativní grafické znázornění funkce *Sphere* nalezne čtenář v sekci *Příloha A*.

³⁸ Účelová funkce *Sphere* má pouze jedno optimum, které je globální a je algoritmem *PSO* nalezeno vždy.

$$\frac{1}{2 \ln(2)} - c + 2 \ln(2) = \frac{1}{2 \ln(2)} 2 \ln(2)$$

$$c = 2 \ln(2) + \frac{1}{2 \ln(2)} - 1 \cong 1,108$$

$$\begin{aligned} v_n^i(t + \Delta t) = & 0,72 v_n^i(t) + 1,108 u_{1n}^i(t) [l_n^i(t) - x_n^i(t)] \Delta t \\ & + 1,108 u_{2n}^i(t) [g_n(t) - x_n^i(t)] \Delta t \end{aligned} \quad (106)$$

Ve třetí variantě jsou koeficienty $a_l = a_g$ voleny ze dvou hodnot $a_l = a_g = 0,86$ nebo $a_l = a_g = 1,48$, a to náhodným výběrem s pravděpodobností 0,5.

Ve čtvrté variantě jsou použity stejné hodnoty koeficientů jako ve druhé variantě, ale mění se způsob výběru $g_n(t)$ pro nejlepší částici. Pokud je detekováno $l_n^i(t) = g_n(t)$, zvolí se za $g_n(t)$ jiná, náhodně vybraná pozice ze zbývajících nejlepších lokálních pozic, ale pouze pro tuto částici. Tím má být odvrácena stagnace této částice.

Pátá varianta algoritmu *PSO* se zaměřuje na “šum” D . Jsou použity stejné hodnoty koeficientů jako ve druhé variantě. Je-li detekována stagnace, je rychlost spočtena nikoli podle rovnice (106), ale přímo podle rovnice (97):

$$v_n^i(t + 1) = (\omega - s + q)v_n^i(t) - \omega q v_n^i(t - 1) + d(l - g) = 0$$

Pátá varianta algoritmu je nejkomplikovanější, neboť k výpočtu realizací proměnných S , Q v rovnici (97) si algoritmus musí pamatovat realizace všech náhodných proměnných v časech t , $t - 1$ a také je třeba dosadit realizaci náhodné veličiny D . Simulací bylo zjištěno, že náhodná veličina D má hustotu pravděpodobnosti (*probability density function pdf*) podobnou *Gaussově křivce*, a tedy za d je dosazena realizace náhodné veličiny normálního rozložení $U(0, \sigma)$, kde σ je standardní odchylka, která je spočtena z charakteristiky náhodné veličiny D (95) jako její směrodatná odchylka.

Protože se výkonnost páté varianty algoritmu ve srovnání s variantou druhou nezlepšila, nebude autorka předkládané práce pátou variantu algoritmu více rozvádět. Ze všech testovaných variant vykazovala zlepšenou výkonnost pouze varianta druhá oproti variantě první, další testované varianty již výkonnost druhé varianty více nezlepšily.

Potud stochastická analýza stagnace částice obsažená ve studii [37].

4.4.5 Otázky k zodpovězení

Autor studie [37] odvodil, jaké mají být střední hodnoty realizací náhodných veličin v rovnici (1), aby zaúčinkoval mechanismus odvrácení stagnace, pokud částice není nejlepší částicí v roji, a přitom nedošlo k divergenci částice.

Autorka předkládané disertační práce zaznamenává, že mezi hodnotami koeficientů $a_l = a_g = 0,86$ získanými ze stochastické analýzy stagnace částice a hodnotami koeficientů $a_l = a_g = 1,48$ získanými z deterministické analýzy stability systému částice je patrný značný rozdíl, přesto je hodnota $a_l = a_g = 1,48$ navrhována jako součást *standardu* algoritmu *PSO* [4].

Naskýtá se rovněž otázka, proč se autor [37] snaží odvracet stagnaci nejlepší částice v roji pomocí čtvrté varianty algoritmu. Je zřejmé, že použitý mechanismus vyvede částici ze stagnující pozice, která nemusí být globálním, ale dokonce ani lokálním optimelem.³⁹ Pokud se však výkonnost čtvrté varianty algoritmu oproti druhé variantě nezlepšila, znamená to snad, že popsáný mechanismus vyvedení částice ze stagnující pozice je nadbytečný, protože je prohledávací schopnost testované varianty algoritmu dostatečná? Vysvětlení, proč se výkonnost čtvrté varianty algoritmu nezlepšila, ve studii [37] chybí.

Chování systému částice popsaného v kapitole 4.2 diferenčními rovnicemi (52), (59), (61), (63) a (64) bylo podmiňováno tím, že systém částice dospěje do stádia stagnace, kdy se pozice částice již nemění (57). Všechny rovnice mají stejnou charakteristickou rovnici, a tedy podmínky vnitřní stability pro stagnující částici platí i pro částici v pohybu.

Stagnace částice je totiž úkaz pro algoritmus *PSO* přirozený a také žádoucí, pokud pozice, k níž nejlepší částice směřuje a u níž se zastavila, je skutečnou globální nejlepší pozicí. *Jsou však podmínky mezní stability prvního řádu pro správnou konvergenci částice dostačující?*

Zůstává rovněž otázkou, proč mají být hodnoty koeficientů a_l , a_g v rovnici (1) shodné. *Je snad možno zajistit stabilitu chování systému částice nejlépe při shodných hodnotách uvedených koeficientů, nebo se jedná o záměrné zjednodušení problému?*

Na vnitřní stabilitu systému částice má vliv i hodnota časového okamžiku Δt . Časový okamžik Δt se vyskytuje ve druhé mocnině v diferenční rovnici (96):

$$v_n^i(t + \Delta t) = (\omega - s\Delta^2 t - q)v_n^i(t) - \omega q v_n^i(t - \Delta t) + d\Delta t(l - g) = 0$$

Proměnná hodnota Δt má tedy vliv na očekávané hodnoty náhodných veličin a je třeba dalšího, podrobnějšího zkoumání vnitřní stability systému částice algoritmu *PSO* jako dynamického systému i z pohledu Δt . *Vliv časového okamžiku na vnitřní stabilitu systému částice by měl být znám* přinejmenším k posouzení možnosti zavedení Δt jako případného dalšího ladícího parametru do algoritmu *PSO*.

³⁹ To by znamenalo, že algoritmus *PSO* vlastně není optimalizátorem, neboť nemá zaručenu konvergenci (nelze stanovit podmínky, za kterých algoritmus *PSO* vždy konverguje).

4.5 Vztahy mezi koeficienty plynoucí ze stochastické analýzy standardní odchylky distribuce vzorků pozic stagnující částice *PSO*

Stochastická analýza stagnace částice obsažená ve studii [37], jež je předmětem kapitoly 4.4, zkoumá mezní podmínky vnitřní stability systému částice zohledňující očekávané hodnoty náhodných veličin $E[u_n^i]$, $E[u_n^i]$ v rovnicích (1), (18) a (19). Očekávané hodnoty náhodných veličin jsou momenty prvního řádu, jejichž realizace musí splňovat diferenční rovnice algoritmu *PSO*. Autorka předkládané disertační práce nyní bude hledat odpověď na otázku, *zda je zkoumání momentů prvního řádu tak, aby se systém blížil k mezi stability, pro zajištění správné konvergence částic ke globálnímu optimu dostačující.*

Každá částice se při svém průletu hyperprostorem pohybuje po trajektorii, která je popsána posloupností rychlostí a pozic této částice. Trajektorie částic jsou z hlediska stochastické analýzy stability chápány jako náhodné funkce a teoreticky je z uvedeného hlediska možno zkoumat i momenty druhého a vyšších řádů.

Na možnost šetření dynamiky momentů vyššího řádu jako první poukazuje studie s názvem “*Mean and Variance of the Sampling Distribution of the Particle Swarm Optimizers During Stagnation*” [38] vydaná v časopise *IEEE Transactions on Evolutionary Computation* v roce 2009. Autor studie [38] *Ricardo Poli* analyzuje standardní odchylku distribuce vzorků pozic částice během stagnace a ukazuje, že pro splnění podmínek vnitřní stability systému částice musí být splněny dvě podmínky, a sice: (i) střední hodnota pozic trajektorie částice $E[r_n^i(t)]$ musí v budoucím vývoji systému konvergovat k nějaké stacionární hodnotě (56), (57) podle podmínek vnitřní systémové stability systému částice, uvedených v kapitole 4.3.1, ale navíc (ii) směrodatná odchylka pozic trajektorie $\sigma(r_n^i(t))$ musí konvergovat k nule.

Autor studie [38] vychází z popisu trajektorie částice jako stochastického systému, zkoumá dynamiku momentů druhého řádu a hledá stacionární body soustavy diferenčních rovnic pro momenty druhého řádu. Odtud vyvozuje podmínku pro směrodatnou odchylku pozic trajektorie.

Popis trajektorie částice jako stochastického systému je obsahem kapitoly 4.5.1. Dynamikou momentů druhého řádu se zabývá kapitola 4.5.2 s podrobnostmi v dodatku D11. Postup odvození podmínky pro směrodatnou odchylku pozic trajektorie částice je nastíněn v kapitole 4.5.3. Podrobné pojednání je v dodatcích D12 a D13.

4.5.1 Trajektorie částice *PSO* jako stochastický systém popsany polohovým vektorem a dynamika momentů prvního řádu

Popis trajektorie částice *PSO* jako stochastického systému je ve studii [38] odvozen z diferenční rovnice obsahující pozici částice v časech $(t - 1)$, (t) , a $(t + 1)$ ve dvou po sobě jdoucích iteracích, v níž figurují proměnné φ_{1n}^i , φ_{2n}^i jako realizace náhodných veličin rovnoměrného rozložení (náhodné proměnné). Na tuto diferenční rovnici je potom aplikován operátor očekávaných hodnot E .

Diferenční rovnici uvedenou v [38] jako rovnici popisující trajektorii i -té částice v její n -té dimenzi odvodila autorka předkládané práce z rovnice (53), kam dosadila $o_n^i(t)$ (19). Obě rovnice zde pro přehlednost autorka opakovaně uvádí:

$$r_n^i(t+1) + (\varphi_n^i - (\omega + 1))r_n^i(t) + \omega r_n^i(t-1) = \varphi_n^i o_n^i(t)$$

$$o_n^i(t) = \frac{\varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)}{\varphi_{1n}^i + \varphi_{2n}^i}$$

Po dosazení a úpravách vznikne rovnice popisující trajektorii částice ve třech po sobě jdoucích časech:

$$r_n^i(t+1) + (\varphi_{1n}^i + \varphi_{2n}^i - (\omega + 1))r_n^i(t) + \omega r_n^i(t-1) = \varphi_n^i \frac{\varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)}{\varphi_{1n}^i + \varphi_{2n}^i}$$

$$r_n^i(t+1) + (\varphi_{1n}^i + \varphi_{2n}^i)r_n^i(t) - (\omega + 1)r_n^i(t) + \omega r_n^i(t-1) = \varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)$$

$$r_n^i(t+1) = (\omega + 1)r_n^i(t) - (\varphi_{1n}^i + \varphi_{2n}^i)r_n^i(t) - \omega r_n^i(t-1) + \varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)$$

Protože na tuto rovnici nyní bude ve shodě s autorem [38] aplikován operátor očekávaných hodnot E a následně budou vyjadřovány také momenty druhého řádu, a tedy budou používány mocniny, bude vhodné pro přehlednost zjednodušit notaci zápisu tak, že pro libovolnou částici i v její libovolné dimenzi n budou použity následující substituce a označení:

$$\begin{aligned} r_n^i(t+1) &\approx r_{t+1} & r_n^i(t) &\approx r_t & r_n^i(t-1) &\approx r_{t-1} \\ \varphi_{1n}^i &\approx \varphi_1 & \varphi_{2n}^i &\approx \varphi_2 \\ l_n^i(t) &\approx l & g_n(t) &\approx g \\ (\omega + 1) &\approx \tilde{\omega} \end{aligned} \tag{107a}$$

Zápis diferenční rovnice trajektorie částice s použitím substitucí je následující:

$$r_{t+1} = \tilde{\omega}r_t - (\varphi_1 + \varphi_2)r_t - \omega r_{t-1} + \varphi_1 l + \varphi_2 g \tag{107}$$

Dynamiku momentů prvního řádu vyjádřil autor [38] aplikací operátoru očekávaných hodnot E na rovnici (107):

$$E[r_{t+1}] = \tilde{\omega}E[r_t] - E[\varphi_1 r_t] - E[\varphi_2 r_t] - \omega E[r_{t-1}] + E[\varphi_1]l + E[\varphi_2]g$$

Náhodné proměnné φ_1 , φ_2 , r_t jsou statisticky nezávislé, proto lze psát:

$$E[r_{t+1}] = \tilde{\omega}E[r_t] - E[\varphi_1]E[r_t] - E[\varphi_2]E[r_t] - \omega E[r_{t-1}] + E[\varphi_1]l + E[\varphi_2]g$$

Proměnná φ_1 je realizací náhodné veličiny rovnoměrného rozložení $(0, a_l)$, a tedy $E[\varphi_1] = a_l/2$, analogicky $E[\varphi_2] = a_g/2$:

$$E[r_{t+1}] = \tilde{\omega}E[r_t] - \frac{a_l}{2}E[r_t] - \frac{a_g}{2}E[r_t] - \omega E[r_{t-1}] + \frac{a_l}{2}l + \frac{a_g}{2}g \tag{108}$$

Autor studie [38], podobně jako autor studie [37], předpokládá, že koeficienty a_l , a_g v rovnici (108) jsou shodné a jsou rovny nějaké maximální hodnotě c .

Autorka předkládané práce poznamenává, že stabilita rovnice (107) byla již zkoumána ve studii [35], která je obsahem kapitoly 4.3.2, a ve studii [36], která je obsahem kapitoly 4.3.3. Autor studie [38] na tyto předcházející výzkumy navazuje a zabývá se pouze podmínkami stability druhého řádu.

4.5.2 Dynamika momentů druhého řádu

Pro zkoumání podmínek stability druhého řádu popisuje autor [38] systém částice jako rozšířený systém prvního řádu, a to formálně pomocí soustavy diferenčních rovnic

$$\mathbf{z}(t+1) = \mathbf{M}\mathbf{z}(t) + \mathbf{b} \quad (109)$$

kde vektory $\mathbf{z}(t+1)$, $\mathbf{z}(t)$ obsahují nejen momenty prvního, ale i druhého řádu:

$$\begin{aligned} \mathbf{z}(t) &= (E[r_t] \ E[r_{t-1}] \ E[r_t^2] \ E[r_t r_{t-1}] \ E[r_{t-1}^2])^T \\ \mathbf{z}(t+1) &= (E[r_{t+1}] \ E[r_t] \ E[r_{t+1}^2] \ E[r_{t+1} r_t] \ E[r_t^2])^T \end{aligned} \quad (110)$$

a vektor $\mathbf{b}$ sdružuje zbylé výrazy, tj. výrazy, které neobsahují žádný z momentů E , jimiž je tvořen vektor $\mathbf{z}(t)$. Jedná se o nehomogenní, afinní systém diferenčních rovnic, kde vektor $\mathbf{b}$ představuje vstup. K popisu systému pomocí matice systému $\mathbf{M}$ (109) potřeboval autor studie [38] získat diferenční rovnice pro momenty obsažené ve vektoru $\mathbf{z}(t+1)$.

Podrobný postup získání diferenčních rovnic, který autor [38] ve své studii použil, je autorkou předkládané práce rekapitulován v dodatku D11. Stručný popis uvedeného postupu je následující:

Diferenční rovnicí pro moment $E[r_{t+1}]$ obsahující $E[r_t]$, $E[r_{t-1}]$, je rovnice (108).

Diferenční rovnicí pro moment druhého řádu $E[r_{t+1} r_t]$ získal autor [38] vynásobením rovnice (107) výrazem r_t a následnou aplikací operátoru E . Vzhledem k předpokladu $a_l = a_g$ dosadil autor v obou případech $E[\varphi_1] = E[\varphi_2] = c/2$.

Pro získání diferenční pro $E[r_{t+1}^2]$ spočetl autor [38] nejdříve výraz r_{t+1}^2 pomocí rovnice (107):

$$r_{t+1}^2 = (\tilde{\omega} r_t - \varphi_1 r_t - \varphi_2 r_t - \omega r_{t-1} + \varphi_1 l + \varphi_2 g)^2$$

a potom seskupil odpovídající výrazy a současně aplikoval operátor E .

Výslednou soustavu rovnic zapsal autor [38] v maticovém tvaru (109). Zápis výsledné soustavy je rovněž uveden v dodatku D11.

Autor studie [38] následně spočetl vlastní hodnoty matice $\mathbf{M}$, která je pátého řádu, a to pomocí blíže neurčeného maticového kalkulátoru poskytujícího výstupy v symbolické podobě. Tyto hodnoty, jež jsou kořeny charakteristického polynomu soustavy, musí být pro splnění podmínek vnitřní stability v absolutní hodnotě menší než jedna.

Autor [38] řešil kořeny charakteristického polynomu soustavy rovnic numericky tak, že pro zvolené hodnoty $[c, \omega]$ spočetl mezní (nejvyšší možnou) hodnotu modulu kořene a výsledky znázornil v grafu [38:pp.715-716]. Takto vzniklou mezní křivku potom porovnal

s mezní křivkou pro diferenční rovnici (108) vymezující oblast stability prvního řádu, kterou stanovil obdobným postupem.

Autor [38] výše uvedeným postupem graficky ukázal, že oblast stability druhého řádu leží kompletně uvnitř oblasti stability prvního řádu. Pokud bude zvolena dvojice $[c, \omega]$ ležící mezi vymezujícími křivkami těchto oblastí, bude splněna pouze podmínka analogická s podmínkou (56), v níž $o_n^i(t)$ je dáno vztahem (19).

Hodnotu stacionárního bodu $E[r_t^*]$ určil autor z rovnice (108), v níž položil $(\omega + 1) \approx \tilde{\omega}$ podle (107a) a $E[r_{t+1}] \approx E[r_t] \approx E[r_{t-1}]$:

$$\begin{aligned} E[r_t^*] &= (\omega + 1)E[r_t^*] - \frac{a_l}{2}E[r_t^*] - \frac{a_g}{2}E[r_t^*] - \omega E[r_t^*] + \frac{a_l}{2}l + \frac{a_g}{2}g \\ E[r_t^*] \left(1 - \left(\omega + 1 - \frac{a_l + a_g}{2} - \omega \right) \right) &= \frac{a_l l + a_g g}{2} \\ E[r_t^*] &= \frac{a_l l + a_g g}{a_l + a_g} \approx c \frac{l + g}{2c} = \frac{l + g}{2} = p \end{aligned} \quad (116)$$

4.5.3 Stacionární body soustavy diferenčních rovnic pro momenty druhého řádu

Hlubšího pochopení dynamického chování částice v závislosti na hodnotách koeficientů c, ω lze dle autora [38] dosáhnout nalezením stacionárního bodu soustavy diferenčních rovnic pro momenty druhého řádu, podobně jako tomu bylo v případě momentů prvního řádu a jak je také popsáno v kapitole 4.2.2.

Pokud systém popsáný maticí M (109) splňuje podmínky stability 2. řádu, potom existuje stacionární bod $\mathbf{z}^*$ (110) takový, že

$$\lim_{t \rightarrow \infty} \mathbf{z}(t) = \mathbf{z}^*(t)$$

Pro stacionární bod $\mathbf{z}^*(t)$ tedy platí:

$$\mathbf{z}^*(t+1) \approx \mathbf{z}^*(t) \approx \mathbf{z}^*(t-1)$$

Dosazením do rovnice (109):

$$\mathbf{z}^*(t) = M\mathbf{z}^*(t) + \mathbf{b}$$

Zavedeme novou proměnnou $\tilde{\mathbf{z}}(t)$ jako vzdálenost vektoru $\mathbf{z}(t)$ od stacionárního bodu:

$$\tilde{\mathbf{z}}(t) = \mathbf{z}(t) - \mathbf{z}^*(t)$$

$$\tilde{\mathbf{z}}(t+1) = \mathbf{z}(t+1) - \mathbf{z}^*(t)$$

Dosadíme-li opět do rovnice (109):

$$\tilde{\mathbf{z}}(t+1) + \mathbf{z}^*(t) = M(\tilde{\mathbf{z}}(t) + \mathbf{z}^*(t)) + \mathbf{b}$$

můžeme dynamiku systému částice znázornit homogenní rovnicí k rovnici (109) pomocí nové proměnné $\tilde{\mathbf{z}}$:

$$\tilde{\mathbf{z}}(t+1) = M\tilde{\mathbf{z}}(t)$$

Tato homogenní soustava má stejný charakteristický polynom jako soustava (109) a platí:

$$\mathbf{z}^*(t) = \left(\lim_{t \rightarrow \infty} E[r_t] \quad \lim_{t \rightarrow \infty} E[r_{t-1}] \quad \lim_{t \rightarrow \infty} E[r_t^2] \quad \lim_{t \rightarrow \infty} E[r_t r_{t-1}] \quad \lim_{t \rightarrow \infty} E[r_{t-1}^2] \right)^T =$$

$$= (p \quad p \quad ? \quad ? \quad ?)^T$$

Otazníky jsou označeny hodnoty, které je třeba spočítat. Autor [38] vychází při výpočtu z úvahy, že pro $t \rightarrow \infty$ jsou hodnoty $E[r_t]$, $E[r_{t-1}]$ téměř nerozlišitelné od stacionárního bodu p . Protože oblast stability druhého řádu leží uvnitř oblasti stability prvního řádu, autor studie [38] usuzuje, že pokud konverguje $E[r_t]$, $E[r_{t-1}]$ k bodu p , bude rovněž konvergovat $E[r_t^2]$, $E[r_t r_{t-1}]$ a $E[r_{t-1}^2]$, a sice k bodům, které označil p_{r^2} a p_{rr} tak, že:

$$\mathbf{z}^*(t) = (p \quad p \quad p_{r^2} \quad p_{rr} \quad p_{r^2})^T$$

Rovněž tak $E[r_{t+1}^2]$ bude konvergovat k p_{r^2} . Za uvedených předpokladů odvodil autor [38] výrazy pro stacionární body p_{r^2} a p_{rr} . Postup autora [38] při odvozování výrazů p_{r^2} a p_{rr} je podrobně rozveden v dodatku D12.

Dále si autor studie [38] položil otázku, jaká je podmínka pro standardní odchylku pozic trajektorie systému částice od jejich očekávaných hodnot

$$\sigma(r_t) = \sqrt{E[r_t^2] - (E[r_t])^2}$$

má-li částice konvergovat, to jest, má-li platit $p_{r^2} = p^2$, a odvodil vztah:

$$\sqrt{p_{r^2} - p^2} = \sqrt{2 \frac{(v - \mu^2)}{\Delta}} |p - l|$$

kde $\mu \approx E[\varphi_1]$ je střední hodnota proměnné φ_1 a $v \approx E[\varphi_1^2]$ je její směrodatná odchylka. Po dosazení za $p - l$, jak plyne ze (116):

$$g = 2p - l \quad p - l = \frac{g - l}{2}$$

$$\sigma(r_t) = \sqrt{p_{r^2} - p^2} = \frac{1}{2} \sqrt{2 \frac{(v - \mu^2)}{\Delta}} |g - l| \quad (119)$$

Výraz $v - \mu^2$ pod odmocninou na pravé straně rovnice (119) je roven standardní odchylce náhodné proměnné $\varphi_1 \approx \varphi_2$ vzhledem k předpokladu rovnosti koeficientů a_l , a_g :

$$\mu \approx E[\varphi_1] = a_l/2 = c/2 \approx E[\varphi_2] = a_g/2 = c/2 \quad (120)$$

$$v \approx E[\varphi_1^2] = E[\varphi_2^2] = c^2/3 \quad (121)$$

Postup autora [38] při odvozování rovnice (119) je autorkou předkládané práce rozveden v dodatku D12.

Výraz Δ v rovnici (119) je substitucí pro výraz obsahující proměnné ω , μ , v , c . Upravený výraz pro Δ je následujícího tvaru:

$$\Delta = c * \frac{12(1 - \omega^2) + c(5\omega - 7)}{6(\omega + 1)} \quad (122)$$

Postup úpravy výrazu Δ v rovnici (119) je uveden v dodatku D13.

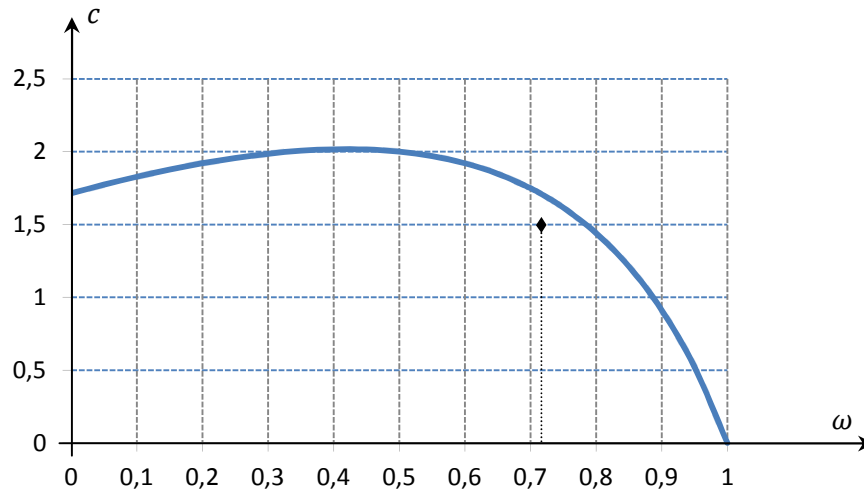
Mezní hodnoty, při které poroste odchylka $\sigma(r_t)$ (119) neomezeně, bude dosaženo při hodnotě $\Delta = 0$:

$$12(1 - \omega^2) + c(5\omega - 7) = 0$$

Vztah mezi c a ω udává křivku, která představuje horní hranici (*bare bone*), pod níž musí ležet dvojice $[c, \omega]$, aby systém částice popsany rovnicemi (1) a (2) s danou dvojicí koeficientů byl stabilní.

$$c = \frac{12(\omega^2 - 1)}{5\omega - 7} \quad (123)$$

Mezní křivka je vykreslena na obrázku Obr. 9:



Obr. 9: Horní hranice stability (*bare bone*) 2. řádu pro dvojice $[c, \omega]$.

Pro dvojici $[c, \omega] = [1.48, 0.72]$ vykazuje systém stabilitu prvního i druhého řádu.

$$c = a_l = a_g \text{ v rovnici (1), } c = 12(\omega^2 - 1)/(5\omega - 7)$$

Dosažením $[c, \omega] = [1.48, 0.72]$ do (123) lze ověřit, že daný bod leží uvnitř oblasti stability 2. řádu:

$$\omega = 0,72$$

$$c = \frac{12(0,72^2 - 1)}{5 * 0,72 - 7} \cong 1,69 > 1,48$$

Aby tedy částice konvergovala, nestačí pouze požadovat, aby očekávané hodnoty pozic trajektorie systému částice se v limitě blížily ke stacionárnímu bodu $p = o_n^i(t)$ analogicky s podmínkou (56):

$$\lim_{t \rightarrow \infty} E[r_t] = p \quad (125)$$

Musí rovněž platit, že standardní odchylka $\sigma(r_t)$ (118) pozic trajektorie systému částice r_t od jejich očekávaných hodnot se v limitě musí blížit k nule:

$$\lim_{t \rightarrow \infty} \sigma(r_t) = 0 \quad (126)$$

Autor [38] dovedl, že tuto podmínku lze splnit pouze v případě, že lokální nejlepší pozice částice je rovna globální nejlepší pozici, neboli $g = l$, a že tedy nelze podat důkaz konvergence při $g \neq l$.

Další otázkou položenou autorem studie [38] je, zda dvojice $[c, \omega]$ plynoucí z analýz stability prvního řádu [35, 36, 37] leží také uvnitř oblasti stability 2. řádu. Bude-li totiž $\Delta = 0$ v rovnici (118), potom $\sigma(r_t)$ vzroste do nekonečna.

Autor [38] spočetl pro dvojici $[c, \omega] = [1.49618, 0.7298]$, že $\sigma(r_t) = 1,0428 * |g - l|$, a $\Delta \neq 0$, a tedy existuje stacionární bod a systém částice vykazuje rovněž stabilitu 2. řádu.

Výpočet je podrobně rozveden v dodatku D13. Autorka předkládané práce zde zaznamenává nepatrný rozdíl ve výsledku, který je způsoben odchylkou v hodnotách $[c, \omega] = [1.47876, 0.7213,]$, které jsou odvozeny z deterministické analýzy a které pro výpočet použila autorka.

4.5.4 Otázky k zodpovězení

Autor [38] ve své studii ukázal, že pro stagnaci částice nestačí požadavek (125), aby střední hodnota pozic trajektorie částice konvergovala k nějaké stacionární hodnotě. Musí rovněž platit, že *standardní odchylka (variance)* pozic trajektorie částice od jejich očekávaných hodnot se musí blížit nule. Avšak stagnace částice vede ke konvergenci pouze za podmínky, že částice je nejlepší částicí v roji, a nelze tedy podat obecný důkaz konvergence. Autor [38] rovněž poukazuje, že v předpokladu stagnace (126) je zcela eliminován vliv účelové funkce a bylo by třeba vyzkoumat, *pro jaký typ účelové funkce systém stagnuje*.

4.6 Vztahy mezi koeficienty plynoucí ze stochastické analýzy stability trajektorií částice zobecněného algoritmu *PSO*

Dalšími stěženy otázkami položenými v kapitole 4.4.5, které je třeba zodpovědět, je otázka rozdílnosti hodnot koeficientů a_l, a_g v rovnici (1) a otázka vlivu časového kroku Δt na vnitřní stabilitu systému částice.

Hodnota koeficientů je při stanovování mezních podmínek stability prvního řádu [36, 37] a podmínek stability 2. řádu [38] uvažována jako shodná, rovna určité maximální hodnotě c . Otázka zněla, zda je možno zajistit stabilitu chování systému částice nejlépe při shodných hodnotách uvedených koeficientů, nebo se jedná o záměrné zjednodušení problému.

Vliv na očekávané hodnoty náhodných veličin, a tedy na vnitřní stabilitu systému částice algoritmu *PSO*, má i časový krok Δt . Tento vliv by měl být prozkoumán vzhledem k posouzení možnosti zavedení Δt jako případného dalšího ladícího parametru do algoritmu *PSO*.

Pro další, podrobnější zkoumání vnitřní stability systému částice algoritmu *PSO* jako dynamického systému byl navržen model algoritmu *PSO* nazvaný zobecněný model *PSO* (*Generalized PSO*, *GPSO*), který se od originálního algoritmu *PSO* [2] odlišuje tím, že připouští rozdílné hodnoty koeficientů a_l, a_g (1) i jiný časový krok než $\Delta t = 1$.

Zobecněná forma algoritmu *PSO* [39] připouštějící rozdílné koeficienty a zavádějící časový krok jako další ladící parametr byla uveřejněna v časopise *Journal of Artificial Evolution and Application* v roce 2008 pod názvem “*The Generalized PSO: A New Door to PSO Evolution*”, jejímiž autory jsou J. L. Fernández Martínez a E. García Gonzalo. Algoritmus autoři podrobněji rozpracovali a výsledky testování vydali v roce 2009 v časopise *Swarm Intelligence* pod názvem “*The PSO family: deduction, stochastic analysis and comparison*”.

Chování systému částice ve studiích [39, 40] je podobně jako ve studiích [37, 38] podmínované tím, že systém částice dospěje do stádia stagnace a chování je shodně popisováno jako *stabilita trajektorie částice*. Zobecněný model *GPSO* slouží zejména pro vyšetřování oblastí stability vyšších řádů. Smyslem tohoto dalšího zkoumání je objevit další souvislosti, a sice mezi koeficienty ω, a_l, a_g , a současně podchytit vliv časového kroku Δt . Za tímto účelem je zaveden nový parametr α jako poměr mezi globálním zrychlením a celkovým středním zrychlením, kdy $\alpha \in [0,2]$:

$$\alpha = \frac{a_g}{E(\varphi_n^i)} = \frac{2a_g}{a_l + a_g} \quad (127a)$$

Ve studiích [39, 40] je stejně jako ve studiích [36, 37, 38] zkoumána vzdálenost mezi trajektoriemi $o_n^i(t)$ a $r_n^i(t)$, která se má s postupujícím časem zmenšovat. Tato vzdálenost je shodně s předpoklady ve studiích [37, 38] ovlivňována tím, že trajektorie $o_n^i(t), r_n^i(t)$ jsou stochastické procesy, jejichž statistické momenty musí splňovat diferenční rovnice algoritmu *PSO*. Oproti studii [37], která zkoumá podmínky stability pro momenty prvního řádu, a studii [38] zkoumající momenty druhého řádu, autoři [39, 40] zkoumají podmínky stability i pro momenty třetího a čtvrtého řádu.

Zápis algoritmu *PSO* pomocí diferenčních rovnic obsahujících $t, \Delta t$ je v [39, 40] chápán jako zobecnění *spojitého modelu PSO* (*PSO continuous model*), v němž proměnné $t, \Delta t$

mohou být jakákoli kladná reálná čísla. Spojitý model *PSO* charakterizovaný jako tlumený pružinový systém o jednotkové hmotnosti byl autory [39, 40] převzat z [41].

Zobecněný model algoritmu *PSO* je charakterizován v kapitole 4.6.1. Odpovídající analogie spojitého modelu je popsána v kapitole 4.6.2. Vztahy mezi trajektoriemi zobecněného modelu jsou předmětem kapitoly 4.6.3.

4.6.1 Zobecněný model algoritmu *PSO* (*Generalized PSO, GPSO*)

Popis zobecněného modelu algoritmu *PSO* (*Generalized PSO, GPSO*) pomocí diferenčních rovnic analogických k rovnicím (21) a (2), z nichž studie [39, 40] vycházejí, odvodila autorka předkládané práce zpětnou transformací rovnice (51), která je zde pro přehlednost uvedena opakovaně:

$$a_n^i(k\Delta t) - \beta v_n^i(k\Delta t) - \varphi_n^i[o_n^i(k\Delta t) - r_n^i(k\Delta t)] = 0$$

Zpětná transformace spočívá ve vyjádření zrychlení a_n^i v rovnici (51) jako přírůstku rychlosti v_n^i podle (23) a diskretizaci času $t = k\Delta t$:

$$\frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} - \beta v_n^i(t) - \varphi_n^i[o_n^i(t) - r_n^i(t)] = 0$$

$$v_n^i(t + \Delta t) - v_n^i(t) - \beta v_n^i(t)\Delta t - \varphi_n^i\Delta t[o_n^i(t) - r_n^i(t)] = 0$$

V této rovnici je podle (27) koeficient $\beta = (\omega - 1)/\Delta t$, jak byl zaveden autory fyzikálního modelu [35]. Zpětným dosazením:

$$v_n^i(t + \Delta t) - v_n^i(t) - \frac{(\omega - 1)}{\Delta t} v_n^i(t)\Delta t - \varphi_n^i\Delta t[o_n^i(t) - r_n^i(t)] = 0$$

Do této rovnice dosadila autorka nikoli za koeficient $\beta = (\omega - 1)/\Delta t$ podle (27), jak jej zavedli autoři fyzikálního modelu [35], ale použila koeficient

$$\beta_{\Delta t} = (w - 1) \tag{127}$$

preferovaný autory [39] tak, aby zůstala zachována rovnost $\beta\Delta t = \beta_{\Delta t}\Delta t$:

$$\frac{(\omega - 1)}{\Delta t} \Delta t = (w - 1)\Delta t$$

Odtud:

$$(\omega - 1) = (w - 1)\Delta t$$

Dosazením za $(\omega - 1)$ a následnými úpravami

$$v_n^i(t + \Delta t) - v_n^i(t) - (w - 1)v_n^i(t)\Delta t - \varphi_n^i\Delta t[o_n^i(t) - r_n^i(t)] = 0$$

$$v_n^i(t + \Delta t) - v_n^i(t) - wv_n^i(t)\Delta t + v_n^i(t)\Delta t - \varphi_n^i\Delta t[o_n^i(t) - r_n^i(t)] = 0$$

získala autorka rovnici pro rychlost částice zobecněného modelu algoritmu *PSO*:

$$v_n^i(t + \Delta t) - (1 + w\Delta t - \Delta t)v_n^i(t) - \varphi_n^i\Delta t[o_n^i(t) - r_n^i(t)] = 0$$

Zápis diferenčních rovnic zobecněného modelu algoritmu *PSO* je následující:

$$v_n^i(t + \Delta t) = (1 - (1 - w)\Delta t)v_n^i(t) + \varphi_n^i \Delta t [o_n^i(t) - r_n^i(t)] \quad (128)$$

$$r_n^i(t + \Delta t) = r_n^i(t) + v_n^i(t + \Delta t)\Delta t \quad (129)$$

Rovnice (129) je analogií rovnice (2). Rovnice (128) je analogií rovnice (21), avšak přechází do rovnice (21) pouze při hodnotě $\Delta t = 1$. Je též evidentní, že fyzikální model zobecněného algoritmu popsaného rovnicemi (128) a (129) má jinou definici tlumícího faktoru než fyzikální model popsaný rovnicemi (21), (2), což autorka ukáže tak, že rovnici (128) podrobí opačné transformaci:

$$\begin{aligned} v_n^i(t + \Delta t) - v_n^i(t) + \Delta t v_n^i(t) - w \Delta t v_n^i(t) + \varphi_n^i \Delta t r_n^i(t) &= \varphi_n^i \Delta t o_n^i(t) \\ \frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} + (1 - w)v_n^i(t) + \varphi_n^i r_n^i(t) &= \varphi_n^i o_n^i(t) \\ \frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} &\approx a_n^i(t) \\ a_n^i(t) + (1 - w)v_n^i(t) + \varphi_n^i r_n^i(t) &= \varphi_n^i o_n^i(t) \end{aligned} \quad (130)$$

Rovnici (130) by bylo možno dále rozvinout dosazením za $\varphi_n^i = \varphi_{1n}^i + \varphi_{2n}^i$.

4.6.2 Algoritmus *GPSO* jako analogie tlumeného pružinového systému

Autoři studie [39, 40] použili k vyšetřování stability trajektorie částice rovnici (130), která je z fyzikálního hlediska analogií tlumeného pružinového systému o jednotkové hmotnosti převzatého z [41]:

$$x''(t) + (1 - w)x'(t) + \varphi x(t) = \varphi_1 l(t) + \varphi_2 g(t) \quad (131)$$

kde $x(t)$ je trajektorie libovolné částice v libovolné dimenzi, $l(t)$ a $g(t)$ jsou lokální a globální nejlepší trajektorie, přičemž $l(t)$ je různá pro různé částice, $(1 - w)$ je tlumící faktor (*damping factor*), $\varphi = \varphi_1 + \varphi_2$ je celková tuhost pružiny (*stiffness constant*) a φ , φ_1 , φ_2 jsou náhodné proměnné.

Trajektorie generované částicemi spojitěho modelu jsou stochastické procesy splňující diferenciální rovnice druhého řádu (130), případně (131).

Porovnáním tlumícího faktoru v rovnici (51) s tlumícím faktorem v rovnicích (130) a (131) získáme vztah mezi koeficienty ω , w vyjádřený pomocí (127):

$$\begin{aligned} w - 1 &= \frac{\omega - 1}{\Delta t} \\ \omega &\approx \omega_{\Delta t} = 1 - (1 - w)\Delta t \end{aligned} \quad (132)$$

Zobecněný model algoritmu *PSO* odpovídá spojitěmu modelu v tom smyslu, že pro rovnice (128), (129) lze zvolit libovolný diskrétní časový krok Δt .

4.6.3 Vztahy mezi trajektoriemi v modelu algoritmu *GPSO*

Vztahy mezi trajektoriemi $o_n^i(t)$, $r_n^i(t)$ v modelu *GPSO* určuje rovnice analogická k rovnici (52), v níž $\omega \approx \omega_{\Delta t}$ podle (132):

$$r_n^i(t + \Delta t) + (\varphi_n^i \Delta^2 t - \omega_{\Delta t} - 1)r_n^i(t) + \omega_{\Delta t} r_n^i(t - \Delta t) = \varphi_n^i \Delta^2 t o_n^i(t) \quad (133)$$

Lze nahlédnout, že rovnice (133) je formálně shodná s rovnicí (52), pokud nahradíme $\omega \approx \omega_{\Delta t}$. Z odvození rovnice (53) lze vysledovat, že pro časový krok $\Delta t = 1$ je uplatněna substituce

$$\varphi_n^i \Delta^2 t \approx \varphi_n^i$$

Podobně, jako byla odvozena rovnice (52) a další vztahy v kapitole 5.2, lze odvodit analogické vztahy pro model *GPSO*, které byly autory [39, 40] využity pro vyšetřování stability trajektorií částic modelu *GPSO*.

Ve studii [40] je rovnice (133) uvedena ve tvaru:

$$r_n^i(t + \Delta t) - A r_n^i(t) - B r_n^i(t - \Delta t) = \varphi_n^i \Delta^2 t o_n^i(t) \quad (134)$$

Ve studii [39] je rovnice (133) uvedena ve tvaru:

$$r_n^i(t + \Delta t) = A r_n^i(t) + B r_n^i(t - \Delta t) + C(t) \quad (135)$$

A je v obou případech náhodná proměnná nezávislá na $r_n^i(t)$, $r_n^i(t - \Delta t)$:

$$A = -(\varphi_n^i \Delta^2 t - \omega_{\Delta t} - 1) = 2 - (1 - w)\Delta t - \varphi_n^i \Delta^2 t \quad (136)$$

B je v obou případech proměnná nezávislá na $r_n^i(t)$, $r_n^i(t - \Delta t)$:

$$B = -\omega_{\Delta t} = (1 - w)\Delta t - 1 \quad (137)$$

a $C(t)$ je stochastický proces, který závisí na proměnné $o_n^i(t)$:

$$C(t) = \varphi_n^i \Delta^2 t o_n^i(t) \quad (138)$$

$$o_n^i(t) = \frac{\varphi_{1n}^i l(t) + \varphi_{2n}^i g(t)}{\varphi_{1n}^i + \varphi_{2n}^i}$$

$$C(t) = (\varphi_{1n}^i l(t) + \varphi_{2n}^i g(t)) \Delta^2 t$$

Polohový vektor r_n^i popisují autoři [39, 40] též pomocí proměnné ζ_n^i zavedené analogicky s proměnnou ξ_n^i (58)⁴⁰

$$\zeta_n^i(t) = r_n^i(t) - o_n^i(t) \quad (139)$$

Vztahy mezi trajektoriemi $\zeta_n^i(t + \Delta t)$, $\zeta_n^i(t)$, $\zeta_n^i(t - \Delta t)$ v modelu *GPSO* potom určuje rovnice analogická k rovnici (59), která byla odvozena za podmínky (56), což je limitní podmínka navozující stav stagnace:

$$\zeta_n^i(t + \Delta t) + (\varphi_n^i \Delta^2 t - (\omega_{\Delta t} + 1))\zeta_n^i(t) + \omega_{\Delta t} \zeta_n^i(t - \Delta t) = 0 \quad (140)$$

⁴⁰ Proměnná ξ_n^i je zavedena jako rozdíl $o_n^i(t) - r_n^i(t)$, to jest v obráceném pořadí proměnných, avšak na rovnice určující vztahy mezi trajektoriemi tato okolnost nemá vliv.

4.6.3.1 Deterministická oblast stability trajektorie částice zobecněného algoritmu *GPSO*

Deterministická oblast stability prvního řádu zobecněného modelu algoritmu *PSO* je ve studiích [39, 40] vyšetřována v závislosti na parametrech w , φ_n^i a Δt . Pozice ζ_n^i je zavedena podle (139). Spojitý algoritmus *PSO* je jako zjednodušený dynamický systém popsán soustavou rovnic analogicky s rovnicemi (63), (64):

$$v_n^i(t + \Delta t) = (1 - (1 - w)\Delta t)v_n^i(t) + \varphi \xi_n^i(t)\Delta t \quad (141)$$

$$\xi_n^i(t + \Delta t) = -(1 - (1 - w)\Delta t)v_n^i(t)\Delta t + (1 - \varphi\Delta^2 t)\xi_n^i(t) \quad (142)$$

kde $\Delta t \in (0, 0.25, \dots, 1)$ a $\varphi_n^i = \varphi_{1n}^i + \varphi_{2n}^i$ je náhodná proměnná nabývající reálných hodnot, jejíž střední hodnota je:

$$E(\varphi_n^i) = \frac{a_l + a_g}{2}$$

Pro stanovení deterministické oblasti stability autoři [39] předpokládají, že náhodná proměnná φ_n^i nabývá středních hodnot v závislosti na koeficientech a_l, a_g . Záměrem autorů je získání informace, jak se projevuje závislost na měnícím se časovém kroku Δt .

Systém popsáný rovnicemi (141) a (142) je kompletně definován maticí M podle (65), v níž je zahrnut i časový krok, neboť časový krok v případě zobecněného algoritmu již není implicitně roven jedné:

$$M = \begin{bmatrix} 1 - (1 - w)\Delta t & \varphi_n^i \Delta t \\ -(1 - (1 - w)\Delta t)\Delta t & 1 - \varphi_n^i \Delta^2 t \end{bmatrix}$$

s charakteristickou rovnicí spočtenou podle (62):

$$\lambda^2 + (\varphi_n^i \Delta^2 t - 2 + (1 - w))\lambda + (1 - (1 - w)\Delta t) = 0$$

Deterministická oblast stability algoritmu je pro tento případ:

$$S_D = \left\{ (w, \varphi_n^i) : 1 - \frac{2}{\Delta t} < w < 1, 0 < \varphi_n^i < \frac{1}{\Delta^2 t} (2\Delta t w - 2\Delta t + 4) \right\}$$

Rozhraní mezi komplexními a reálnými kořeny určuje parabola:

$$(\varphi_n^i \Delta^2 t + (1 - w)\Delta t - 2)^2 = 4(1 - (1 - w)\Delta t)$$

Úplné odvození závislosti $[w, \varphi_n^i]$ je obsaženo v dodatku D14.

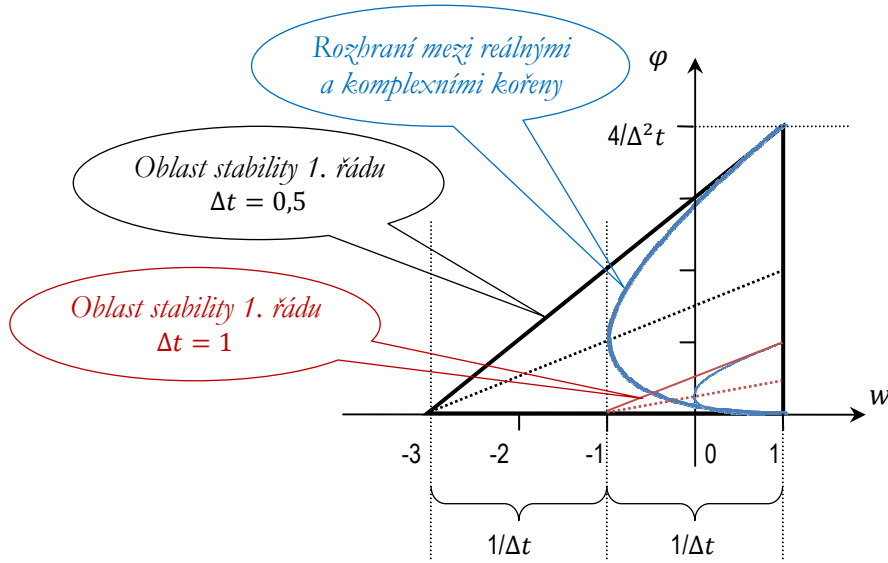
Lze nahlédnout, že pro $\Delta t = 1$ tato deterministická oblast stability zobecněného algoritmu *GPSO* splyne s deterministickou oblastí stability původního algoritmu *PSO*:

$$S = \{(\omega, \varphi_n^i) : |\omega| < 1, 0 < \varphi_n^i < 2(\omega + 1)\}$$

kdy $w \approx \omega$ podle (132) a rozhraní mezi komplexními a reálnými kořeny určuje parabola:

$$(\varphi_n^i - (\omega + 1))^2 - 4\omega = 0$$

Pro zmenšující se časový krok Δt se oblast stability zvětšuje. Situaci znázorňuje obrázek Obr. 9:



Obr. 10: Změna deterministické oblasti stability zobecněného algoritmu *GPSO* v prostoru koeficientů $[w, \varphi]$ v závislosti na Δt .

4.6.3.2 Oblast stability momentů prvního řádu trajektorie částice zobecněného algoritmu *GPSO*

Východiskem pro stanovení oblasti stability momentů prvního řádu trajektorie částice zobecněného algoritmu *GPSO* [40] je rovnice (135), na niž je aplikován operátor E , a momenty prvního řádu:

$$E(r_n^i(t + \Delta t)) = E(A)E(r_n^i(t)) + E(B)E(r_n^i(t - \Delta t)) + E(C(t))$$

Odpovídající nehomogenní (afinní) dynamický systém má formální zápis pomocí soustavy diferenčních rovnic

$$\mu(t + \Delta t) = A_\mu \mu(t) + b(t)$$

kde vektory $\mu(t + \Delta t)$, $\mu(t)$ obsahují momenty prvního řádu, A_μ je matice systému:

$$\begin{bmatrix} E(r_n^i(t + \Delta t)) \\ E(r_n^i(t)) \end{bmatrix} = \begin{bmatrix} E(A) & E(B) \\ 1 & 0 \end{bmatrix} \begin{bmatrix} E(r_n^i(t)) \\ E(r_n^i(t - \Delta t)) \end{bmatrix} + \begin{bmatrix} E(C(t)) \\ 0 \end{bmatrix} \quad (143)$$

Spočteme charakteristickou rovnici podle (62):

$$\det \left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} E(A) & E(B) \\ 1 & 0 \end{bmatrix} \right) = \begin{vmatrix} \lambda - E(A) & -B \\ -1 & \lambda \end{vmatrix} = \lambda^2 - E(A)\lambda - B$$

$$\lambda^2 - E(A)\lambda - B = 0$$

Dosazením A (136), B (137):

$$\lambda^2 + (\varphi_n^i \Delta^2 t - \omega_{\Delta t} - 1)\lambda + \omega_{\Delta t} = 0$$

Tutéž charakteristickou rovnicí obdrželi autoři [39] pro diferenční rovnici (140).

Oblast stability momentů prvního řádu je vymezena:

$$S = \{(\omega_{\Delta t}, \varphi_n^i \Delta^2 t) : |\omega_{\Delta t}| < 1, 0 < E(\varphi_n^i \Delta^2 t) < 2(\omega_{\Delta t} + 1)\}$$

Po dosazení za $\omega_{\Delta t} = 1 - (1 - w)\Delta t$ (132):

$$S = \left\{ \left(w, E(\varphi_n^i) \right) : 1 - \frac{2}{\Delta t} < w < 1, 0 < E(\varphi_n^i) < \frac{1}{\Delta^2 t} (2\Delta t w - 2\Delta t + 4) \right\}$$

Rozhraní mezi reálnými a komplexními kořeny udává limitní parabola:

$$(\varphi_n^i \Delta^2 t - \omega_{\Delta t} - 1)^2 = 4\omega_{\Delta t}$$

Po dosazení za $\omega_{\Delta t} = 1 - (1 - w)\Delta t$ (132):

$$(\varphi_n^i \Delta^2 t + (1 - w)\Delta t - 2)^2 = 4(1 - (1 - w)\Delta t)$$

Úplné odvození závislosti $[\omega_{\Delta t}, \varphi_n^i] \approx [w, \varphi_n^i]$ je v dodatku D14. Lze nahlédnout, že oblast stability momentů prvního řádu zobecněného algoritmu odpovídá deterministické oblasti, jak ji znázorňuje obrázek Obr. 10.

4.6.3.3 Oblast stability momentů druhého řádu trajektorie částice zobecněného algoritmu *GPSO*

Východiskem pro stanovení oblasti stability druhého řádu je ve studii [40] rovnice (135), na niž je aplikován operátor E , a dále necentralizovaný moment druhého řádu $\mu_2(t + \Delta t)$, jehož složkami jsou druhé mocniny a/nebo součin momentů prvního řádu:

$$E(r_n^i(t + \Delta t)) = E(A)E(r_n^i(t)) + E(B)E(r_n^i(t - \Delta t)) + E(C(t))$$

$$\mu_2(t + \Delta t) = A_\sigma \mu_2(t) + b(t)$$

$$\mu_2(t + \Delta t) = \begin{bmatrix} E(r_n^{i^2}(t + \Delta t)) \\ E(r_n^i(t + \Delta t)r_n^i(t)) \\ E(r_n^{i^2}(t)) \end{bmatrix}$$

$$A_\sigma = \begin{bmatrix} E(A^2) & 2E(AB) & E(B^2) \\ E(A) & E(B) & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (144)$$

$$\boldsymbol{\mu}_2(t) = \begin{bmatrix} E(r_n^{i^2}(t)) \\ E(r_n^i(t)r_n^i(t - \Delta t)) \\ E(r_n^{i^2}(t - \Delta t)) \end{bmatrix}$$

$$\mathbf{b}(t) = \begin{bmatrix} E(C^2(t)) + 2E(AC(t)r_n^i(t)) + 2E(BC(t)r_n^i(t - \Delta t)) \\ E(C(t)r_n^i(t)) \\ 0 \end{bmatrix}$$

Pro zkoumání podmínek stability druhého řádu popsali autoři [39, 40] systém částice jako rozšířený systém prvního řádu popsany pomocí matice A pátého řádu, sestavené z matic A_μ (143), A_σ (144):

$$M = \begin{bmatrix} E(A) & E(B) & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & E(A^2) & 2E(AB) & E(B^2) \\ 0 & 0 & E(A) & E(B) & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Autoři dosadili za prvky matice M výrazy (136), (137) a (138), v nichž nejdříve nahradili $E(\varphi_n^i) \approx \varphi$ proměnnou a_g/α podle (127a) následně získali vlastní hodnoty matice M . Tyto hodnoty, jež jsou kořeny charakteristického polynomu soustavy, musí být pro splnění podmínek vnitřní stability menší než jedna. Řešením se autoři dopracovali ke stanovení oblasti stability druhého řádu zobecněného algoritmu *GPSO*:

$$S^2 = \left\{ (\omega, E(\varphi_n^i)) : 1 - \frac{2}{\Delta t} < w < 1, 0 < E(\varphi_n^i) < \varphi(w, \alpha, \Delta t) \right\} \quad (145)$$

kde $\varphi_n^i(w, \alpha, \Delta t)$ je analytické vyjádření limitní hyperboly [39:p.9], [40:p.255]. Tuto křivku pro parametry $\alpha = 1$ (127a) a časový krok $\Delta t = 1$ sestrojil rovněž autor [38]. Limitní hyperboly vymezují oblasti stability druhého řádu pro dvojice $[w, \varphi]$ pro různé parametry α , podobně jako přímka $\varphi = 2(\omega + 1)$ na obrázku Obr. 7(b) vymezuje oblast stability 1. řádu pro $\alpha = 1$, to jest při rovnosti koeficientů v rovnici (1) $a_l = a_g = c$. Oblast stability 2. řádu je obsažena v oblasti stability 1. řádu, závisí symetricky na α a dosahuje maximální velikosti pro $\alpha = 1$, kdy $a_l = a_g$.

4.6.4 Převzaté závěry experimentálních ověření výsledků stochastické analýzy stability trajektorie částice algoritmu *GPSO*

Teoretické výsledky zkoumání stability trajektorií zobecněného algoritmu *GPSO* ukazují, že oblasti stability druhého, ale i třetího a čtvrtého řádu závisí na setrvačnosti, lokálním a globálním zrychlení a na velikosti časového kroku $\Delta t \in (0,1)$ a jsou pro tutéž velikost časového kroku obsaženy v oblasti stability prvního řádu. Oblasti stability 3. i 4. řádu dosahují pro daný časový krok Δt svého maxima rovněž pro $\alpha = 1$, to jest při rovnosti koeficientů $a_l = a_g$. Pro nižší časový krok se oblast stability zvětšuje. Tímto jsou zodpovězeny otázky rovnosti hodnot koeficientů a vlivu časového kroku na vnitřní

stabilitu systému částice, položené v úvodu kapitoly. Autorka předkládané disertační práce podotýká, že autoři předcházejících studií důvod pro volbu koeficientů $a_l = a_g$ explicitně neuvádějí.

Vyvstává další otázka, a sice *zda je volba koeficientů zaručující maximální velikost oblasti vnitřní stability systému částice též zárukou nejlepší prohledávací schopnosti*. Autoři [39, 40] provedli sérii experimentů s účelovými funkcemi *Rosenbrock*, *Rastrigin* a *Griewank* [39] a později rozšířili experimenty na funkce *Ackley* a *DeJong* [40].⁴¹ Výsledky experimentů *pro uvedené funkce* potvrzují, že nízké hodnoty α pro tutéž hodnotu φ_n^l značí, že lokální zrychlení je větší než globální zrychlení, a tedy že algoritmus je více explorativní. Autoři doporučují pro $\Delta t = 1$ volbu koeficientu $w \approx \omega$ v rozmezí 0,5 ... 0,8. V počáteční fázi prohledávání doporučují volbu hodnoty $\alpha < 1$, to jest volbu $a_l > a_g$. Přestože je oblast stability pro $a_l = a_g$ největší, prohledávací schopnost je lepší pro volbu $a_l > a_g$ a závisí na testované účelové funkci. V závěrečné fázi prohledávání doporučují autoři [39, 40] dvě strategie, a sice (i) volbu hodnoty Δt v rozmezí 0,8 ... 0,9, (ii) volbu hodnoty Δt střídavě pro lichou a sudou iteraci 1,2 a 0,8.

Autoři [39, 40] tedy potvrzují výsledky raných experimentů s metodami *TVAC* a *TVIW*, jak jsou popsány v kapitole 3.3.2 s tím, že ve všech případech mají hodnoty $E(\varphi)$, w , α , Δt splňovat podmínky stability dané oblastí S^2 (145).

Pro praktické aplikace bude volba $\Delta t = 1$. Autorka předkládané disertační práce postrádá odvození jednoduchého vztahu mezi $[\varphi, w]$ s parametrem α v podobě mezní křivky obdobné křivce na obrázku Obr. 9, avšak pro dvojici $[\varphi, w]$. Vztah bude autorkou odvozen a použit v návrhu řešení.

⁴¹ Informativní grafické znázornění zmiňovaných účelových funkcí nalezne čtenář v sekci *Příloha A*.

4.7 Kritické zhodnocení současného stavu poznání v oblasti analytického stanovení hodnot koeficientů

Dle poznatků autorky předkládané disertační práce získaných z *meta-analýzy* současného stavu poznání ve sledované oblasti, prezentovaných v kapitolách 4.1 až 4.6, se analytické postupy stanovení koeficientů algoritmu *PSO* obsažené v literatuře odvíjejí z fyzikálního modelu chování systému částice jako tlumeného harmonického oscilátoru periodicky kmitajícího na mezi stability, a také všechny popisy systému částice použité v analýzách vycházejí z tohoto fyzikálního modelu [33].

Vodítkem pro závěrečné kritické zhodnocení současného stavu poznání v oblasti analytického stanovení koeficientů algoritmu *PSO* budou pro autorku otázky, které si položila v kapitole 4.2.5, a sice (i) zda výše uvedená představa chování systému částice odpovídá skutečnému chování částice *PSO*, (ii) zda je zde záruka, že bod v hyperprostoru, v němž se ustálí trajektorie pomyslného středu přitažlivosti $\sigma_n^i(t)$ (19), je jediným atraktorem, a sice globálním optimem, (iii) zda by chování systému částice nemělo být interpretováno spíše jako chaotické, (iv) jaký vliv na chování částice mají náhodné koeficienty v rovnici (1) pro aktualizaci rychlosti částice a (v) zda splnění limitní podmínky (56) zaručí, že nedojde k divergenci částice.

V analytickém přístupu ke stanovení hodnot koeficientů a_l, a_g a jejich vztahu ke koeficientu ω vysledovala autorka shodu v základních úvahách analytiků, že algoritmus bude pracovat efektivně, pokud hodnoty ω, a_l, a_g budou zvoleny tak, aby algoritmus pracoval ve stabilním režimu, který se blíží mezi stability.⁴² Příliš tlumený systém bude vykazovat nedostatečnou vyhledávací (explorativní) schopnost a při překročení meze stability dojde k prudkému nárůstu rychlosti částic a explozi roje.

Analytikové se shodují, že oblast stability je největší při shodných hodnotách koeficientů a_l, a_g . Ohledně maximální vyhledávací schopnosti ve vztahu k maximální oblasti stability však mezi analytiky shoda nepanuje. První skupina analytiků ztotožňuje mez maximální oblasti stability s maximální prohledávací schopností a navrhuje tudíž řešení, v nichž jsou koeficienty ω, a_l, a_g neměnné po celou dobu simulace a koeficienty a_l, a_g mají navíc stejnou, a to maximální možnou hodnotu. Druhá skupina analytiků připouští, že prohledávací schopnost může být lepší pro volbu $a_l \neq a_g$ i za cenu menší oblasti stability, a tedy většího omezení pro hodnoty koeficientů.

Postupy stanovení hodnot koeficientů spadající do první skupiny zahrnují (i) deterministickou analýzu stability chování systému částice [36] (kapitola 4.3.3) a (ii) stochastickou analýzu stagnace nejlepší částice [37] (kapitola (4.4) a [38] (kapitola 4.5).

Do druhé skupiny spadají (i) deterministická analýza konvergence [35] (kapitola 4.3.2) a (ii) stochastická analýza chování systému částice zobecněného algoritmu *GPSO* [39, 40] (kapitola 4.6), jež navazují na výsledky stochastických analýz [37] a [38], avšak připouštějí proměnné koeficienty.

Kritické připomínky k verzi algoritmu *PSO* s konstantními koeficienty jsou předmětem kapitoly 4.7.1, zhodnocení verze algoritmu s proměnnými koeficienty je obsahem kapitoly 4.7.2, otázky interpretace dynamického modelu částice jsou pojednány v kapitole 4.7.3.

⁴² Dynamický systém pracující na mezi stability je ještě pokládán za stabilní.

4.7.1 Stabilita a konvergence verze algoritmu *PSO* s konstantními koeficienty

Ve verzi algoritmu *PSO* [4] s konstantními koeficienty odvozenými z deterministické analýzy [36], jak je popsáno v kapitole 4.3.3, má být rychlost částice aktualizována podle rovnice (88):

$$v_n^i(t + \Delta t) = \chi v_n^i(t) + \chi \psi_{1n}^i [l_n^i(t) - r_n^i] \Delta t + \chi \psi_{2n}^i [g_n(t) - r_n^i] \Delta t$$

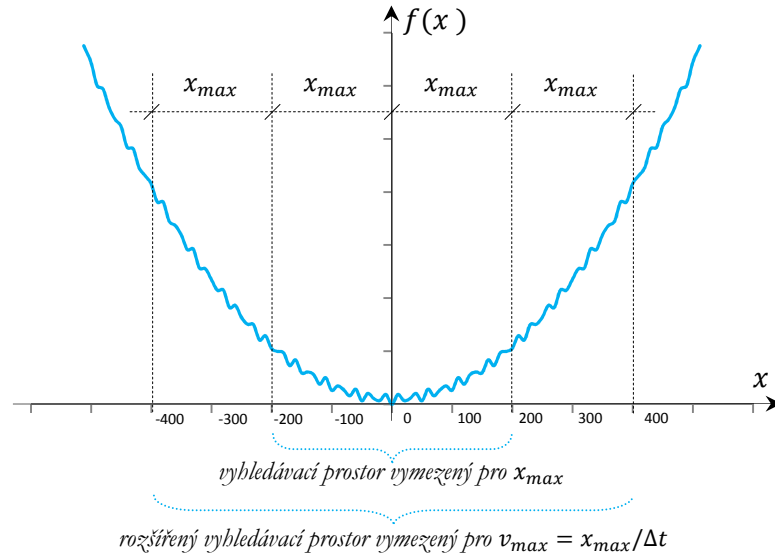
kde $\psi_n^i = (\psi_{1n}^i + \psi_{2n}^i) = (a_l u_{1n}^i + a_g u_{2n}^i)$ (1). Pro $\psi_n^i \leq 4$ má být zvoleno $0 < \chi < 1$ (86) a pro $\psi_n^i > 4$ se hodnota χ spočte podle vzorce (87):

$$\chi = 2k / \left(\psi_n^i - 2 + \sqrt{\psi_n^{i^2} - 4\psi_n^i} \right)$$

Konstanta k má být volena v intervalu $0 < k \leq 1$. Koeficient χ má maximum pro $k = 1$. Této maximální konfiguraci odpovídá trojice koeficientů $(\chi, a_l, a_g) = (0.72, 1.48, 1.48)$.

Model algoritmu *PSO* s konstantními koeficienty by měl dle autorů [36], [4] vykazovat stabilitu tak, že částice má dovoleno opustit vyhledávací prostor, nemělo by však dojít k explozi roje. R. C. Eberhart a Y. Shi provedli již v roce 1999⁴³ experimenty s uvedenou verzí algoritmu [5], a sice s účelovými funkcemi *Sphere*, *Rosenbrock*, *Rastrigin*, *Griewank* a *Schaffer's f6* (Příloha A). Jelikož algoritmus vykazoval nestabilitu, rozhodli se autoři [5] pro zavedení dodatečného omezení pro rychlost ve tvaru $v_{max} = x_{max}/\Delta t$, kde pro x_{max} byla nejprve zvolena relativně vysoká hodnota a poté hodnota odpovídající vyhledávacímu prostoru testované účelové funkce. Při hodnotě x_{max} odpovídající vyhledávacímu prostoru bylo dle autorů [5] pozorováno výrazné zlepšení stability.

Situaci znázorňuje obrázek Obr. 11:



Obr. 11: Rozšíření vyhledávacího prostoru $(-2x_{max} < x < 2x_{max})$ pro účelovou funkci $f(x)$ *Griewank* (Příloha A).

⁴³ Publikace [36] je z roku 2002, ale metoda konstrikce později označená jako “Constriction Type 1” byla známa a publikována v prezentacích již v roce 1998.

Kritické připomínky autorky předkládané disertační práce k variantě algoritmu PSO s konstantními koeficienty jsou následující:

- (i) Z průběhu testovaných funkcí, jejichž profily jsou znázorněny v sekci *Příloha A*, je patrné, že k testování stabilních a prohledávacích vlastností algoritmu byly použity funkce, jejichž tendenční křivky⁴⁴ jsou podobné parabole, s globálním optemem v ose křivky, a tedy mechanismus použitý k omezování rychlosti částice by u funkcí, které tuto vlastnost nesplňují (jako je např. funkce *Schwefel*), nemohl vrátit částici zpět. Závislost volby koeficientů na testované účelové funkci analytikové nezkoumají.
- (ii) V analýzách a návrhu standardu [4] je testování prováděno na algoritmu, jenž narušuje *pseudo-paralelní* režim, jak je tento režim popsán v kapitole 3.3.5. Algoritmus [4] tedy mění původní algoritmus [2]. Porovnání algoritmů je na obrázku Obr. 12:

1. Initialize an array of particles with random positions and velocities on D dimensions, 2. Evaluate the desired minimization function in D variables, 3. Compare evaluation with particle's previous best value $PBEST[]$: If current value $< PBEST[]$ then $PBEST[] =$ current value and $PBESTx[] [d] =$ current position in D dimensional hyperspace, 4. Compare evaluation with group's previous best $PBEST[GBEST]$: If current value $< PBEST[GBEST]$ then $GBEST =$ particle's array index, 5. Change velocity by the following formula: $v[] [d] = v[] [d] +$ $ACC \ CONST * rand() * (PBESTx[] [d] - PresentX[] [d]) +$ $ACC \ CONST * rand() * (PBESTx[GBEST][d] - PresentX[] [d]),$ and, 6. Move to $PresentX[] [d] + v[] [d]$: Loop to step 2 and repeat until a criterion is met.	(a)
for each time step t do for each particle i in the swarm do update position $\mathbf{x}(t)$ using equations (1) & (2) calculate particle fitness $f(\mathbf{x}(t))$ update $\mathbf{p}^i, \mathbf{p}_g$ end for end for	(b)
Create and initialize an n-dimensional S repeat for each particle $i \in [1..m]$ if $f(S.\mathbf{x}^i) < f(S.\mathbf{y}^i)$ then $S.\mathbf{y}^i = S.\mathbf{x}^i$ if $f(S.\mathbf{y}^i) < f(S.\hat{\mathbf{y}})$ then $S.\hat{\mathbf{y}} = S.\mathbf{y}^i$ end perform updates on S using equations (1) (2) until stopping condition is true.	(c)

$PBEST[] \approx \mathbf{p}_g \approx S.\hat{\mathbf{y}} \approx \mathbf{g}(t)$ je globální nejlepší pozice (rovnice (1))

$PBESTx[] [d] \approx l_n^i$ je lokální nejlepší pozice libovolné částice i v dimenzi $d \approx n$ (rovnice (1))

$S.\mathbf{y}^i \approx \mathbf{l}^i$ je lokální nejlepší pozice libovolné částice (rovnice (1)), m je počet částic

$ACC \ CONST \approx a_t \approx a_g$ jsou koeficienty zrychlení (rovnice (1))

Obr. 12: (a) Doslovný přepis původního algoritmu PSO [2].

(b) *Sekvenční* algoritmus PSO použitý v návrhu standardu [4] s koeficienty dle (87) a (88).

(c) *Pseudo-paralelní* přepis původního algoritmu [2] převzatý z [26:p.24].

- (iii) U testů provedených v [5] není uvedeno, jak je testovací algoritmus implementován, a z experimentální výsledků obsažených v [5] jednoznačně nevyplývá, zda jsou úniky mimo vyhledávací prostor ze statistického hlediska pouze ojedinělé, anebo se ve skutečnosti jedná o nestabilitu.

⁴⁴ *Tendenční křivkou* míní autorka tzv. *spojnici trendu*, která je pro danou funkci spočtena polynomicky podle implicitního vzorce obsaženého v programu *MS Excel*. Tendenční křivky jsou vykresleny v sekci *Příloha A*.

Změna *pseudo-paralelního* režimu algoritmu PSO na *sekvenci* [4] dle obrázku Obr. 12(b) totiž vede k přednostnímu zpracování informací od částic, které byly při inicializaci vygenerovány dříve, a tedy mají nižší přidělené pořadí. Tato změna algoritmu u funkcí s tendencí k parabole sice povede k rychlejší konvergenci, zrychlené prohledávání vedoucí k případné předčasné konvergenci bude však nežádoucí u funkcí s tendencí přímky, jako je funkce *Schwefel* (Příloha A). Autoři [4] zmírňují riziko předčasné konvergence tím, že navrhují použití modelu *lbest* zpomalujícího konvergenci (kapitola 3.3.2), to však problém dle autorky předkládané disertační práce systémově neřeší, ale pouze zmírňuje.⁴⁵

Nestabilita, kterou algoritmus vykazuje i při použití hodnot koeficientů, jež mají stabilitu zajišťovat, je analytiky shodně přičítána působení náhodných čísel. Analytikové tedy hledají další podmínky stability, a to právě s ohledem na působení náhodných čísel, jejichž zapojení do rozhodování částice je z hlediska prohledávací schopnosti částice nezbytné [35] (*Trelea*, kapitola 4.3.2). Z uvedeného pohledu je další podmínkou stability systému částice, aby střední hodnota pozice trajektorie částice konvergovala k nějaké stacionární hodnotě, a musí rovněž platit, že standardní odchylka (*variance*) pozic trajektorie částice od jejich očekávaných hodnot se musí blížit nule [37] (*Polí*, kapitola 4.5.2). Ačkoli trojice koeficientů $(\chi, a_l, a_g) = (0.72, 1.48, 1.48)$ [36] [4] (*Clerc*, *Kennedy*, kapitola 4.3.3) uvedenou podmínku dle diagramu na obrázku Obr. 9 splňuje, algoritmus [4] přesto vykazuje nestabilitu [5] (*Eberhart*, *Shi*).

4.7.2 Stabilita a konvergence verze algoritmu *PSO* s proměnnými koeficienty

Schopnost prohledávání a stabilita systému částice jsou protichůdné požadavky, neboť nejstabilnější by byl algoritmus s nulovou prohledávací schopností. Snaha analytiků posunout dynamické chování systému částice k mezi stability je tedy logická. Při zapojení náhodných prvků však trojice koeficientů $(\chi, a_l, a_g) = (0.72, 1.48, 1.48)$ (splňující dle [4] maximální mez stability systému částice) dle autorů zobecněného algoritmu *GPSO* [39, 40] nejen stabilitu nezajistí, nezajistí ani nejlepší prohledávací schopnost, jež může být lepší i při rozdílných hodnotách koeficientů $a_l \neq a_g$, a tedy pro menší oblast stability. Autoři [39, 40] současně zkoumají podmínky stability vyšších řádů, neboť nestabilitu systému částice i při splnění podmínek stability 2. řádu [37] přičítají působení náhodných prvků.

Autorka podotýká, že autoři algoritmu *GPSO* [39, 40] rovněž použili k testování stabilních a prohledávacích vlastností algoritmu účelové funkce s tendencí podobnou parabole, a sice funkce *Rosenbrock*, *Rastrigin*, *Griewank*, *Ackley* a *DeJong* (Příloha A). Přesto jsou teoretické závěry i experimentální výsledky ohledně podmínek stability vyšších řádů pro autorku přesvědčivé. Avšak zodpovězení otázky, *čím jsou způsobeny úniky mimo vyhledávací prostor*, zda je lze přičítat pouze vlivu náhodných prvků a zda jsou tyto úniky ze statistického hlediska pouze ojedinělé, anebo se ve skutečnosti jedná o nestabilitu, je dle autorky předkládané disertační práce složitější a souvisí s interpretací modelu částice *PSO*.

⁴⁵ Autorka shledala pouze dva případy verze algoritmu podobné verzi na obrázku Obr. 12(b). Prvním případem je skupina robotů majících za úkol nalezení nejsilnějšího zdroje etanolu [42], avšak zde je paralelní režim realizován přímo fyzickým pohybem robotů ve vyhledávacím prostoru. Druhým případem jsou roje *CPSO-S* [26:p.144] (kapitola 3.3.5), kde *pseudo-paralelní* režim na úrovni jednotlivých rojů je završen sekvenci zpracování přes jednotlivé roje, to je však *deklarovaným* výzkumným záměrem autora [26].

4.7.3 Interpretace dynamického modelu systému částice *PSO*

Interpretace dynamického systému částice podle obrázku Obr. 6, který je grafickým vyjádřením rovnice (51), odpovídá představě tlumeného harmonického oscilátoru, který není buzen (*forced*) či řízen (*driven*) žádnou vnější silou. Rovnovážná poloha $\mathbf{o}_n^i(t)$, jako součást elastické síly, je vyjádřena jako lineární kombinace (superpozice) proměnných $\mathbf{l}_n^i(t)$ a $\mathbf{g}_n(t)$ (19). Tyto proměnné však představují dílčí pozice v hyperprostoru, které jsou výsledkem vyhodnocení hodnot účelové funkce $f(t)$ v pozicích $\mathbf{l}^i(t)$ a $\mathbf{g}(t)$, přičemž *neplatí*, že hodnota účelové funkce v pozicích $\mathbf{l}^i(t) + \mathbf{g}(t)$ je rovna součtu hodnot účelové funkce v jednotlivých pozicích, neboť předpokládáme, že účelová funkce $f(t)$ (*Příloha A*) *není lineární funkcionál*, neboli:

$$f(\mathbf{l}^i(t)) + f(\mathbf{g}(t)) \neq f(\mathbf{l}^i(t) + \mathbf{g}(t))$$

a tedy řešíme (nekorektní) inverzní problém určený operátorem h takový, že

$$h(f(\mathbf{l}^i(t))) = \mathbf{l}^i(t) \quad h(f(\mathbf{g}(t))) = \mathbf{g}(t) \quad (146)$$

Vztahy (146) chtěla autorka předkládané práce naznačit, že pozice $\mathbf{o}_n^i(t)$ závisí na hodnotách účelové funkce, i když tato závislost není ve výchozích rovnicích (51) (52) (53) explicitně matematicky vyjádřena. Dle autorky je nutno modelovat dynamický systém částice nikoli jako tlumený harmonický oscilátor, ale jako buzený oscilátor (*forced oscillator*) popsáný nehomogenní diferencí rovnicí (28), v níž vstupní veličina $\mathbf{m} \cdot \boldsymbol{\varphi}_n^i \mathbf{o}_n^i(t)$ představuje budící sílu. Právě působení budící síly $F_b = \mathbf{m} \cdot \boldsymbol{\varphi}_n^i \mathbf{o}_n^i(t)$ může dle autorky předkládané práce být dalším potenciálním zdrojem nestability, tak jak ji popisují autoři [5].⁴⁶ Problém dle autorky spočívá v tom, že oscilace systému částice nejsou silou F_b řízeny (*driven oscillator*), nýbrž neovladatelně buzeny (*forced oscillator*), neboť účelová funkce, od níž se odvozuje poloha pozice $\mathbf{o}_n^i(t)$, je vstupní, nezávislou veličinou.

Při působení síly F_b je zřejmé, že splnění limitní podmínky (56) nezaručí, že nedojde k divergenci částice, i když podmínky vnitřní stability systému částice, zaručující též přenosovou stabilitu typu *BIBO*, budou splněny.⁴⁷

Chování systému částice vykazuje atributy deterministického chaotického chování,⁴⁸ a tedy není zde záruka, že bod v hyperprostoru, v němž se ustálí trajektorie pomyslného středu přitažlivosti $\mathbf{o}_n^i(t)$, je jediným atraktorem, a sice globálním optimem.

Za přítomnosti náhodných prvků a neznámého průběhu účelové funkce na vstupu do systému (z pozice vnitřního pozorovatele) je dle autorky jediným řešením nežádoucí účinek budící síly F_b předvídat a ovládat systém částice jako komplexní (*complex*) systém [43] tak, aby byl udržován “na hraně (pokraji) chaosu” (*on edge of chaos*).

⁴⁶ Autoři [5] popisují svá pozorování doslovně takto: “Bylo to jako pozorování kosmické lodi prohledávající galaxii Mléčné dráhy s cílem nalezení objektu, o němž je předpokládáno, že se nachází ve Sluneční soustavě.” V souvislosti s nelineárními buzenými oscilátory je podobný jev v literatuře označován idiomem *blue sky catastrophe*.

⁴⁷ I při katastrofickém scénáři [5] stále platí, že pro omezený vstup je výstup též omezen. Vnitřní stabilita a konvergence tedy není totéž.

⁴⁸ Systém částice (i) je citlivý na počáteční podmínky, (ii) jako dynamický systém je popsán třemi stavovými proměnnými, jimiž jsou pozice, rychlost a čas (jako implicitní proměnná), (iii) na vstupu do systému je skryta nelinearita a (iv) mezi vstupem do systému a pozicí částice je potenciální zpětná vazba.

Kapitola 5

Návrh řešení

Účelem *meta-analýzy současného stavu poznání v oblasti analytického stanovení koeficientů algoritmu PSO*, jež byla obsahem čtvrté kapitoly předkládané disertační práce, bylo hlubší pochopení problému *zvýšení skupinové inteligence roje PSO posílením autonomní činnosti jednotlivých částic roje*, kdy rychlosti $\mathbf{v}(t)$ a pozice $\mathbf{x}(t)$ částic jsou i nadále aktualizovány podle rovnic (1) a (2).

Globálním cílem je *zlepšení výkonnosti algoritmu, jehož modelem je originální algoritmus PSO* [1], [2]. Autorka předkládané práce věří, že tohoto cíle dosáhne rozšířením originálního algoritmu PSO o *koncept vnitřního pozorovatele*, který bude vyhodnocovat prostředí a činnost částic a určovat hodnoty koeficientů (ω, a_l, a_g) v rovnici aktualizace rychlosti (1).

Koncept vnitřního pozorovatele byl nastíněn v kapitole 2.1.2. Řešený problém *posílení autonomní činnosti* částic byl podrobně formulován v kapitole 3.6. Postup (strategie) řešení problému bude směřovat k dosažení dílčích cílů stanovených v kapitole 3.7 a zahrnuje následující úlohy vnitřního pozorovatele, jejichž řešení bude autorkou navrženo:

- (i) zajištění dynamické stability systému částice při zachování stochastického charakteru operace roje (ve smyslu kapitoly 1.2.2),
- (ii) s tím související způsob stanovení hodnot koeficientů (ω, a_l, a_g) v rovnici (1),
- (iii) způsob detekce výlučnosti hodnoty účelové funkce vzhledem k ostatním hodnotám, tak jak byla výlučnost specifikována v kapitolách 3.5 a 3.6, a také
- (iv) posouzení *úspěšnosti řešení* jako metriky, dle konceptuálního schématu na obrázku Obr. 3, v souladu s definicí v kapitole 3.6.

V návrhu řešení budou zachovány základní premisy originálního algoritmu PSO, za něž autorka pokládá *pseudo-paralelní způsob zpracování* a rovnice (1) a (2).

5.1 Zajištění dynamické stability systému částice PSO

Z výsledků *meta-analýzy současného stavu poznání v oblasti analytického stanovení koeficientů algoritmu PSO* vyplynulo, že pro udržení dynamické stability systému částice při zachování stochastického charakteru operace roje je zapotřebí udržení *vnitřní stability* systému částice, a to stability prvního i druhého řádu. V kritickém hodnocení výsledků *meta-analýzy* v kapitole 4.7.3 je potom objasněna potřeba odvrácení katastrofického scénáře, který vede k extenzívním únikům částic mimo vyhledávací prostor, kdy udržení vnitřní stability systému částice je podmínka nutná, avšak zjevně, jak ukazují experimenty provedené autory studie [5], nikoli podmínka dostačující.

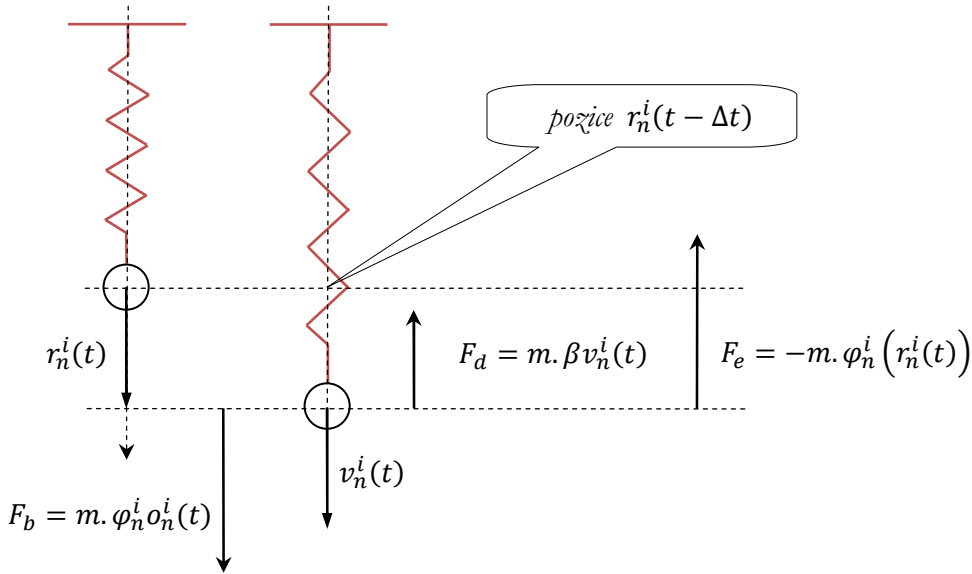
Návrh řešení odvrácení katastrofického scénáře je obsahem kapitoly 5.1.1, kde je současně objasněno, proč se autorka práce přiklání k variantě algoritmu s proměnnými koeficienty. Udržení vnitřní stability pomocí hodnot koeficientů (ω, a_l, a_g) je obsahem kapitoly 5.1.2, kde se autorka inspirovuje výsledky stochastické analýzy v kapitole 4.5.

5.1.1 Odvrácení katastrofického scénáře

Jako první krok k nalezení řešení pro odvrácení katastrofického scénáře navrhuje autorka předkládané disertační práce změnu interpretace dynamického modelu systému částice *PSO* tak, že tlumený harmonický oscilátor dle obrázku Obr. 6 bude nahrazen buzeným tlumeným harmonickým oscilátorem (*forced oscillator*) podle obrázku Obr. 13, a to stále ve shodě s rovnicí (28). Rovnici (28) lze formálně přepsat do tvaru diferenční rovnice 2. řádu, která je vyjádřením rovnice buzeného tlumeného harmonického oscilátoru:

$$m \cdot a_n^i(k\Delta t) - m \cdot \beta v_n^i(k\Delta t) + m \cdot \varphi_n^i r_n^i(k\Delta t) = m \cdot \varphi_n^i o_n^i(k\Delta t) \quad (147)$$

Analogicky s definicí buzeného tlumeného harmonického oscilátoru působí na i -tou částici v její n -té dimenzi odporující síla⁴⁹ $F_d = m \cdot \beta v_n^i(k\Delta t)$ a elastická síla $F_e = -m \cdot \varphi_n^i r_n^i(k\Delta t)$. Oscilátor je buzen (*forced*) silou $F_b = m \cdot \varphi_n^i o_n^i(t)$. Síla $F = m \cdot a_n^i(k\Delta t)$ uděluje částici výsledné zrychlení.



Obr. 13: Částice *PSO* jako buzený oscilátor (*forced oscillator*).

Hmotný bod připojený k elastickému objektu zanedbatelné hmotnosti.

Na i -tou částici *PSO* v její n -té dimenzi působí odporující síla F_d , elastická síla F_e a budící síla F_b , $t = t_0 + k\Delta t$, $\Delta t = 1$, $\Delta r_n^i(t) = r_n^i(t) - r_n^i(t - \Delta t)$

Změna interpretace dynamického modelu systému částice je pouze formální změnou nazírání na řešený problém, neboť podmínky vnitřní stability odvozené ve čtvrté kapitole pro výchozí rovnici (28) zůstávají v platnosti i pro rovnici (147). Jinou interpretací modelu chtěla autorka ukázat, že pozornost je třeba věnovat působení budící síly F_b : (147a)

$$F_b = m \cdot \varphi_n^i o_n^i(t) = m \cdot (\varphi_{1n}^i + \varphi_{2n}^i) \frac{\varphi_{1n}^i l_n^i(t) + \varphi_{2n}^i g_n(t)}{\varphi_{1n}^i + \varphi_{2n}^i} = m \cdot (a_l u_{1n}^i l_n^i(t) + a_g u_{2n}^i g_n(t))$$

⁴⁹ Koeficient $\beta = (\omega(t) - 1)/\Delta t$ (27) je v praxi záporný, neboť $0 < \omega < 1$, odtud záporné znaménko v rovnici (147), oproti obvyklé konvenci.

Velikost budící síly F_b je ovlivňována hodnotou účelové funkce, jež se promítá do pozic $l_n^i(t)$ a $g_n(t)$. Pozice mohou v dané iteraci oproti předchozí iteraci zůstat nezměněny, nastane-li však změna, může tato změna vyvolat výraznou odezvu,⁵⁰ kdy částice opustí vyhledávací prostor. Hodnota účelové funkce tedy způsobuje žádoucí pertubace, které urychlují vývoj systému částice a celého roje, avšak působení silných pertubací, které by vedly částici k opuštění vyhledávacího prostoru, by mohlo způsobit destrukci roje. Řešením je, aby vnitřní pozorovatel silnou pertubaci odhalil a destrukci odvrátil. Pokud vnitřní pozorovatel zjistí, že se částice chystá opustit vyhledávací prostor, tj. že $|x_n^i(t+1)| > x_{n_{max}}^i(t+1)$, zapojí omezovač rychlosti částice $v_n^i(t+1)$ (1) tak, aby pozice částice $x_n^i(t+1)$ (2), vypočtená vzápětí z této rychlosti, zůstala ve vyhledávacím prostoru:

$$\begin{aligned} x_n^i(t+1) &= x_n^i(t) + [v_n^i(t+1) \pm k v_{n_{max}}^i(t+1)]\Delta t \\ x_n^i(t+1) &= x_n^i(t) + v_n^i(t+1)\Delta t \pm k x_{n_{max}}^i(t+1) \end{aligned} \quad (148)$$

kde x_n^i , v_n^i jsou libovolné souřadnice pozice, resp. rychlosti i -té částice, k je konstanta volená tak, aby $0 < |x_n^i(t+1)| < x_{n_{max}}^i$.

Používání omezovače rychlosti částice v algoritmu PSO je zmíněno v kapitole 3.2.2 v souvislosti s vlastností systému označenou jako *existence všeobecného povědomí (awareness)*. Autorka předkládané práce nyní podala *vysvětlení, proč je použití omezovače v algoritmu PSO nezbytné*. Ze vztahu pro budící sílu F_b plyne, že k odvrácení katastrofického scénáře [5], jenž je pojednán v kapitole 4.7.3, *musí být skutečné chování systému částice PSO usměrňováno omezovačem rychlosti* tak, aby chování systému částice odpovídalo chování komplexního systému (*complex system*)⁵¹ a rovněž koeficienty a_l , a_g musí být proměnné, neboť volba konstantních koeficientů $a_l = a_g$ by znamenala významnou ztrátu možnosti regulace budící síly F_b na vstupu.

5.1.2 Vztahy mezi koeficienty

Pro vyjádření souvislosti mezi proměnnými koeficienty $a_l \neq a_g$ a koeficientem ω v rovnici (1) odvodila autorka předkládané práce vztah mezi střední hodnotou proměnné $\varphi_n^i = \varphi_{1n}^i + \varphi_{2n}^i$ a koeficienty ω , a_l , a_g , jenž je analogický rovnici (123). Při výpočtu vycházela

⁵⁰ Oscilátor má obecně tendenci “vychytávat” frekvence blízké jeho rezonanční frekvenci. Má-li oscilátor na vstupu jiný než harmonický signál, kterým je řízen, tj. např. obdélníkový signál, který má mnoho frekvenčních složek dle Fourierova rozvoje, odezva oscilátoru může zaznamenat náhlý, značný vzrůst, pokud dojde k souběhu se složkou, která je v blízkosti rezonančního kmitočtu oscilátoru. Pro diskrétní případ systému částice a nepravidelný vstupní signál lze náhlé odezvy očekávat zcela jistě.

⁵¹ Komplexní systém [43] je charakterizován tím, že je dostatečně flexibilní (dostatečně otevřený pro změnu), ale ne natolik, aby se choval chaoticky. Nebrání šumu a pertubacím, které urychlují jeho vývoj, avšak působení silných pertubací, které by mohly způsobit jeho destrukci, sám odhalí a odvrátí v procesu samo-organizace.

Komplexní systém podporuje nelineární, *samo-organizační (self-organization)* proces [43], kdy drobné odchylky mohou rozhodnout, kterého z více atraktorů systém dosáhne. Po dosažení atraktoru mají tyto odchylky malý, nebo žádný vliv. Samo-organizační proces eliminuje rozdíly mezi dvěma počátečními stavy, které skončí ve stejném atraktoru. Stejně tak nepatrně odlišný počáteční stav může skončit v jiném atraktoru. Samo-organizaci nejen nebrání, nýbrž prospívají šum a pertubace, náhodnost nebo neurčitost: čím více odchylek nebo výkyvů, tím rychleji systém dosáhne atraktoru a tím stabilnější atraktor bude. Tento poznatek je označován jako *order from noise principle* (von Foerster, 1960).

autorka z rovnice (119) pro standardní odchylku náhodné proměnné $\varphi_n^i \approx \varphi$ s tím, že náhodná proměnná φ již nemá trojúhelníkové (*triangular*), ale lichoběžníkové (*trapezoidal*) rozložení (neboť $a_l \neq a_g$). Autorka tudíž nahradila vztahy (120) a (121) vztahy pro lichoběžníkové rozložení:

$$a = E[\varphi] = E[\varphi_1] + E[\varphi_2] = a_l/2 + a_g/2 \quad (149)$$

$$\vartheta = E[\varphi^2] = E(\varphi_1^2) + E(\varphi_2^2) + 2E(\varphi_1\varphi_2) = \frac{a_l^2}{3} + \frac{a_g^2}{3} + \frac{a_la_g}{2} \quad (150)$$

Po vzoru autorů [39, 40] zavedla autorka koeficient α (127a) jako poměr $a_g/E[\varphi]$ a výsledkem je vztah pro střední hodnotu standardní odchylky pozic trajektorie systému částice a (149), která udává křivku představující horní hranici (*bare bone*), pod níž musí ležet dvojice $[a, \omega]$, aby systém částice popsany rovnicemi (1) a (2) vykazoval stabilitu prvního i druhého řádu:

$$a = E[\varphi] = \frac{12(1-\omega^2)}{4(2-\omega) + (\alpha^2 - 2\alpha)(1+\omega)} \quad (151)$$

Křivka $a(\omega, \alpha)$ (151) je parametrická, s parametrem α (127a). Při překročení hodnoty a pro danou hodnotu $0 < \omega < 1$ a zvolený parametr α by standardní odchylka pozic trajektorie systému částice od jejich očekávaných hodnot rostla neomezeně. Pro $\alpha = 1$, tedy pro $a_l = a_g$, přejde vztah (151) ve vztah (123).

Vnitřní pozorovatel vypočte pro zvolenou hodnotu ω ($0 < \omega < 1$) a zvolený parametr α hodnotu $E[\varphi]$ (151), dále vypočte hodnoty $a_g = \alpha * E[\varphi]$, $a_l = a_g(2 - \alpha)/\alpha$ podle (127a). Konkrétní hodnotu koeficientu ω bude pozorovatel volit v závislosti na rovnoměrnosti prozkoumaných pozic v dynamické historii roje a na fázi prohledávání. Přesný způsob bude uveden v kapitole 5.4.

Podrobný postup odvození vztahu (151) je vyčleněn do dodatku D15. Při odvozování se autorka inspirovala úvahami uvedenými v kapitole 4.5.3 a postupy v dodatcích D12 a D13. Autorka vztahem (151) odvodila případ limitní hyperboly (145), uvedený autory [39:p.9], [40:p.255], po dosazení $\Delta t = 1$:⁵²

$$E[\varphi] = \frac{12}{\Delta t} * \frac{(1-\omega^2)}{4 - 4(\omega - 1)\Delta t + (\alpha^2 - 2\alpha)(2 + (\omega - 1)\Delta t)}$$

5.2 Způsob detekce výlučnosti hodnoty účelové funkce

Algoritmus *PSO* ohodnocuje *fitness* nalezené pozice hodnotou účelové funkce pro tuto pozici. Jelikož zlepšení nebo zhoršení *fitness* je měřeno pouhou detekcí změny hodnoty účelové funkce a nezkoumá se velikost změny, autorka předkládané práce zavedla v kapitole 3.6 metriku označenou jako *vylučnost hodnoty fitness*, která má sloužit vnitřnímu pozorovateli jako kvalitativní ukazatel (číslo) vyjadřující relativní poměr dané hodnoty *fitness* k ostatním hodnotám *fitness* v ostatních pozicích. Bude-li hodnota účelové funkce

⁵² Vztah je autory [39, 40] uveden jako výsledek řešení podmínek stability analýzou vlastních hodnot matice pátého řádu popisující systém, tak jak autorka zmiňuje v kapitole 4.6.3.3. Analýza [39, 40] byla provedena využitím blíže neurčeného maticového kalkulátoru se symbolickým výstupem. Autorka v analogii s postupem v [38] upřednostnila přímé odvození vztahu (150) v dodatku D15.

v dosud objevené globální nejlepší pozici odhalena jako výlučná, autorka předpokládá, že s určitou pravděpodobností detekováno buď lokální, nebo (hledané) globální optimum.

Při hledání způsobu detekce *vylučnosti hodnoty fitness* autorka vycházela ze dvou skutečností, a sice: (i) že účelová funkce, od níž se odvozuje poloha pozic $l_n^i(t)$ a $g_n(t)$ v rovnici (1), je vstupní, nezávislou veličinou, jejíž působení na rychlost (1) a pozici částice (2) je z pohledu vnějšího pozorovatele deterministické, avšak z pohledu vnitřního pozorovatele se změny pozic v jednotlivých iteracích musí nutně jevit jako náhodné, a (ii) nachází-li se i -tá částice v určité pozici $\mathbf{x}^i(t)$, potom vnitřní pozorovatel, vzhledem k působení náhodných prvků, může pouze předpokládat s určitou pravděpodobností, že se v následující iteraci bude částice nacházet v pozici $\mathbf{x}^i(t+1)$.

Autorka tedy usoudila, že množina nejlepších dosud dosažených pozic $\mathbf{l}^i(t)$ pro jednotlivé částice, která je udržována v dynamické historii roje, má charakter statistického vzorku, jehož rozsah je dán počtem částic. Základním souborem je nekonečný soubor všech možných, teoreticky dosažitelných pozic.

5.2.1 Pravděpodobnostní model roje částic *PSO*

K dalšímu zkoumání chování částic navrhuje autorka předkládané práce vytvoření pravděpodobnostního modelu roje. Jako první krok definuje autorka trajektorii i -té částice jako posloupnost jednotlivých pozic

$$\mathbf{x}^i(t_0) \mathbf{x}^i(t_1) \mathbf{x}^i(t_2) \cdots \mathbf{x}^i(t_k),$$

které jsou realizacemi náhodných veličin X_k v časech $t_0 \cdots t_k$, $t_k = t_0 + k\Delta t$.

Z popisu algoritmu *PSO* vyjádřeného rovnicemi (1) a (2) lze dále vysledovat, že pozice částice v čase $t+1$ závisí pouze na pozici v čase t a rychlosti v čase $t+1$. Rychlost v čase $t+1$ je odvozena z rychlosti v čase t a pozic $\mathbf{x}_n^i(t)$, $l_n^i(t)$ a $g_n(t)$. Pohyb i -té částice po trajektorii lze tedy popsat jako dynamický diskretní (vzhledem k času) stochastický proces následovně:

$$P\{X_{k+1} = \mathbf{x}_{k+1}^i | X_k = \mathbf{x}_k^i, X_{k-1} = \mathbf{x}_{k-1}^i, \dots, X_0 = \mathbf{x}_0^i\} = P\{X_{k+1} = \mathbf{x}_{k+1}^i | X_k = \mathbf{x}_k^i\} \quad (152)$$

kde $\mathbf{x}_k^i$ je zkrácený zápis pro $\mathbf{x}^i(t_k)$. Vztahem (152) chtěla autorka vyjádřit skutečnost, že pravděpodobnost, že částice zaujme v čase $t_0 + (k+1)\Delta t$ pozici $\mathbf{x}_{k+1}^i$, je dána pouze pravděpodobností zaujetí pozice $\mathbf{x}_k^i$ v čase předcházejícím, tj. $t_0 + k\Delta t$. Časový index k zavádí autorka proto, že pravděpodobnostní rozložení (*probability density function pdf*) veličiny X_k může být v každé iteraci odlišné.⁵³ Volíme pouze (počáteční) rozložení veličiny X_0 , a to zpravidla jako rovnoměrné.

Autorka dále zavádí náhodnou proměnnou Y_k jejíž realizace v čase $t_0 + k\Delta t$ vyjadřuje hodnotu *fitness* v pozici $\mathbf{x}_k^i$:

$$Y_k = y_k^i: X_k = \mathbf{x}_k^i: y_k^i = f(\mathbf{x}_k^i)$$

⁵³ Analytikové [33:pp.35-38] podávají postup, kterým odvozují rozložení veličiny X v k -té iteraci z počátečního rozložení. Ukazují, že rovnoměrné rozložení přechází v normální rozložení a mění se pouze parametry tohoto rozložení. Autorka předkládané práce je však obezřetnější a upozorňuje, že tyto postupy neberou v potaz působení účelové funkce na pozice $\mathbf{l}^i$ (a tedy $\mathbf{x}_k^i$), které je vyjádřeno vztahem (146).

kde f je účelová funkce. Každá hodnota (*fitness*) $f(\mathbf{x}_k^i)$ je porovnávána s hodnotou $f(\mathbf{l}^i)$, kde $\mathbf{l}^i$ je dosud nejlepší objevená pozice i -té částice v dynamické historii roje. Pokud je hodnota $f(\mathbf{x}_k^i)$ lepší než hodnota $f(\mathbf{l}^i)$, potom se $\mathbf{x}_k^i$ stává novou nejlepší objevenou pozicí i -té částice v dynamické historii roje. Ač je hodnota *fitness* pro každou pozici určena (vypočtena nebo změřena) deterministicky, je přiřazena náhodné pozici, a proto je množina nejlepších *fitness* v dynamické historii roje

$$\{y^i | y^i = f(\mathbf{l}^i), i = 1, \dots, M\} \quad (153)$$

autorkou interpretována jako statistický vzorek, jehož rozsah je dán počtem částic M . Základním souborem je nekonečný soubor všech možných, teoreticky dosažitelných hodnot účelové funkce.

5.2.2 Testování výlučnosti globální nejlepší hodnoty účelové funkce

Autorka definovala množinu *fitness* (153) udržovanou vnitřním pozorovatelem v průběhu dynamické historie roje. Hodnoty *fitness* y^i (153) (kde $i = 1, 2, \dots, M$, M je počet částic) jsou realizacemi náhodných proměnných Y^i ⁵⁴ a úkolem je zjistit, zda hodnota *fitness* $y^i = f(\mathbf{l}^i) \approx f(\mathbf{g})$, kde $\mathbf{l}^i$ je lokální nejlepší pozice i -té proměnné a $\mathbf{g}$ je nejlepší globální pozice, je *vylučná*, a definovat způsob stanovení výlučnosti.

Množina realizací proměnných Y^i (153) tvoří statistický vzorek, jehož rozsah je malý, neboť odpovídá počtu částic. Rozsah vzorku tudíž není možno zvětšit zvětšením počtu částic, jak by odpovídalo obvyklému (statistickému) přístupu. Pravděpodobnostní charakteristiky (tj. distribuční funkce (*cumulative distribution function cdf*) a hustota pravděpodobnosti (*probability density function pdf*)) náhodných proměnných Y^i nejsou známy. Autorka bude předpokládat, že všechny proměnné Y^i mají stejná pravděpodobnostní rozložení a tato rozložení jsou spojitá.⁵⁵ Pokud předpoklad shody neplatí, měl by existovat způsob, jak tuto skutečnost detekovat.

5.2.2.1 Aplikace teorie vylučování odlehlých výsledků (*outliers*)

Na problém *definice výlučnosti* autorka aplikovala *Dixonovu* teorii vylučování odlehlých výsledků (*outliers*), publikovanou v časopise *Biometrics* v roce 1953 [44] [45].

Teorie vychází ze zkušenosti, že převážná část souborů výsledků opakovaných pokusů má jedno-vrcholové rozložení, jež se jen výjimečně blíží rozložení *Gaussovu*. Každé jedno-vrcholové rozložení lze charakterizovat (i) parametrem centroidní tendence (tj. střední hodnotou, mediánem, aritmetický průměrem, ...) a (ii) parametrem variability (standardní směrodatnou odchylkou). Problém je, že pokud počítáme střední hodnotu z výběru, máme pouze její odhad. Vyskytne-li se zde ojedinělý výsledek, může to značit přítomnost hrubé chyby, ale též vliv nového, neuvažovaného faktoru. Úkolem je rozpoznání situace, kdy se vyskytne ojedinělé pozorování, které se vzdaluje od zbytku populace. Tato

⁵⁴ Autorka rozlišuje veličiny Y^i a Y_k . Veličiny Y^i odpovídají nejlepším hodnotám, zatímco veličiny Y_k odpovídají hodnotám zjištěným v k -té iteraci. Y^i jsou pořadové statistiky (*order statistics*).

⁵⁵ Autorka tak soudí proto, že částice může zaujmout jakoukoli pozici a k této pozici vždy existuje hodnota účelové funkce.

pozorovaná extrémní hodnota se může významně lišit, aniž nějak významně ovlivní analýzu standardní odchylky.

Teorie [44] zavádí tzv. modely kontaminace dvojího typu, takové, že výskyt “chyby” ve střední hodnotě může být důsledkem začlenění dvou a více položek, které mají zcela odlišný původ. Tyto položky (i) jsou realizacemi proměnné s odlišným pravděpodobnostním rozložením (*gross errors*):⁵⁶

$$\mathcal{N}(\mu + \lambda\sigma, \sigma^2)$$

nebo (ii) jsou z populace s toutéž střední hodnotou, ale jinou odchylkou (*scalar errors*):

$$\mathcal{N}(\mu, \lambda^2\sigma^2),$$

přičemž je velmi pravděpodobné, že chyba bude kombinací obou typů. Aplikaci teorie vylučování odlehlých výsledků na množinu nejlepších *fitness* v dynamické historii roje (153) provedla autorka postupem, jehož základní myšlenka je následující:

Necht' množina náhodných proměnných $Y = \{Y^1, Y^2, \dots, Y^{M-1}, Y^M\}$ má konkrétní realizaci $\{y^1, y^2, \dots, y^{M-1}, y^M\}$ (153). Jednotlivé hodnoty lze uspořádat vzestupně takto:

$$y_{(1)} < y_{(2)} < \dots < y_{(M-1)} < y_{(M)} \quad (154)$$

kde index v závorce již neznačí identifikátor částice, ale pořadí ve vzestupném uspořádání hodnot *fitness*. Je zřejmé, že existuje $M!$ (permutací) konkrétních realizací, jimž odpovídá totéž uspořádání (*order statistics*).⁵⁷ Spočteme podíl

$$z = \frac{y_{(2)} - y_{(1)}}{y_{(M)} - y_{(1)}} \quad (155)$$

který je realizací náhodné proměnné Z . Platí $0 < z < 1$. Necht' r je reálné číslo, $0 < r < 1$. Potom pravděpodobnost

$$P(Z \leq r) = F_z(r) = \int_0^r f(z) dz \quad (156)$$

kde F_z je distribuční funkce (*cumulative distribution function cdf*) a $f(z)$ je hustota pravděpodobnosti (*probability density function pdf*) náhodné proměnné Z . Za předpokladu, že veličiny Y^i mají normální rozložení, lze vzorec pro $f(z)$ odvodit, a lze spočítat i tzv. *kritické hodnoty* r pro zvolené pravděpodobnosti $P(Z \leq r)$.

Kritické hodnoty r pro zvolené pravděpodobnosti byly autorem [44] spočteny a jsou pro zvolené hodnoty $\alpha_z = P(Z > r)$ tabelovány. Postup při testování, zda je hodnota $y_{(1)}$ výlučná (*outlier*) s pravděpodobností $P(Z \leq r) = 1 - \alpha_z$, je následující:

Spočteme empirický podíl (155) a vycházíme z hypotézy $\mathcal{H}_0$: Hodnota $y_{(1)}$ se s pravděpodobností $(1 - \alpha_z)$ významně neliší od zbytku populace (není *outlier*). Pokud platí $z < r$, nulovou hypotézu $\mathcal{H}_0$ přijímáme, v opačném případě ji zamítáme a přijímáme alternativní hypotézu $\mathcal{H}_1$: Hodnota $y_{(1)}$ se s pravděpodobností $(1 - \alpha_z)$ významně liší od zbytku populace (je *outlier*).

⁵⁶ Pojem *gross error* zavedl v roce 1850 profesor matematiky *William Chauvenet*, činný na *U. S. Naval Academy*.

⁵⁷ Je-li např. počet částic $M = 3$, fitness $y^1 = 0.25$, $y^2 = 2.15$, $y^3 = 1.20$, potom uspořádání je $\{0.25, 1.20, 2.15\}$. Stejně uspořádání má realizace $y^1 = 1.20$, $y^2 = 0.25$, $y^3 = 2.15$. Existuje celkem $3! = 3 * 2 * 1 = 6$ realizací, jimž odpovídá totéž uspořádání.

Tabelované hodnoty vypočítal autor [44] pro různé podíly v závislosti na velikosti zkoumaného vzorku. Statistikové (uživatelé tabulek)⁵⁸ obecně uvádějí, že podíl (155) je vhodné použít u vzorků malého rozsahu, $3 \leq M \leq 7$. Pro $8 \leq M \leq 10$ je doporučováno opomenutí $y_{(M)}$ při výpočtu empirického podílu, pro $11 \leq M \leq 13$ je doporučováno opomenutí $y_{(2)}$ a $y_{(M)}$, pro $14 \leq M \leq 30$ je doporučováno opomenutí $y_{(2)}$, $y_{(M-1)}$ a $y_{(M)}$. Autor [44] oproti tomu uvádí, že analytik si musí sám zvážit, zda a jakým způsobem mohou být jeho data zatížena (*biased*), a podle toho zvolit odpovídající přístup.

Příslušný integrál pro výpočet $F_z(r)$ (156) odvodil autor [44] analyticky a tabulky byly vyčísleny přímými výpočty. V roce 2006 byla v časopise *Journal of Statistical Software* uveřejněna publikace [48], v níž její autor validoval výsledky v tabulkách tak, že spočetl hodnoty integrálů numericky v jazyce *Fortran*.⁵⁹ Odvození vztahu pro pravděpodobnost (156) je naznačeno v dodatku D16.

5.2.2.2 Postup při testování výlučnosti

Autorka předkládané práce bude experimentovat s počtem částic $M = 20$ a s variantami podílů, jež autor [44] označuje jako r_{20} , r_{21} , r_{22} :

$$r_{20} = \frac{y_{(3)} - y_{(1)}}{y_{(M)} - y_{(1)}} \quad r_{21} = \frac{y_{(3)} - y_{(1)}}{y_{(M-1)} - y_{(1)}} \quad r_{22} = \frac{y_{(3)} - y_{(1)}}{y_{(M-2)} - y_{(1)}} \quad (157)$$

Odpovídající tabulky pro dvojice (*kritická hodnota* r_{xx} , *pravděpodobnost* $1 - \alpha_z$) a parametr M budou implementovány přímo v algoritmu *PSO*. Vnitřní pozorovatel bude v každé iteraci počítat odpovídající empirické podíly r_{20} , r_{21} , r_{22} a aproximacemi v odpovídajících tabulkách získá hodnoty pravděpodobností, s jakými se $y_{(1)}$ významně liší od zbytku populace (tj. je *outlier*). Hodnota pravděpodobnosti ($1 - \alpha_z$) bude tedy indikovat míru výlučnosti hodnoty $y_{(1)}$.⁶⁰ Experimentováno bude s hodnotami pravděpodobnosti pro r_{20} , r_{21} , r_{22} [44] a různými účelovými funkcemi (*Příloha A*) a bude zkoumáno, jaká míra výlučnosti byla zjištěna a zda tak byla indikována některá lokální nebo (hledaná) globální nejlepší pozice. (Po počátečním testování byl pro další experimentování zvolen podíl r_{22} .)

Příklad Lokální hodnoty *fitness* pro $M = 20$ částic seřazené ve vzestupném pořadí:

k	1	2	3	4	5	6	7	8	9	10
$y_{(k)}$	118	124	187	224	237	238	255	263	283	299
k	11	12	13	14	15	16	17	18	19	20
$y_{(k)}$	313	346	347	397	419	421	443	448	455	508

Spočtený podíl r_{22} podle (157):

⁵⁸ Autorka shledává, že tzv. *Dixonův Q-test* je používán výhradně k vyloučení hodnot z výpočtu, typicky např. aritmetického průměru, a tím se původní, širší význam *teorie vylučování odlehlých výsledků* z povědomí uživatelů vytrácí.

⁵⁹ Autorka předkládané práce narazila při studiu publikace [44] na určité nepřesnosti ve vzorcích, a proto pátrala, zda se tyto nepřesnosti nemohly promítnout do výpočtů a zda se někdo tímto problémem zabýval, a objevila publikaci [48]. Autor [48] též tyto nepřesnosti zjišťuje, označuje je jako tiskové chyby a ujišťuje, že tabulky jsou vyčísleny správně.

⁶⁰ Autorka řeší minimalizační problémy. Jsou definovány i vztahy reciproké k (156) v případě, že by byla zkoumána výlučnost *fitness* $y_{(M)}$ pro maximalizační problém.

$$r_{22} = \frac{y_{(3)} - y_{(1)}}{y_{(18)} - y_{(1)}} = \frac{187 - 118}{448 - 118} = 0.209$$

V tabulce [44] *TABLE VI*. pro 20 prvků aproximujeme hodnotu $\alpha = 0.6$, a tedy hodnota *fitness* $y_{(1)} = 118$ se od zbylých hodnot liší (je výlučná) s pravděpodobností nejvýše 0.4, což je hodnota, která není důvodem k posílení globálního prohledávání. (*Konec příkladu.*)

5.3 Sledování chování částice v závislosti na účelové funkci a úspěšnost řešení

Podkladem pro posouzení *úspěšnosti řešení* jako metriky dle konceptuálního schématu na obrázku Obr. 3, v souladu s definicí v kapitole 3.6, bude zjištěná hodnota rozptylu prozkoumaných pozic v dynamické historii roje. Vnitřní pozorovatel si bude pro každý rozměr n polohového vektoru částice $\mathbf{r}_n^i$ (16) vytvářet *histogram* v podobě grafického znázornění distribuce dat pomocí sloupcového grafu se sloupci stejné šířky (interval zvaný *bin*), v němž výška sloupce bude vyjadřovat četnost sledované veličiny v intervalu daném šířkou sloupce. Histogram bude vytvořen pro každý rozměr $n = 1 \dots N$ částice a bude sloužit jako mapa vymezeného hyperprostoru ve smyslu bodu (iii) v kapitole 3.7 a sledovanou veličinou budou prozkoumané pozice.

Vnitřní pozorovatel bude sledovat rovnoměrnost rozložení prozkoumaných pozic aplikací *Pearsonova testu dobré shody* χ^2 . Vymezený prostor bude pro každý rozměr rozdělen na počet intervalů odpovídající počtu částic M , což odpovídá teoretické četnosti jedné pozice v každém intervalu (sloupci histogramu) po inicializaci roje. Po každé iteraci zvýší vnitřní pozorovatel očekávanou, teoretickou četnost o jednu pozici, v každé iteraci bude měřena empirická četnost a spočtena hodnota testovacího kritéria z :

$$z = \sum_{m=1}^M \frac{[o_m - e]^2}{e} \quad (158)$$

kde o_m je pozorovaná (*observed*) četnost v intervalu m a e je očekávaná (*expected*) četnost, která je průběžně zvyšována a která je pro každý interval táž (rozložení je rovnoměrné).

Příklad Pro 20 dvourozměrných částic budou vytvořeny dva histogramy, každý histogram bude rozdělen na 20 intervalů. Očekává se, že po inicializaci bude v každém intervalu obsažena jedna pozice, a tedy $e = 1$. Po první iteraci se očekávají dvě pozice obsažené v každém intervalu, atd. (*Konec příkladu.*)

Je-li hodnota testovacího kritéria z (158) realizací náhodné veličiny Z a je-li hodnota četnosti o_m realizací náhodné veličiny O a má-li veličina O rovnoměrné rozdělení, pak má veličina Z přibližně rozdělení χ^2 (*chí-kvadrát*) s $k = (M - 1) = (20 - 1) = 19$ stupni volnosti. Vnitřní pozorovatel tedy spočte pravděpodobnost

$$p = P(Z \leq z) = F(k, z) \quad (159)$$

kde z je testovací kritérium (158), $F(z, k)$ je distribuční funkce proměnné Z s rozdělením χ^2 a k stupni volnosti. Pravděpodobnost p (159), jež udává míru rovnoměrnosti rozdělení

pozic, spočte pozorovatel podle příslušného vzorce pro distribuční funkci proměnné s rozložením χ^2 :

$$F(z, k) = \frac{\gamma\left(\frac{k}{2}, \frac{z}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \quad (160)$$

kde z je hodnota testovacího kritéria spočtená dle (158), $k = M - 1$ je počet stupňů volnosti, γ je dolní nekompletní funkce gama (*lower incomplete gamma function*) a Γ je horní nekompletní funkce gama (*upper incomplete gamma function*).

Vzhledem k nízkému počtu částic ($M = 20$) nelze očekávat, že by se pozorovaná četnost o_m po inicializaci rovnala očekávané, teoretické četnosti e . Bylo by také zapotřebí většího počtu (minimálně 100) částic, aby teoretická četnost při počtu intervalů 20 dosáhla pěti, pro niž je *Pearsonův test dobré shody* χ^2 spolehlivější. Počet částic však nelze zvyšovat, neboť rovnoměrnosti rozložení pozic má být dosaženo prohledávací schopností částic na počátku prohledávání.

Další otázkou je vzájemný vztah počtu sloupců histogramu, šířky sloupce histogramu a velikosti vyhledávacího prostoru. Pro volbu šířky sloupce při konstrukci histogramu neexistuje pravidlo, jež by bylo “správné”. Nejčastěji citované pravidlo pro určení šířky sloupce histogramu odvodil v roce 1979 *David W. Scott* [47] pro normální rozložení sledované veličiny:

$$h = \frac{3,49}{\sqrt[3]{m}} \sigma$$

kde h je šířka sloupce, σ je standardní odchylka sledované veličiny, m je počet realizací náhodné veličiny, jejichž četnost se zaznamenává. Počet sloupců se určí jako podíl rozsahu hodnot sledované veličiny a šířky sloupce. Spočítá-li autorka šířku sloupce pro účelovou funkci *Schweifel* (*Příloha A*) s rozsahem $(-512 - 512)$ a pro $m \approx M \approx 100$ částic:

$$h = \frac{3,49}{\sqrt[3]{100}} \sigma = \frac{3,49}{\sqrt[3]{100}} \sqrt{\frac{(-512 - 512)^2}{12}} \approx 222$$

a tedy počet sloupců $1024/222 \approx 5$, pro $M \approx 20$ částic vychází počet sloupců ještě nižší. Jedná se zjevně o výpočet optimálního počtu sloupců, který je volen tak, aby byl histogram dobře graficky zobrazitelný, což je přístup, který není pro autorčin záměr sledování rovnoměrnosti rozložení pozic vhodný. Autorka proto volila počet sloupců, jenž je roven zvolenému počtu částic $M = 20$, a všechny účelové funkce budou pro srovnatelnost výsledků testovány (také) v rozsahu $[-512, 512]$.

5.4 Výsledný algoritmus *PSO*

Výsledný algoritmus *PSO* je navržen pro *pozorování chování částic* v závislosti na (i) počtu částic vzhledem k velikosti vyhledávacího prostoru, (ii) působení účelové funkce a (iii) vlivu náhodných koeficientů. Chování částic pro různé účelové funkce bude testováno za srovnatelných podmínek pro počet částic $M = 20$ a testy budou prováděny na počtu 100 různých posloupností náhodných čísel dostatečné délky, předem vygenerovaných.

Pro pozorování rovnoměrnosti prozkoumaných pozic v závislosti na velikosti vyhledávacího prostoru slouží N histogramů, kde N je počet stupňů volnosti optimalizačního problému (rozměr částice). Každý z histogramů je nadefinován jako pole délky $M = 20$, kde každý prvek pole ($o_m^n | n = 1 \dots N, m = 1 \dots M$) obsahuje hodnotu pozorovaných (*observed*) četností. Hodnoty teoretických (*expected*) četností v každém prvku obsahuje pomocná proměnná *expected*, jejíž hodnota se v každé iteraci navyšuje o 1.

Hodnota koeficientu ω v rovnici (1) bude odvozována od rovnoměrnosti pozic určené pravděpodobností p (159), (160). Pro rovnoměrnější rozdělení, a tedy vyšší p , bude zvolen koeficient nižší, podle vztahu $\omega \approx 1 - p$.

Autorka očekává, že po inicializaci bude v prvních iteracích pravděpodobnost p (159), (160) stoupat, a to vlivem lokálního prohledávání, které bude posíleno, zatímco globální prohledávání bude potlačeno. Ustane-li nárůst pravděpodobnosti p , mělo by to značit, že částice začínají konvergovat. Algoritmus proto přejde z fáze lokálního prohledávání (*local searching*), kdy $a_l \gg a_g$, do fáze vyváženého prohledávání (*balanced searching*), kdy $a_l = a_g$. Byla-li identifikována *fitness* dosud objevené globální nejlepší pozice jako výlučná, algoritmus by mohl přejít do fáze globálního prohledávání (*global searching*), kdy $a_g \gg a_l$, avšak vnitřní pozorovatel musí sledovat i rovnoměrnost prohledaných pozic.

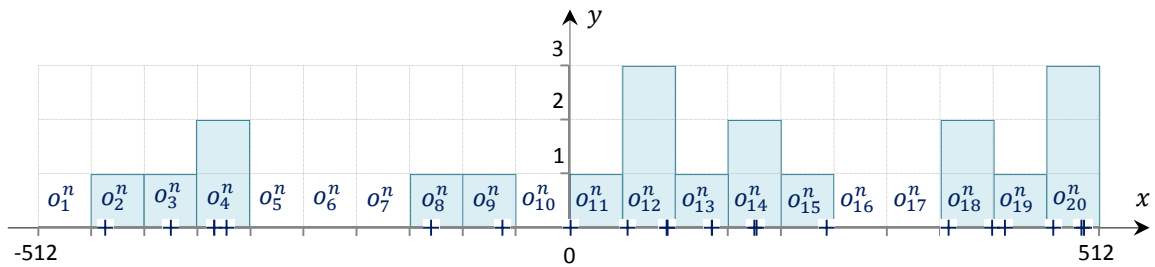
Koeficient $\alpha = 2a_g/(a_l + a_g)$ (127a) bude volen úměrně pravděpodobnosti p ($\alpha \approx p$). Výlučnost *fitness* globální nejlepší pozice bude posuzována výpočtem podílů (157) a aproximací z tabulek implementovaných přímo v algoritmu *PSO*.

Protože lze očekávat, že pro každý rozměr částice bude po inicializaci naměřena jiná pravděpodobnost p , budou koeficienty (ω, a_l, a_g) nadefinovány jako pole

$$(\omega^n | n = 1 \dots N), (a_l^n | n = 1 \dots N), (a_g^n | n = 1 \dots N).$$

Rovněž tak fáze prohledávání budou nadefinovány jako pole ($d^n | n = 1 \dots N$), kde N je rozměr částice.

Příklad Postup při stanovení hodnot koeficientů:



Obrázek znázorňuje příklad histogramu pro 20 částic, pro libovolný rozměr částice n , a znázorňuje pozorované četnosti pozic (osa y) pro sloupce histogramu $o_1^n \dots o_{20}^n$ (osa x).

Po inicializaci, kdy očekávaná četnost e (*expected*) v každém z intervalů je rovna jedné, spočte pozorovatel hodnoty koeficientů následujícím postupem: Vnitřní pozorovatel nejdříve provede výpočet testovacího kritéria z podle (158):

$$z = \sum_{m=1}^{20} \frac{[o_m^n - e]^2}{e} = \frac{[0 - 1]^2}{1} + 2 * \frac{[1 - 1]^2}{1} + \frac{[2 - 1]^2}{1} + 3 * \frac{[0 - 1]^2}{1} + 2 * \frac{[1 - 1]^2}{1} +$$

$$\begin{aligned} & \frac{[0-1]^2}{1} + \frac{[1-1]^2}{1} + \frac{[3-1]^2}{1} + \frac{[1-1]^2}{1} + \frac{[2-1]^2}{1} + \frac{[1-1]^2}{1} + 2 * \frac{[0-1]^2}{1} + \frac{[2-1]^2}{1} + \\ & \quad \frac{[1-1]^2}{1} + \frac{[3-1]^2}{1} + \frac{[3-1]^2}{1} + \frac{[1-1]^2}{1} + \frac{[3-1]^2}{1} = \\ & 1 + 0 + 1 + 3 + 0 + 1 + 0 + 4 + 0 + 1 + 0 + 2 + 1 + 0 + 4 = 18 \end{aligned}$$

Následně pozorovatel spočte pravděpodobnost rovnoměrného rozmístění pozic p (159) dosazením hodnoty udávající počet stupňů volnosti $k = 19$ a hodnoty testovacího kritéria $z = 18$ do (160):⁶¹

$$p = P(Z \leq z) = F(k, z) = \frac{\gamma\left(\frac{k}{2}, \frac{z}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} = \frac{\gamma\left(\frac{19}{2}, \frac{18}{2}\right)}{\Gamma\left(\frac{19}{2}\right)} \approx 0.5224382739862669$$

Vnitřní pozorovatel může pracovat v různých režimech prohledávání v závislosti na detekci odlehlých hodnot. Pro režim lokálního prohledávání ($mood = 0$) dle Obr. 14(b) zvolí:

$$\begin{aligned} \omega & \approx 1 - p = 1 - 0.5224382739862669 = 0.47756172601373315, \\ \alpha & \approx p = 0.5224382739862669, \end{aligned}$$

dále spočte střední hodnotu koeficientů $a = E(\varphi)$ dosazením ω, α do rovnice (151):

$$\begin{aligned} a & = E[\varphi] = \frac{12(1-\omega^2)}{4(2-\omega) + (\alpha^2 - 2\alpha)(1+\omega)} \cong \\ a & \cong \frac{12(1-0.4775617260^2)}{4(2-0.4775617260) + (0.5224382739^2 - 2 * 0.5224382739)(1+0.4775617260)} \\ a & \cong 1.8716702469 \end{aligned}$$

a spočte a_g, a_l podle (127a):

$$\begin{aligned} a_g & = \alpha * a = 0.5224382739 * 1.8716702469 = 0.9778321732 \\ a_l & = a_g(2 - \alpha)/\alpha = 0.9778321732(2 - 0.5224382739)/0.5224382739 \\ a_l & = 2.7655083204. \end{aligned}$$

Pro režim lokálního prohledávání ($mood = 2$) dle Obr. 14(b) vnitřní pozorovatel zvolí:

$$\begin{aligned} \omega & \approx p = 0.5224382739 \\ \alpha & \approx 1 - p = 1 - 0.5224382739 = 0.4775617260 \end{aligned}$$

a spočte a_g, a_l podle (127a):

$$\begin{aligned} a_g & = 0.8674338127 \\ a_l & = 2.7653258371 \end{aligned}$$

Pro režim rovnoměrného prohledávání ($mood = 1$) dle Obr. 14(b) vnitřní pozorovatel zvolí:

$$\begin{aligned} \omega & \approx p = 0.5224382739862669 \\ \alpha & = 1 \end{aligned}$$

⁶¹ Autorka naprogramovala kalkulátor funkcí γ a Γ , na jehož přepis odkazuje [Příloha B](#).

a spočte a_g, a_l podle (127a):

$$a_g = a_l = 1.98839551449.$$

Hodnoty koeficientů spočte pozorovatel uvedeným postupem pro každý rozměr částice. Hodnoty koeficientů bude pozorovatel počítat s nejvyšší možnou přesností, neboť se předběžně ukázalo, že navrhovaný algoritmus je na přesnost velmi citlivý. Volba režimů bude záviset na dosažené rovnoměrnosti prohledávání a detekci výlučné *fitness* (*outlier*). Detailní nastavení podmínek bude předmětem experimentování. (*Konec příkladu.*)

Pokud platí předpoklad působení budící síly F_b (obrázek Obr. 13), potom autorka usuzuje, že rozdílné posloupnosti pravděpodobností p (159), (160) lze očekávat také pro různé účelové funkce (*Příloha A*), a to i při konstantních koeficientech (0.72, 1.48, 1.48). Proto budou pro srovnání sledovány i posloupnosti pravděpodobností p (159), (160) pro variantu algoritmu *PSO* s konstantními koeficienty [5] a bude měřen počet úniků mimo vymezený vyhledávací prostor. Počet úniků bude dán počtem zapojení omezovače rychlosti částice podle (148).

Návrh výsledného algoritmu *PSO* v *pseudo*-jazyce je na obrázcích Obr. 14(a), Obr. 14(b). Experimentální ověření algoritmu je obsahem následující, šesté kapitoly, kde je rovněž diskutován navržený způsob nastavení hodnot koeficientů (ω, a_l, a_g) .

```

Create and initialize  $M$   $N$ -dimensional particles  $(\mathbf{x}^i, \mathbf{v}^i)$ :  $(\mathbf{x}_n^i | i = 1 \dots M, n = 1 \dots N)$ 
 $(\mathbf{v}_n^i | i = 1 \dots M, n = 1 \dots N)$ 

Create  $M$  local best positions  $(\mathbf{l}^i | i = 1 \dots M)$ 
Create  $N$   $M$ -dimensional histograms  $(o_m^n | n = 1 \dots N, m = 1 \dots M)$ 
Create and initialize  $\omega, a_l, a_g$  as  $N$ -dimensional arrays
 $(\omega^n | n = 1 \dots N), (a_l^n | n = 1 \dots N), (a_g^n | n = 1 \dots N) \approx [0, 0, \dots, 0]$ 
Create and initialize  $N$ -dimensional phase indicators  $(d^n | n = 1 \dots N) \approx [1, 1, \dots, 1]$ 

/* Initialize histograms */
expected = 1
for each dimension  $n \in [1..N]$ 
  for each bin  $o_m^n \in [1..M]$ 
     $o_m^n = 0$ 
    for each particle  $i \in [1..M]$ 
      if  $\mathbf{x}_n^i$  lies within the bin  $o_m^n$  then  $o_m^n = o_m^n + 1$ 
    end
  end
end

/* Initialize  $M$  local best positions and the global best position */
for each particle  $i \in [1..M]$ 
   $\mathbf{l}^i = \mathbf{x}^i$ 
end
 $\mathbf{g} = \mathbf{x}^1$ 
for each particle  $i \in [2..M]$ 
  if  $f(\mathbf{l}^i) < f(\mathbf{g})$  then  $\mathbf{g} = \mathbf{l}^i$ 
end

```

Obr. 14(a): Návrh inicializační fáze výsledného algoritmu *PSO*.

```

/* Start main PSO */
Boolean outlier = false
repeat
  /* Find out whether there is an outlier */
  compute  $r_{xx}$  according equation (157)
  elicit probability  $(1 - \alpha_{xx})$ 
  if  $y_{(1)}$  (154) is an outlier then outlier = true

  /* Set values of coefficients  $\omega, a_l, a_g$  */
  for each dimension  $n \in [1..N]$ 
    compute quotient  $z$  using equation (158)
    get probability  $p^n$  using equations (159) (160)
    if  $(p_{max}^n \approx 1) \wedge (outlier = true)$  then  $(d^n = 3)$ 
    if  $(d^n = 2) \wedge (outlier = false)$  then  $(d^n = 0)$ 
    if  $(d^n = 1) \wedge (outlier = true)$  then  $(d^n = 2)$ 
    if  $(d^n = 0) \wedge (p_{previous}^n < p^n)$  then  $(d^n = 1)$  (beginning of convergence)
    case mood  $d^n$  of:
      0:  $(\omega^n \approx 1 - p^n); (\alpha^n \approx p^n);$  (local searching  $\alpha^n < 1$ )
      1:  $(\omega^n \approx p^n); (\alpha^n \approx 1);$  (balanced searching  $\alpha^n = 1$ )
      2:  $(\omega^n \approx p^n); (\alpha^n \approx 1 - p^n);$  (local searching  $\alpha^n < 1$ )
      3:  $(\omega^n \approx p^n); (\alpha^n \approx 2 - p^n);$  (global searching  $\alpha^n > 1$ )
    end
    compute  $a = E(\varphi)$  using equation (151)
    compute  $a_g^n = \alpha^n * a$  resulting from (127a)
    compute  $a_l^n = a_g^n(2 - \alpha)/\alpha$  resulting from (127a)
  end

  /* Update positions of particles in the search space */
  for each particle  $i \in [1..M]$ 
    for each dimension  $n \in [1..N]$ 
      update velocity  $v_n^i$  using equation (1)
      update position  $x_n^i$  using equation (2)
      if position  $x_n^i$  is out of range then involve limiter (148)
    end
  end

  /* Update observed positions in the  $n^{th}$  histogram */
  for each dimension  $n \in [1..N]$ 
    for each histogram.interval  $o^n \in [1..M]$ 
      for each particle  $i \in [1..M]$ 
        if  $x_n^i$  lies within the histogram.interval  $o_m^n$  then  $o_m^n = o_m^n + 1$ 
      end
    end
  end

  /* Update number of expected positions in each bin */
  expected = expected + 1

  /* Update global best and  $M$ -dimensional local bests */
  for each particle  $i \in [1..M]$ 
    if  $f(x^i) < f(l^i)$  then  $l^i = x^i$ 
    if  $f(l^i) < f(g)$  then  $g = l^i$ 
  end

until stopping condition is true.

```

Obr. 14(b): Návrh fáze prohledávání vymezeného vyhledávacího prostoru výsledného algoritmu *PSO*.

Kapitola 6

Experimentální ověření navrhovaného řešení a dosažené výsledky

Pro testování navrhovaného algoritmu *PSO* vytvořila autorka parametrický testovací model originálního algoritmu [4] a výsledného algoritmu podle návrhu na obrázcích Obr. 14(a), Obr. 14(b). Testovací model byl vytvořen v jazyce *Java* v prostředí *NetBeans IDE 6.9.1*. Dokumentace algoritmu je v příloze *Příloha B*.

Experimentováno bylo s originální i navrhovanou verzí algoritmu *PSO*. U každého experimentu je uvedeno, proč byl experiment prováděn, co bylo jeho cílem a jaké jsou výsledky experimentu. Každý experiment se opírá o určité tvrzení a cílem experimentu je tvrzení vyvrátit. Pro úspěšné případy, kdy tvrzení vyvráceno či zamítnuto nebylo, bylo rovněž analyzováno, zda příčiny úspěchu závisí pouze na výchozích předpokladech (premisách), a nikoli na dalších okolnostech, jež nebyly původně vzaty v úvahu. Výsledky experimentů byly průběžně diskutovány.

6.1 Experimenty s originální verzí algoritmu *PSO* s konstantními koeficienty

Smyslem experimentování s originální verzí algoritmu *PSO* s konstantními koeficienty [4] byla simulace chování částic dle katastrofického scénáře [5] a ověření konvergence ke globálnímu optimu, zkoumání vlivu inicializace na výsledek prohledávání, zkoumání vlivu účelové funkce na průběh prohledávání a získání výsledků pro srovnání s výslednou, navrhovanou verzí algoritmu *PSO* na obrázcích Obr. 14(a), Obr. 14(b).

6.1.1 Simulace chování částic dle katastrofického scénáře

6.1.1.1 Proč byl experiment prováděn

Důvodem experimentování s katastrofickým scénářem jsou kritické připomínky autorky předkládané práce k variantě algoritmu *PSO* s konstantními koeficienty [4] odvozenými z deterministické analýzy [36], tak jak jsou uvedeny v kapitole 4.7.1. Závěr autorce vyznívá tak, že při použití koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ v rovnici (1) není podle autorů [4] zapotřebí použití omezovače rychlosti částice. Experimenty s verzí algoritmu *PSO* [4] provedli autoři [5] a popisují svá pozorování doslovně takto: “Bylo to jako pozorování kosmické lodi prohledávající galaxii Mléčné dráhy s cílem nalezení objektu, o němž je předpokládáno, že se nachází ve Sluneční soustavě.” Autoři [5] tak popisují jev, který je, v souvislosti s nelineárními buzenými oscilátory, dle poznatků autorky práce v literatuře označován idiomem *blue sky catastrophe*. Autoři [5] doporučují ke zvýšení stability algoritmu

použití omezovače, kdy částicím je omezeně dovoleno opuštění vyhledávacího prostoru způsobem znázorněným na obrázku Obr. 11.

Autorka předkládané práce soudí, že u funkcí, jejichž tendenční křivka⁶² není podobná parabole, mechanismus dle obrázku Obr. 11 nezaručí návrat částice zpět do vymezeného vyhledávacího prostoru, algoritmus [4] selže a částice diverguje. Je rovněž otázkou, jak se změní stabilita algoritmu v případě, že částicím bude dovolen pohyb pouze v původním vymezeném vyhledávacím prostoru, tj. v prostoru, který je na počátku inicializován.

6.1.1.2 Popis experimentu

Pro simulaci chování částic dle katastrofického scénáře zvolila autorka originální, *pseudo*-paralelní verzi algoritmu *PSO* dle rovnic (1) a (2), s dvaceti jednorozměrnými částicemi, s konstantními koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$. *Stabilita algoritmu*, jako další metrika, byla měřena počtem (četností) úniků částic, nebo pokusů částic o únik mimo původní vyhledávací prostor, vymezený při inicializaci.

Pro první fázi experimentování zvolila autorka *uni*-modální⁶³ funkci *Sphere* (parabola) a *multi*-modální funkci s tendencí paraboly *Rastrigin*. Vzorce, grafy a jednodimenzionální profily obou funkcí jsou v příloze *Příloha A*. Vyhledávací prostor byl v obou případech inicializován v intervalu $[-5.12, +5.12]$. K zajištění shodných výchozích podmínek bylo předem vytvořeno sto souborů s náhodnými čísly z generátoru *pseudo*-náhodných čísel, a s každou funkcí bylo provedeno třikrát sto simulačních běhů, každý v počtu 500 iterací. Částicím byl (i) v první sérii povolen pohyb bez omezení, (ii) ve druhé sérii byl pohyb omezen do intervalu $[-2 * 5.12, +2 * 5.12]$ a (iii) ve třetí sérii byl pohyb omezen do intervalu $[-5.12, +5.12]$. Pro každý simulační běh (simulaci) byla měřena četnost úniků v případě (i) a v případech (ii), (iii) četnost pokusů o únik mimo interval $[-5.12, +5.12]$.

Ve druhé fázi byl experiment zopakován s vymezením vyhledávacího prostoru do intervalu $[-5.12, +5.12]$, tj. stokrát většího, a k testovaným funkcím byla za shodných podmínek připojena *multi*-modální funkce *Schweffel* s tendencí přímky, jejíž vzorec, graf a jednodimenzionální profil je v příloze *Příloha A*.

6.1.1.3 Cíl experimentu

Cílem experimentu bylo ukázat, že výsledky měření podporují zamítnutí hypotéz $\mathcal{H}_0, \mathcal{H}_0''$ a vedou k přijetí alternativních hypotéz $\mathcal{H}_1, \mathcal{H}_1''$:

$\mathcal{H}_0$: Při volbě koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ a při omezování pohybu částic stabilita originálního algoritmu (1), (2) *vzrůstá*.

$\mathcal{H}_1$: Při volbě koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ a při omezování pohybu částic stabilita originálního algoritmu (1), (2) *nevrůstá*.

$\mathcal{H}_0''$: Při volbě koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$, při působení účelové funkce *Schweffel* a pohybu bez omezení nebo s omezením do intervalu $[-2 * 5.12, +2 * 5.12]$ originální algoritmus (1), (2) *neselže*.

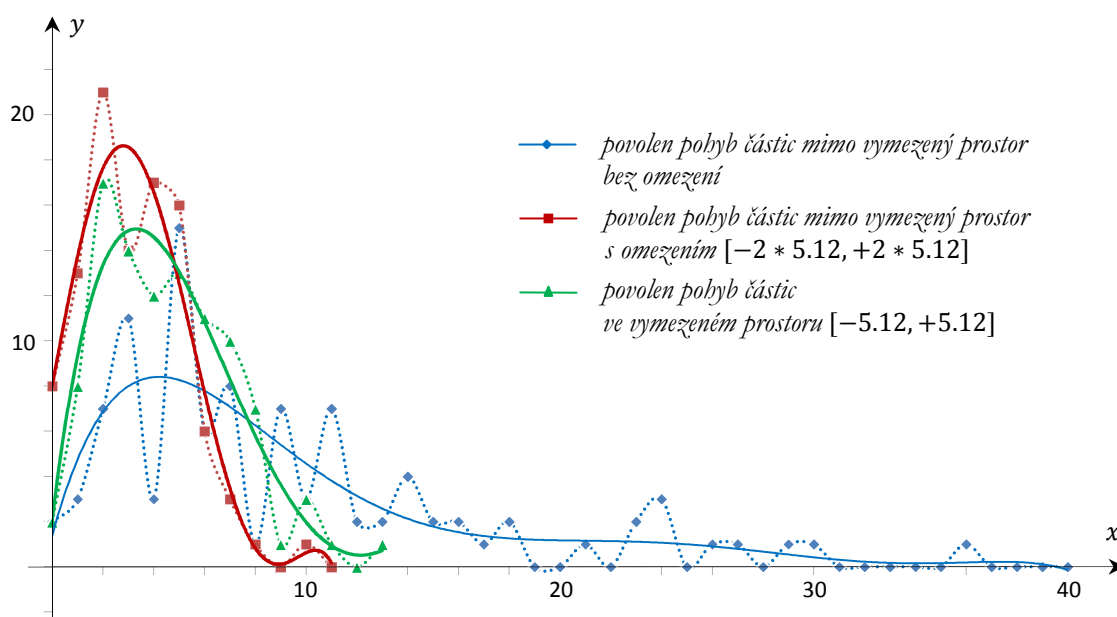
⁶² Tendenční křivkou míní autorka tzv. *spojnici trendu*, která je pro danou funkci spočtena polynomicky podle implicitního vzorce obsaženého v programu *MS Excel*. Tendenční křivky jsou vykresleny v sekci *Příloha A*.

⁶³ Jako *uni*-modální funkce je označována funkce s jedním (globálním) optimem, jako *multi*-modální funkce je označována funkce s více (lokálními) optimy.

$\mathcal{H}_1''$: Při volbě koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$, při působení účelové funkce *Schwefel* a pohybu bez omezení nebo s omezením do intervalu $[-2 * 512, +2 * 512]$ originální algoritmus (1), (2) *selže*.

6.1.1.4 Výsledky experimentu

Výsledky experimentu pro účelové funkce *Sphere* a *Rastrigin* jsou znázorněny graficky na obrázcích Obr. 15(a), Obr. 15(b) a Obr. 16. Výsledky pro funkci *Schwefel* jsou graficky znázorněny na obrázku Obr. 17. Naměřené absolutní četnosti počtu úniků nebo pokusů o únik byly vyneseny jako jednotlivé body a polynomicke regrese polynomy 4. až 6. řádu s využitím implicitního vzorce obsaženého v programu *MS Excel* byly získány tendenční křivky k naměřeným četnostem.

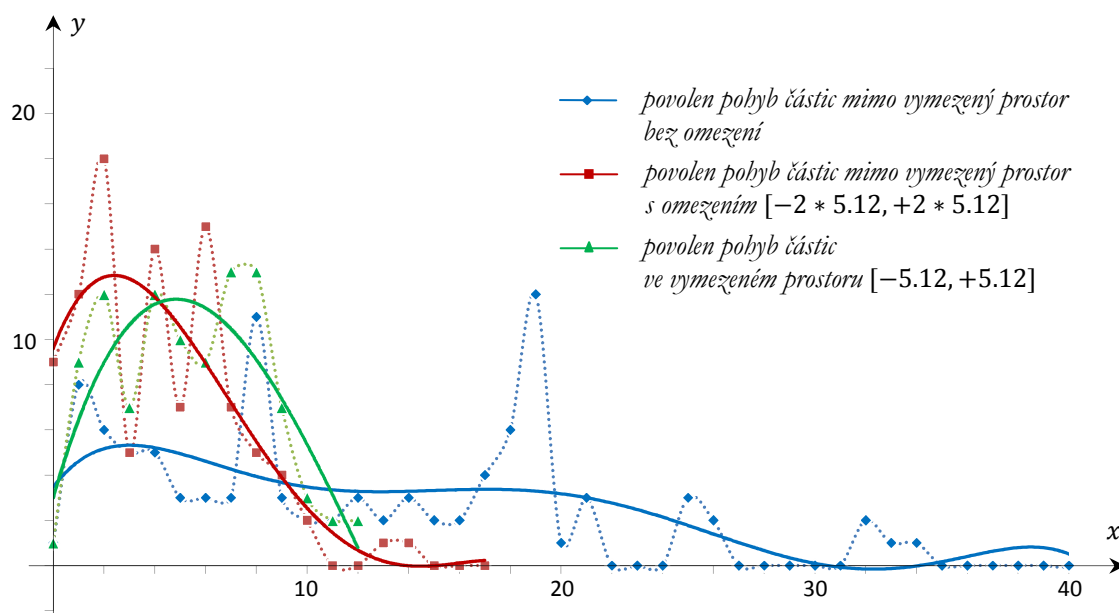


Obr. 15(a): Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k počtům úniků jednorozměrných částic mimo vymezený prostor $[-5.12, +5.12]$ (osa x) (tečkované spojnice) a odpovídající tendenční křivky pro originální algoritmus při působení funkce *Sphere*.

Příklad Počet pokusů o únik mimo prostor $[-2 * 5.12, +2 * 5.12]$ na obrázku Obr. 15(a) činí celkem 2 pokusy u 21 simulací ze 100, 3 pokusy u 14 simulací, atd.

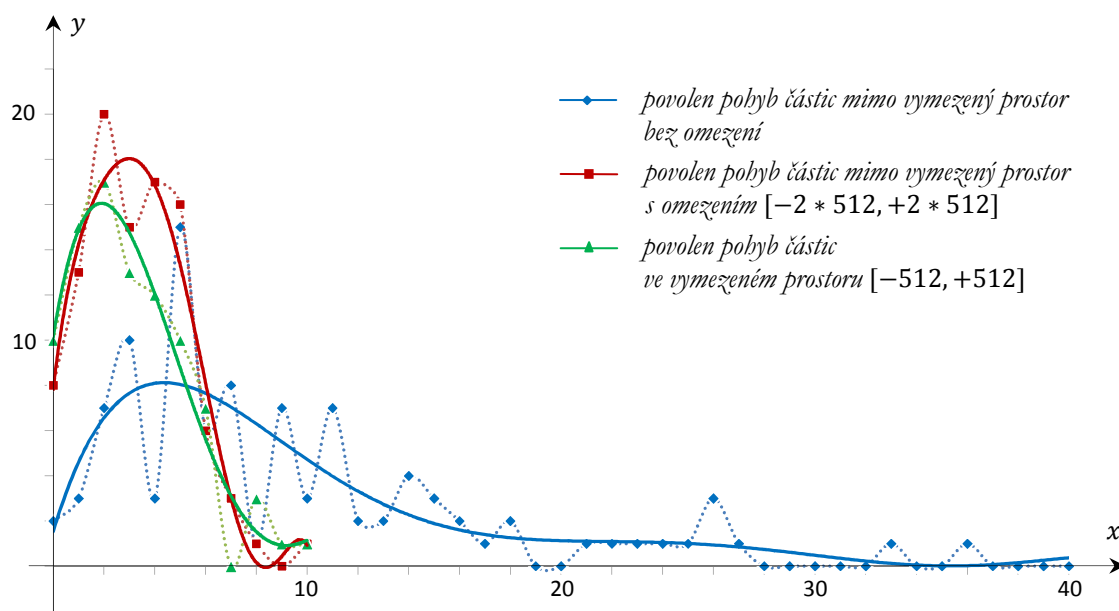
Stabilita algoritmu u varianty s povoleným pohybem do dvojnásobku vymezeného, inicializovaného vyhledávacího prostoru (—■—) se oproti neomezenému pohybu částic (—●—) zvýšila, avšak při dalším omezení pohybu (—▲—) naopak klesla. Tendenční křivky pro funkci *Sphere* vykreslené na obrázku Obr. 15(a) tedy podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$.

Obdobný vývoj byl za daných podmínek zaznamenán pro funkci *Rastrigin*. Tendenční křivky vykreslené na obrázku Obr. 15(b) pro funkci *Rastrigin* rovněž podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$.



Obr. 15(b): Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k počtům úniků jednorozměrných částic mimo vymezený prostor $[-5.12, +5.12]$ (osa x) (tečkované spojnice) a odpovídající tendenční křivky pro originální algoritmus při působení funkce *Rastrigin*.

Výsledkem experimentu s funkcemi *Sphere* a *Rastrigin* za obdobných podmínek, avšak s vymezením inicializovaného vyhledávacího prostoru do intervalu $[-512, 512]$, byly zcela shodné tendenční křivky pro funkci *Sphere*, tendenční křivky pro funkci *Rastrigin* jsou vykresleny na obrázku Obr. 16.

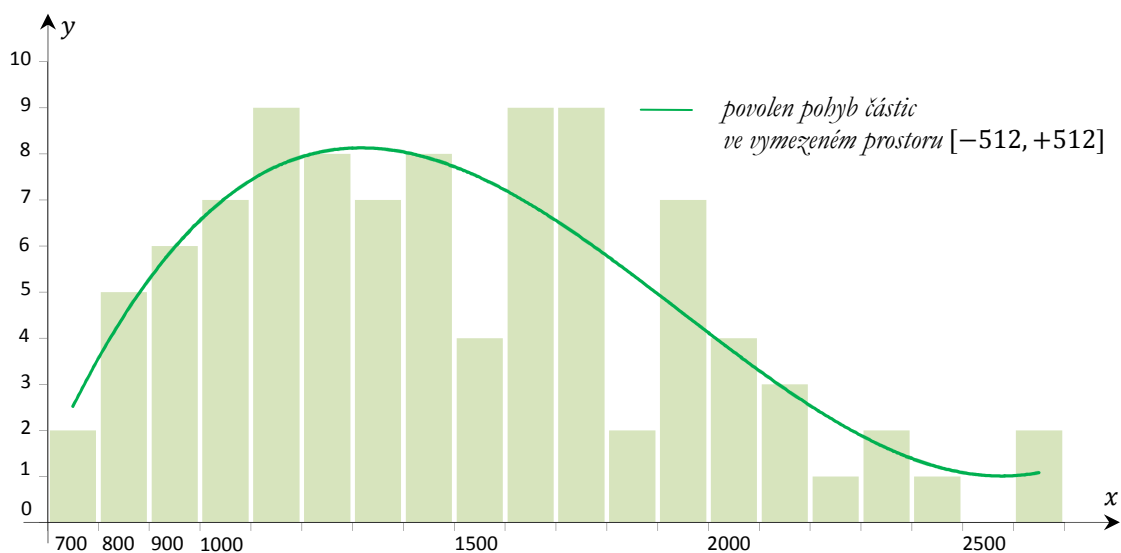


Obr. 16: Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k počtům úniků jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) (tečkované spojnice) a odpovídající tendenční křivky pro originální algoritmus při působení funkce *Rastrigin*.

Tendenční křivky na obrázku Obr. 16 pro funkci *Rastrigin* neposkytují dostatečnou informaci pro zamítnutí hypotézy $\mathcal{H}_0$, ale nevedou ani k jejímu přijetí. Autorka se spíše přiklání k alternativní hypotéze $\mathcal{H}_1$, že při dalším omezení pohybu ($\rightarrow$) stabilita nevzrostla. Shodu v naměřených absolutních četnostech u funkce *Sphere* vysvětluje autorka tím, že funkce *Sphere* je *uni*-modální symetrická funkce (parabola), a zvětší-li se vyhledávací prostor stonásobně, zvětší se stonásobně i počáteční rychlosti částic a poměr mezi naměřenými hodnotami tak zůstane zachován. To ovšem neplatí pro *multi*-modální funkci *Rastrigin*, kde poměry nejsou dány pouze vymezením prostoru při inicializaci, ale též přítomností lokálních optim, která urychlují konvergenci ke globálnímu optimu. Rychlejší konvergenci vysvětluje autorka i nárůst stability při zvětšení vyhledávacího prostoru u funkce *Rastrigin*, jak ukazuje graf na obrázku Obr. 16 v porovnání s grafem na obrázku Obr. 15(b).

Zůstává otázkou, čím je způsoben pokles stability, přesněji, proč stabilita dále nevzrůstá, při dalším omezení pohybu částic ($\rightarrow$), jak ukazují grafy na obrázcích Obr. 15(a) a Obr. 15(b). Autorka soudí, že při daných, neměnných hodnotách koeficientů (ω, a_l, a_g) (1) se vymezený prostor stává částicím příliš těsný, a proto částice více narážejí na mantinely, jež jsou jim dány. Scénář však ještě nelze označit jako katastrofický, což ovšem neplatí pro funkci *Schweffel*.

Tendenční křivka na obrázku Obr. 17 odrážející četnosti pokusů o únik při omezení pohybu na inicializovaný vyhledávací prostor $[-512, +512]$ při působení funkce *Schweffel* byla odvozena z histogramu absolutních četností, jenž vyjadřuje počty simulací na ose y vzhledem k pokusům o únik částic mimo vymezený prostor na ose x .



Obr. 17: Histogram absolutních četností vyjadřující počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) a odpovídající tendenční křivka pro originální algoritmus při působení funkce *Schweffel*.

Příklad Počet pokusů o únik mimo prostor $[-512, +512]$ na obrázku Obr. 17 je mezi 700 až 799 pokusy u 2 simulací ze sta, 800 až 899 pokusů u 5 simulací, 2600 až 2699 pokusů u 2 simulací, atd.

Oproti funkcím *Sphere* a *Rastrigin*, kdy nejvyšší počet úniků mimo inicializovaný vyhledávací prostor při neomezeném pohybu nepřesáhl hodnotu 42, neomezený pohyb částic při působení funkce *Schwefel* vedl ke kolapsu vyhledávání, rovněž tak při omezení pohybu na dvojnásobek inicializovaného vyhledávacího prostoru, neboť vyhledávání záhy přecházelo do pozic se zápornou hodnotou *fitness*. Při omezení pohybu na inicializovaný vyhledávací prostor $[-512, +512]$ byl nárůst hodnoty u dvou simulací značný a ve dvou případech přesáhl hodnotu 2600.

Tendenční křivka vykreslená na obrázku Obr. 17 pro funkci *Schwefel* tedy podporuje zamítnutí hypotézy $\mathcal{H}_0''$ a vede k přijetí alternativní hypotézy $\mathcal{H}_1''$.

6.1.2 Ověření konvergence ke globálnímu optimu

6.1.2.1 Proč byl experiment prováděn

V předchozím experimentování s funkcí *Schwefel* byly zaznamenány simulace, v nichž algoritmus konvergoval předčasně a nedosáhl globálního optima. Neúspěšné simulace byly zaznamenány u 4 simulací ze 100, v nichž došlo k relativně nízkým hodnotám pokusů o únik (konkrétně 38, 48, 56 a 47 pokusů, do histogramu na obrázku Obr. 17 nejsou tyto simulace zahrnuty). Účelem experimentu bylo ověření, zda přítomnost lokálních optim urychluje konvergenci a zda takto zrychlená konvergence, ať již ke globálnímu či lokálnímu optimu, je vysvětlením nárůstu stability. Účelem experimentu bylo tedy zkoumání úspěšnosti řešení a rychlosti konvergence.

6.1.2.2 Popis experimentu

Pro ověření konvergence ke globálnímu optimu zvolila autorka originální, *pseudo*-paralelní verzi algoritmu *PSO* dle rovnic (1) a (2), s dvaceti jednorozměrnými částicemi, s konstantními koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$.

V první fázi bylo experimentováno s funkcí *Sphere* (parabola), funkcí *Rastrigin* a funkcí *Schwefel*, jejichž vzorce, grafy a jednodimenzionální profily jsou v příloze *Příloha A*. Vyhledávací prostor byl ve všech případech inicializován v intervalu $[-512, +512]$ a částicím byl pohyb omezen shodně, do intervalu $[-512, +512]$. K zajištění shodných výchozích podmínek bylo použito sto předem vytvořených souborů s *pseudo*-náhodnými čísly z předchozího experimentu a s každou funkcí bylo provedeno sto simulačních běhů, každý v počtu 500 iterací. Ve druhé fázi bylo experimentováno pouze s funkcemi *Sphere* a *Rastrigin*. Experiment byl doplněn o sérii měření bez omezení vyhledávacího prostoru a sérii s vymezením vyhledávacího prostoru do intervalu $[-2 * 512, +2 * 512]$ a byla doplněna tři odpovídající srovnávací měření pro prostor stokrát menší, tj. $[-5.12, +5.12]$.

Pro každou simulaci byla vyhodnocena úspěšnost řešení a rychlost konvergence. *Úspěšnost řešení*, jako metrika na nejvyšší úrovni dle obrázku Obr. 3, byla měřena počtem simulací (z celkového počtu 100 simulací), v nichž algoritmus konvergoval ke správnému, globálnímu optimu. *Rychlost konvergence* byla zavedena jako další metrika a byla pro každou simulaci měřena jako pořadí iterace, v níž byla naposledy zaznamenána změna v globální nejlepší pozici. Iterace byly číslovány 0 až 499. Výsledky druhé fáze měření byly dány do souvislosti s výsledky měření stability v předchozím experimentu.

6.1.2.3 Cíl experimentu

Cílem experimentu bylo ukázat, že výsledky měření podporují zamítnutí hypotéz $\mathcal{H}_0$, $\mathcal{H}_0'$, $\mathcal{H}_0''$ a vedou k přijetí alternativních hypotéz $\mathcal{H}_1$, $\mathcal{H}_1'$, $\mathcal{H}_1''$:

$\mathcal{H}_0$: Přítomnost lokálních optim *neurčuje* konvergenci.

$\mathcal{H}_1$: Přítomnost lokálních optim *urychluje* konvergenci.

$\mathcal{H}_0'$: Při postupném omezování pohybu částic do inicializovaného vyhledávacího prostoru rychlost konvergence *vzrůstá*.

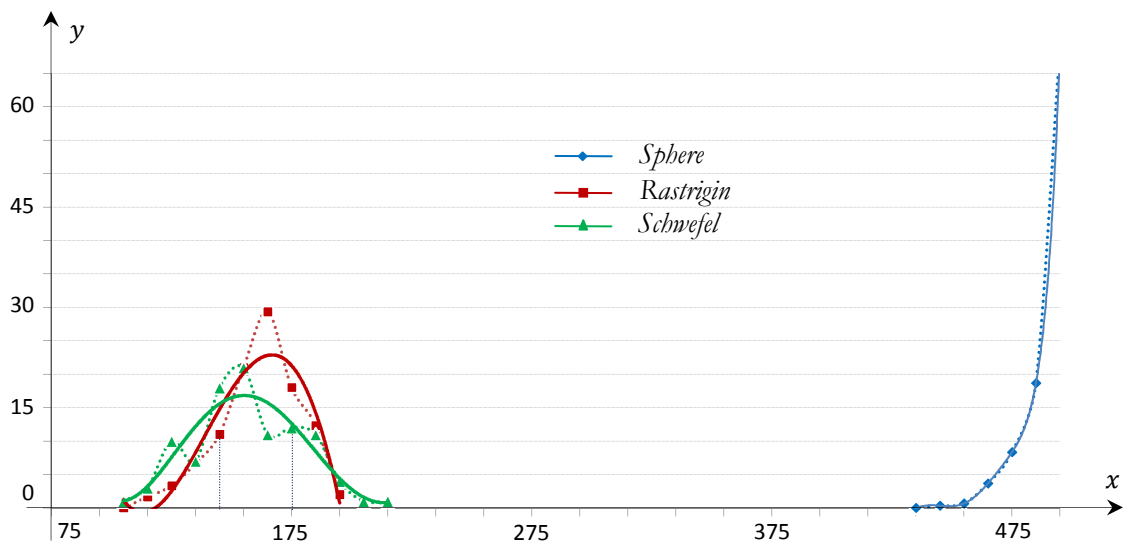
$\mathcal{H}_1'$: Při postupném omezování pohybu částic do inicializovaného vyhledávacího prostoru rychlost konvergence *nevrůstá*.

$\mathcal{H}_0''$: Při volbě koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ konverguje úspěšně více než 95% případů.

$\mathcal{H}_1''$: Volba koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ nezaručuje 95% úspěšnost.

6.1.2.4 Výsledky experimentu

Výsledky první fáze experimentování jsou graficky znázorněny na obrázku Obr. 18.



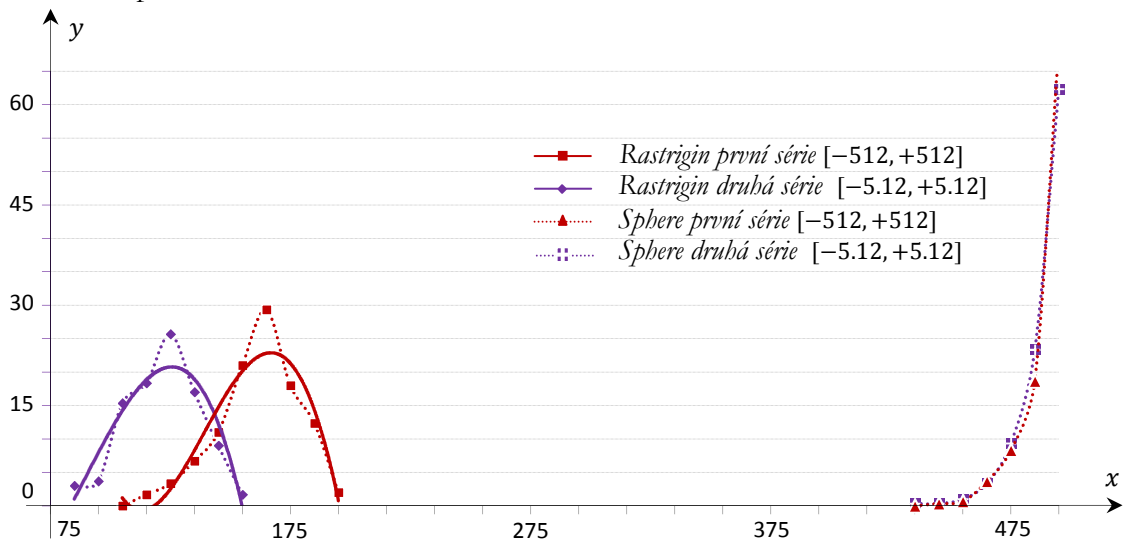
Obr. 18: Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k rychlosti konvergence (osa x) (tečkované spojnice) a odpovídající tendenční křivky pro originální algoritmus, při inicializaci a vymezení vyhledávacího prostoru $[-512, +512]$, při působení funkcí *Sphere*, *Rastrigin* a *Schwefel*.

Příklad Při působení funkce *Rastrigin* byla zaznamenána poslední změna u globální nejlepší pozice (přibližně) při 145. iteraci u 10 simulací, při působení funkce *Schwefel* konvergovalo 12 simulací při 175. iteraci, atd.

Z grafu na obrázku Obr. 18 lze vysledovat, že při působení *uni*-modální funkce *Sphere* se algoritmus vyznačoval relativně pomalou konvergencí, neboť měl tendenci zpřesňovat globální nejlepší pozici až do posledních iterací, i když při nižší požadované přesnosti sledované změny v globální nejlepší pozici by bylo možno algoritmus ukončit dříve.

Přítomnost lokálních optim u *multi*-modálních funkcí *Rastrigin* a *Schweffel* konvergenci algoritmu urychluje, což autorka vysvětluje tím, že lokální optima působí jako “můstky”, které mohou částice při svém pohybu ke globálnímu optimu využít. Tendenční křivky vykreslené na obrázku Obr. 18 tedy podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$.

Graf na obrázku Obr. 19 zobrazuje výsledky druhé fáze experimentu, v níž byl zkoumán vliv omezení pohybu částic (pohyb bez omezení, s omezením na dvojnásobek vyhledávacího prostoru, s omezením do vyhledávacího prostoru) na rychlost konvergence algoritmu při působení funkcí *Sphere* a *Rastrigin*. Pro první sérii měření byl vyhledávací prostor inicializován do intervalu $[-512, +512]$, pro druhou sérii do intervalu stokrát zmenšeného $[-5.12, +5.12]$. Počáteční rychlosti částic byly voleny úměrně velikosti vymezeného prostoru.



Obr. 19: Absolutní četnosti vyjadřující počty simulací (osa y) vzhledem k rychlosti konvergence (osa x) (tečkované spojnice) a odpovídající tendenční křivky pro originální algoritmus, při inicializaci a vymezení vyhledávacího prostoru $[-512, +512]$ a $[-5.12, +5.12]$, při působení funkcí *Sphere* a *Rastrigin*.

Povolení pohybu částic mimo inicializovaný prostor nemělo na rychlost konvergence vliv. Pro každou z funkcí byly v první sérii naměřeny tři téměř shodné křivky, které byly sloučeny v jednu (—■—). Obdobná situace nastala ve druhé sérii měření (—■—). Získané výsledky tedy podporují zamítnutí hypotézy $\mathcal{H}_0'$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1'$.

Rychlost konvergence algoritmu při působení funkce *Rastrigin* však byla ovlivněna velikostí inicializovaného vyhledávacího prostoru. Graf na obrázku Obr. 19 ukazuje rychlejší konvergenci pro funkci *Rastrigin* v druhé sérii měření (—■—), pro menší inicializovaný vyhledávací prostor, v porovnání s první sérií (—■—), pro stonásobně větší inicializovaný vyhledávací prostor. Rychlejší konvergenci vysvětluje autorka hodnotami koeficientů zrychlení a setrvačnosti (ω, a_l, a_g) , které byly v obou sériích shodné, a tedy menší prostor v přítomnosti lokálních optim byl prohledán rychleji.

Úspěšnost řešení, měřena počtem simulací, v nichž algoritmus konvergoval ke správnému, globálnímu optimu, byla u funkcí *Sphere* a *Rastrigin* v první i druhé fázi

experimentů 100% a u funkce *Schwefel* v první fázi experimentů 96%. Výsledky měření úspěšnosti řešení vedou k přijetí hypotézy $\mathcal{H}_0$, avšak autorka se přesto brání přijetí hypotézy $\mathcal{H}_0$ a přiklání se k alternativní hypotéze $\mathcal{H}_1$, že volba koeficientů $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ úspěšnost 95% nezaručuje,⁶⁴ neboť neúspěšné simulace při působení funkce *Schwefel*, jež činily 4% a v nichž byl zaznamenán nízký počet úniků, zůstaly neobjasněny.

Výsledky provedených experimentů totiž neposkytly dostatečnou informaci pro odpověď na otázku, zda existuje příčinný vztah mezi rychlostí konvergence měřenou pořadím iterace, v níž byla naposledy zaznamenána změna v globální nejlepší pozici,⁶⁵ a stabilitou, měřenou počtem úniků či pokusů o únik mimo inicializovaný vyhledávací prostor. Autorka se pokusila o sledování vývoje počtu úniků mimo vymezený prostor v závislosti na čase měřeném pořadím iterace⁶⁶ při působení funkce *Schwefel*, očekávajíc, že úniky budou častější v počátcích vyhledávání, jednotlivé simulace se však vyvíjely zcela neporovnatelně. Autorka tedy usoudila, že zde musí působit ještě další prvek, ovlivňující budoucí vývoj prohledávání v čase, a tímto prvkem je *inicializace vyhledávacího prostoru*.

6.1.3 Vliv inicializace na výsledek prohledávání

6.1.3.1 Proč byl experiment prováděn

Důvodem zkoumání vlivu rozmístění částic ve vyhledávacím prostoru po inicializaci na budoucí vývoj prohledávání v čase bylo zjištění, že rychlost konvergence, měřená pořadím iterace, v níž byla naposledy zaznamenána změna v globální nejlepší pozici, ani stabilita, měřená počtem úniků či pokusů o únik mimo inicializovaný vyhledávací prostor, nevypovídají nic o konvergenci algoritmu ke správnému, globálnímu optimu.

Na rychlost konvergence a stabilitu u (zkoumaných) *multi*-modálních funkcí měla určitý vliv velikost vyhledávacího prostoru, nelze však usoudit, zda rychlejší konvergence vedla či nevedla k nárůstu stability nebo zda nižší stabilita (vyšší nestabilita projevující se vyšším počtem úniků mimo inicializovaný vyhledávací prostor) zaručovala lepší prohledávací schopnost. Nejlepší prohledávací schopnost by měl vykazovat originální algoritmus [4] s koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ při působení funkce *Schwefel*, s nestabilitou o dva řády vyšší oproti funkci *Rastrigin*, přesto algoritmus u funkce *Schwefel* selhal v úspěšnosti vyhledávání ve čtyřech případech ze sta.⁶⁷

Jelikož je systém částice citlivý na počáteční podmínky, jak bylo blíže pojednáno v kapitolách 4.2.5 a 4.7.3, rozhodla se autorka zkoumat, nakolik je rozmístění částic ve vyhledávacím prostoru po inicializaci rovnoměrné a jaký vliv má počáteční rozmístění částic na úspěšnost řešení, která je měřena (poměrným) počtem simulací, v nichž algoritmus konvergoval ke správnému, globálnímu optimu.

⁶⁴ Autorka důvodně předpokládala, že větší neúspěšnost nastane u vícerozměrného vyhledávacího prostoru, jak též ukázaly experimenty provedené později.

⁶⁵ Algoritmus lze ukončit při dosažení určitého počtu iterací, nebo dosáhne-li rozdíl mezi globální nejlepší pozicí vzhledem k předchozí iteraci určité hodnoty ε . Autorka zvolila nejnižší možnou hodnotu ε , aby mohla pozorovat vývoj algoritmu co nejdéle.

⁶⁶ Simulace je diskrétní proces s časovým krokem $\Delta t = 1$, a proto lze čas měřit přímo v iteracích.

⁶⁷ Další zhoršení autorka důvodně předpokládala u vícerozměrného vyhledávacího prostoru.

6.1.3.2 Popis experimentu

Pro zkoumání vlivu inicializace vyhledávacího prostoru na konvergenci algoritmu ke globálnímu či lokálnímu optimu zvolila autorka originální, *pseudo*-paralelní verzi algoritmu *PSO* dle rovnic (1) a (2), s dvaceti jednorozměrnými částicemi, s konstantními koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$.

Bylo experimentováno s funkcí *Schwefel*, jejíž vzorec, graf a jednodimenzionální profil jsou v příloze *Příloha A*. Vyhledávací prostor byl inicializován v intervalu $[-512, +512]$ a částicím byl pohyb omezen shodně do intervalu $[-512, +512]$. K zajištění srovnatelných výchozích podmínek bylo použito sto předem vytvořených souborů s *pseudo*-náhodnými čísly z předchozích experimentů, které byly označeny *Random_xxx*, kde $xxx \approx 010 \dots 109$, a bylo provedeno sto simulačních běhů, každý běh s jedinou iterací, a bylo pozorováno rozmístění částic ve vyhledávacím prostoru po inicializaci. Získané informace byly porovnávány s výsledky předchozích experimentů s funkcí *Schwefel* a byl posouzen vliv každého jednotlivého rozmístění na budoucí vývoj a konvergenci.

6.1.3.3 Cíl experimentu

Cílem experimentu bylo ukázat, že výsledky pozorování podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$:

$\mathcal{H}_0$: Rozmístění částic ve vyhledávacím prostoru po inicializaci *nemá zásadní vliv* na úspěšnost řešení měřenou poměrným počtem (%) simulací, v nichž algoritmus konvergoval ke globálnímu optimu.

$\mathcal{H}_1$: Rozmístění částic ve vyhledávacím prostoru po inicializaci *má zásadní vliv* na úspěšnost řešení měřenou poměrným počtem (%) simulací, v nichž algoritmus konvergoval ke globálnímu optimu.

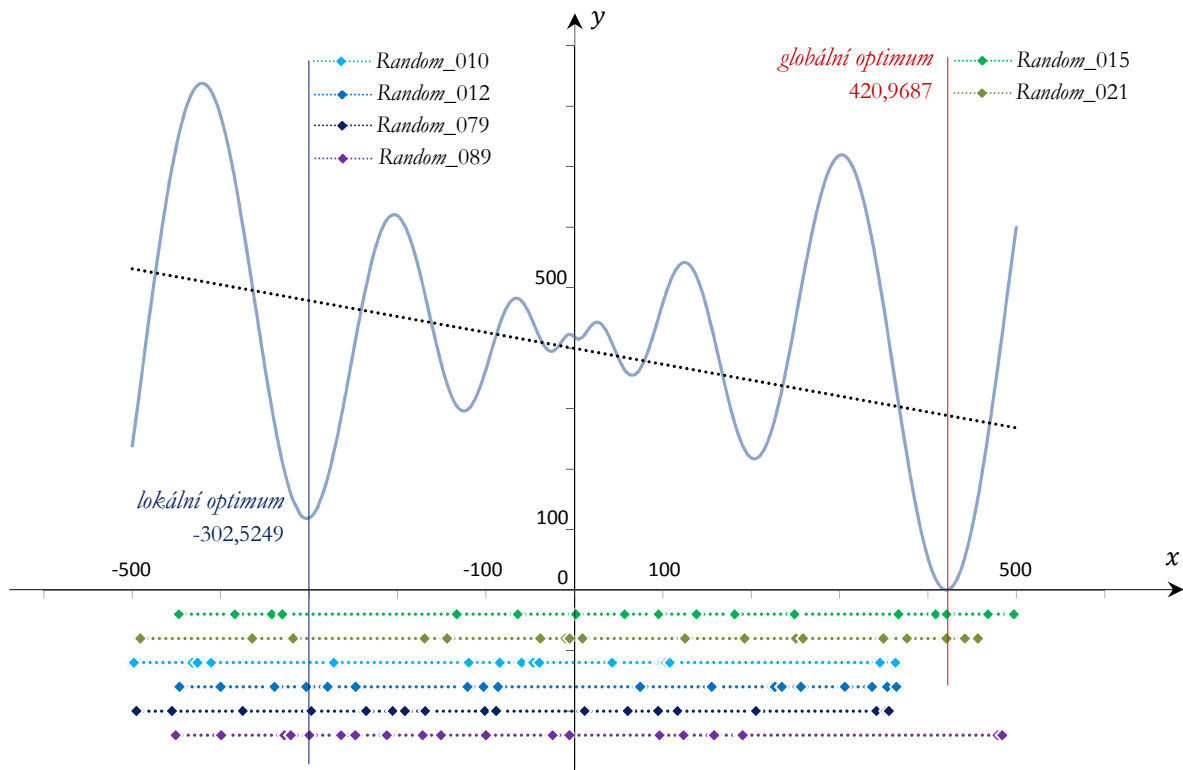
6.1.3.4 Výsledky experimentu

Prezentaci výsledků pozorování vlivu rozmístění částic po inicializaci na konvergenci algoritmu pro vybrané simulace ukazuje obrázek Obr. 20.

Jelikož je počet částic relativně malý, počáteční rozmístění částic je nerovnoměrné. Úspěšný výsledek prohledávání tudíž závisí na dostatečné prohledávací schopnosti algoritmu. Z výsledků experimentů je patrné, že prohledávací schopnost algoritmu není dostatečná. V simulacích s použitím souborů *Random_010*, *Random_012*, *Random_079* a *Random_089* se algoritmus zjevně nepokusil prohledat oblast kolem globálního optima, v níž na počátku prohledávání nebyly rozmístěny žádné částice, a nechal se strhnout k předčasné konvergenci k lokálnímu optimu. Správný výsledek simulací používajících soubory *Random_015* a *Random_021* rovněž nemusí být zásluhou dostatečné prohledávací schopnosti algoritmu, neboť počáteční pozice $x = 420,87$ při použití souboru *Random_015* a $x = 420,73$ při použití *Random_021* způsobují, že algoritmus již lepší pozice ve zbylých oblastech nenalezne, a tedy dospěje ke správnému řešení.

Výsledky pozorování tedy podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$ a navíc ukazují na nedostatečnou prohledávací schopnost algoritmu, kterou není možno zlepšit dalším snižováním stability (zvyšováním nestability)

pomocí koeficientů (ω, a_l, a_g) , neboť nestabilita algoritmu při působení funkce *Schwefel*, jak ji znázorňuje předchozí obrázek Obr. 17, je již nyní značná.



Obr. 20: Jednodimenzionální profil funkce *Schwefel* při inicializaci vyhledávacího prostoru $[-512, +512]$ a rozmístění částic ve vyhledávacím prostoru po inicializaci pro simulace s vybranými, předem vytvořenými soubory s *pseudo*-náhodnými čísly.

Chování částic ve vyhledávacím prostoru projevující se vysokým počtem pokusů o únik z vymezeného vyhledávacího prostoru při působení funkce *Schwefel*, jež je nesrovnatelné s umírněným chováním částic při působení funkcí *Rastrigin* nebo *Sphere*, však dle autorky nelze přičítat pouze režimu na mezi vnitřní stability systému, ale též značně odlišnému charakteru působení funkce *Schwefel*. Skutečnost, že různé účelové funkce mají odlišný vliv na budoucí vývoj prohledávání, byla demonstrována v následujícím experimentu.

6.1.4 Vliv účelové funkce na průběh prohledávání

6.1.4.1 Proč byl experiment prováděn

Jednoduchý experiment má ukázat, že různé účelové funkce mají různý vliv na budoucí vývoj prohledávání, a to za zcela srovnatelných podmínek daných velikostí vyhledávacího prostoru, hodnotami koeficientů (ω, a_l, a_g) i tím, že pohyb částic mimo inicializovaný vyhledávací prostor není nijak omezen.

6.1.4.2 Popis experimentu

Pro zkoumání vlivu různých účelových funkcí na vývoj prohledávání zvolila autorka originální, *pseudo*-paralelní verzi algoritmu *PSO* dle rovnic (1) a (2), s dvaceti jednorozměrnými částicemi, s konstantními koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$.

Bylo experimentováno s funkcemi *Rastrigin*, *Sphere* a *Schweffel*, jejichž vzorce, grafy a jednodimenzionální profily jsou v příloze *Příloha A*. Vyhledávací prostor byl inicializován v intervalu $[-512, +512]$ a částicím nebyl pohyb nijak omezen. Částicím bylo na počátku prohledávání přiřazeno pořadí $0 \approx 19$ shodné ve všech simulacích. K zajištění shodných výchozích podmínek byl použit soubor s *pseudo*-náhodnými čísly *Random_010* a byla pozorována trajektorie (náhodně vybrané) částice, s pořadím 16, ve všech iteracích. Trajektorie částic byly následně porovnány.

6.1.4.3 Cíl experimentu

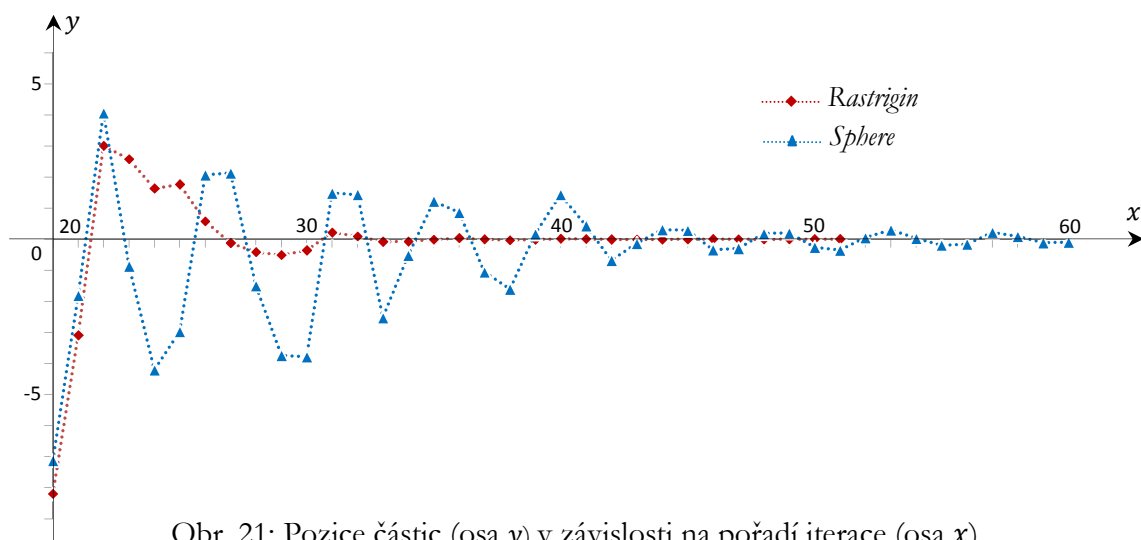
Cílem experimentu bylo ukázat, že výsledky porovnání trajektorií podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$:

$\mathcal{H}_0$: Působení různých účelových funkcí *nemá vliv* na průběh prohledávání.

$\mathcal{H}_1$: Působení různých účelových funkcí *má vliv* na průběh prohledávání.

6.1.4.4 Výsledky experimentu

Porovnání trajektorií při působení funkcí *Rastrigin* a *Sphere* znázorňuje graf závislosti pozice částice (y) na pořadí iterace (x) na obrázku Obr. 21. Výsledky jsou znázorněny od 20. iterace, kdy se rozdíl mezi trajektoriemi začaly projevovat zcela znatelně. Průběh trajektorie pro funkci *Schweffel* není zakreslen, neboť z výchozí pozice $x = -59.88$ se částice dopracovala ve 20. iteraci na pozici s hodnotou $x = -42295.66$ a ve 110. iteraci na pozici s hodnotou $x = -4.28 \cdot 10^{10}$. Výsledky porovnání podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$.



Obr. 21: Pozice částic (osa y) v závislosti na pořadí iterace (osa x).

Porovnání odpovídajících trajektorií částic při působení funkcí *Rastrigin* a *Sphere* pro vybranou částici a vybraný soubor s *pseudo*-náhodnými čísly, při inicializaci vyhledávacího prostoru $[-512, +512]$ a neomezeném pohybu částic.

6.1.5 Vyhodnocení výsledků experimentování s originální verzí algoritmu *PSO* s konstantními koeficienty

Výsledky experimentování s originální verzí algoritmu *PSO* s konstantními koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ potvrdily a/nebo přinesly nové, následující poznatky:

- (i) Není-li pohyb částic omezen do inicializovaného vyhledávacího prostoru, algoritmus selhává (částice divergují) při působení účelových funkcí, jejichž tendenční křivka se nepodobá parabole.
- (ii) Stabilita algoritmu s konstantními koeficienty, měřená počtem úniků nebo pokusů o únik částic mimo vyhledávací prostor vymezený při inicializaci, je při neomezeném pohybu částic nízká (nestabilita je vysoká), při omezování pohybu částic do určité velikosti prostoru stabilita vzrůstá (nestabilita klesá), při dalším omezování pohybu již stabilita nevzrůstá a začíná opět klesat (tj. nestabilita vzrůstá).
- (iii) Je-li je počáteční rychlost částic úměrná velikosti vyhledávacího prostoru vymezeného při inicializaci, potom velikost tohoto prostoru nemá vliv na rychlost konvergence algoritmu, měřenou pořadím iterace, v níž byla naposledy zaznamenána změna v globální nejlepší pozici.
- (iv) Přítomnost lokálních optim konvergenci algoritmu urychluje. Za přítomnosti lokálních optim je menší vyhledávací prostor při konstantních hodnotách koeficientů prohledán rychleji.
- (v) Stabilita algoritmu, měřená počtem úniků nebo pokusů o únik částic mimo vyhledávací prostor vymezený při inicializaci, není ukazatelem prohledávacích schopností algoritmu ani konvergence algoritmu ke správnému, globálnímu optimu.
- (vi) Prohledávací schopnost algoritmu s konstantními koeficienty je ovlivněna též počátečním rozmístěním částic ve vyhledávacím prostoru.
- (vii) Průběh trajektorií částic je ovlivňován působením účelové funkce, přičemž vliv účelové funkce není nevýznamný a může vést k divergenci částic.

Zásadní nedostatky originální verze algoritmu *PSO* s konstantními koeficienty $(\omega, a_l, a_g) = (0.72, 1.48, 1.48)$ spatřuje autorka předkládané práce v tom, že algoritmus nebere v úvahu omezený počet částic, jimiž disponuje, a nemění strategii vzhledem k velikosti vyhledávacího prostoru, ani vzhledem k rovnoměrnosti rozmístění částic ve vyhledávacím prostoru po inicializaci. Algoritmus rovněž nesleduje vliv působení účelové funkce na trajektorie částic. Velikosti vyhledávacího prostoru a působení účelové funkce nejsou přizpůsobeny ani hodnoty koeficientů, jež vyjadřují setrvačnost částice a lokální a globální zrychlení. Pro relativně malý vyhledávací prostor mohou být, v přítomnosti lokálních optim, hodnoty příliš vysoké a mohou tak (neúměrně) zvyšovat nestabilitu.

Vysoká nestabilita algoritmu přitom není zárukou dostatečné prohledávací schopnosti algoritmu, neboť prohledávací schopnost algoritmu je ovlivňována i rovnoměrností rozmístění částic ve vyhledávacím prostoru po inicializaci. Při nerovnoměrném rozmístění částic musí být posíleno lokální prohledávání vymezeného vyhledávacího prostoru a při menší velikosti vyhledávacího prostoru musí být hodnoty všech koeficientů nižší. Hodnoty koeficientů je tedy nutno přizpůsobit i (ne)rovnoměrnosti rozmístění částic ve vyhledávacím prostoru po inicializaci a průběžným přizpůsobováním hodnot koeficientů

během simulace je nutno usměrňovat vývoj prohledávání tak, aby bylo dosaženo dostatečně vysoké pravděpodobnosti, že algoritmus konverguje ke globálnímu optimu.

6.2 Experimenty s navrhovanou verzí algoritmu *PSO*

Navržená verze algoritmu *PSO* podle obrázků Obr. 14(a) a Obr. 14(b) přináší změnu strategie prohledávání, jejímž záměrem je odstranění nedostatků originální verze algoritmu s konstantními koeficienty [4], plynoucích z vyhodnocení výsledků experimentování v kapitole 6.1.5.

Navrhovaný algoritmus vychází ze strategie proměnných koeficientů [11] [23] [39, 40], v níž koeficienty ω , a_l (1), které se vztahují k setrvačné a kognitivní složce, v průběhu prohledávání lineárně klesají, a koeficient a_g (1), který se vztahuje k sociální složce, v průběhu prohledávání lineárně stoupá. Oproti těmto původním strategiím zavádí nová strategie čtyři režimy (*mood*) prohledávání dle Obr. 14(b), v nichž jsou hodnoty koeficientů odvozeny přímo z charakteristik prostředí a nemají lineární průběh:

- (i) lokální prohledávání (*mood* = 0),
- (ii) rovnoměrné prohledávání s vyváženou kognitivní a sociální složkou (*mood* = 1),
- (iii) lokální prohledávání (*mood* = 2) a
- (iv) globální prohledávání s potlačenou setrvačnou složkou a posílenou globální složkou (*mood* = 3).

Navrhovaný algoritmus *PSO* byl podroben počátečnímu testování, jehož účelem bylo ověření způsobu sledování rovnoměrnosti prohledaných pozic, dále pak ověření navrženého způsobu detekce výlučnosti globální nejlepší *fitness* a zpřesnění či nalezení nových podmínek přechodu mezi jednotlivými režimy prohledávání (*mood* na obrázku Obr. 14(b)). Následné experimenty byly prováděny s již upravenou verzí navrhovaného algoritmu a byly zaměřeny na měření stability, ověření konvergence ke globálnímu optimu a porovnání výsledků s výsledky předchozích experimentů s originální verzí algoritmu [4] prezentovanými v kapitole 6.1.

6.2.1 Vliv inicializace na výsledek prohledávání, sledování rovnoměrnosti a detekce výlučnosti globální nejlepší *fitness*

6.2.1.1 Proč byl experiment prováděn

Experiment realizuje počáteční testování navrhované verze algoritmu *PSO*. Návrh nové strategie prohledávání aplikuje zcela nové postupy a bylo nutno připustit, že se při experimentování mohou objevit nové poznatky, že dosavadní předpoklady nemusí být naplněny a že původní návrh může doznat změn. Počáteční experimenty s navrženou verzí algoritmu *PSO* byly proto založeny na detailním pozorování, jakých hodnot bude nabývat pravděpodobnost rovnoměrnosti prozkoumaných pozic, jak se budou vyvíjet naměřené hodnoty, zda budou detekovány výlučné hodnoty *fitness* a jaké hodnoty pravděpodobností tomu napovídají.⁶⁸

⁶⁸ Obvyklý způsob experimentování bývá založen na statistickém vyhodnocení výsledků velkého počtu simulací (v počtu 10000) s mnoha iteracemi (10000) a vícerozměrnými (10) částicemi.

6.2.1.2 Popis experimentu

Experimentu, spočívajícímu ve zkoumání vlivu inicializace na výsledek prohledávání, sledování rovnoměrnosti prozkoumaných pozic a detekci výlučnosti globální nejlepší *fitness*, byl podroben *pseudo*-paralelní algoritmus *PSO* na obrázcích Obr. 14(a), Obr. 14(b).

Pro pozorování byl zvolen roj s dvaceti jednorozměrnými částicemi a *multi*-modální funkce *Schwefel*, jejíž vzorec, graf a jednodimenzionální profil je v příloze *Příloha A*. K zajištění shodných výchozích podmínek s předchozími experimenty byly použity tytéž, již vytvořené soubory s *pseudo*-náhodnými čísly z předchozích experimentů. S funkcí *Schwefel* bylo provedeno sto simulačních běhů, každý v počtu 500 iterací.

Částicím byl pohyb omezen do intervalu $[-512, +512]$ a byly porovnávány iterace jednotlivých simulací, měřeny pravděpodobnosti rovnoměrného rozmístění dosud prohledaných pozic dle (158), (159), (160), z nich počítány koeficienty $(\omega[n], a_l[n], a_g[n])$ pro $n = 1$ (případ jednorozměrné částice) dle postupu, jehož konkrétní příklad je uveden v kapitole 5.4. Průběžně byly detekovány podmínky, z nichž bylo možno usoudit na *vylučnost* dosud nejlepší globální pozice dle postupu popsaneho v kapitole 5.2.2.2, to vše ve vztahu k *úspěšnosti řešení*.

Pro první fázi experimentování s *pravděpodobností vylučnosti dosud nejlepší globální pozice* bylo zvoleno testovací kritérium r_{22} (157). Analogické (reciproké) testovací kritérium bylo aplikováno též pro stanovení *pravděpodobnosti vylučnosti dosud nejhorší globální pozice*, neboť předběžné výsledky experimentů napověděly, že obě kritéria mají rovnocennou vypovídací schopnost:

$$r_{22min} = \frac{y_{(3)} - y_{(1)}}{y_{(18)} - y_{(1)}} \quad r_{22max} = \frac{y_{(20)} - y_{(18)}}{y_{(20)} - y_{(3)}} \quad (161)$$

6.2.1.3 Cíl experimentu

Cílem experimentu bylo ukázat, že výsledky měření podporují zamítnutí hypotéz $\mathcal{H}_0, \mathcal{H}_0''$ a vedou k přijetí alternativních hypotéz $\mathcal{H}_1, \mathcal{H}_1''$:

$\mathcal{H}_0$: Mezi hodnotou pravděpodobností rovnoměrného rozmístění pozic ve vyhledávacím prostoru a volbou hodnot koeficientů *není souvislost*.

$\mathcal{H}_1$: Mezi hodnotou pravděpodobností rovnoměrného rozmístění pozic ve vyhledávacím prostoru a volbou hodnot koeficientů *lze nalézt souvislost*.

$\mathcal{H}_0''$: Vyhodnocením pravděpodobností výlučnosti globální nejlepší a globální nejhorší *fitness* *nelze indikovat konvergenci* k lokálnímu nebo globálnímu optimu a volit režim prohledávání.

$\mathcal{H}_1''$: Vyhodnocením pravděpodobností výlučnosti globální nejlepší a globální nejhorší *fitness* *lze indikovat konvergenci* k lokálnímu nebo globálnímu optimu a volit režim prohledávání.

6.2.1.4 Výsledky experimentu

Průběžným vyhodnocováním pravděpodobností *vylučnosti* globální nejlepší a globální nejhorší *fitness*, označených *probMin* a *probMax*, a to v každé iteraci, ve všech simulacích,

byly vypočítávány hodnoty pravděpodobností výlučnosti $p \in \{0.5, 0.8, 0.95, 0.995\}$ (156), jež představují určité mezní hodnoty, které mají zvláštní význam pro stanovení podmínek přechodu mezi jednotlivými režimy prohledávání. Byly zavedeny logické proměnné, pro něž byly vnitřním pozorovatelem na počátku každé iterace vyhodnoceny pravdivostní hodnoty následujícím způsobem:

$$\begin{aligned}
 LocMin &= (probMin > 0.5) \wedge (probMin < 0.8) \\
 LocMax &= (probMax > 0.5) \wedge (probMax < 0.8) \\
 GlobMin &= (probMin \geq 0.8) \\
 GlobMinMin &= (probMin \geq 0.95) \\
 GlobMax &= (probMax \geq 0.8) \\
 GlobMaxMax &= (probMax \geq 0.995) \\
 MinZero &= (probMin \geq 0.0) \wedge (probMin \leq 0.5) \\
 MaxZero &= (probMax \geq 0.0) \wedge (probMax \leq 0.5)
 \end{aligned} \tag{162}$$

Vnitřní pozorovatel si rovněž udržoval pro (každý) rozměr částice n hodnotu dosud nejvyšší dosažené pravděpodobnosti $pMax[n]$ v průběhu simulace a v proměnné $probability[n]$ si průběžně pamatoval hodnotu pravděpodobnosti rovnoměrného rozmístění pozic z předchozí iterace, spočtenou dle rovnic (158), (159), (160). Aktuální hodnotu pravděpodobnosti obsahovala proměnná $histogram[n].ProbabilityOfUniform$. V první fázi bylo experimentováno s částicemi jednorozměrnými.

Hodnota proměnné $probability[n] = 1$ byla nastavena jako stav po inicializaci, a tedy $(probability[n] < 1)$ značí, že se nejedná o první iteraci. V průběhu dalších iterací nastávaly nepříznivé situace, které musel vnitřní pozorovatel rozpoznat a řešit, a naopak, nastávaly příznivé situace, které musel vnitřní pozorovatel rozpoznat a využít. Využití všech příznivých situací a odvrácení všech nepříznivých situací bylo podmínkou přechodu do režimu globální konvergence ($mood = 3$). Vnitřní pozorovatel tak musel být schopen rozpoznat a odvrátit konvergenci k lokálnímu minimu v režimu lokálního prohledávání ($mood = 0$), musel zajistit přechod na rovnoměrné prohledávání ($mood = 1$) při zlepšení rovnoměrnosti rozmístění dosud prohledaných pozic, a naopak, musel zanechat rovnoměrné prohledávání a zajistit přechod na lokální prohledávání ($mood = 2$) k odvrácení předčasné konvergence i v režimu rovnoměrného prohledávání. V režimu lokálního prohledávání ($mood = 0$) se pokoušel o větší rozpořívání částic v případě, že nebylo indikováno optimum. Zajišťoval střídání režimů lokálního prohledávání, pokud se rovnoměrnost rozmístění dosud prohledaných pozic nezlepšovala.

Následuje výčet úloh vnitřního pozorovatele vyjádřených v pseudokódu. Modul je naprogramován jako proces řízený událostmi.⁶⁹ Oproti návrhu algoritmu na obrázku Obr. 14(b) nastalo upřesnění logických podmínek přechodu mezi jednotlivými režimy, které je průběžně komentováno.

$$\underline{\text{Stanovení logických podmínek pro povolení režimu globální konvergence (mood = 3)}} \tag{163}$$

$$\begin{aligned}
 &\text{if } ((GlobMinMin \vee GlobMaxMax) \wedge (pMax[n] > 0.5)) \text{ then } mood[n] = 3 \\
 &\text{if } (GlobMin \wedge GlobMax) \text{ then } mood[n] = 3 \\
 &\text{if } ((MinZero \vee LocMin) \wedge GlobMaxMax \wedge (histogram[n].ProbabilityOfUniform < 0.0001)) \\
 &\quad \text{then } mood[n] = 3
 \end{aligned}$$

⁶⁹ V aktuální iteraci jsou provedeny pouze některé z úloh v závislosti na splnění podmínek (na událostech).

Podmínky pro povolení režimu globálního prohledávání nahrazují logickou proměnnou *outlier* na obrázku Obr. 14(b) a lze je slovně popsat následovně:

Má-li být globální prohledávání povoleno pouze na základě detekce výlučnosti, musí být alespoň jedna z pravděpodobností výlučnosti *fitness* hodně vysoká a nejvyšší hodnota pravděpodobnosti rovnoměrnosti rozmístění dosud prohledaných pozic musí být vyšší než 0.5. Globální prohledávání lze povolit i při nižších hodnotách pravděpodobnosti výlučnosti, avšak v tomto případě musí nastat jejich souběh.

Poslední z podmínek indikuje případ, kdy nejhorší hodnota *fitness* je výlučná s nejvyšší pravděpodobností, a to proto, že dochází ke konvergenci k pozici s nejlepší *fitness* po předchozím odvrácení všech detekovaných nepříznivých situací.

Odvracení konvergence k lokálnímu minimu záměnou režimů lokálního prohledávání (mood 0 a 2)

```

if ((LocMin v LocMax)  $\wedge$  (probability[n] < 1)) then
    switch mood[n]
        case 0:
            mood[n] = 2;
            break;
        case 2:
            mood[n] = 0;
    end
end

```

(164)

Je-li indikována globální nejlepší nebo globální nejhorší hodnota *fitness* jako výlučná s pravděpodobností větší než 0.5 a menší než 0.8, pozorovatel zajistí vzájemnou výměnu lokálních režimů prohledávání, (*probability*[n] < 1) značí, že se nejedná o první iteraci.

V dané situaci nesmí být povolen režim globálního prohledávání, neboť se může jednat o indikaci lokálního optima. Režim rovnoměrného prohledávání by v dané situaci rovněž mohl přispět k předčasné konvergenci.

Přechod na rovnoměrné prohledávání (mood = 1) při zlepšení rovnoměrnosti rozmístění pozic

```

if ((mood[n] = 0)  $\wedge$  (probability[n] < histogram[n].ProbabilityOfUniform)  $\wedge$ 
    (not LocMin)  $\wedge$  (not LocMax))
    then mood[n] = 1

```

(165)

Do režimu rovnoměrného prohledávání lze přejít z režimu lokálního prohledávání (*mood* = 0) pouze tehdy, nejsou-li indikovány globální nejlepší a globální nejhorší hodnoty *fitness* jako výlučné s pravděpodobností větší než 0.5 a menší než 0.8.

Zanechání rovnoměrného prohledávání (mood = 1)

```

if ((mood[n] = 1)  $\wedge$  (MinZero  $\wedge$  MaxZero)) then mood[n] = 2
if ((mood[n] = 1)  $\wedge$  (LocMin v LocMax v GlobMin v GlobMax)) then mood[n] = 2

```

(166)

Příkazy slouží k obnovení režimu lokálního prohledávání po pokusu o rozpohybování částic nebo po pokusu k odvrácení předčasné konvergence v předešlé iteraci.

Pokus o rozpohybování částic v režimu lokálního prohledávání (mood = 0), když algoritmus neindikuje optimum

```

if ((mood[n] = 0)  $\wedge$  (MinZero  $\wedge$  MaxZero)  $\wedge$  (probability[n] < 1)) then mood[n] = 1

```

(167)

Jsou-li v režimu lokálního prohledávání opakovaně indikovány globální nejlepší a globální nejhorší hodnoty *fitness* jako výlučné s pravděpodobností menší nebo rovné 0.5, pozorovatel může zvolit progresivnější způsob prohledávání, (*probability*[*n*] < 1) značí, že se nejedná o první iteraci. V následující iteraci může být obnoven režim lokálního prohledávání.

Pokus o rozpohybování částic v režimu lokálního prohledávání (*mood* = 0), když algoritmus indikuje optimum předčasně

```
if ((mood[n] = 0) ∧ (LocMin ∨ LocMax ∨ GlobMin ∨ GlobMax) ∧ (probability[n] < 1))
  then mood[n] = 1
```

 (168)

Jsou-li v režimu lokálního prohledávání indikovány globální nejlepší nebo globální nejhorší hodnoty *fitness* jako výlučné s pravděpodobností větší než 0.5, může se jednat o předčasnou konvergenci, (*probability*[*n*] < 1) značí, že se nejedná o první iteraci. Pozorovatel zvolí režim rovnoměrného prohledávání, ale pokud situace v následující iteraci trvá, bude obnoven režim lokálního prohledávání.

Vzájemná výměna režimů lokálního prohledávání (*mood* = 0, 2) při zhoršení rozmístění pozic

```
if ((probability[n] > histogram[n].ProbabilityOfUniform) ∧ (probability[n] < 1) ∧
  (not LocMin ∧ not LocMax)) then
  switch mood[n]
    case 0:
      mood[n] = 2;
      break;
    case 2:
      mood[n] = 0;
  end
```

 (169)

Zmenší-li se pravděpodobnost rovnoměrného rozmístění dosud prozkoumaných pozic v režimu lokálního prohledávání (a nejedná-li se o první iteraci), může to značit, že nastává konvergence. Pozorovatel proto vzájemně vymění režimy lokálního prohledávání, aby částice přinutil obrátit svou pozornost jinam a tím konvergenci zpomalil nebo zcela odvrátil. Pokud přesto konvergence trvá, jedná se patrně o konvergenci ke globálnímu optimu a částice mají možnost potvrdit tuto tendenci v následujících iteracích.

Nastavení režimu prohledávání podle hodnoty *mood*[*n*]

```
switch mood[n]
  case 0: (lokální prohledávání)
     $\omega$ [n] = 1 − histogram[n].getProbabilityOfUniform;
     $\alpha$ [n] = histogram[n].getProbabilityOfUniform
    break;
  case 1: (rovnoměrné prohledávání)
     $\omega$ [n] = histogram[n].getProbabilityOfUniform;
     $\alpha$ [n] = 1
    break;
  case 2: (lokální prohledávání)
     $\omega$ [n] = histogram[n].getProbabilityOfUniform;
     $\alpha$ [n] = 1 − histogram[n].getProbabilityOfUniform;
    break;
```

 (170)

```

case 3: (globální prohledávání)
     $\omega[n] = 0.1$ ; // histogram[n].getProbabilityOfUniform;
     $\alpha[n] = 2 - \omega[n]$ ; // histogram[n].getProbabilityOfUniform
    break;
end

```

Sekce nastavení režimu prohledávání podle hodnoty *mood[n]* (170) zůstala vzhledem k návrhu fáze prohledávání na obrázku Obr. 14(b) bez podstatných změn, pouze v režimu globálního prohledávání byla hodnota koeficientu $\omega[n]$ nastavena pevně na hodnotu 0.1. Hodnota pravděpodobnosti rovnoměrného rozmístění dosud prohledaných pozic při konvergenci ke globálnímu optimu totiž nezadržitelně klesala a algoritmus měl tudíž tendenci hodnotu $\omega[n]$ v režimu globálního prohledávání neustále zmenšovat. Následkem tohoto poklesu klesala i hodnota koeficientu $a_l[n]$ až k nule, což nebylo žádoucí.

6.2.2 Simulace chování částic dle katastrofického scénáře

6.2.2.1 Proč byl experiment prováděn

Z výsledků experimentů s originální verzí algoritmu *PSO* [4] s konstantními koeficienty, jejichž vyhodnocení je obsahem kapitoly 6.1.5, vyplynulo, že k odvrácení katastrofického scénáře [5] mají být hodnoty koeficientů přizpůsobeny velikosti vyhledávacího prostoru, působení účelové funkce a (ne)rovnoměrnosti rozmístění částic ve vyhledávacím prostoru po inicializaci.

Účelem experimentu s navrhovanou verzí algoritmu *PSO* bylo zjištění, jak průběžné přizpůsobování hodnot koeficientů během simulace ovlivňuje vývoj prohledávání a stabilitu algoritmu, měřenou množstvím pokusů o únik z vyhledávacího prostoru.

6.2.2.2 Popis experimentu

Experimentu byl podroben *pseudo*-paralelní algoritmus *PSO* na obrázcích Obr. 14(a), Obr. 14(b), s dvaceti jednorozměrnými částicemi, upravený dle výsledků počátečního testování uvedených v kapitole 6.2.1.4. *Stabilita* algoritmu byla měřena počtem (četností) pokusů částic o únik mimo původní vyhledávací prostor, vymezený při inicializaci.

Experimentováno bylo s funkcemi *Schwefel*, *Rastrigin* a *Sphere*, jejichž vzorce, grafy a jednodimenzionální profily jsou v příloze *Příloha A*. K zajištění srovnatelných výchozích podmínek s předchozími experimenty byly shodně použity tytéž, již vytvořené soubory s *pseudo*-náhodnými čísly. S každou funkcí bylo provedeno sto simulačních běhů, každý v počtu 500 iterací. Částicím byl pohyb omezen do intervalu $[-512, +512]$. Výsledky experimentování byly porovnány s odpovídajícími výsledky v kapitole 6.1.1.4.

6.2.2.3 Cíl experimentu

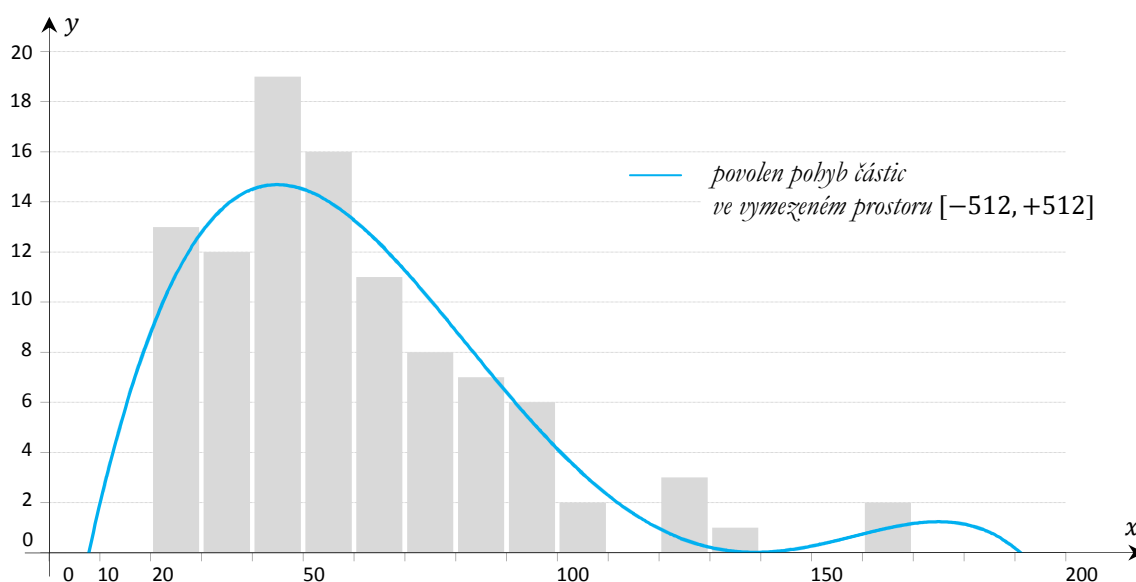
Cílem experimentu bylo ukázat, že výsledky měření podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$:

$\mathcal{H}_0$: Průběžné přizpůsobování hodnot koeficientů (ω, a_l, a_g) nemá vliv na stabilitu algoritmu.

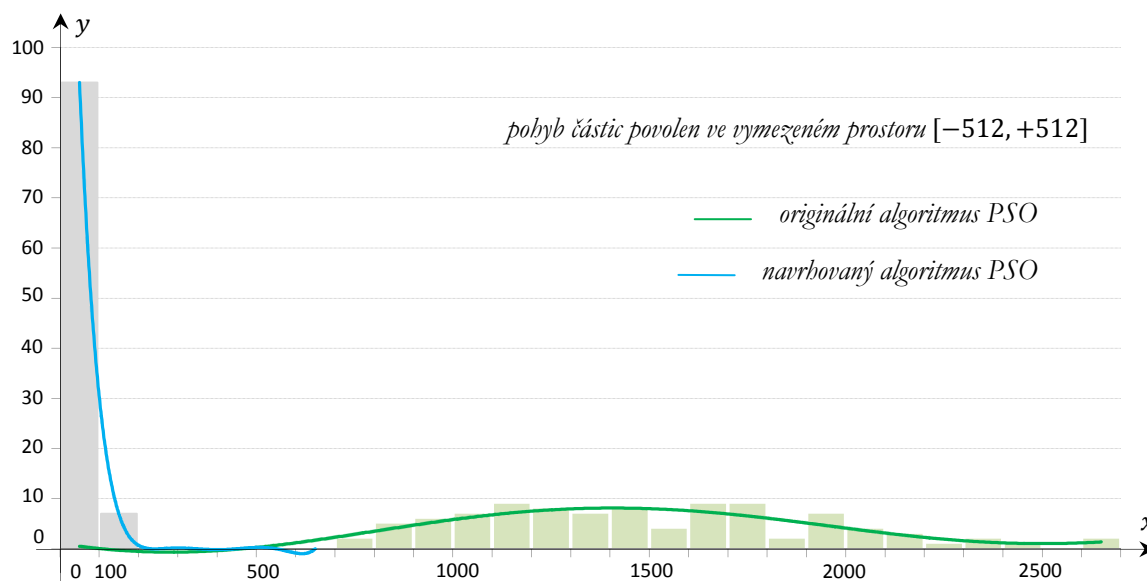
$\mathcal{H}_1$: Průběžné přizpůsobování hodnot koeficientů (ω, a_l, a_g) má vliv na stabilitu algoritmu.

6.2.2.4 Výsledky experimentu

Výsledky experimentu pro funkci *Schweffel* jsou graficky znázorněny na obrázku Obr. 22 a Obr. 23. Tendenční křivka odrážející četnosti pokusů o únik při omezení pohybu na inicializovaný vyhledávací prostor $[-512, +512]$ byla odvozena z histogramu absolutních četností, jenž vyjadřuje počty simulací na ose y vzhledem k pokusům o únik částic mimo vymezený prostor na ose x .



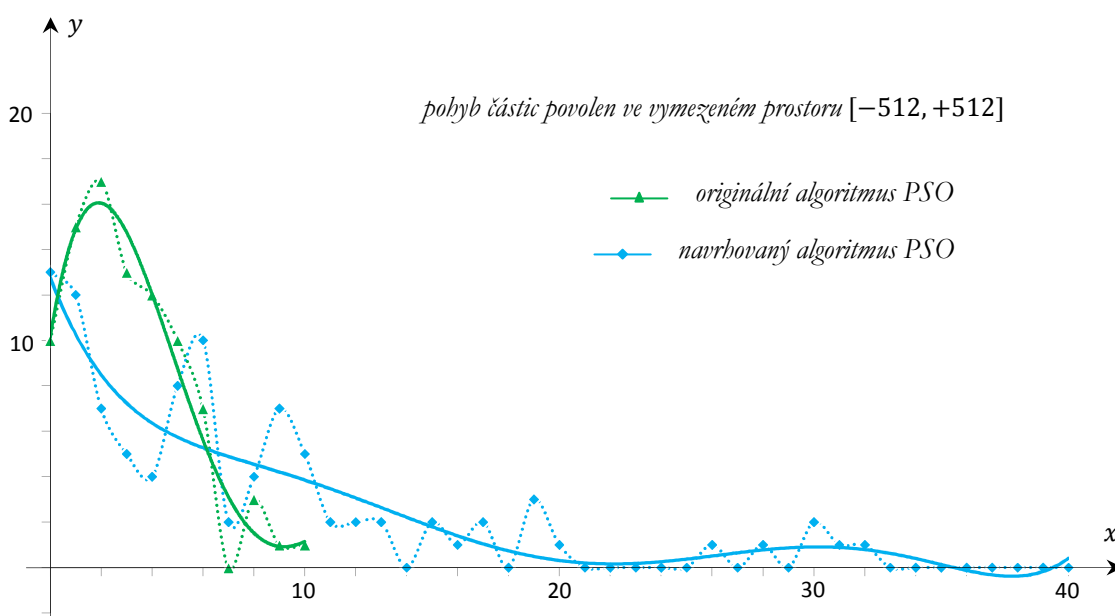
Obr. 22: Histogram absolutních četností vyjadřující počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) a odpovídající tendenční křivka pro navrhovaný algoritmus při působení funkce *Schweffel*.



Obr. 23: Porovnání histogramů absolutních četností vyjadřujících počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) navrhovaného algoritmu a originálního algoritmu při působení funkce *Schweffel*.

Obrázek Obr. 23 znázorňuje porovnání histogramů absolutních četností originálního algoritmu a navrhovaného algoritmu při působení funkce *Schwefel*. Z porovnání je zjevné, že počet pokusů o únik mimo vymezený vyhledávací prostor při působení funkce *Schwefel* klesl z původních 700 až 2600 pokusů u originálního algoritmu na hodnoty pod 100 pokusů (92% případů) u navrhovaného algoritmu.

Výsledky experimentu pro funkce *Rastrigin* a *Sphere* jsou graficky znázorněny na obrázcích Obr. 24 a Obr. 25. Naměřené absolutní četnosti počtu pokusů o únik byly vyneseny jako jednotlivé body a polynomicou regresí s využitím implicitních vzorců obsažených v programu *MS Excel* byly získány tendenční křivky k naměřeným četnostem. Tendenční křivky byly porovnány s tendenčními křivkami četností u originálního algoritmu, získanými v předcházejících experimentech.

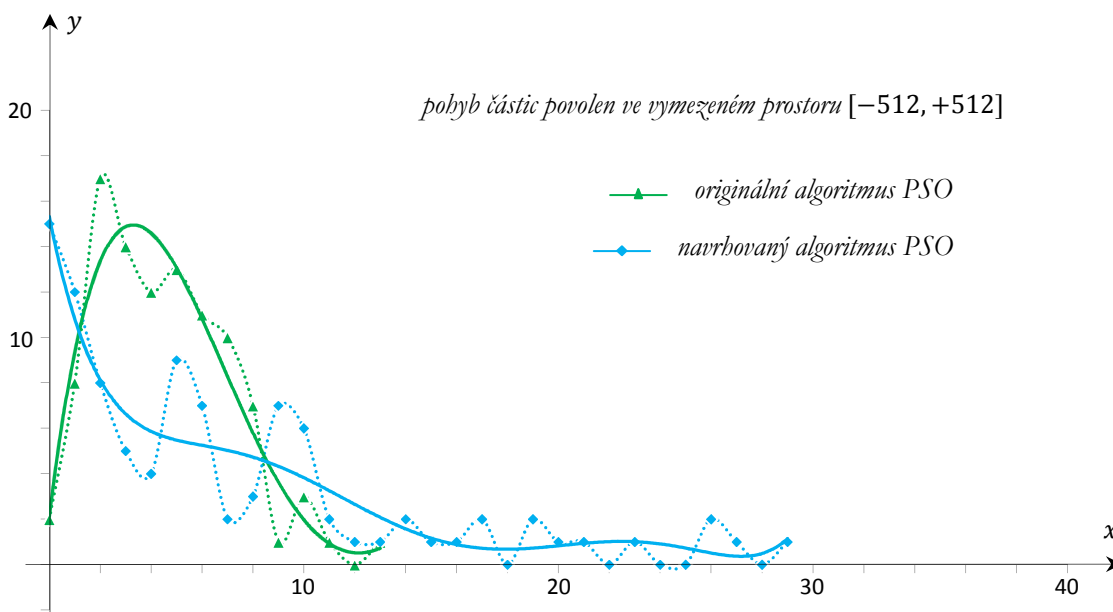


Obr. 24: Porovnání absolutních četností vyjadřujících počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) navrhovaného algoritmu a originálního algoritmu při působení funkce *Rastrigin*.

Tendenční křivka absolutních četností počtu simulací vzhledem k pokusům o únik z vymezeného prostoru u navrhovaného algoritmu při působení funkce *Rastrigin* (Obr. 24) se průběhem podobá tendenční křivce při neomezeném pohybu částic (Obr. 16), u navrhovaného algoritmu je však zaznamenán vyšší počet simulací, u nichž byl zaznamenán nulový nebo minimální počet pokusů o únik. Tato tendence je ještě výraznější při působení funkce *Sphere*, jak ukazuje graf na obrázku Obr. 25.

Tendenční křivky četností vyjadřujících počty simulací vzhledem k pokusům o únik z vymezeného vyhledávacího prostoru na obrázcích Obr. 23, Obr. 24 a Obr. 25 tedy podporují zamítnutí hypotézy $\mathcal{H}_0$ a podporují přijetí hypotézy $\mathcal{H}_1$, neboť průběžné přizpůsobování hodnot koeficientů u navrhovaného algoritmu má vliv na stabilitu měřenou počtem pokusů o únik. Při působení funkce *Schwefel* byl jednoznačně zaznamenán výrazný nárůst stability. Při působení funkcí *Rastrigin* a *Sphere* není (očekávaný)

nárůst stability zcela jednoznačný, a proto byly následně zkoumány příčiny ojedinělých případů, u nichž vzrostla nestabilita, a to v souvislosti se zkoumáním *úspěšnosti řešení*.



Obr. 25: Porovnání absolutních četností vyjadřujících počty simulací (osa y) vzhledem k pokusům o únik jednorozměrných částic mimo vymezený prostor $[-512, +512]$ (osa x) navrhovaného algoritmu a originálního algoritmu při působení funkce *Sphere*.

6.2.3 Ověření konvergence ke globálnímu optimu

6.2.3.1 Proč byl experiment prováděn

Experimenty s originální verzí algoritmu při působení funkce *Schwefel* ukázaly, že stabilita algoritmu, měřená počtem úniků nebo pokusů o únik částic mimo vyhledávací prostor vymezený při inicializaci, je určitou charakteristikou chování, není však ukazatelem prohledávacích schopností algoritmu, jejichž výsledkem má být konvergence algoritmu ke správnému, globálnímu optimu. Nápadný nárůst stability (kapitola 6.1.2.1) u jednotlivých případech simulací, provázený relativně nízkým počtem pokusů o únik u originálního algoritmu [4] při působení funkce *Schwefel*, byl naopak ukazatelem neúspěšných simulací, kdy algoritmus konvergoval předčasně a nedosáhl globálního optima.

Účelem experimentu na ověření konvergence ke globálnímu optimu s navrhovanou verzí algoritmu *PSO* bylo ověření, zda průběžné přizpůsobování hodnot koeficientů (ω, a_l, a_g) během simulace, které ovlivňuje stabilitu algoritmu, současně usměrňuje vývoj prohledávání tak, že vzrůstá pravděpodobnost *úspěšnosti řešení*.

6.2.3.2 Popis experimentu

Experimentu na ověření konvergence ke globálnímu optimu byl podroben *pseudo*-paralelní algoritmus *PSO* na obrázcích Obr. 14(a), Obr. 14(b), s dvaceti jednorozměrnými částicemi, upravený dle výsledků počátečního testování, uvedených v kapitole 6.2.1.4.

Úspěšnost řešení, jako metrika na nejvyšší úrovni dle obrázku Obr. 3, byla měřena počtem simulací (z celkového počtu 100 simulací), v nichž algoritmus konvergoval ke správnému, globálnímu optimu.

Experimentováno bylo s funkcemi *Schwefel*, *Rastrigin* a *Sphere*, jejichž vzorce, grafy a jednodimenzionální profily jsou v příloze *Příloha A*. K zajištění srovnatelných výchozích podmínek s předchozími experimenty byly shodně použity tytéž, již vytvořené soubory s *pseudo*-náhodnými čísly. S každou funkcí bylo provedeno sto simulačních běhů, každý v počtu 500 iterací. Částicím byl pohyb omezen do intervalu $[-512, +512]$. Výsledky experimentování byly porovnány s odpovídajícími výsledky v kapitole 6.1.2.4.

6.2.3.3 Cíl experimentu

Cílem experimentu bylo ukázat, že výsledky měření podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$:

$\mathcal{H}_0$: Průběžné přizpůsobování hodnot koeficientů *neovlivňuje úspěšnost řešení*.

$\mathcal{H}_1$: Průběžné přizpůsobování hodnot koeficientů *povede k úspěšnosti 99% a vyšší*.

6.2.3.4 Výsledky experimentu

Strategie průběžného přizpůsobování hodnot koeficientů algoritmu *PSO* specifikovaná v kapitole 6.2.1 dle návrhu řešení v kapitole 5, vedla u všech 100 simulací a u všech tří testovaných funkcí *Schwefel*, *Rastrigin* a *Sphere* k nalezení globálního optima.

Při působení funkce *Rastrigin* algoritmus neindikoval hodnoty pravděpodobnosti p (156) pro nastavení proměnných *LocMin*, *GlobMin*, *GlobMinMin* (162) a přepínání mezi jednotlivými režimy (*mood*) (170) se odvíjelo z hodnot proměnných indikujících *výlučnost globální nejhorší fitness*. Tuto konfiguraci však vykazovala i řada iterací při působení funkce *Schwefel*.

Podobně jako při působení funkce *Rastrigin* se algoritmus choval při působení funkce *Sphere*. Ve shodě s navrhovanou strategií prohledávání byla i při působení funkce *Sphere* vyhodnocována podezření na lokální minima, neboť vnitřní pozorovatel nemohl tušit, že se v případě funkce *Sphere* jednalo o *uni*-modální funkci.

Předběžné výsledky testování konvergence ke globálnímu optimu u navrhovaného algoritmu *PSO* tedy podporují přijetí hypotézy $\mathcal{H}_1$, s následujícími doporučeními:

Mezní hodnoty pravděpodobností $p \in \{0.5, 0.8, 0.95, 0.995\}$ figurující v proměnných (162) byly pozorovány pro všechny tři účelové funkce tytéž, avšak problém činila přesnost určení hodnot p (156), jež jsou získávány aproximací z tabelovaných hodnot převzatých z [44] a implementovaných přímo v navrhovaném algoritmu. Některé jednotlivé simulace by pro správnou konvergenci vyžadovaly vyhodnocení $p = 0.5$ pro indikaci lokálního minima, avšak většina simulací vycházela z hodnoty $p > 0.5$. Podobně tomu bylo i u hodnoty $p = 0.8$. Hodnoty pravděpodobností (156) by proto bylo třeba počítat přesně dle vzorců [44] přímo v algoritmu a experimentální pozorování by bylo třeba opakovat.⁷⁰

Lze si též povšimnout určité asymetrie u proměnných *GlobMinMin* a *GlobMaxMax* (162), v jejichž vyhodnocení se mezní hodnoty $p = 0.95$ a $p = 0.995$ liší. Možným

⁷⁰ Autorka soudí, že mezní hodnoty pravděpodobností p mohou být určitými pravděpodobnostními (momentovými) charakteristikami. Problém bude vyžadovat novou analýzu s novým, pravděpodobnostním modelem roje.

důvodem je dle autorky skutečnost, že algoritmus zvýhodňuje pozici s globální nejlepší *fitness*, k níž částice směřují a kde dochází k větší kumulaci částic, před pozicí s globální nejhorší *fitness*, a proto se globální nejlepší *fitness* jeví jako méně výlučná. Pro stanovení výlučnosti globální nejlepší a globální nejhorší *fitness* by proto bylo vhodné zavést testovací kritérium r_{32} namísto testovacího kritéria r_{22} (157), které bylo při testování použito. Testovací kritérium r_{22} opomíjí hodnoty $y_{(2)}$ a $y_{(19)}$. Testovací kritérium r_{32} by opomíjelo hodnoty $y_{(2)}$, $y_{(3)}$ a $y_{(19)}$ a bylo by spočteno následovně:

$$r_{32min} = \frac{y_{(4)} - y_{(1)}}{y_{(18)} - y_{(1)}} \quad r_{32max} = \frac{y_{(20)} - y_{(18)}}{y_{(20)} - y_{(4)}}$$

Hodnoty pravděpodobností (156) pro r_{32} sice nejsou v [44] tabelovány, avšak z důvodů již zmíněných je třeba hodnoty pravděpodobností přesně spočítat dle vzorců odvozených v [44]. Experimenty též ukázaly, že vysoká přesnost musí být zachována i při stanovení hodnoty pravděpodobnosti rovnoměrného rozmístění (159) dle testovacího kritéria (158).⁷¹

6.2.4 Vliv účelové funkce a inicializace na hodnoty koeficientů

6.2.4.1 Proč byl experiment prováděn

Experiment má ukázat, že počáteční iterace jsou pro úspěšný vývoj prohledávání rozhodující a že hodnoty koeficientů v počátečních iteracích mohou být při působení různých funkcí jak shodné, tak i rozdílné, a to za srovnatelných podmínek daných velikostí vyhledávacího prostoru, při použití téže posloupnosti náhodných čísel a strategie prohledávání.

6.2.4.2 Popis experimentu

Experimentu byl podroben *pseudo*-paralelní algoritmus *PSO* na obrázcích Obr. 14(a), Obr. 14(b), s dvaceti jednorozměrnými částicemi, upravený dle výsledků počátečního testování, uvedených v kapitole 6.2.1.4.

Experimentováno bylo s funkcemi *Schwefel*, *Rastrigin* a *Sphere*, jejichž vzorce, grafy a jednodimenzionální profily jsou v příloze *Příloha A*. K zajištění srovnatelných výchozích podmínek byly použity již vytvořené soubory s *pseudo*-náhodnými čísly. Vyhledávací prostor byl inicializován v intervalu $[-512, +512]$, částicím byl pohyb do tohoto intervalu omezen a byl pozorován vývoj koeficientů pro různé soubory pro jednotlivé funkce.

6.2.4.3 Cíl experimentu

Cílem experimentu bylo ukázat, že výsledky porovnání hodnot koeficientů podporují zamítnutí hypotézy $\mathcal{H}_0$ a vedou k přijetí alternativní hypotézy $\mathcal{H}_1$:

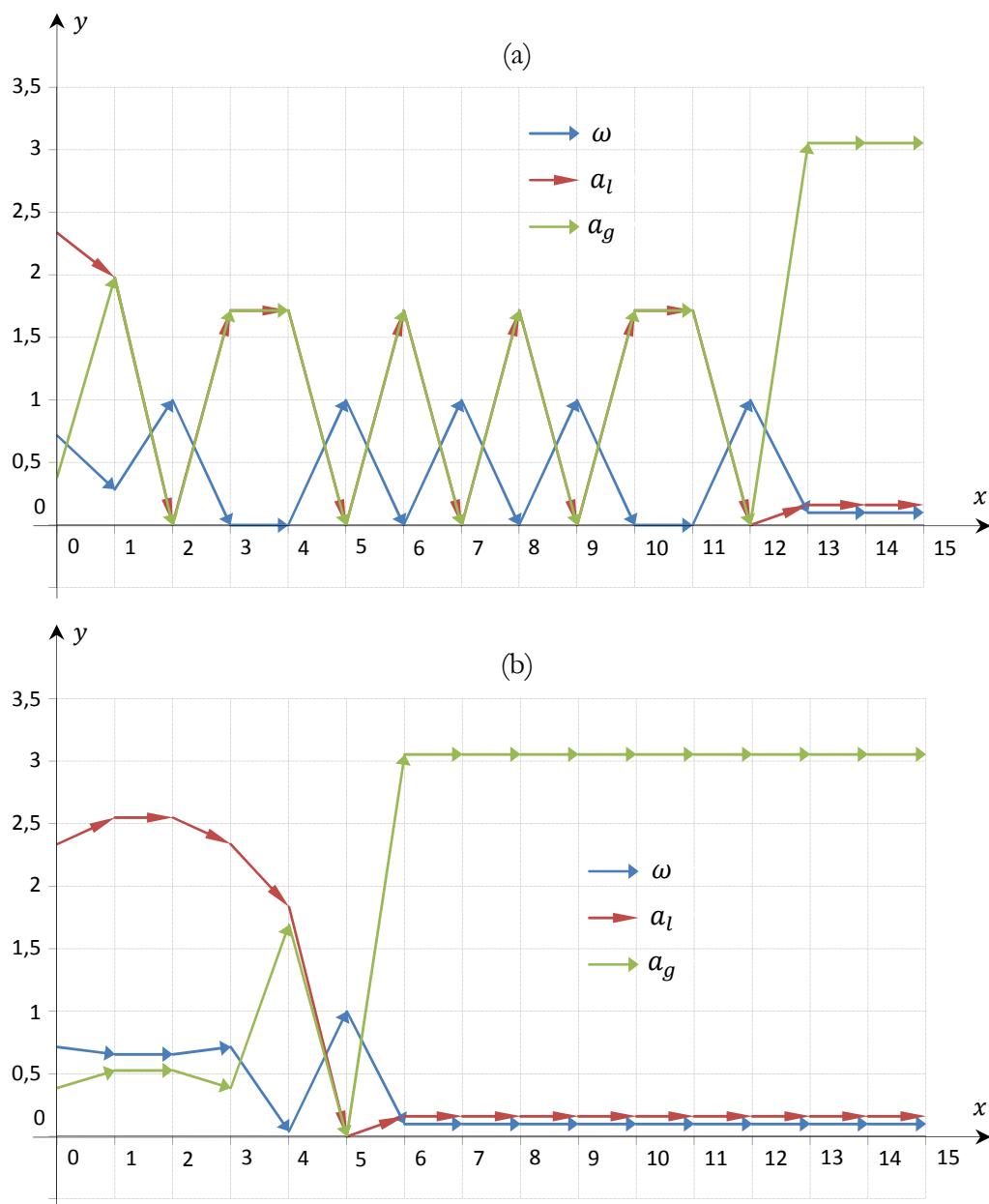
$\mathcal{H}_0$: *Hodnoty jednotlivých koeficientů jsou i při působení funkcí s rozdílnými tendenčními křivkami vždy shodné.*

$\mathcal{H}_1$: *Hodnoty jednotlivých koeficientů se budou při působení funkcí s rozdílnými tendenčními křivkami odlišovat.*

⁷¹ Jakékoli pokusy o zaokrouhlení při stanovení koeficientů ω nebo α v sekci (170) vedly ke kolapsu prohledávání. Uvedené chování přičítá autorka chaotickým rysům chování systému částice.

6.2.4.4 Výsledky experimentu

Pro demonstraci výsledků možného vývoje hodnot koeficientů byly vybrány simulace se souborem s *pseudo*-náhodnými čísly *Random_010*, což je jeden ze souborů, u něhož originální algoritmus [4] při působení funkce *Schwefel* skončil v lokálním optimu. Vývoj koeficientů ω , a_l , a_g v prvních 15 iteracích při působení funkce *Schwefel* znázorňuje graf na obrázku Obr. 26(a). Vývoj koeficientů ω , a_l , a_g při působení funkce *Rastrigin* a *Sphere* znázorňuje graf na obrázku Obr. 26(b).

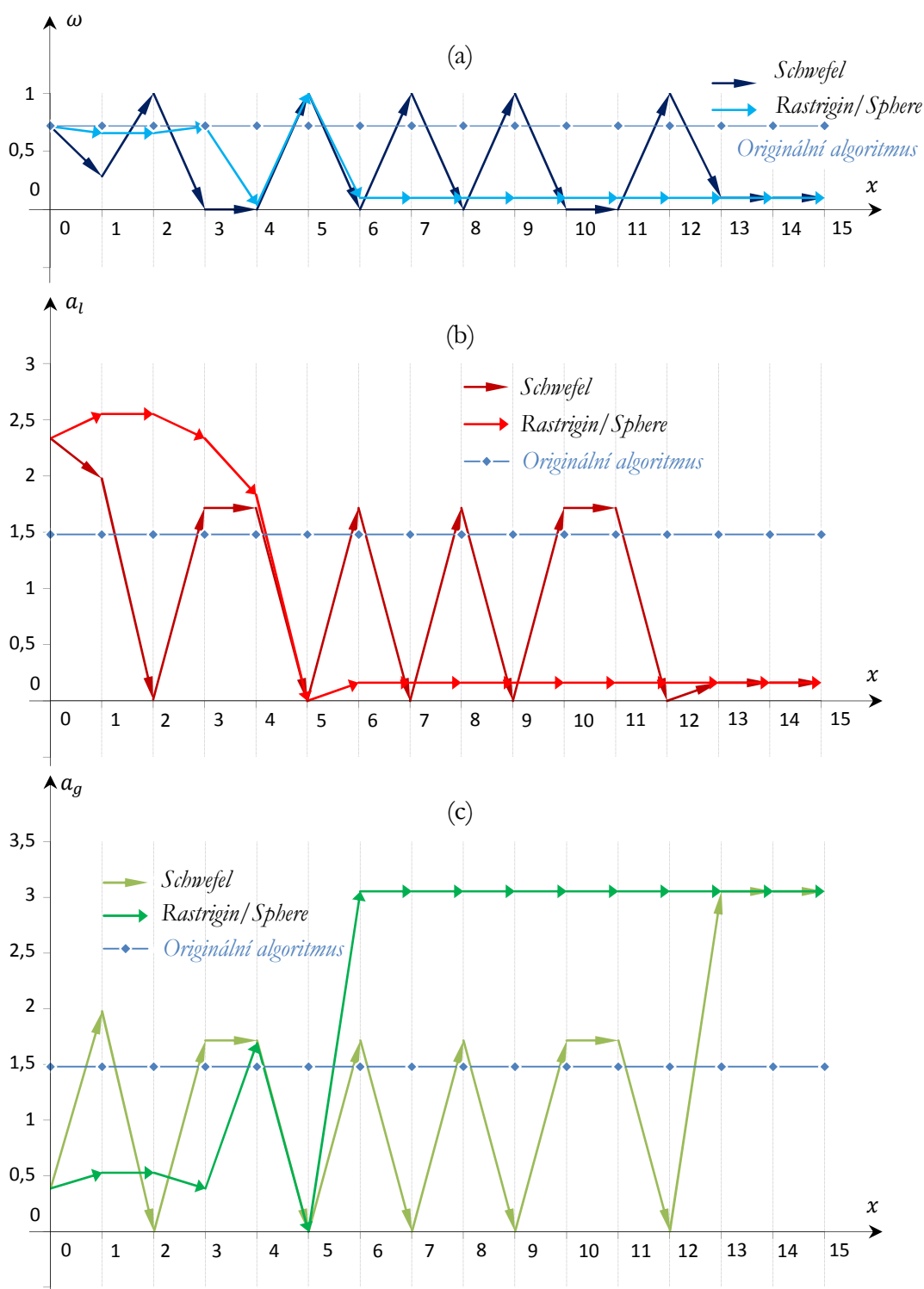


Obr. 26: Hodnoty koeficientů (osa y) v závislosti na pořadí iterace (osa x).

Porovnání vývoje hodnot koeficientů ω , a_l , a_g pro vybraný soubor s *pseudo*-náhodnými čísly při omezení pohybu částic do vyhledávacího prostoru $[-512, +512]$

(a) při působení funkce *Schwefel* (b) při působení funkcí *Rastrigin* a *Sphere*.

Vzájemné porovnání průběhů hodnot jednotlivých koeficientů ω , a_l , a_g při působení funkcí *Schwefel*, *Rastrigin* a *Sphere* je na obrázcích Obr. 27 (a), (b), (c):



Obr. 27: Hodnoty koeficientů (osa y) v závislosti na pořadí iterace (osa x).

(a) koeficient setrvačnosti ω , (b) koeficient lokálního zrychlení a_l ,
(c) koeficient globálního zrychlení a_g . Vzájemné porovnání průběhů hodnot koeficientů pro vybraný soubor s *pseudo*-náhodnými čísly při omezení pohybu částic do vyhledávacího prostoru $[-512, +512]$ při působení funkcí *Schwefel*, *Rastrigin* a *Sphere*.

Je důležité poznamenat, že hodnoty koeficientů na obrázcích Obr. 26 (a), (b) a Obr. 27 (a), (b), (c) splňují podmínky stability druhého řádu a svými hodnotami se pouze přibližují k nule (týká se koeficientů ω , a_l , a_g), případně k hodnotě jedna (koeficient ω).

Z průběhu hodnot koeficientů v závislosti na pořadí iterace na obrázku Obr. 26(a) je patrné, že pozorovatel v navrhovaném algoritmu volil v případě působení funkce *Schwefel* strategii střídavého přecházení z režimů lokálního prohledávání do režimu rovnoměrného prohledávání po dobu prvních 13 iterací (tj. 0 až 12), poté přešel do režimu globálního prohledávání. Prohledávání bylo započato v režimu 0 (*mood* 0), dále pokračovalo režimy 1, 0, 2, 1, 0, 2, 0, 1, 0, 1, 2, 0 a 3 a vedlo k úspěšnému řešení. Bylo zaznamenáno 76 pokusů o únik z vymezeného vyhledávacího prostoru, což je více než v případě originálního algoritmu s 38 zaznamenanými pokusy o únik, kde se však jednalo o neúspěšnou simulaci (u úspěšných simulací se počet pohyboval v rozmezí 700 až 2600, jak znázorňují obrázky Obr. 17 a Obr. 23).

Při působení funkcí *Rastrigin* a *Sphere* byl dán pokyn k přechodu do režimu globálního prohledávání již po prvních šesti iteracích, tj. iteracích 0 až 5 na obrázku Obr. 26(b). Prohledávání bylo započato v režimu 0 (*mood* 0), dále pokračovalo režimy 2, 2, 0, 2, 0 a 3. Bylo zaznamenáno 12 pokusů o únik z vyhledávacího prostoru, což je více než v případě originálního algoritmu, kde byly zaznamenány 2 pokusy při omezeném pohybu a 9 pokusů při neomezeném pohybu.

Nárůst nestability v uvedených konkrétních případech vysvětluje autorka tím, že simulace s použitím souboru *Random_010* vede po inicializaci k značně nerovnoměrnému rozmístění pozic (jak je patrné z obrázku Obr. 20) a pozorovatel se snaží o rovnoměrnější prohledání vymezeného prostoru i za cenu zvýšení nestability. Zvýšený počet pokusů o únik byl zaznamenán též u simulací s použitím souborů, jako je *Random_015* (Obr. 20), kdy byl zaznamenán pokus o předčasnou konvergenci, a vnitřní pozorovatel se snažil prostor lépe prohledat.

Shodný průběh hodnot koeficientů v zaznamenaných počátečních iteracích u funkcí *Rastrigin* a *Sphere* vysvětluje autorka tím, že spolu s rovnoměrností rozmístění pozic má na hodnoty koeficientů vliv též *vylučnost fitness*, která není s hodnotou *fitness* totožná a která může být tudíž pozorovatelem posouzena shodně u dvou rozdílných funkcí, jejichž průběh má podobnou tendenci (*Příloha A*).

Porovnání výsledků experimentování na obrázcích Obr. 27 (a), (b), (c) podporuje zamítnutí hypotézy $\mathcal{H}_0$ a vede k přijetí alternativní hypotézy $\mathcal{H}_1$, že hodnoty jednotlivých koeficientů se budou při působení funkcí s rozdílnými tendenčními křivkami odlišovat.

6.2.5 Vyhodnocení výsledků experimentování s navrhovanou verzí algoritmu *PSO*

Výsledky experimentování s *pseudo*-paralelní verzí algoritmu *PSO* s průběžným přizpůsobováním hodnot koeficientů dle návrhu v kapitole 5, jež byla specifikována v kapitole 6.2.1, přinesly a/nebo potvrdily následující poznatky:

- (i) Strategie průběžného přizpůsobování hodnot koeficientů algoritmu *PSO* vedla u všech testovaných případů k nalezení globálního optima s tím, že je autorkou doporučováno zavedení nového testovacího kritéria (161), jež by nahradilo kritérium

- (157) a které by (při řešení minimalizačního problému) nebralo v úvahu druhou, třetí a předposlední hodnotu *fitness* ve vzestupném uspořádání hodnot lokálních nejlepších *fitness* (154), a hodnoty pravděpodobnosti výlučnosti *fitness* (156) by měly být počítány přesně dle vzorců uvedených v [44].
- (ii) Nadále platilo, že pohyb částic bylo třeba omezit do inicializovaného vyhledávacího prostoru, neboť algoritmus musel čelit perturbacím vlivem působení účelové funkce a nebyl-li pohyb omezen, algoritmus selhal (částice divergovaly) při působení účelové funkce s tendenční křivkou nepodobnou parabole.
 - (iii) Experimenty opakovaně potvrdily, že *stabilita algoritmu*, měřená počtem pokusů o únik mimo vyhledávací prostor, není ukazatelem prohledávacích schopností algoritmu. Navrhovaný algoritmus průběžným přizpůsobováním hodnot koeficientů reguloval stabilitu (ve smyslu snižování nebo zvyšování), čímž současně usměrňoval vývoj prohledávání tak, že vzrostla pravděpodobnost *úspěšnosti řešení* dle metriky na obrázku Obr. 3, posuzovaná konvergencí do správného, tj. globálního optima.
 - (iv) Experimenty ukázaly, že pro úspěšný vývoj prohledávání byly rozhodující počáteční iterace a že hodnoty koeficientů při působení různých funkcí byly v počátečních iteracích rozdílné. Pokud byla zaznamenána shoda koeficientů i pro různé účelové funkce, pak tato shoda byla důsledkem způsobu určování pravděpodobnosti *výlučnosti hodnoty fitness* v případech účelových funkcí s podobným průběhem.
 - (v) Rozdílné hodnoty koeficientů byly pozorovány i při působení téže účelové funkce, a to pro různé testované náhodné posloupnosti. Pozorovatel v navrhované verzi algoritmu zohledňuje relativně malý počet částic a mění strategii prohledávání též vzhledem k velikosti vyhledávacího prostoru.
 - (vi) Hodnoty pravděpodobnosti rovnoměrného rozmístění (159), z nichž se odvozují koeficienty $\omega[n]$ a $\alpha[n]$ (170), musí být vyčísleny s dostatečnou přesností, neboť pokusy o zaokrouhlování při výpočtech koeficientů $\omega[n]$ či $\alpha[n]$ vedly ke kolapsu prohledávání, což autorka přičítá chaotickým rysům v chování systému částice.
 - (vii) Z důvodu přesnosti, jež je rozhodující pro úspěšnost konvergence, je autorkou doporučováno, aby byly hodnoty pravděpodobnosti *výlučnosti fitness* (156) získávané aproximací z tabelovaných hodnot nahrazeny přímým výpočtem dle vzorců odvozených v [44].
 - (viii) Experimentálně zjištěné mezní hodnoty pravděpodobností výlučnosti *fitness* (156) $p \in \{0.5, 0.8, 0.95, 0.995\}$ mohou být hodnotami určitých pravděpodobnostních (momentových) charakteristik, jež mohou mít pro jinou skupinu účelových funkcí jiné hodnoty. Analytické zdůvodnění tohoto pozorování však přesahuje původní záměr předkládané disertační práce, uvedený v kapitole 1.3, jímž byl návrh vhodné strategie prohledávání, jejíž součástí bylo *experimentální zjištění*, jak se vliv řídicí účelové funkce zastupující řešený problém projevuje na chování částic.

Kapitola 7

Posouzení rizik, návazné problémy a doporučení dalšího směru výzkumu

Zadání disertační práce v kapitole 2.1, spočívající v (i) návrhu programového systému pro demonstraci použitelnosti skupinové inteligence, (ii) implementaci navrženého systému a (iii) provedení experimentů, bylo v daných bodech naplněno. Pro demonstraci použitelnosti skupinové inteligence byl zkoumán algoritmus *PSO* a jeho varianty a byla navržena strategie průběžného přizpůsobování koeficientů (ω, a_l, a_g) , poskytujících informaci potřebnou pro aktualizaci rychlosti částice v rovnici (1), charakteristikám prostoru řešeného problému. Navržený systém byl implementován v jazyce *Java* v prostředí *NetBeans IDE 6.9.1* a byly provedeny experimenty s originální i navrženou verzí algoritmu.

Výsledky experimentů budou nyní porovnány a zhodnoceny vzhledem k dosažení cíle disertační práce. Budou posouzena rizika navrhovaného řešení, specifikovány návazné problémy a bude doporučen směr dalšího výzkumu.

7.1 Posouzení rizik navrhovaného řešení a návazné problémy

7.1.1 Zhodnocení výsledků experimentů vzhledem k dosažení globálního cíle disertační práce

Experimenty s originální verzí algoritmu s konstantními koeficienty [4], jejichž vyhodnocení je v kapitole 6.5.1, opakovaně potvrdily, že stabilita algoritmu, měřená počtem úniků nebo pokusů o únik částic mimo vyhledávací prostor vymezený při inicializaci, není ukazatelem prohledávacích schopností algoritmu. Byly zaznamenány simulace, v nichž algoritmus konvergoval předčasně a nedosáhl globálního optima a v nichž došlo k relativně nízkým počtům pokusů o únik (kolem 50 pokusů, kapitola 6.1.2.1) ve srovnání s úspěšnými simulacemi (řádově 10^3 pokusů), avšak při působení jiné účelové funkce za zcela srovnatelných podmínek byl počet pokusů výrazně nižší (pouhé dva pokusy).

Další experimenty s originální verzí algoritmu [4] prokázaly, že průběh trajektorie je působením účelové funkce ovlivňován, přičemž vliv účelové funkce není zanedbatelný (kapitola 6.1.4). Výsledky experimentů tak svědčí pro změnu interpretace modelu systému částice a pro nezbytnost použití omezovače rychlosti částice k odvrácení katastrofického scénáře, jak je popsáno v kapitole 5.1.1. Experimenty s originální verzí algoritmu [4] též odhalily, že prohledávací schopnost algoritmu s konstantními koeficienty je ovlivněna počátečním rozmístěním částic ve vyhledávacím prostoru, jak je uvedeno v kapitole 6.1.3.4 a znázorněno na obrázku Obr. 21.

Návrh způsobu odstranění nedostatků originální verze algoritmu PSO [4] s konstantními koeficienty, jenž je obsažen v kapitole 5, vychází z úvahy, že prohledávací schopnost algoritmu není možno zlepšovat pouhým zvyšováním nestability pomocí koeficientů (ω, a_l, a_g) až na mez stability $\approx (0.72, 1.48, 1.48)$.⁷² Experimenty realizujícími počáteční testování navrhované verze algoritmu PSO, jejichž výsledky jsou uvedeny v kapitole 6.2.1, bylo ověřeno, že budou-li hodnoty koeficientů splňovat podmínku stability 2. řádu (151) a budou-li průběžně přizpůsobovány rovnoměrnosti rozmístění prozkoumaných pozic se zřetelem na hodnotu pravděpodobnosti *vylučnosti účelové funkce* p (156), může být tímto způsobem regulována stabilita trajektorie částice ve smyslu jejího průběžného snižování nebo zvyšování, čímž bude usměrňován vývoj prohledávání tak, že vzroste pravděpodobnost *úspěšnosti řešení*.

Strategie průběžného přizpůsobování hodnot koeficientů algoritmu PSO vedla u všech testovaných případů k nalezení globálního optima. Bylo tak dosaženo cíle disertační práce specifikovaného v kapitole 3.7, jímž bylo *zlepšení výkonnosti algoritmu, jehož modelem je originální algoritmus PSO* [1], [2], a to *posílením autonomní činnosti částic roje*, kde rychlosti $\mathbf{v}$ a pozice $\mathbf{x}$ jsou nadále aktualizovány podle rovnic (1) a (2).

Nadále platilo, že pohyb částic bylo třeba omezit do inicializovaného vyhledávacího prostoru. Při působení účelových funkcí s odlišnými tendenčními křivkami (*Příloha A*) byly hodnoty koeficientů rozdílné. Rozdílné hodnoty koeficientů byly pozorovány i při působení těžce účelové funkce, a to pro různé testované náhodné posloupnosti, kdy pro úspěšný vývoj prohledávání byly rozhodující počáteční iterace.

Kompletní vyhodnocení výsledků experimentování s navrhovanou verzí algoritmu je uvedeno v kapitole 6.2.5.

7.1.2 Zhodnocení navrhovaného řešení vzhledem ke splnění dílčích cílů disertační práce

Globálního cíle disertační práce bylo dosaženo splněním dílčích cílů specifikovaných v kapitole 3.7. Originální algoritmus byl rozšířen o koncept vnitřního pozorovatele, jenž v každé iteraci poskytuje částicím sadu koeficientů k regulaci rychlosti podle rovnice (1), a to nezávisle pro každý rozměr částice. Hodnoty byly odvozovány z charakteristik prostředí bez nutnosti vnějšího řízení, při zachování stability trajektorie částic 2. řádu a stochastického charakteru operace roje. Pozorovatel volil režim prohledávání a tomu odpovídající sadu koeficientů podle pravděpodobností *vylučnosti* nejlepší a nejhorší detekované hodnoty *fitness*, stanovených aplikací *Dixonovy* teorie pomocí zvoleného testovacího kritéria. Uvedeným postupem dokázal pozorovatel rozpoznat a odvrátit konvergenci k lokálnímu optimu, a to s vysokou pravděpodobností. Pozorovatel rovněž vyhodnocoval a zohledňoval nejvyšší dosaženou hodnotu pravděpodobnosti *rovnoměrnosti rozmístění dosud prohledaných pozic* pro každý rozměr částice. Jednotlivé režimy prohledávání (170) volil pozorovatel dle splnění specifických podmínek (162) až (169).

⁷² Jak znázorňuje obrázek Obr. 17, nestabilita algoritmu při působení funkce *Schweffel* byla sice při úspěšných simulacích značná, ale u simulací s předčasnou konvergencí nestabilita poklesla (vzrostla stabilita).

Počáteční experimenty byly prováděny s jednorozměrnými částicemi. Předběžné experimenty s dvourozměrnými částicemi ukázaly, že částice mohou prohledávat každý rozměr n nezávisle v jiném režimu prohledávání (*mood* 0,1,2 3) (170), aniž se režimy navzájem rušivě ovlivňují, a předběžné výsledky byly rovněž příznivé.⁷³ Systém však bude třeba testovat i na vícerozměrných částicích.

Experimentování s navrhovanou verzí algoritmu rovněž upozornilo na rizika *úspěšnosti řešení* (tj. konvergence ke správnému, globálnímu optimu) spojená jednak s volbou vhodného testovacího kritéria pro určení pravděpodobnosti výlučnosti hodnot *fitness* (157) (161) (171) a jednak s přesností výpočtů hodnot pravděpodobnosti výlučnosti *fitness* (156) a pravděpodobnosti rovnoměrného rozmístění prohledaných pozic (159).

Specifikace rizik spolu s návrhy opatření, jež mohou uvedená rizika eliminovat, budou nyní pojednány podrobněji.

7.1.2.1 Volba testovacího kritéria a stanovení pravděpodobnosti výlučnosti hodnoty *fitness*

Pro stanovení hodnoty pravděpodobností nejlepší a nejhorší dosud naměřené hodnoty *fitness* byla použita testovací kritéria r_{22min} a r_{22max} (161), jež (záměrně) neberou v potaz v pořadí druhou a předposlední pořadovou statistiku (*order statistics*) (154).

Při experimentování byla naměřena hodnota $p = 0.995$ (156) jako maximální hodnota pravděpodobnosti *výlučnosti nejhorší fitness*, zatímco maximální hodnota pravděpodobnosti *výlučnosti nejlepší fitness* byla naměřena pouze $p = 0.95$ (kapitola 6.1.1).

Jak již bylo zmíněno ve vyhodnocení experimentu v kapitole 6.2.3.4, je otázkou, zda je tato asymetrie žádoucí a zda by ji nebylo možno odstranit volbou kritéria r_{32} (171), navrženého autorkou práce, které by ignorovalo kromě druhé a předposlední pořadové statistiky navíc též třetí pořadovou statistiku. Experimenty rovněž ukázaly, že hodnotu pravděpodobnosti p (156) bude třeba počítat přesně dle vzorců odvozených v [44].⁷⁴

Autorka předkládané práce věří, že zpřesnění výpočtu pravděpodobnosti *výlučnosti fitness* p (156) vyřeší kolizní situace způsobené mezními hodnotami pravděpodobností, jež jsou zmíněny v kapitole 6.2.3.4. Přesné vyčíslení hodnoty pravděpodobnosti p (156) je rozhodující k vyhodnocení podmínek pro přepínání mezi jednotlivými režimy prohledávání (170), ale okamžikem přepnutí je navíc ihned ovlivněn výpočet hodnot koeficientů ω a α (170), z nichž se následně dopočítávají hodnoty zbylých koeficientů a_l a a_g (151) (127a).

Zpřesnění výpočtu a opakování experimentálního testování může přinést upřesnění hodnot v množině mezních hodnot pravděpodobností výlučnosti p (156), jež byla empiricky stanovena jako $p = \{0.5, 0.8, 0.95, 0.995\}$, a také odhalení, zda tyto mezní hodnoty jsou závislé na působící účelové funkci. Autorka soudí, že tyto mezní hodnoty mohou být hodnotami určitých (momentových) pravděpodobnostních charakteristik naměřených hodnot *fitness*. Řešení tohoto návazného problému bude vyžadovat

⁷³ Experimenty s dvourozměrnými částicemi byly provedeny orientačně.

⁷⁴ Je třeba provést náhradu stanovení hodnoty pravděpodobnosti p (156), která je získávána aproximací z tabulek, a to výpočtem příslušných integrálů přímo v implementaci navrženého systému nebo zvolit jiné, dokonalejší vývojové prostředí disponující statistickými výpočty.

odpovídající analýzu pravděpodobnostního modelu roje částic PSO navrženého autorkou v kapitole 5.2.1.

7.1.2.2 Odvození hodnot koeficientů z pravděpodobnosti rovnoměrného rozmístění prozkoumaných pozic

Způsob stanovení hodnot koeficientů ω a α z pravděpodobnosti rovnoměrného rozložení dosud prozkoumaných pozic (159) (170), navržený autorkou předkládané práce, je intuitivní, i když vychází z logické úvahy, že lokální prohledávání má být posíleno při menší pravděpodobnosti rovnoměrného rozmístění dosud prozkoumaných pozic ve vymezeném vyhledávacím prostoru a oslabeno při větší pravděpodobnosti rovnoměrného rozmístění. Koeficienty a_l a a_g jsou následně dopočteny ze vztahů mezi koeficienty $(\omega, a_l, a_g, \alpha)$ plynoucími ze stochastické analýzy stability trajektorie částice (151) (127a).

Deterministické a stochastické analýzy stability systému částice (kapitola 4) odhalují vztahy mezi koeficienty a poskytují mezní hodnoty pro zachování vnitřní stability systému částice, nedávají však odpověď, jaké konkrétní hodnoty koeficientů mají být voleny. Odvození koeficientů ω a α v navrhované verzi algoritmu z charakteristik prohledávaného prostředí tak, aby částice získala veškerou informaci potřebnou pro aktualizaci své rychlosti učním z vlastní zkušenosti a/nebo jí byla poskytnuta pozorovatelem uvnitř systému, je součástí strategie posílení *autonomní činnosti částic*, jež výše uvedený nedostatek eliminuje a jež má přispět k posílení *všeobecného povědomí (awareness)* v systému roje, jak je požadováno v kapitole 3.4.

Intuitivní řešení způsobu stanovení hodnot koeficientů ω a α (170) představuje určité riziko *(ne)úspěšnosti řešení*, které však má dle autorky zanedbatelný dopad ve srovnání s rizikem plynoucím z nepřesnosti výpočtu pravděpodobnosti výlučnosti hodnoty *fitness*, jež bylo specifikováno v předchozí kapitole 7.1.2.1.

Experimenty ukázaly, že vysoká přesnost výpočtu musí být rovněž zachována při stanovení hodnoty pravděpodobnosti rovnoměrného rozmístění (159) dle testovacího kritéria (158), neboť jakékoli pokusy o zaokrouhlování při stanovení koeficientů ω nebo α v sekci (170) vedly ke kolapsu prohledávání projevujícího se vysokým počtem pokusů o únik (neúměrným zvýšením nestability) a/nebo konverencí k lokálnímu optimu (při působení *multi-modální funkce*). Pozorované chování přičítá autorka chaotickým prvkům chování systému částice.⁷⁵

7.1.3 Návazné problémy k řešení

Rozpoznaná rizika *úspěšnosti řešení*, specifikovaná v předchozí kapitole 7.1.2, vedla k identifikaci problémů, jež jsou technického rázu, a také problémů, týkajících se dalšího směru výzkumu. Problémy technického rázu se vztahují (i) k dostupnosti statistických funkcí ve vývojovém prostředí algoritmu a (ii) k přesnosti výpočtů a budou nyní zmíněny podrobněji. Problémům vztahujícím se k doporučení dalšího směru výzkumu bude věnována další, samostatná kapitola.

⁷⁵ Byl by jistě možný i jiný způsob stanovení koeficientů ω a α , než je způsob použitý autorkou práce. Zvolený systém by však bylo nutno (dle autorky) též striktně dodržovat, včetně vysoké přesnosti výpočtů.

Vývojové prostředí *NetBeans IDE 6.9.1*, v němž byl implementován navržený algoritmus *PSO*, obsahuje knihovnu (*unit*) *Math* poskytující matematické funkce k výpočtu hodnoty *fitness* účelových funkcí používaných k testování optimalizátorů (*Příloha A* obsahuje výběr z těchto funkcí), ale nedisponuje knihovnou statistických funkcí, a tedy histogramy a kalkulátor funkcí γ a Γ pro výpočet hodnoty distribuční funkce proměnné s rozložením χ^2 (160) k určení pravděpodobnosti p (159) naprogramovala autorka práce pro pevně zvolený počet částic $M = 20$. Pravděpodobnost *vylučnosti* p (156) získala autorka aproximací z tabulky pro testovací kritérium r_{22} (157) a počet částic $M = 20$, kterou převzala z *Dixonovy* publikace [44] a kterou implementovala přímo v algoritmu *PSO*.

Pro další vývoj algoritmu *PSO* doporučuje autorka doplnění vývojového prostředí *NetBeans* o statistické funkce obsažené v knihovně [org.apache.commons.math3.stat](http://commons.apache.org/math3.stat), jež je k dispozici jako tzv. *Open Source* a jež je dostupná na adrese

<http://commons.apache.org/proper/commons-math/userguide/stat.html>

Spočtení pravděpodobnosti *vylučnosti* p (156) tak, jak ji aplikuje autorka v navrhované verzi algoritmu, není standardní statistickou funkcí. Autorka shledala, že tato funkce je jako tzv. *Dixonův Q-test* používána nikoli k pouhé detekci odlehlých hodnot (*outliers*) dle *Dixonovy teorie* [44], ale v užším významu, konkrétně k vylučování hodnot, které jsou detekovány jako odlehlé s pravděpodobností 95% nebo vyšší, a to z prováděných výpočtů (typicky aritmetického průměru).⁷⁶ Funkci tedy bude třeba pro další experimentování s navrhovanou verzí algoritmu v požadované podobě naprogramovat.⁷⁷

7.2 Doporučení dalšího směru výzkumu

K hlubšímu pochopení problému *zvyšování skupinové inteligence roje* by měl být další výzkum zaměřen na modelování roje jako *komplexního systému* (*complex system*) a na související analýzu pravděpodobnostního modelu navrženého autorkou v kapitole 5.2.1.

7.2.1 Roj jako komplexní systém

Chování systému částice *nelze* pojímat čistě deterministicky s tím, že by působení náhodných prvků bylo vnímáno pouze jako šum. Jak bylo dokumentováno v závěru kapitoly 4.3.2, náhodné prvky mají v algoritmu *PSO* svou nezastupitelnou úlohu, neboť jejich přítomnost zvyšuje kmitání, zpomaluje konvergenci, zabráňuje stagnaci částice a zaručí, že nastane pohyb částice i v případě, kdy se lokální nejlepší pozice částice ani globální nejlepší pozice roje průběžně nezlepšují, neboť tento pohyb by jinak nenastal.

Úplný popis systému neposkytuje ani interpretace trajektorie částice jako stochastické funkce, neboť dalším potenciálním zdrojem perturbací je působení budící síly F_b (147a) na vstupu systému částice (kapitola 5.1.1), přičemž velikost budící síly je ovlivňována hodnotami účelové funkce. Průběh účelové funkce z pozice vnitřního pozorovatele není znám a nemáme ani přímo matematicky vyjádřenu uvedenou závislost. Umíme pouze změřit nebo spočítat hodnotu účelové funkce (*fitness*) v daném místě hyperprostoru.

⁷⁶ *Dixonův Q-test* je k dispozici např. na adrese <http://www.ee.ucl.ac.uk/~mflanaga/java/Outliers.html> jako *Open Source*.

⁷⁷ Funkci stanovení pravděpodobnosti *vylučnosti* by bylo (dle autorky) vhodné naprogramovat, otestovat a verifikovat jako tzv. "inženýrské dílo".

Odchyly způsobené budící silou mohou být relativně velké, ale protože uvedená závislost není v rovnici (1) matematicky podchycena,⁷⁸ odchyly se do podmínek vnitřní stability nepromítají.

Teorie systémů označuje systém s výše popsanými vlastnostmi jako *otevřený*. Chování systému částice odpovídá charakteristice *komplexního systému* (*complex system*), jak ho charakterizuje Francis Paul Heylighen [43].⁷⁹

Komplexní systém dle [43] je charakterizován tím, že je dostatečně *flexibilní* (dostatečně otevřený pro změnu), ale ne natolik, aby se choval chaoticky. Nebrání šumu a pertubacím, které urychlují jeho vývoj, avšak působení silných pertubací, které by mohly způsobit jeho destrukci, sám odhalí a odvrátí v procesu *samo-organizace*. *Komplexní systém* tedy podporuje *nelineární, samo-organizační proces*, kdy drobné odchyly mohou rozhodnout, kterého z více atraktorů systém dosáhne. Po dosažení atraktoru mají tyto odchyly malý, nebo žádný vliv. *Samo-organizační proces* eliminuje rozdíly mezi dvěma počátečními stavy, které skončí ve stejném atraktoru. Stejně tak nepatrně odlišný počáteční stav může skončit v jiném atraktoru. *Samo-organizační* nejen nebrání, nýbrž prospívají šum a pertubace, náhodnost nebo neurčitost: čím více odchylek nebo výkyvů, tím rychleji systém dosáhne atraktoru a tím stabilnější atraktor bude. Tento poznatek je označován jako *order from noise principle*.⁸⁰

Systém částice roje vykazuje prvky *chaotického chování*, neboť (i) je citlivý na počáteční podmínky, (ii) jako dynamický systém je popsán třemi stavovými proměnnými, jimiž jsou pozice, rychlost a čas (jako implicitní proměnná u diskrétního systému), (iii) na vstupu do systému je skryta nelinearita a (iv) mezi vstupem do systému a pozicí částice je potenciální zpětná vazba. Dle F. P. Heylighena však dokonalý chaos, stejně jako dokonalý řád, není *komplexní*, neboť obojí lze úplně deterministicky popsat.

Systém částice však úplně deterministicky popsat nelze, neboť nelze popsat již zmíněnou závislost mezi účelovou funkcí a budící silou. Autorka předkládané práce proto změnila interpretaci modelu systému částice, a sice na buzený tlumený harmonický oscilátor (*forced oscillator*) (kapitola 5.1.1), aby interpretace chování částice odpovídala chování komplexního systému (*complex system*), a existencí budící síly zdůvodnila, proč musí být skutečné chování systému částice usměrňováno omezovačem rychlosti částice (148) v případě, že pozice částice (2), vypočtená z této rychlosti, by opustila vymezený prostor. Částice je tak vrácena zpět do vyhledávacího prostoru.

Částice totiž (dle autorky) nesmí opustit prostor ani takovým způsobem, jaký je znázorněn na obrázku Obr. 11, neboť nás nezajímají optima vně vyhledávacího prostoru. Takové vyhledávání je dle autorky nesystematické, neboť prostor vně vyhledávacího prostoru nebyl inicializován (nebyly v něm na počátku rovnoměrně rozmístěny částice). Pokud by byl prostor vymezen na začátku nepřesně (nesprávně), bylo by nutno vymezení prostoru opravit a provést novou inicializaci a nové prohledávání, případně nadefinovat nový problém, jímž by byl *způsob vymezení vyhledávacího prostoru pro daný problém*.

Návrh pravděpodobnostního modelu roje v kapitole 5.2.1 byl krokem k *modelování roje jako komplexního systému*. Úvahy autorky práce, související s modelováním roje, jsou uvedeny v následující kapitole 7.2.2.

⁷⁸ Řešíme tzv. nekorektní inverzní problém (146) (kapitola 4.7.3).

⁷⁹ Francis Paul Heylighen (nar. 1960) je belgický kybernetik zkoumající vznik a vývoj inteligentní organizace.

⁸⁰ Heinz von Foerster, 1960, *On Self-Organizing Systems and Their Environments* [49]

7.2.2 Model roje

Fyzikální model [33], popsáný v kapitole 4.1, modelující roj jako soustavu hmotných bodů, je dle autorky předkládané práce nutno chápat pouze jako jednu z možných interpretací rovnic (1) a (2), jež umožňuje posouzení roje z hlediska mechaniky hmotného bodu. Uvedený přístup však současně odhaluje určité zjednodušení při matematickém modelování roje jako sociálního simulátoru [1].

Roj je totiž v původním, originálním konceptu algoritmu *PSO* [1] modelován jako soustava nehmotných entit, a to proto, aby nehmotné částice mohly zaujmout v prostoru stejnou pozici, čímž má být dle autorů modelu roje *PSO* [1] vyjádřeno, že dvě nehmotné entity mohou zaujmout zcela shodný názor, zatímco při pokusu dvou hmotných entit zaujmout shodnou fyzickou pozici musí nutně dojít ke kolizi. Autoři modelu roje ve studii [1] deklarují, že vytvořili prostor, který je prost kolizí (*collision free space*).

Autorka předkládané práce však nezastává názor autorů [1] [2], že nehmotné částice, přestože nemají žádný objem, mohou zaujmout v hyperprostoru reálných čísel tutéž pozici. Pozice jsou v algoritmu *PSO* modelovány jako realizace náhodných proměnných a dvě pozice jako takové tudíž (teoreticky) nikdy nemohou nabýt stejných reálných hodnot, neboť realizace náhodných proměnných nabývají hodnot z určitého intervalu a pozice se tedy mohou k sobě pouze přiblížit na (nekonečně) malou vzdálenost.

Je tedy lhostejné, zda jsou částice *PSO* hmotné či nehmotné povahy, neboť svým názorem, či pozicí, se k sobě mohou pouze přiblížit. Přínosem modelování roje jako soustavy hmotných bodů naopak je, že model umožňuje posouzení roje jako nekonzervativního, disipativního (ztrátového) systému, který je efektivnější než konzervativní systém, což potvrzují i obsáhlé experimenty, z jejichž výsledků vychází obsah kapitoly 3.3.

Nekonzervativní systém vznikl z roje *PSO* zavedením koeficientu ω do originálního algoritmu [2]. Ztrátová složka síly, která byla identifikována autory [33] jako síla vnitřního tření podle *Stokesova* zákona, však předpokládá, že částice bude mít tvar koule o určitém objemu. Jak již bylo podotknuto v kapitole 4.1.6, částice žádný tvar nemá, nemá ani objem, neboť je pokládána za hmotný bod. Na otázku, *proč je nekonzervativní systém efektivnější*, nepodávají autoři fyzikálního modelu [33] vyčerpávající objasnění, ale pouze naznačují možný rozpor z hlediska druhé věty termodynamiky a doporučují, aby hodnota koeficientu $\omega < 1$ v rovnici (1) byla v počátcích prohledávání volena “prozíravě” tak, aby ztráty třením nezpůsobovaly úbytek energie potřebné k prohledávání příliš rychle.⁸¹

⁸¹ Autorka předkládané disertační práce k tomuto tématu poznamenává, že na uvedenou otázku, a sice *proč je z hlediska termodynamiky nekonzervativní systém efektivnější než systém konzervativní*, dává odpověď *Prigožinova teorie* (Илья Романович Пригожин) (I. R. Prigogin) disipativních systémů z roku 1977. Zmíněná teorie má však i v současnosti oponenty, kteří poukazují na skutečnost, že tato teorie je v rozporu s druhou větou termodynamiky, konkrétně že entropie systému nemůže při vývoji systému v čase klesnout, ale zůstává konstantní nebo narůstá, přičemž nejvyšší je tehdy, kdy systém dosáhne *ekvilibrria* (v kontextu roje *globálního optima*). *Prigožinova teorie* říká, že druhá věta termodynamiky platí pouze pro uzavřené systémy (izolované a neizolované). Otevřené systémy (z termodynamického hlediska) mají schopnost se “zbavovat” entropie jakožto nevyužitelné energie za podmínky, že mohou čerpat energii ze svého okolí, přičemž tato čerpaná energie je vyšší nebo nejvýše rovna uvolněné entropii.

Z hlediska modelování systému částice není dle autorky předkládané práce pochyb o tom, že částice musí být modelována jako dynamický systém, tj. systém, který mění své chování v čase, a že dynamické systémy vznikají z disipativních systémů (kapitola 4.2.2). Dynamický model systému částice použitý k vyšetřování vnitřní stability systému částice, jak je znázorněn na obrázku Obr. 6, je dle autorky akceptovatelný, autorka pouze navrhl, aby byl model systému částice interpretován tak, jak je znázorněno na obrázku Obr. 13, kde kromě elastické a odporující síly působí na částici i budící síla (kapitola 5.1.1),⁸² což umožní zkoumat systém částice a celý roj jako *komplexní systém* (kapitola 7.2.1).

K dalšímu zkoumání chování částic navrhl autorka model roje jako celku, a to modelu pravděpodobnostního. Pravděpodobnostní přístupy k modelování roje lze vysledovat též v literatuře a autorka v dané souvislosti odkazuje na publikaci [33].

Analytikové [33:pp.35-38] definují náhodnou veličinu X , jejímiž realizacemi jsou pozice částic v určité iteraci, a podávají postup, kterým odvozují pravděpodobnostní rozložení veličiny X v k -té iteraci z počátečního rozložení. Ukazují, že počáteční rovnoměrné rozložení přechází v normální rozložení a mění se pouze parametry tohoto rozložení. Autorka předkládané práce však upozorňuje, že tyto postupy neberou v úvahu působení budící síly, které je vyjádřeno vztahem (146), přesněji, neberou v úvahu působení účelové funkce na lokální nejlepší pozice $\mathbf{l}^i$, a tím i na pozice částic $\mathbf{x}^i$, jež jsou z pozic $\mathbf{l}^i$ odvozeny (1).

Autorka vytvořila návrh pravděpodobnostního modelu roje v kapitole 5.2.1. V modelu jsou kromě prozkoumaných pozic zahrnuty i náhodné veličiny, jež odpovídají lokálním nejlepším hodnotám $\mathbf{l}^i$ (1). Množina lokálních nejlepších *fitness* je autorkou interpretována jako náhodný vzorek, jehož rozsah je dán počtem částic. Základním souborem je nekonečný soubor všech možných, teoreticky dosažitelných hodnot účelové funkce. Pravděpodobnostní model roje navržený autorkou je tedy inovativní v tom, že zahrnuje i nekorektní inverzní problém (146) (kapitola 4.7.3).

Pravděpodobnostní model navržený autorkou předkládané práce je zamýšlen k dalšímu výzkumu zaměřenému na hlubší pochopení *problému dalšího zvýšení inteligence roje částic (swarm intelligence)*. Po vyřešení technických problémů specifikovaných v kapitole 7.1.3 by další výzkum měl být zaměřen na nalezení nejvhodnějšího testovacího kritéria (155) pro výpočet hodnoty pravděpodobnosti výlučnosti hodnoty *fitness*, na zpřesnění mezních hodnot pravděpodobností výlučnosti hodnoty *fitness* p (156) a na odhalení, zda tyto mezní hodnoty jsou závislé na působící účelové funkci, a na analytické zdůvodnění uvedeného pozorování.

⁸² Podmínky vnitřní stability systému částice zůstávají v platnosti, neboť rovnice (28) i (147) mají stejnou charakteristickou rovnici (54).

Závěr

Základní záměr předkládané disertační práce, specifikovaný v kapitole 1.3, jímž byl *aplikovaný výzkum skupinové (kolektivní) inteligence*, byl splněn. K prokázání použitelnosti inteligence skupiny byl zkoumán algoritmus na bázi roje částic (*Particle Swarm Optimization PSO*) [1], [2], jenž byl inspirován biologickými komplexními systémy a v němž je řešený problém převeden na matematickou optimalizaci, kdy roj částic (*particle swarm*) hledá globální optimum ve vymezeném prostoru problému a prohledávání je řízeno podle předem nadefinované účelové funkce (*objective function*), která zastupuje řešený problém. Byla navržena a experimentálně ověřena strategie prohledávání, v níž částice průběžně přizpůsobují své chování charakteristikám prostoru řešeného problému, a bylo experimentálně zjištěno, jak se vliv řídicí účelové funkce zastupující řešený problém projevuje v chování částic. Výsledky experimentování s navrženou strategií prohledávání byly porovnány s výsledky experimentů s referenční verzí algoritmu *PSO*.

Strategie průběžného přizpůsobování chování částic charakteristikám prostoru řešeného problému vedla u všech testovaných případů k úspěšnému řešení. Experimenty ukázaly, že klasické řešení, kde jedinou podmínkou je stabilní trajektorie, po níž se částice pohybuje v prostoru řešeného problému, a kde vliv řídicí účelové funkce je ve výsledku eliminován, může selhat, a že dynamická stabilita trajektorií částic sama o sobě není ukazatelem prohledávacích schopností algoritmu ani konvergence algoritmu ke správnému řešení. Navrhovaný algoritmus reguloval stabilitu algoritmu průběžným přizpůsobováním chování částic charakteristikám prostoru řešeného problému, čímž současně usměrňoval vývoj prohledávání tak, že vzrostla pravděpodobnost úspěšnosti řešení.

Součástí návrhu strategie průběžného přizpůsobování chování částic charakteristikám prostoru řešeného problému byla změna interpretace dynamického modelu systému částice, a sice z tlumeného harmonického oscilátoru na buzený tlumený harmonický oscilátor (*forced oscillator*) a pozornost byla věnována působení budící síly, jež je ovlivňována hodnotou účelové funkce. Existenci budící síly bylo zdůvodněno, proč musí být skutečné chování systému částice usměrňováno omezovačem rychlosti částice tak, aby chování částice odpovídalo chování komplexního systému (*complex system*).

Ke zkoumání chování roje částic jako celku a testování výlučnosti dosud nejlepší objevené hodnoty účelové funkce (*fitness*) byl navržen pravděpodobnostní model roje. Inovativním prvkem modelování roje je integrace množiny lokálních nejlepších hodnot účelové funkce (*fitness*) dosažených jednotlivými částicemi do pravděpodobnostního modelu a zavedení konceptu vnitřního pozorovatele, jenž v průběhu dynamické historie roje tuto množinu udržuje a jeho úkolem je zjistit, s jakou pravděpodobností se nejlepší či nejhorší hodnota *fitness* z množiny lokálních nejlepších *fitness* významně odlišuje od ostatních hodnot, tj. je výlučná. Na problém pravděpodobnosti výlučnosti byla úspěšně aplikována *Dixonova teorie vylučování odlehlých výsledků (outliers)* [44] [45].

Dalším úkolem pozorovatele bylo odvození hodnot vah ovlivňujících rychlost částice, jež odpovídají rovnoměrnosti rozmístění pozic v prostoru řešeného problému dosud prohledaného částicemi, a dále volba režimu prohledávání vyhodnocením podmínek

zahrnujících hodnoty pravděpodobnosti výlučnosti nejlepší a nejhorší hodnoty z množiny lokálních *fitness*.

Dixonova teorie umožňuje detekovat odlehlý výsledek již při počtu tří prvků, a tedy aplikace navržené strategie průběžného přizpůsobování hodnot vah ovlivňujících rychlost částice by umožnila snižování počtu částic algoritmu *PSO*.⁸³ Počet částic by teoreticky bylo možno snižovat na méně než (minimálních) 20 částic, což by mělo význam nejen pro snižování výpočetní složitosti při řešení optimalizačních problémů s více stupni volnosti (s vícerozměrnými částicemi), ale i při vývoji aplikací, jako jsou senzorické robotické roje (*olfactory robot swarms*). Z teoretického hlediska může algoritmus *PSO* sloužit jako laboratorní model pro výzkum chování *otevřeného, komplexního systému*.

Pro převod navrhované verze algoritmu *PSO* do praxe, jeho podporu a udržitelnost je třeba disponovat vývojovým prostředím, které kromě matematických funkcí bude poskytovat i statistické funkce, včetně aplikace *Dixonovy* teorie. Pro vnitřního pozorovatele nesmí být výpočet hodnot charakteristik prostředí, z nichž odvozuje hodnoty vah ovlivňujících rychlost částice, složitější, než např. výpočet hodnoty účelové funkce.

Pravděpodobnostní model roje navržený autorkou předkládané disertační práce je zamýšlen k dalšímu výzkumu zaměřenému na hlubší pochopení *problému inteligence roje částic* a možnosti jejího dalšího zvyšování. Autorka obdivuje invenci analytiků, jejichž pracemi se v souvislosti se zhodnocením současného stavu poznání ve sledované oblasti podrobně zabývala, a je přesvědčena, že jednoduchý, geniální algoritmus *PSO*, který i po 20 letech od svého uvedení přináší nové podněty, si další analytický vklad bezpochyby заслужuje.

⁸³ Současná tendence vede spíše k navyšování počtu částic. Autoři originálního algoritmu s konstantními koeficienty [4] doporučují 50 částic v roji.

Odkazy

- [1] Kennedy, J., Eberhart, R. C.: *Particle swarm optimization*. In: Proceedings of the IEEE International Conference on Neural Networks, vol. IV, pp. 1942-1948. Perth, Australia 1995, pp. 1942–1948.
- [2] Eberhart, R. C., Kennedy, J.: *A new optimizer using particle swarm theory*. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, pp. 39–43. Nagoya, Japan, 1995.
- [3] Gardner, A., Grafen, A.: *Capturing the superorganism: a formal theory of group adaptation*. Journal of Evolutionary Biology, Volume 22, Number 4, April 2009, pp. 659–671.
- [4] Bratton, D., Kennedy, J.: *Defining a Standard for Particle Swarm Optimization*. IEEE Swarm Intelligence Symposium, Proc., pp. 120–127, April 2007. Print ISBN 1-4244-0708-7, DOI 10.1109/SIS.2007.368035.
- [5] Eberhart, R. C., Shi, Y.: *Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization*. IEEE Congress on Evolutionary Computation, Proc., Vol.1, pp. 84–88, Jul 2000. ISBN 0-7803-6375-2, DOI 10.1109/CEC.2000.870279.
- [6] Brown, B. B.: *Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts*. An earlier paper published by RAND, Document No: P-3925, 1968, 15 pages.
- [7] Alberts, D. S.: *NATO Code of Best Practice for C2 Assessment*. Product of a NATO RTO sponsored Research Group SAS-026. Command and Control Research Program (CCRP), ISBN 1-893723-09-7, Oct 2002, Reprint July 2004.
- [8] Pěchouček, M., Mařík, V.: *Social Models and their Applications*. Technical Report, Gerstner Laboratory, Czech Technical University in Prague, Jan 2005.
- [9] Cao, W., Bian, C. G., Hartvigsen, G.: *Achieving efficient cooperation in multi-agent system: The twin-base modeling*. In Kandzia and M. Klusch, editors, *Cooperative Information Agents*, number 1202 in LNAI, pages 210–221. Springer-Verlag, Heidelberg, 1997.
- [10] Pěchouček, M., Mařík, V., Štěpánková, O.: *Role of Acquaintance Models in Agent-Based Production Planning Systems*. In M. Klusch L. Kerschberg, editors, *Cooperative Information Agents IV*, LNAI No. 1860, Springer-Verlag, Heidelberg 2000, 179–190.
- [11] Ratnaweera, A., Halgamuge, S. K., Watson, H. C.: *Self-Organizing Hierarchical Particle Swarm Optimizer With Time Varying Acceleration Coefficients*. IEEE Transactions on Evolutionary Computation, pp. 240–255, Vol. 8, No. 3, June 2004
- [12] Kennedy, J.: *The particle swarm: Social adaptation of knowledge*. In: Proceedings of International Conference on Evolutionary computation, pp.303–308. Indianapolis, IN, USA, 1997.
- [13] Clearwater, S. H., Hogg, T., Hubermann, B. A.: *Cooperative Problem Solving*. In Computation: The Micro and Macro View, pp. 33-70. Singapore: World Scientific, 1992.
- [14] Moffat, J., Alberts, D. S.: *Maturity Levels for NATO NEC*. TR21958 V 2.0, Defence Science & Technology Laboratory, U.K., Dec 2006.

-
- [15] Alberts, D. S., Huber, R. K., Moffat, J.: *NATO NEC C2 maturity model*. DoD Command and Control Research Program. ISBN 978-1-893723-21-4, February 2010.
 - [16] Alberts, D. S., Hayes, R. E.: *Power to the Edge Command... Control... in the Information Age*. CCRP Publication Series, Washington, D.C., ISBN 1-893723-13-5. 1st printing June 2003. Reprint June 2004.
 - [17] Llinas, J., Bowman, Ch., Rogova, G., Walz, E., White, F.: *Revisiting the JDL Data Fusion Model*. In: Proceedings of the 7th International Conference on Information Fusion, Stockholm, Sweden, Jun 28 - Jul 1, 2004.
 - [18] Corkill Daniel D.: *Collaborating Software: Blackboard and Multi-Agent Systems & the Future*, In: Proceedings of the International Lisp Conference, New York, October 2003.
 - [19] Kennedy, J.: *Small Worlds and Mega-Minds: Effect of Neighborhood Topology on Particle Swarm Performance*. In: Proceedings of the Congress on Evolutionary Computation, pp. 1931–1938. Washington DC, USA, July 1999.
 - [20] Kennedy, J., Mendes, R.: *Population structure and particle swarm performance*. In: Proceedings of the IEEE Congress on Evolutionary Computation, pp. 1671–1676. Hawaii, USA, 2002.
 - [21] Peer, E. S., F. van den Bergh, Engelbrecht, A. P.: *Using Neighborhoods with the Guaranteed Convergence PSO*. Department of Computer Science, University of Pretoria, South Africa, 2002
 - [22] Shi, Y., Eberhart, R. C.: *A Modified Particle Swarm Optimiser*. In: Proceedings of the IEEE International Conference of Evolutionary Computation. Anchorage, Alaska, May 1998, pp. 69–73.
 - [23] Shi, Y., Eberhart, R. C.: *Empirical study of particle swarm optimization*. In Proc. of IEEE International Congress on Evolutionary Computation, vol. 3, 1999, pp. 101–106.
 - [24] Eberhart, R. C., Shi, Y.: *Tracking and optimizing dynamic systems with particle swarms*. In Proc. of IEEE Congress on Evolutionary Computation 2001, Seoul, Korea, 2001, pp. 94–97.
 - [25] Angeline, P. J.: *Evolutionary optimization versus particle swarm optimization: Philosophy and the performance difference*. In Lecture Notes in Computer Science, vol. 1447, Proceedings of 7th International Conference on Evolutionary Programming, Evolutionary Programming VII, Mar 1998, pp. 600–610.
 - [26] F. van den Bergh: *An Analysis of Particle Swarm Optimizers*. PhD thesis. Department of Computer Science, University of Pretoria. South Africa, 2002.
 - [27] Ismail, A., Engelbrecht, A. P., *Measuring Diversity in the Cooperative Particle Swarm Optimizer*. Proceedings of 8th International Conference on Swarm Intelligence, ANTS 2012, Brussels, Belgium, September 12-14, 2012, Springer-Verlag Berlin Heidelberg, LNCS 7461, ISSN 0302-9743, DOI 10.1007/978-3-642-32650-9_9, Print ISBN 978-3-642-32649-3, Online ISBN 978-3-642-32650-9.

- [28] Liang, J. J., Qin, A. K., Suganthan, P. N., Baskar S.: *Comprehensive Learning Particle Swarm Optimizer for Global Optimization of Multimodal Functions*. IEEE Transactions on Evolutionary Computation, pp. 281–295, Vol. 10, No. 3, June 2006.
- [29] Ismail, A., Engelbrecht, A. P.: *The Self-adaptive Comprehensive Learning Particle Swarm Optimizer*. Proceedings of 8th International Conference on Swarm Intelligence, ANTS 2012, Brussels, Belgium, September 12–14, 2012, Springer-Verlag Berlin Heidelberg, LNCS 7461, ISSN 0302-9743, DOI 10.1007/978-3-642-32650-9_9, Print ISBN 978-3-642-32649-3, Online ISBN 978-3-642-32650-9.
- [30] Mendes, J., Kennedy, J., Neves, J.: *The Fully Informed Particle Swarm: Simpler, Maybe Better*. IEEE Transactions on Evolutionary Computation, pp. 204–210, Vol. 8, No. 3, Jun 2004.
- [31] Torney, C., Neufeld, Z., Couzin, I. D.: *Context – dependent interaction leads to emergent search behavior in social aggregates*. PNAS, pp. 22055–22060, Vol. 106, No. 52, December 29, 2009.
- [32] Couzin, I. D.: *Collective cognition in animal groups*. Trends in Cognitive Sciences, Vol. 13, pp. 36–43, January 2009.
- [33] Mikki, S. M., Kishk, A. A.: *Particle Swarm Optimization: A Physics-Based Approach*. Synthesis Lectures on Computational Electromagnetics, Lecture 20. Morgan & Claypool, 2008.
- [34] Wang, Q., A., Wang, R.: *Is it possible to formulate least action principle for dissipative systems?* SAO/NASA Astrophysics Data System Database, Bibcode 2012arXiv1201.6309W.
- [35] Trelea, I. C.: *The particle swarm optimization algorithm: convergence analysis and parameter selection*. Information Processing Letters, pp. 317–325, Vol. 85, No. 6, 2003.
- [36] Clerc, M., Kennedy, J.: *The particle swarm – explosion, stability, and convergence in a multidimensional complex space*. IEEE Transactions on Evolutionary Computation, pp. 58–73, Vol. 6, No. 1, Feb 2002.
- [37] Clerc, M.: *Stagnation Analysis in Particle Swarm Optimisation or What Happens when Nothing Happens*. Technical Report CSM-460, Department of Computer Science, University of Essex. ISSN: 1744-8050, August 2006.
- [38] Poli, R.: *Mean and Variance of the Sampling Distribution of Particle Swarm Optimizers during Stagnation*. IEEE Transactions on Evolutionary Computation, Vol. 13, No. 4, August 2009.
- [39] Fernández Martínez, J. L., García Gonzalo, E.: *The generalized PSO: a new door for PSO evolution*. Journal of Artificial Evolution and Applications, DOI 10.1155/2008/861275, Article ID 861275, 15 pages, 2008.
- [40] Fernández Martínez, J. L., García Gonzalo, E.: *The PSO family: deduction, stochastic analysis and comparison*. Swarm Intelligence (2009) 3: pp. 245–273, 2009. DOI 10.1007/s11721-009-0034-8.
- [41] Fernández Martínez, J. L., García Gonzalo, E., Fernández Alvarez, J. P.: *Theoretical analysis of particle swarm trajectories through a mechanical analogy*. International Journal of Computational Intelligence Research, 4(2), 2008.

-
- [42] Akat, S. B., Gazi, V., Marques, L.: *Asynchronous particle swarm optimization-based search with a multi-robot system: simulation and implementation on a real robotic system*. Turk J Elec Eng & Comp Sci, Vol.18, No.5, 2010, DOI:10.3906/elk-0906-7.
 - [43] Heylighen F.: *Self-organization in Communicating Groups: the emergence of coordination, shared references and collective intelligence*. In: Language and Complexity, Barcelona University Press, 2011.
 - [44] Dixon, W. J.: *Ratios Involving Extreme Values*. The Annals of Mathematical Statistics, 22 (1951), No. 1, pp. 68–78.
 - [46] Dixon, W. J.: *Processing Data for Outliers*. Biometrics, March 1953, Vol. 9, No. 1, pp. 74–89.
 - [47] Dixon, W. J.: *Analysis of Extreme Values*. The Annals of Mathematical Statistics, 21 (1950), No. 4, pp. 488–506. DOI:10.1214/aoms/1177729747.
 - [48] McBane, G.: *Programs to Compute Distribution Functions and Critical Values for Extreme Value Ratios for Outlier Detection*. Journal of Statistical Software, May 2006, Volume 16, Issue 3.
 - [49] von Foerster, H.: *On Self-Organizing Systems and Their Environments*. Self-Organizing Systems, M.C. Yovits and S. Cameron (eds.), Pergamon Press, London, pp. 31–50, (1960).
 - [50] Eberhart, R. C., Simpson, P., Dobbins, R.: *Computational Intelligence PC Tools*, pp. 212–226. Academic Press Professional, 1996.
 - [51] Zheng, Y. L., Ma, L. H., Zhang, L. Y., Qian, J. X.: *On the convergence analysis and parameter selection in particle swarm optimisation*. In: Proceedings of the 2nd International Conference on Machine, Learning and Cybernetics (ICMLC'03), Vol. 3, pp. 1802–1807, China 2003.
 - [52] Van den Bergh, F., Engelbrecht, A. P.: *A study of particle swarm optimization particle trajectories*. Information Sciences, Vol. 176, No. 8, pp. 937–971, 2006.
 - [53] Yamille del Valle, Venayagamoorthy, G. K., Mohagheghi, S., Hernandez, J. C., Harley, R. G.: *Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems*. IEEE Transactions on Evolutionary Computation, pp. 171–195, Vol. 12, No. 2, April 2008.
 - [54] Corne, D., Dorigo, M., Glover, F., editors: *New Ideas in Optimization*, pp. 379–387. McGraw Hill, 1999.
 - [55] F. van den Bergh, Engelbrecht, A. P.: *A new locally convergent particle swarm optimizer*. In: Proceedings of the IEEE Conference on Systems, Man and Cybernetics. Hammamet, Tunisia, Oct. 2002.
 - [56] Eberhart, R. C., Simpson, P., Dobbins, R.: *Computational Intelligence PC Tools*, pp. 212–226. Academic Press Professional, 1996.

Dodatky

D1 Odvození *Eulerovy-Lagrangeovy* rovnice (9)

Mějme funkcionál ve tvaru:

$$A = \int_{t_0}^{t_1} L(t, f(t, r), f'(t, r')) dt \quad (6)$$

kde t_0 je konstanta značící počáteční čas vývoje systému, t_1 je konstanta značící koncový čas vývoje systému, funkce $f'(t, r') = df/dt$ a funkce $L(t, f(t, r), f'(t, r'))$ je dvakrát spojitě diferencovatelná s ohledem na argumenty t, f, f' . Pro uvedený typ funkcionálu lze problém hledání nulové variace převést na řešení *Eulerovy-Lagrangeovy* rovnice podle následujícího postupu:

Předpokládejme, že $A(f)$ dosahuje svého lokálního minima v bodě f_0 a $\eta(t, r)$ je libovolná testovací funkce, která má alespoň první dvě derivace podle času a vymizí v okrajových bodech t_0 a t_1 . Funkce $f(t, r)$ musí mít alespoň první derivaci podle času. Potom platí

$$A(f_0) \leq A(f_0 + \varepsilon \eta)$$

pro libovolné ε blíží se k nule, kde $\varepsilon \eta$ se nazývá *variace* funkce f a značí se δf .

Variace $A(f)$ je definována jako:

$$\delta(A(f, \eta)) = \lim_{\varepsilon \rightarrow 0} \frac{A(f + \varepsilon \eta) - A(f)}{\varepsilon} = \left. \frac{d}{d\varepsilon} A(f + \varepsilon \eta) \right|_{\varepsilon \rightarrow 0}$$

Vektor δA znamená, že uvažujeme infinitesimální posunutí (první variaci) nahrazující vektor $f(t, r)$ vektorem $(f(t, r) + \varepsilon \eta(t, r))$.⁸⁴ Substitucí $f = f_0 + \varepsilon \eta$ do (6) dostáváme:

$$A(f_0 + \varepsilon \eta) = \int_{t_0}^{t_1} L(t, (f_0 + \varepsilon \eta), (f'_0 + \varepsilon \eta')) dt \quad (7)$$

Pokud má funkcionál (6) minimum pro $f = f_0$, musí mít funkcionál (7), který je funkcí ε , minimum pro $\varepsilon = 0$, což zapíšeme:

$$0 = \left. \frac{d}{d\varepsilon} A(f_0 + \varepsilon \eta) \right|_{\varepsilon \rightarrow 0} = \int_{t_0}^{t_1} \frac{d}{d\varepsilon} (L(t, (f_0 + \varepsilon \eta), ((f'_0 + \varepsilon \eta')))) dt \quad (8)$$

Použijeme substituci $y = f_0 + \varepsilon \eta$, $y' = f'_0 + \varepsilon \eta'$ do (8) a provedeme totální derivaci funkce $L(t, y, y')$ podle ε :

⁸⁴ Posunutí $\varepsilon \eta(t, r)$ lze chápat jako nepatrnou deformaci trajektorie v omezeném časovém úseku, kdy v počátečním a konečném bodě trajektorie je tato deformace nulová.

$$\int_{t_0}^{t_1} \frac{d}{d\varepsilon} (L(t, y, y')) dt = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial y} \frac{dy}{d\varepsilon} + \frac{\partial L}{\partial y'} \frac{dy'}{d\varepsilon} \right) dt = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial y} \eta + \frac{\partial L}{\partial y'} \eta' \right) dt$$

Integrál součtu rozdělíme na součet integrálů, druhý z integrálů řešíme využitím metody *per-partes* podle vzorce $\int uv' = uv - \int u'v$, kam dosadíme $u = \partial L / (\partial y')$, $v = \eta$, a dostáváme:

$$\int_{t_0}^{t_1} \frac{\partial L}{\partial y} \eta dt + \int_{t_0}^{t_1} \frac{\partial L}{\partial y'} \eta' dt = \int_{t_0}^{t_1} \frac{\partial L}{\partial y} \eta dt + \left[\frac{\partial L}{\partial y'} \eta \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \frac{d}{dt} \frac{\partial L}{\partial y'} \eta dt$$

Jelikož řešíme úlohu s tzv. pevnými konci, neboli předpokládáme, že $\eta(t_0) = \eta(t_1) = 0$, prostřední člen vymizí, zbylé integrály sloučíme v jeden a dostáváme:

$$0 = \delta(A(f_0), (\eta)) = \frac{d}{d\varepsilon} A(f_0 + \varepsilon\eta) \Big|_{\varepsilon \rightarrow 0} = \int_{t_0}^{t_1} \eta \left(\frac{\partial L}{\partial y} - \frac{d}{dt} \frac{\partial L}{\partial y'} \right) dt$$

Pokud má být výše uvedený integrál roven nule, potom podle fundamentálního lemmatu variačního počtu⁸⁵ musí platit:

$$\frac{\partial L}{\partial y} - \frac{d}{dt} \frac{\partial L}{\partial y'} = 0 \quad (9)$$

Vztah (9) se nazývá *Eulerova-Lagrangeova rovnice*, v níž L je *Lagrangeova funkce*. Pokud známe Lagrangeovu funkci, můžeme aplikací principu stacionární akce získat pohybové rovnice fyzikálního systému.⁸⁶

D2 Odvození pohybové rovnice nabité částice v nekonzervativním visko-elastickém silovém poli (50) (Reakční síla radiačního vyzařování a Abrahamova-Lorentzova formule)

K odvození úplného ztrátového modelu částice v nekonzervativním silovém poli autorka předkládané disertační práce uvažovala, že oscilující (periodicky pohybující se) částice nese elektrický náboj a visko-elastický model sil působících na částici autorka rozšířila o tzv. reakční sílu radiačního vyzařování (*radiation reaction force*) F_r^i

$$F^i = ma^i = F_r^i + F_s^i + F_c^i$$

a odvodila obecný vzorec pro sílu odpovídající ztrátám vyzařováním *pro periodický děj* (oscilátor) v non-relativistickém pojetí⁸⁷ pomocí následujících úvah:

⁸⁵ Fundamentální lemma variačního počtu (*fundamental lemma of calculus of variations*) stanoví, že pokud je určitý integrál součinu dvou spojitých funkcí $f(x)$ a $b(x)$ nulový pro všechny spojitě funkce $b(x)$, které vymizí v okrajových bodech integrálu a mají první dvě derivace spojitě, potom je $f(x)$ rovno nule.

⁸⁶ Jinými slovy, pokud známe dynamický zákon, podle něhož se těleso pohybuje (dynamiku pohybu), můžeme získat vztah mezi silami působícími na těleso. V konzervativním silovém poli je to zákon zachování mechanické energie, který stanoví, že rozdíl kinetické a potenciální energie tělesa zůstává konstantní.

⁸⁷ Pojetí před teorií relativity a kvantovou fyzikou, pro rychlosti mnohem menší než je rychlost světla.

Uvažujme periodicky se pohybující elektrický náboj q (jednotka *Coulomb*) se zrychlením $a(t)$ (jednotka $m.s^{-2}$). Tento náboj vyzáří během jedné časové periody (τ_0, τ_1) energii W (jednotka *Joule*) výkonem $P(t) = dW/dt$ (jednotka $Joule.s^{-1} = N.m.s^{-1} = kg.m^2.s^{-3}$) podle *Larmorova* vzorce:

$$P(t) = \frac{2q^2}{3c^3} a^2(t) \quad (38)$$

kde c je rychlost světla ve vakuu (jednotka $m.s^{-1}$). Náboj q tudíž nedosáhne očekávané hodnoty kinetické energie, kterou bychom odvodili z jeho zrychlení, neboť část této energie je vyzářena. Situace se jeví tak, jako by na náboj působila zvláštní síla F_r (jednotka *Newton*, N), působící proti jeho zrychlení. Jelikož předpokládáme, že radiační výkon je výsledkem vynaložené práce W odpovídající vyzářené energii E_r , můžeme stanovit velikost této síly z následujícího vztahu, kam jsme dosadili z (38):

$$W = - \int_{\tau_0}^{\tau_1} F_r v(t) dt = \int_{\tau_0}^{\tau_1} P(t) dt = \frac{2q^2}{3c^3} \int_{\tau_0}^{\tau_1} a^2(t) dt = \frac{2q^2}{3c^3} \int_{\tau_0}^{\tau_1} \frac{dv(t)}{dt} \frac{dv(t)}{dt} dt$$

Integrál spočteme metodou *per-partes* použitím substituce $x = dv/dt$, $dy/dt = dv/dt$, $y = v$:

$$- \int_{\tau_0}^{\tau_1} F_r v(t) dt = \frac{2q^2}{3c^3} \left(\left[v(t) \frac{dv(t)}{dt} \right]_{\tau_0}^{\tau_1} - \int_{\tau_0}^{\tau_1} v(t) \frac{d^2v(t)}{dt^2} dt \right) \quad (39)$$

Veličina $J = d^2v(t)/dt^2 = da(t)/dt$ se nazývá *jerk* a je rovna časové změně zrychlení (jednotka $m.s^{-3}$). K časové změně zrychlení náboje může dojít pouze při změně velikosti síly F způsobující pohyb náboje. Je-li tato síla konstantní, konstantní bude rovněž zrychlení náboje, a tedy ($J = 0$).

Počáteční zrychlení je náboji vždy uděleno působící silou F . Počáteční zrychlení způsobí, že náboj začne působit silou F_r proti vlastnímu pohybu. To znamená, že časová změna zrychlení J bude nenulová a záporná. Člen

$$\left[v(t) \frac{dv(t)}{dt} \right]_{\tau_0}^{\tau_1} = 0$$

z rovnice (39) vymizí, neboť předpokládáme periodický děj v průběhu jedné periody. Porovnáním levé a pravé strany rovnice (39) dostáváme *Abrahamovu-Lorentzovu* formuli pro reakční sílu radiačního vyzařování:

$$F_r = \frac{2q^2}{3c^3} \frac{d^2v(t)}{dt^2} = \frac{2q^2}{3c^3} \frac{da(t)}{dt} = \gamma \frac{da(t)}{dt} \quad (40)$$

Vidíme, že reakční síla radiačního vyzařování F_r je funkcí gradientu zrychlení částice da/dt .⁸⁸ Spočteme proto rozptýlenou energii E_r a její přírůstek v závislosti na zrychlení:

⁸⁸ Rovnoměrně zrychlená částice nevykazuje podle (40), (41) žádnou radiační reakci, což je v rozporu s tvrzením, že nabitá částice vždy vyzařuje elektromagnetické vlnění. Vysvětlení se dotýká fyzikálního významu zrychlení částice a není zcela jednoznačné (tj. existuje několik možných vysvětlení).

$$\begin{aligned}
E_r &= - \int_{\tau_0}^{\tau_1} F_r \cdot v \cdot dt = - \int_{\tau_0}^{\tau_1} \gamma \frac{da}{dt} \cdot v \cdot dt = -\gamma \int_{\tau_0}^{\tau_1} v \cdot da \\
\frac{dE_r}{da} &= - \frac{d}{da} \gamma \int_{\tau_0}^{\tau_1} v \cdot da = -\gamma \int_{\tau_0}^{\tau_1} \frac{dv}{da} da = -\gamma \int_{\tau_0}^{\tau_1} dv = -\gamma v
\end{aligned} \tag{41}$$

Abychom mohli formálně odvodit pohybovou rovnici sil působících na nabitou částici, musíme použít modifikovanou *Eulerovu-Lagrangeovu* rovnici pro příslušnou *Lagrangeovu* funkci L_r , kdy L_r je nejen funkcí času, polohy a rychlosti, ale i zrychlení. Modifikovaná *Eulerova-Lagrangeova* rovnice je odvozena v dodatku D3 a má tvar:

$$\frac{\partial L_r}{\partial r} - \frac{d}{dt} \frac{\partial L_r}{\partial r'} + \frac{d^2}{dt^2} \frac{\partial L_r}{\partial r''} = 0 \tag{45}$$

Využitím uvedených vztahů odvodila autorka předkládané disertační práce pohybovou rovnici nabitě částice ve visko-elastickém nekonzervativním poli tak, že nejdříve spočetla jednotlivé členy rovnice (45) a z těchto členů potom sestavila výslednou rovnici:

$$\begin{aligned}
\frac{\partial L_r}{\partial r_n^i} &= \frac{\partial}{\partial r_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) - E_{sn}^i(r_n^i) - E_{rn}^i(a_n^i) \right) = - \frac{\partial U_n^i}{\partial r_n^i} - \frac{\partial E_{sn}^i}{\partial r_n^i} \\
\frac{\partial L_r}{\partial r_n^i} &= - \frac{\partial U_n^i}{\partial r_n^i} - \mu \mathcal{M} v_n^i
\end{aligned} \tag{46}$$

$$\begin{aligned}
\frac{d}{dt} \frac{\partial L_r}{\partial v_n^i} &= \frac{d}{dt} \frac{\partial}{\partial v_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) - E_{sn}^i(r_n^i) - E_{rn}^i(a_n^i) \right) = \frac{d}{dt} \frac{\partial}{\partial v_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i \right) \\
\frac{d}{dt} \frac{\partial L_r}{\partial v_n^i} &= \frac{d}{dt} m v_n^i = m \frac{d^2}{dt^2} r_n^i = m a_n^i
\end{aligned} \tag{47}$$

$$\frac{d^2}{dt^2} \frac{\partial L_r}{\partial a_n^i} = \frac{d^2}{dt^2} \frac{\partial}{\partial a_n^i} \left(\frac{1}{2} m v_n^i \cdot v_n^i - U_n^i(r_n^i) - E_{sn}^i(r_n^i) - E_{rn}^i(a_n^i) \right) = - \frac{d^2}{dt^2} \frac{\partial E_{rn}^i}{\partial a_n^i}$$

Za $\partial E_{rn}^i / \partial a_n^i$ autorka dosadila vztah (41):

$$\begin{aligned}
\frac{d^2}{dt^2} \frac{\partial E_{rn}^i}{\partial a_n^i} &= \frac{d^2}{dt^2} (-\gamma v_n^i) = -\gamma \frac{d^2 v_n^i}{dt^2} = -\gamma \frac{da_n^i}{dt} \\
\frac{d^2}{dt^2} \frac{\partial L_r}{\partial a_n^i} &= +\gamma \frac{da_n^i}{dt}
\end{aligned} \tag{48}$$

Sestavením (46), (47) a (48) vznikne pohybová rovnice i -té částice v n -té dimenzi nesoucí náboj v nekonzervativním visko-elastickém silovém poli:

$$- \frac{\partial U_n^i}{\partial r_n^i} - \mu \mathcal{M} v_n^i - m a_n^i + \gamma \frac{d}{dt} a_n^i = 0 \tag{49}$$

Protože se síly působící na částici v jejích jednotlivých dimenzích sčítají, vyjádříme výslednici sil jako

$$-\frac{\partial U^i}{\partial r^i} - \mu \mathcal{M} v^i - m a^i + \gamma \frac{d}{dt} a^i = 0 \quad (50)$$

a získáme pohybovou rovnici i -té nabitě částice ve visko-elastickém silovém poli, kterou můžeme následně porovnat s pohybovou rovnicí i -té částice *PSO* (29).

D3 Odvození modifikované *Eulerovy-Lagrangeovy* rovnice (45)

Máme odvodit modifikovanou *Eulerovu-Lagrangeovu* rovnici pro příslušnou *Lagrangeovu* funkci L_r , kdy L_r je nejen funkcí času, polohy a rychlosti, ale i zrychlení. Akce A_r systému bude:

$$A_r = \int_{t_0}^{t_1} L_r(t, f(t, r), f'(t, r'), f''(t, r'')) dt \quad (42)$$

kde t_0 je konstanta značící počáteční čas vývoje systému, t_1 je konstanta značící koncový čas vývoje systému, funkce $f'(t, r') = df/dt$, funkce $f''(t, r'') = d^2f/dt^2$ a funkce $L_r(t, f(t, r), f'(t, r'), f''(t, r''))$ je třikrát spojitě diferencovatelná s ohledem na argumenty t, f, f', f'' . Problém hledání nulové variace převedeme na řešení *Eulerovy-Lagrangeovy* rovnice podle postupu analogického jako v případě (6), (9) v dodatku D1 následovně:

Předpokládejme, že $A_r(f)$ dosahuje svého lokálního minima v bodě f_0 a $\eta(t, r)$ je libovolná testovací funkce, která má alespoň první tři derivace podle času a vymizí v okrajových bodech t_0 a t_1 . Funkce $f(t, r)$ musí mít alespoň první dvě derivace podle času. Potom platí

$$A_r(f_0) \leq A_r(f_0 + \varepsilon \eta)$$

pro libovolné ε blíží se k nule, kde $\varepsilon \eta$ se nazývá *variace* funkce f a značí se δf . Variace $A_r(f)$ je definována jako:

$$\delta(A_r(f, \eta)) = \lim_{\varepsilon \rightarrow 0} \frac{A_r(f + \varepsilon \eta) - A_r(f)}{\varepsilon} = \frac{d}{d\varepsilon} A_r(f + \varepsilon \eta) \Big|_{\varepsilon \rightarrow 0}$$

Druhá variace $A(f)$ je definována analogicky jako:

$$\delta^2(A_r(f, \eta)) = \lim_{\varepsilon \rightarrow 0} \frac{A_r(f' + \varepsilon \eta') - A_r(f')}{\varepsilon} = \frac{d}{d\varepsilon} A_r(f' + \varepsilon \eta') \Big|_{\varepsilon \rightarrow 0}$$

Vektor δA_r znamená, že uvažujeme infinitesimální posunutí (první variaci) nahrazující vektor $f(t, r)$ vektorem $(f(t, r) + \varepsilon \eta(t, r))$. Vektor $\delta^2 A_r$ znamená, že uvažujeme infinitesimální posunutí nahrazující vektor $f'(t, r)$ vektorem $(f'(t, r) + \varepsilon \eta'(t, r))$.

Substitucí $f = f_0 + \varepsilon \eta$ do (42) dostáváme:

$$A_r(f_0 + \varepsilon \eta) = \int_{t_0}^{t_1} L_r(t, (f_0 + \varepsilon \eta), (f_0' + \varepsilon \eta'), (f_0'' + \varepsilon \eta'')) dt \quad (43)$$

Pokud má funkcionál (42) minimum pro $f = f_0$, musí mít funkcionál (43), který je funkcí ε , minimum pro $\varepsilon = 0$, což zapíšeme:

$$0 = \frac{d}{d\varepsilon} A_r(f_0 + \varepsilon\eta) \Big|_{\varepsilon \rightarrow 0} = \int_{t_0}^{t_1} \frac{d}{d\varepsilon} (L_r(t, (f_0 + \varepsilon\eta), (f'_0 + \varepsilon\eta'), (f''_0 + \varepsilon\eta''))) dt \quad (44)$$

Použijeme substituci $y = f_0 + \varepsilon\eta$, $y' = f'_0 + \varepsilon\eta'$, $y'' = f''_0 + \varepsilon\eta''$ do (44) a provedeme totální derivaci funkce $L_r(t, y, y', y'')$ podle ε :

$$\begin{aligned} \int_{t_0}^{t_1} \frac{d}{d\varepsilon} (L_r(t, y, y')) dt &= \int_{t_0}^{t_1} \left(\frac{\partial L_r}{\partial y} \frac{dy}{d\varepsilon} + \frac{\partial L_r}{\partial y'} \frac{dy'}{d\varepsilon} + \frac{\partial L_r}{\partial y''} \frac{dy''}{d\varepsilon} \right) dt \\ \int_{t_0}^{t_1} \frac{d}{d\varepsilon} (L_r(t, y, y')) dt &= \int_{t_0}^{t_1} \left(\frac{\partial L_r}{\partial y} \eta + \frac{\partial L_r}{\partial y'} \eta' + \frac{\partial L_r}{\partial y''} \eta'' \right) dt \end{aligned}$$

Integrál součtu rozdělíme na součet integrálů, druhý z integrálů řešíme využitím metody per-partes podle vzorce $\int uv' = uv - \int u'v$, kam dosadíme $u = \partial L_r / (\partial y')$, $v = \eta$, třetí z integrálů řešíme využitím metody per-partes podle vzorce $\int uv' = uv - \int u'v$, kam dosadíme $u = \partial L_r / (\partial y'')$, $v = \eta'$, a dostáváme:

$$\begin{aligned} \int_{t_0}^{t_1} \frac{d}{d\varepsilon} (L_r(t, y, y')) dt &= \int_{t_0}^{t_1} \frac{\partial L_r}{\partial y} \eta dt + \int_{t_0}^{t_1} \frac{\partial L_r}{\partial y'} \eta' dt + \int_{t_0}^{t_1} \frac{\partial L_r}{\partial y''} \eta'' dt = \\ &= \int_{t_0}^{t_1} \frac{\partial L_r}{\partial y} \eta dt + \left[\frac{\partial L_r}{\partial y'} \eta \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt} \frac{\partial L_r}{\partial y'} \eta dt + \left[\frac{\partial L_r}{\partial y''} \eta' \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt} \frac{\partial L_r}{\partial y''} \eta' dt \end{aligned}$$

Jelikož řešíme úlohu s tzv. pevnými konci, neboli předpokládáme, že $\eta(t_0) = \eta(t_1) = 0$, $\eta'(t_0) = \eta'(t_1) = 0$, příslušné členy vymizí. Poslední integrál řešíme opět metodou *per-partes* podle vzorce $\int uv' = uv - \int u'v$, kam dosadíme $u = d(\partial L_r / (\partial y'')) / dt$, $v = \eta'$:

$$\int_{t_0}^{t_1} \frac{d}{dt} \frac{\partial L_r}{\partial y''} \eta' dt = \left[\frac{\partial L_r}{\partial y''} \eta' \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d^2}{dt^2} \frac{\partial L_r}{\partial y''} \eta dt$$

kde první člen rovněž vymizí.

Nyní sestavíme spočtené integrály, vytkneme η , dosadíme do (44) a dostáváme:

$$0 = \delta(A_r(f_0), (\eta)) = \frac{d}{d\varepsilon} A_r(f_0 + \varepsilon\eta) \Big|_{\varepsilon \rightarrow 0} = \int_{t_0}^{t_1} \eta \left(\frac{\partial L_r}{\partial y} - \frac{d}{dt} \frac{\partial L_r}{\partial y'} + \frac{d^2}{dt^2} \frac{\partial L_r}{\partial y''} \right) dt$$

Pokud má být uvedený integrál roven nule, potom podle fundamentálního lemmatu variačního počtu musí platit:

$$\frac{\partial L_r}{\partial y} - \frac{d}{dt} \frac{\partial L_r}{\partial y'} + \frac{d^2}{dt^2} \frac{\partial L_r}{\partial y''} = 0 \quad (45)$$

Vztah (45) je hledaná *Eulerova-Lagrangeova* rovnice.

D4 Odvození rovnice (52) charakterizující částici *PSO* jako dynamický systém popsaný polohovým vektorem

Rovnici charakterizující částici *PSO* jako dynamický systém popsaný polohovým vektorem $r_n^i(t)$ (16) odvodila autorka tak, že do rovnice (51) zpětně dosadila za $\beta = (\omega - 1)/\Delta t$ podle (27), dále vyjádřila zrychlení $a_n^i(k\Delta t)$ jako přírůstek rychlosti $v_n^i(k\Delta t)$ podle (22):

$$\begin{aligned} \frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} - \frac{\omega - 1}{\Delta t} v_n^i(t) - \varphi_n^i o_n^i(t) + \varphi_n^i r_n^i(t) &= 0 \\ v_n^i(t + \Delta t) - v_n^i(t) - (\omega - 1)v_n^i(t) + \varphi_n^i r_n^i(t)\Delta t &= \varphi_n^i o_n^i(t)\Delta t \\ v_n^i(t + \Delta t) - v_n^i(t) - \omega v_n^i(t) + v_n^i(t) + \varphi_n^i r_n^i(t)\Delta t &= \varphi_n^i o_n^i(t)\Delta t \\ v_n^i(t + \Delta t) - \omega v_n^i(t) + \varphi_n^i r_n^i(t)\Delta t &= \varphi_n^i o_n^i(t)\Delta t \end{aligned}$$

Rychlosti autorka nahradila analogicky, jako nahradila zrychlení:

$$\begin{aligned} r_n^i(t + \Delta t) - r_n^i(t) - \omega (r_n^i(t) - r_n^i(t - \Delta t)) + \varphi_n^i r_n^i(t)\Delta^2 t &= \varphi_n^i o_n^i(t)\Delta^2 t \\ r_n^i(t + \Delta t) - r_n^i(t) - \omega r_n^i(t) + \omega r_n^i(t - \Delta t) + \varphi_n^i r_n^i(t)\Delta^2 t &= \varphi_n^i o_n^i(t)\Delta^2 t \end{aligned}$$

Výsledkem je diferenční rovnice:

$$r_n^i(t + \Delta t) + (\varphi_n^i \Delta^2 t - \omega - 1)r_n^i(t) + \omega r_n^i(t - \Delta t) = \varphi_n^i o_n^i(t)\Delta^2 t \quad (52)$$

D5 Odvození rovnice (61a) charakterizující částici *PSO* jako dynamický systém popsaný vektorem rychlosti

Při odvození rovnice charakterizující částici *PSO* jako dynamický systém popsaný vektorem rychlosti vycházela autorka opět z rovnice (51), kam dosadila zpětně za $\beta = (\omega - 1)/\Delta t$ podle (27), dále dosadila za $\xi_n^i(t) = o_n^i(t) - r_n^i(t)$ podle (58), vyjádřila zrychlení a_n^i jako přírůstek rychlosti v_n^i a provedla úpravy:

$$\begin{aligned} a_n^i(k\Delta t) - \beta v_n^i(k\Delta t) - \varphi_n^i [o_n^i(k\Delta t) - r_n^i(k\Delta t)] &= 0 \\ \frac{v_n^i(t + \Delta t) - v_n^i(t)}{\Delta t} - \frac{\omega - 1}{\Delta t} v_n^i(t) - \varphi_n^i \xi_n^i(t) &= 0 \\ v_n^i(t + \Delta t) - v_n^i(t) - (\omega - 1)v_n^i(t) - \varphi_n^i \xi_n^i(t)\Delta t &= 0 \\ v_n^i(t + \Delta t) - v_n^i(t) - \omega v_n^i(t) + v_n^i(t) - \varphi_n^i \xi_n^i(t)\Delta t &= 0 \\ v_n^i(t + \Delta t) - \omega v_n^i(t) - \varphi_n^i \xi_n^i(t)\Delta t &= 0 \end{aligned} \quad (60)$$

Rovnice (60) je pouze jinou podobou rovnice (21). Nyní autorka vyjádřila pozici určenou proměnnou $\xi_n^i(t)$ pomocí rychlosti $v_n^i(t)$ a pozice $\xi_n^i(t - \Delta t)$. Pozici $r_n^i(t)$ vyjádřila analogicky s rovnicí (2):

$$\xi_n^i(t) = o_n^i(t) - r_n^i(t)$$

$$r_n^i(t) = r_n^i(t - \Delta t) + v_n^i(t)\Delta t$$

$$-o_n^i(t) + r_n^i(t) = -o_n^i(t) + r_n^i(t - \Delta t) + v_n^i(t)\Delta t$$

$$\xi_n^i(t) = \xi_n^i(t - \Delta t) - v_n^i(t)\Delta t$$

Za $v_n^i(t)$ dosadila autorka podle rovnic (1), (21) a (58):

$$v_n^i(t) = \omega v_n^i(t - \Delta t) + \varphi_n^i \xi_n^i(t - \Delta t)\Delta t$$

$$\xi_n^i(t) = \xi_n^i(t - \Delta t) - \omega v_n^i(t - \Delta t)\Delta t - \varphi_n^i \xi_n^i(t - \Delta t)\Delta^2 t$$

$$\xi_n^i(t) = -\omega v_n^i(t - \Delta t)\Delta t + (1 - \varphi\Delta^2 t)\xi_n^i(t - \Delta t)$$

Získaný vztah pro $\xi_n^i(t)$ autorka dosadila do (60) a provedla úpravy:

$$v_n^i(t + \Delta t) - \omega v_n^i(t) - \varphi_n^i \left(-\omega v_n^i(t - \Delta t)\Delta t + (1 - \varphi\Delta^2 t)\xi_n^i(t - \Delta t) \right) \Delta t = 0$$

$$v_n^i(t + \Delta t) - \omega v_n^i(t) + \varphi_n^i \Delta^2 t \omega v_n^i(t - \Delta t) - (1 - \varphi\Delta^2 t)\varphi_n^i \xi_n^i(t - \Delta t)\Delta t = 0$$

Za $(\varphi_n^i \xi_n^i(t - \Delta t)\Delta t)$ autorka dosadila podle rovnic (21) a (60) a provedla úpravy:

$$v_n^i(t) - \omega v_n^i(t - \Delta t) = \varphi_n^i \xi_n^i(t - \Delta t)\Delta t$$

$$v_n^i(t + \Delta t) - \omega v_n^i(t) + \varphi_n^i \Delta^2 t \omega v_n^i(t - \Delta t) - (1 - \varphi\Delta^2 t) \left(v_n^i(t) - \omega v_n^i(t - \Delta t) \right) = 0$$

$$\begin{aligned} v_n^i(t + \Delta t) - \omega v_n^i(t) + \varphi_n^i \Delta^2 t \omega v_n^i(t - \Delta t) - v_n^i(t) + \varphi\Delta^2 t v_n^i(t) + \omega v_n^i(t - \Delta t) \\ - \varphi\Delta^2 t \omega v_n^i(t - \Delta t) = 0 \end{aligned}$$

Výsledkem je diferenční rovnice popisující dynamické podmínky chování částice pomocí rychlosti:

$$v_n^i(t + \Delta t) + \left(\varphi_n^i \Delta^2 t - (\omega + 1) \right) v_n^i(t) + \omega v_n^i(t - \Delta t) = 0 \quad (61a)$$

D6 Odvození podmínky (73) vnitřní stability systému částice

Pro odvození podmínky vnitřní stability systému pro reálné řešení $\xi_n^i(t)$ pro reálné kořeny $\lambda_{1,2}$ (72) charakteristické rovnice (54) postačí zkoumat podmínku $|\lambda_2| < 1$:

$$\left| -\frac{b}{2} - \sqrt{\frac{b^2}{4} - \omega} \right| < 1$$

Protože pro absolutní hodnoty platí:

$$\left| -\frac{b}{2} - \sqrt{\frac{b^2}{4} - \omega} \right| \leq \left| -\frac{b}{2} \right| + \left| -\sqrt{\frac{b^2}{4} - \omega} \right|$$

podmínku zesílíme:

$$\left| -\frac{b}{2} \right| + \left| -\sqrt{\frac{b^2}{4} - \omega} \right| < 1$$

a provedeme úpravy:

$$\begin{aligned} \left| \sqrt{\frac{b^2}{4} - \omega} \right| &< 1 - \left| \frac{b}{2} \right| \\ \frac{b^2}{4} - \omega &< 1 - 2 \left| \frac{b}{2} \right| + \frac{b^2}{4} \\ |b| &< \omega + 1 \end{aligned}$$

Dosadíme zpět za $b = \varphi_n^i - (\omega + 1)$ a výsledná podmínka vnitřní stability systému pro reálné řešení $\xi_n^i(t)$ pro reálné kořeny $\lambda_{1,2}$ (72) charakteristické rovnice (54) je:

$$|\varphi_n^i - (\omega + 1)| < \omega + 1 \quad (73)$$

D7 Úprava diskriminantu ve výrazu (85)

Aby autorka mohla snáze porovnat výraz (84) s výrazem pro $|\vartheta_1|$, nejdříve si ve výrazu pro $|\vartheta_1|$ upraví výraz pro diskriminant:

$$\begin{aligned} |\vartheta_1| &= \left| \frac{(\chi + 1) - \chi\psi_n^i}{2} + \frac{\sqrt{(\chi\psi_n^i - (\chi + 1))^2 - 4\chi}}{2} \right| = \\ &= \left| \frac{(\chi + 1) - \chi\psi_n^i}{2} + \frac{\sqrt{\chi^2\psi_n^{i^2} - 2\chi\psi_n^i(\chi + 1) + (\chi + 1)^2 - 4\chi}}{2} \right| = \\ &= \left| \frac{(\chi + 1) - \chi\psi_n^i}{2} + \frac{\sqrt{\chi^2 \left(\psi_n^{i^2} - 4\frac{1}{\chi} - 2\psi_n^i - 2\psi_n^i\frac{1}{\chi} + 1 + 2\frac{1}{\chi} + \frac{1}{\chi^2} \right)}}{2} \right| = \\ &= \left| \frac{(\chi + 1) - \chi\psi_n^i}{2} + \frac{\sqrt{\chi^2 \left(\left(\psi_n^i - \frac{1}{\chi} \right)^2 - 2 \left(\psi_n^i - \frac{1}{\chi} \right) + 1 - 4\frac{1}{\chi} \right)}}{2} \right| \end{aligned}$$

Upravený výraz pro $|\vartheta_1|$ je roven:

$$|\vartheta_1| = \left| \frac{(\chi + 1) - \chi\psi_n^i}{2} + \frac{\sqrt{\chi^2 \left(\left(\psi_n^i - \frac{1}{\chi} + 1 \right)^2 - 4\frac{1}{\chi} \right)}}{2} \right| \quad (85)$$

D8 Odvození rovnice (97) pro stochastickou analýzu stagnace nejlepší částice

Diferenční rovnici odvodila autorka předkládané práce následujícím postupem:

Vyjdeme z rovnic (1) a (2) a popíšeme rychlost částice v časech $(t - \Delta t)$, t a $(t + \Delta t)$ ve dvou po sobě jdoucích iteracích tak, že budeme sledovat i realizace náhodných veličin v rovnici (1). Ve shodě s autorem studie [37] předpokládáme, že pro koeficient ω platí $0 < \omega < 1$. Jelikož autor [37] hledá maximální hodnotu koeficientů $a_l = a_g = a_{max}$, použil v rovnici (1) pro hodnotu a_{max} společné označení c .

Jako první krok spočteme iterativním způsobem, analogicky s rovnicí (1), polohový vektor $v_n^i(t)$:

$$\begin{aligned} v_n^i(t) = & \omega v_n^i(t - \Delta t) + cu_{1n}^i(t - \Delta t)[l_n^i(t - \Delta t) - r_n^i(t - \Delta t)]\Delta t \\ & + cu_{2n}^i(t - \Delta t)[g_n(t - \Delta t) - r_n^i(t - \Delta t)]\Delta t \end{aligned} \quad (89)$$

Výraz $cu_{1n}^i(t - 1)$ v rovnici (89) je realizací náhodné proměnné, kterou ve shodě s autorem [37] označíme C , výraz $cu_{2n}^i(t - 1)$ je další realizací této náhodné proměnné C . Nyní z rovnice (89) odstraníme hranaté závorky:

$$\begin{aligned} v_n^i(t) = & \omega v_n^i(t - \Delta t) + cu_{1n}^i(t - \Delta t)l_n^i(t - \Delta t)\Delta t - cu_{1n}^i(t - \Delta t)r_n^i(t - \Delta t)\Delta t \\ & + cu_{2n}^i(t - \Delta t)g_n(t - \Delta t)\Delta t - cu_{2n}^i(t - \Delta t)r_n^i(t - \Delta t)\Delta t \end{aligned} \quad (90)$$

Spočteme iterativním způsobem, analogicky s rovnicí (2), polohový vektor $r_n^i(t)$:

$$r_n^i(t) = r_n^i(t - \Delta t) + v_n^i(t)\Delta t \quad (91)$$

Do rovnice (1) pro $v_n^i(t + \Delta t)$, která je zde opakovaně uvedena:

$$\begin{aligned} v_n^i(t + \Delta t) = & \omega v_n^i(t) + cu_{1n}^i(t)[l_n^i(t) - r_n^i(t)]\Delta t \\ & + cu_{2n}^i(t)[g_n(t) - r_n^i(t)]\Delta t \end{aligned}$$

dosadíme za $r_n^i(t)$ výraz z rovnice (91):

$$\begin{aligned} v_n^i(t + \Delta t) = & \omega v_n^i(t) + cu_{1n}^i(t)[l_n^i(t) - r_n^i(t - \Delta t) - v_n^i(t)\Delta t]\Delta t \\ & + cu_{2n}^i(t)[g_n(t) - r_n^i(t - \Delta t) - v_n^i(t)\Delta t]\Delta t \end{aligned}$$

a vzniklou rovnici upravíme:

$$v_n^i(t + \Delta t) = \omega v_n^i(t) + cu_{1n}^i(t)l_n^i(t)\Delta t - cu_{1n}^i(t)r_n^i(t - \Delta t)\Delta t - cu_{1n}^i(t)v_n^i(t)\Delta t\Delta t$$

$$+ cu_{2n}^i(t)g_n(t)\Delta t - cu_{2n}^i(t)r_n^i(t - \Delta t)\Delta t - cu_{2n}^i(t)v^i(t)\Delta t\Delta t$$

Nyní seskupíme výrazy pro v_n^i , r_n^i :

$$\begin{aligned} v_n^i(t + \Delta t) &= v_n^i(t)(\omega - cu_{1n}^i(t)\Delta^2 t - cu_{2n}^i(t)\Delta^2 t) \\ &\quad + cu_{1n}^i(t)l_n^i(t)\Delta t + cu_{2n}^i(t)g_n(t)\Delta t \\ &\quad - r_n^i(t - \Delta t)(cu_{1n}^i(t)\Delta t + cu_{2n}^i(t)\Delta t) \end{aligned}$$

a zohledníme skutečnost, že se hodnoty lokální a globální nejlepší pozice ve výrazech pro $v_n^i(t)$, $v_n^i(t + \Delta t)$ nezměnily, neboť řešíme případ stagnace, neboli $l_n^i(t) = l_n^i(t - \Delta t)$, $g_n = g_n(t - \Delta t)$, a dosadíme tyto výrazy za $l_n^i(t)$, $g_n^i(t)$ do rovnice pro $v_n^i(t + \Delta t)$:

$$\begin{aligned} v_n^i(t + \Delta t) &= v_n^i(t)(\omega - cu_{1n}^i(t)\Delta^2 t - cu_{2n}^i(t)\Delta^2 t) \\ &\quad + cu_{1n}^i(t)l_n^i(t - \Delta t)\Delta t + cu_{2n}^i(t)g_n(t - \Delta t)\Delta t \\ &\quad - r_n^i(t - \Delta t)(cu_{1n}^i(t)\Delta t + cu_{2n}^i(t)\Delta t) \end{aligned} \quad (92)$$

Výrazy $cu_{1n}^i(t)$, $cu_{2n}^i(t)$ v rovnici (92) jsou dalšími realizacemi náhodné proměnné C (máme již tedy celkem čtyři realizace). Nyní z rovnice (90) osamostatníme polohový vektor $r_n^i(t - \Delta t)$:

$$\begin{aligned} -r_n^i(t - \Delta t) &= \\ &= \frac{v_n^i(t) - \omega v_n^i(t - \Delta t) + cu_{1n}^i(t - \Delta t)l_n^i(t - \Delta t)\Delta t + cu_{2n}^i(t - \Delta t)g_n(t - \Delta t)\Delta t}{cu_{1n}^i(t - \Delta t)\Delta t + cu_{2n}^i(t - \Delta t)\Delta t} \end{aligned}$$

Pro získání výsledné rovnice pro rychlost zbývá dosadit za $r_n^i(t - \Delta t)$ do rovnice (92). Abychom získali přehlednější výsledný zápis, zavedeme následující substituce:

$$\begin{aligned} l_n^i(t - \Delta t) &\approx l \\ g_n(t - \Delta t) &\approx g \\ cu_{1n}^i(t - \Delta t) &\approx c_1 \\ cu_{2n}^i(t - \Delta t) &\approx c_2 \\ cu_{1n}^i(t) &\approx c_3 \\ cu_{2n}^i(t) &\approx c_4 \end{aligned}$$

Zápisy rovnice (92) a výrazu pro $r_n^i(t - \Delta t)$ dostávají s použitím substitucí následující podobu:

$$\begin{aligned} v_n^i(t + \Delta t) &= v_n^i(t)(\omega - c_3\Delta^2 t - c_4\Delta^2 t) + c_3l\Delta t + c_4g\Delta t \\ &\quad - r_n^i(t - \Delta t)(c_3\Delta t + c_4\Delta t) \end{aligned} \quad (93)$$

$$-r_n^i(t - \Delta t) = \frac{v_n^i(t) - \omega v_n^i(t - \Delta t) + c_1l\Delta t + c_2g\Delta t}{c_1\Delta t + c_2\Delta t} \quad (94)$$

Nyní již můžeme do rovnice (93) pro rychlost $v_n^i(t + \Delta t)$ dosadit výraz (94) pro $r_n^i(t - \Delta t)$:

$$v_n^i(t + \Delta t) = v_n^i(t) \left(\omega - c_3 \Delta^2 t - c_4 \Delta^2 t + \frac{c_3 \Delta t + c_4 \Delta t}{c_1 \Delta t + c_2 \Delta t} \right) - v_n^i(t - \Delta t) \omega \frac{c_3 \Delta t + c_4 \Delta t}{c_1 \Delta t + c_2 \Delta t} + c_3 l \Delta t + c_4 g \Delta t + (c_1 l \Delta t + c_2 g \Delta t) \frac{c_3 \Delta t + c_4 \Delta t}{c_1 \Delta t + c_2 \Delta t} = 0$$

Zavedeme další substituci:

$$\frac{c_3 \Delta t + c_4 \Delta t}{c_1 \Delta t + c_2 \Delta t} = q$$

Zápis rovnice pro rychlost $v_n^i(t + \Delta t)$ s použitím další substituce je následující:

$$v_n^i(t + \Delta t) = v_n^i(t) (\omega - c_3 \Delta^2 t - c_4 \Delta^2 t + q) - v_n^i(t - \Delta t) \omega q + c_3 l \Delta t + c_4 g \Delta t + (c_1 l \Delta t + c_2 g \Delta t) q$$

Zavedeme další substituci:

$$c_3 \Delta t + c_4 \Delta t = s \Delta t$$

Zápis rovnice pro rychlost $v_n^i(t + \Delta t)$ s použitím další substituce je následující:

$$v_n^i(t + \Delta t) = v_n^i(t) (\omega - s \Delta^2 t + q) - v_n^i(t - \Delta t) \omega q + l(c_3 \Delta t - c_1 \Delta t q) + g(c_4 \Delta t - c_2 \Delta t q)$$

Lze nahlédnout, že platí rovnost výrazů:

$$c_3 \Delta t - c_1 \Delta t q = \frac{c_2 \Delta t c_3 \Delta t - c_1 \Delta t c_4 \Delta t}{c_1 \Delta t + c_2 \Delta t} = -(c_4 \Delta t - c_2 \Delta t q) \approx d \Delta t \quad (95)$$

Zohledníme-li tuto rovnost, můžeme psát výslednou diferenční rovnici pro rychlost:

$$v_n^i(t + \Delta t) = (\omega - s \Delta^2 t - q) v_n^i(t) - \omega q v_n^i(t - \Delta t) + d \Delta t (l - g) \quad (96)$$

Ve studii [37] je však zkoumána diferenční rovnice pro rychlost za předpokladu $\Delta t = 1$. Časový krok Δt je ve studii [37] implicitně roven jedné od počátku, a tedy jeho vliv na rychlost uvažován není. Rovnici (96) si proto odpovídajícím způsobem přepíšeme:

$$v_n^i(t + 1) = (\omega - s + q) v_n^i(t) - \omega q v_n^i(t - 1) + d(l - g) \quad (97)$$

D9 Charakteristiky náhodných veličin v rovnici (97) a jejich vliv na rychlost částice

Veličiny s , q , d rovnici (97) jsou realizacemi náhodných veličin S , Q , D . Je definována také náhodná veličina

$$Z = \omega - S + Q$$

Náhodná veličina D , jež je ve studii [37] označena jako W a jejíž realizace je obsažena v rovnici (97), je v [37] interpretována jako šum. Čím větší je rozdíl $(l - g)$, tím větší je vliv tohoto šumu na rychlost.

Hodnota realizace náhodné veličiny Q je vždy kladná, a tedy hodnota $(-\omega Q)$ v rovnici (97) je vždy záporná, a proto je veličina ωQ označována v [37] jako odporující síla (*back force*). Její hodnota závisí na ω .

Hodnota realizace náhodné veličiny $Z = \omega - S + Q$ bude častěji kladná, ale může být i záporná. Veličina Z je ve studii [37] označována jako dopředná síla (*forth force*).

Charakteristika náhodné veličiny S

Náhodná veličina $S = C_3 + C_4$ v rovnici (97) je součtem dvou vzájemně nezávislých náhodných veličin rovnoměrného rozložení $U(0, c)$. Hustota pravděpodobnosti každé z veličin C_3, C_4 je $f_c = 1/c$. Necht' u, x jsou reálná čísla. Pravděpodobnost $(S = u)$ je dána pravděpodobností, že $(C_3 = x)$ a současně $(C_4 = u - x)$, což zapíšeme:

$$P(S = u) = f_S(u) = \int_{-\infty}^{+\infty} f_{C_3}(x) f_{C_4}(u - x) dx = \int_{-\infty}^{+\infty} f_c(x) f_c(u - x) dx$$

Hustota pravděpodobnosti $f_c(x) = f_c(u - x) = \left(\frac{1}{c}\right)$.

Za předpokladu $c \neq 0, 0 \leq x \leq c, 0 \leq u \leq 2c$ nastávají dva případy:
 $0 < u \leq c$

$$f_{S_1}(u) = \int_0^u \left(\frac{1}{c}\right) \left(\frac{1}{c}\right) dx = \frac{u}{c^2}$$

$c \leq u < 2c$

$$f_{S_2}(u) = \int_{u-c}^c \left(\frac{1}{c}\right) \left(\frac{1}{c}\right) dx = \frac{2c - u}{c^2}$$

Rozsah realizací náhodné veličiny S je $(0, 2c)$, střední hodnota $\mu(S) = c$. Veličinu S lze zapsat rovněž jako $S = c(U_1 + U_2)$, kde U_1, U_2 jsou náhodné veličiny rovnoměrného rozložení $U(0, 1)$.

Obdobně bychom spočetli hustotu pravděpodobnosti náhodné veličiny dané součtem vzájemně nezávislých náhodných veličin $C_1 + C_2$.

Charakteristika náhodné veličiny Q

Náhodná veličina Q v rovnici (97) je podíl dvou náhodných veličin Q_1, Q_2 kde $Q_1 = S, Q_2 = C_1 + C_2$. Rozsah realizací náhodné veličiny Q je $(0, \infty)$. Abychom mohli určit střední hodnotu $\mu(Q)$ veličiny Q , musíme znát i hustotu pravděpodobnosti veličiny Q .

Studie [37] stanovuje $\mu(Q)$ postupem, který je pouze naznačen. Protože $\mu(Q)$ je ukazatel, který je uplatňován jedním z autorů [38] v navrhované standardní verzi algoritmu *PSO* [4], hodnotu $\mu(Q)$, jejíž odvození je v [37] pouze naznačeno, zde odvodíme podrobně.

$$\mu(Q) = \mu\left(\frac{Q_1}{Q_2}\right) = \mu\left(\frac{C_3 + C_4}{C_1 + C_2}\right) = \mu\left(\frac{cU_3 + cU_4}{cU_1 + cU_2}\right) = \mu\left(\frac{c(U_3 + U_4)}{c(U_1 + U_2)}\right)$$

kde U_1, U_2, U_3, U_4 jsou náhodné veličiny rovnoměrného rozložení $U(0,1)$. Protože Q_1, Q_2 jsou nezávislé náhodné veličiny, můžeme psát:

$$\mu(Q) = \mu(U_3 + U_4) * \mu\left(\frac{1}{U_1 + U_2}\right) = 1 * \mu\left(\frac{1}{U_1 + U_2}\right)$$

Protože $\mu(U_3 + U_4) = 1$, pro určení střední hodnoty $\mu(Q)$ nám postačí, zjistíme-li střední hodnotu veličiny $1/(U_1 + U_2)$.

Zavedeme nejdříve obecně náhodnou veličinu V jako součin dvou nezávislých náhodných veličin $X, 1/Y$ rovnoměrného rozložení $U(0,2)$. Hustoty pravděpodobnosti veličin X, Y

$$f_X(x) = \begin{cases} x & 0 \leq x \leq 1 \\ 2 - x & 1 < x \leq 2 \end{cases}$$

$$f_Y(y) = \begin{cases} y & 0 \leq y \leq 1 \\ 2 - y & 1 < y \leq 2 \end{cases}$$

V ostatních případech jsou hustoty pravděpodobnosti veličin X, Y rovny nule.

Pravděpodobnost ($V = z$), kde z je reálné číslo, vyjádříme jako sdruženou hustotu pravděpodobnosti (*joint probability density*) veličin $X, 1/Y$:

$$P(V = z) = f_V(z) = \left| \frac{d}{dz} P(X/Y \leq z) \right| = \left| \frac{d}{dz} P(X \leq zY) \right|$$

$$P(X \leq zY) = P\left(0 \leq X \leq zY, 0 \leq Y \leq \frac{1}{z}\right) = \int_0^{1/z} \int_0^{zy} f_{XY}(x, y) dx dy \quad (98)$$

Vztah (98) vyjadřuje sdruženou distribuční funkci (*joint distribution*) nezávislých veličin $X, 1/Y$. Pro nezávislé veličiny platí:

$$\int_0^{1/z} \int_0^{zy} f_{XY}(x, y) dx dy = \int_0^{1/z} \int_0^{zy} f_X(x) f_Y(y) dx dy \quad (99)$$

Ve vztahu (99) je $f_X(x)$ hustota pravděpodobnosti veličiny X a $f_Y(y)$ je hustota pravděpodobnosti veličiny Y .

V obecném případě bychom spočetli sdruženou distribuční funkci náhodných veličin $X, 1/Y$ rovnoměrného rozložení $(0,2)$. Derivací bychom získali hustotu pravděpodobnosti $f_V(z)$ proměnné $V = X/Y$ a spočetli bychom střední hodnotu proměnné V .

Můžeme však využít toho, že v našem konkrétním případě hledáme sdruženou distribuční funkci $P(1 \leq zY)$ pro hodnotu ($f_X(x) = x = 1$), a tedy ($zy = 1$), ($z = 1/y$). Rozsah náhodné veličiny Y je $Y(0,2)$, a proto hodnota proměnné z bude z intervalu $(1/2, 1)$.

Spočteme nejdříve vnitřní integrál obsažený ve vztahu (99):

$$\int_0^{zy} f_X(x) dx \int_0^{zy} 1 dx = [x]_0^{zy} = zy$$

Nyní spočteme vnější integrál (99) pro $0 \leq y \leq 1$, to jest pro $1 \leq z \leq +\infty$

$$\int_0^{1/z} f_Y(y) dy = \int_0^{1/z} y dy = \left[\frac{y^2}{2} \right]_0^{1/z} = \frac{1}{2z^2}$$

Spočteme hustotu pravděpodobnosti $f_V(z)$ proměnné $V = X/Y$ v intervalu $1 \leq z \leq +\infty$

$$f_V(z) = \left| \frac{d}{dz} \left(\frac{1}{2z^2} \right) \right| = \left| -\frac{1}{z^3} \right| = \left| \frac{1}{z^3} \right|$$

Dále spočteme vnější integrál (99) pro $1 \leq y \leq 2$, to jest pro $1/2 \leq z \leq 1$

$$\int_0^{1/z} f_Y(y) dy = \int_0^{1/z} (2-y) dy = \left[2y - \frac{y^2}{2} \right]_0^{1/z} = \frac{2}{z} - \frac{1}{2z^2}$$

Spočteme hustotu pravděpodobnosti $f_V(z)$ proměnné $V = X/Y$ v intervalu $1/2 \leq z \leq 1$

$$f_V(z) = \left| \frac{d}{dz} \left(\frac{2}{z} - \frac{1}{2z^2} \right) \right| = \left| -\frac{2}{z^2} + \frac{1}{z^3} \right|$$

Spočteme hledanou střední hodnotu náhodné proměnné Q

$$\begin{aligned} \mu(Q) = \mu(V) &= \int_{-\infty}^{+\infty} z f_V(z) dz = \int_{1/2}^1 z \left(\frac{2}{z^2} - \frac{1}{z^3} \right) dz + \int_1^{+\infty} z \left(\frac{1}{z^3} \right) dz \\ &= \int_{1/2}^1 \left(\frac{2}{z} - \frac{1}{z^2} \right) dz - \int_1^{+\infty} \frac{1}{z^2} dz = \int_{1/2}^1 \frac{2}{z} dz - \int_{1/2}^1 \frac{1}{z^2} dz + \int_1^{+\infty} \frac{1}{z^2} dz \\ \mu(Q) &= 2[\ln(z)]_{1/2}^1 = 2 \left(\ln(1) - \ln\left(\frac{1}{2}\right) \right) = 2 \ln(2) \end{aligned}$$

Hledaná střední hodnota náhodné proměnné Q je $\mu(Q) = 2 \ln(2)$.

Charakteristika náhodné veličiny Z

Náhodnou veličinu $Z = \omega - S + Q$ lze přepsat dosazením za S, Q

$$Z = \omega - (C_3 + C_4) + \frac{C_3 + C_4}{C_1 + C_2} = \omega - c(U_3 + U_4) + \frac{c(U_3 + U_4)}{c(U_1 + U_2)}$$

kde U_1, U_2, U_3, U_4 jsou náhodné veličiny rovnoměrného rozložení $U(0,1)$. Uvážíme-li, že jednotlivé realizace náhodné veličiny C nabývají hodnot z intervalu $(0, c)$, pak minimální hodnota, kterou nabývá veličina Z , je

$$z_{min} = \omega - s_{max} + q_{smax} = \omega - c(1+1) + \frac{c(1+1)}{c(1+1)} = \omega - 2c + 1$$

Rozsah realizací náhodné veličiny Z je $(\omega - 2c + 1, \infty)$. Položíme-li $(\omega - 2c + 1 = 0)$, a to splněním podmínky

$$c = \frac{\omega + 1}{2} \quad (100)$$

dosáhneme toho, že realizace náhodné veličiny Z budou v intervalu $(0, \infty)$.

Střední hodnota veličiny Z je dána součtem obsahujícím střední hodnoty S, Q

$$\mu(Z) = \omega - \mu(S) + \mu(Q) = \omega - c + 2 \ln(2)$$

Charakteristika náhodné veličiny D

Náhodná veličina D je definována podle vztahu (95) jako

$$D = \frac{C_2 C_3 - C_1 C_4}{C_1 + C_2} = c \frac{U_2 U_3 - U_1 U_4}{U_1 + U_2}$$

Velčina D je symetrická se střední hodnotou $\mu(D) = 0$. Rozsah realizací veličiny D je $(-c, +c)$.

D10 Odvození vzorce (103) pro maximální hodnotu koeficientu c_d plynoucího z deterministické analýzy

Vzorec má být odvozen z následujících vztahů plynoucích z deterministické analýzy stability systému částice:

$$\omega_{max} \approx \chi_{max} = \frac{1}{|\mu_2|} = \frac{2}{\psi_n^i - 2 + \sqrt{\psi_n^{i^2} - 4\psi_n^i}} \approx \chi \quad (104)$$

$$\chi_{max} * \psi_n^i = 2c_d \quad (105)$$

Vztah (104) vznikne z rovnice (87) dosazením $k = 1$. Vztah (104) lze odvodit z deterministické verze algoritmu, kdy náhodné proměnné $u_{1n}^i(t)$, $u_{2n}^i(t)$ v rovnicích (1), (18) a (19) nabývají očekávaných hodnot $E[u_{1n}^i(t)] = E[u_{2n}^i(t)] = 0,5$. Tedy:

$$E[\varphi_n^i] = E[\varphi_{1n}^i] + E[\varphi_{2n}^i] = a_l E[u_{1n}^i(t)] + a_g E[u_{2n}^i(t)] = \frac{a_l + a_g}{2}$$

Maximální hodnoty tudíž budou

$$\varphi_{n_{max}}^i = a_{l_{max}} + a_{g_{max}} = c_d + c_d$$

a položíme-li $\varphi_n^i = \chi \psi_n^i$, docházíme ke vztahu (105).

Při odvození vztahu pro c_d vyjdeme ze vztahu (104), kde $\chi_{max} \approx \chi$:

$$\chi = \frac{2}{\psi_n^i - 2 + \sqrt{\psi_n^{i^2} - 4\psi_n^i}}$$

Převrátíme čitatele a jmenovatele na obou stranách rovnice, osamostatníme vztah s odmocninou a obě strany rovnice umocníme:

$$\begin{aligned}
 \frac{2}{\chi} - \psi_n^i + 2 &= \sqrt{\psi_n^{i^2} - 4\psi_n^i} \\
 \left(\frac{2}{\chi} - \psi_n^i + 2\right)^2 &= \psi_n^{i^2} - 4\psi_n^i \\
 \frac{4}{\chi^2} - \frac{2\psi_n^i}{\chi} + \frac{4}{\chi} - \frac{2\psi_n^i}{\chi} + \psi_n^{i^2} - 2\psi_n^i + \frac{4}{\chi} - 2\psi_n^i + 4 &= \psi_n^{i^2} - 4\psi_n^i \\
 \frac{4}{\chi^2} - 2\frac{2\psi_n^i}{\chi} + 2\frac{4}{\chi} + 4 &= 0 \\
 4\left(\frac{1}{\chi} + 1\right)^2 &= 2\frac{2\psi_n^i}{\chi} \\
 \frac{(1 + \chi)^2}{\chi^2} &= \frac{\psi_n^i}{\chi} \\
 (1 + \chi)^2 &= \psi_n^i \chi
 \end{aligned}$$

Využijeme vztah (105) $\chi_{max} * \psi_n^i = 2c_d$ a docházíme k hledanému vzorci (103) pro c_d :

$$c_d = \frac{(1 + \chi_{max})^2}{2}$$

D11 Odvození soustavy diferenčních rovnic (109)

K popisu systému pomocí matice systému M potřebuje autor studie [38] nejdříve získat diferenční rovnice pro momenty $E[r_{t+1}^2]$, $E[r_{t+1}r_t]$. Autor studie [38] nejdříve spočetl výraz r_{t+1}^2 pomocí rovnice (108):

$$\begin{aligned}
 r_{t+1}^2 &= (\tilde{\omega}r_t - \varphi_1r_t - \varphi_2r_t - \omega r_{t-1} + \varphi_1l + \varphi_2g)^2 = \\
 &\tilde{\omega}^2r_t^2 - \tilde{\omega}\varphi_1r_t^2 - \tilde{\omega}\varphi_2r_t^2 - \tilde{\omega}r_t\omega r_{t-1} + \tilde{\omega}r_t\varphi_1l + \tilde{\omega}r_t\varphi_2g \\
 &- \tilde{\omega}\varphi_1r_t^2 + \varphi_1^2r_t^2 + \varphi_1\varphi_2r_t^2 + \varphi_1r_t\omega r_{t-1} - \varphi_1^2r_tl - \varphi_1r_t\varphi_2g \\
 &- \tilde{\omega}\varphi_2r_t^2 + \varphi_1\varphi_2r_t^2 + \varphi_2^2r_t^2 + \varphi_2r_t\omega r_{t-1} - \varphi_2r_t\varphi_1l - \varphi_2^2r_tg \\
 &- \tilde{\omega}r_t\omega r_{t-1} + \varphi_1r_t\omega r_{t-1} + \varphi_2r_t\omega r_{t-1} + \omega^2r_{t-1}^2 - \omega r_{t-1}\varphi_1l - \omega r_{t-1}\varphi_2g \\
 &+ \tilde{\omega}r_t\varphi_1l - \varphi_1r_t\varphi_1l - \varphi_2r_t\varphi_1l - \omega r_{t-1}\varphi_1l + \varphi_1^2l^2 + \varphi_1l\varphi_2g \\
 &+ \tilde{\omega}r_t\varphi_2g - \varphi_1r_t\varphi_2g - \varphi_2^2r_tg - \omega r_{t-1}\varphi_2g + \varphi_1l\varphi_2g + \varphi_2^2g^2
 \end{aligned}$$

Autor [38] potom seskupil odpovídající výrazy a současně aplikoval operátor E :

$$\begin{aligned}
 E[r_{t+1}^2] &= E[r_t^2](\tilde{\omega}^2 - 2\tilde{\omega}E[\varphi_1] - 2\tilde{\omega}E[\varphi_2] + E[\varphi_1^2] + 2E[\varphi_1\varphi_2] + E[\varphi_2^2]) \\
 &\quad + E[r_tr_{t-1}](-2\tilde{\omega}\omega + 2E[\varphi_1]\omega + 2E[\varphi_2]\omega) \\
 &\quad + E[r_{t-1}^2](\omega^2)
 \end{aligned} \tag{111}$$

$$\begin{aligned}
& + E[r_t](2\tilde{\omega}E[\varphi_1]l + 2\tilde{\omega}E[\varphi_2]g - 2E[\varphi_1^2]l - 2E[\varphi_1\varphi_2]g - 2E[\varphi_1\varphi_2]l - 2E[\varphi_2^2]g) \\
& + E[r_{t-1}](-2\omega E[\varphi_1]l - 2\omega E[\varphi_2]g) \\
& + E[\varphi_1^2]l^2 + 2E[\varphi_1]l E[\varphi_2]g + E[\varphi_2^2]g^2
\end{aligned}$$

Nyní již autor [38] mohl psát soustavu diferenčních rovnic pro momenty $E[r_{t+1}]$, $E[r_{t+1}r_t]$, $E[r_{t+1}^2]$.

Diferenční rovnici pro moment $E[r_{t+1}]$ získal autor [38] z rovnice (108).

Diferenční rovnici pro moment druhého řádu $E[r_{t+1}r_t]$ získal autor [38] vynásobením rovnice (107) výrazem r_t a následnou aplikací operátoru E . Vzhledem k předpokladu $a_l = a_g$ dosadil autor v obou případech $E[\varphi_1] = E[\varphi_2] = c/2$.

Diferenční rovnice pro $E[r_{t+1}^2]$ je rovnice (111), v níž autor [38] provedl substituce

$$\begin{aligned}
E[\varphi_1] &= E[\varphi_2] \approx \mu \\
E[\varphi_1^2] &= E[\varphi_2^2] \approx \nu
\end{aligned} \tag{112}$$

Výsledná soustava diferenčních rovnic je následující:

$$E[r_{t+1}] = E[r_t](\tilde{\omega} - c) - \omega E[r_{t-1}] + c \frac{l+g}{2} \tag{113}$$

$$E[r_{t+1}^2] = E[r_t^2](\tilde{\omega}^2 - 2\tilde{\omega}\mu - 2\tilde{\omega}\mu + \nu + 2\mu^2 + \nu) \tag{114}$$

$$+ E[r_t r_{t-1}](-2\tilde{\omega}\omega + 2\mu\omega + 2\mu\omega)$$

$$+ E[r_{t-1}^2](\omega^2)$$

$$+ E[r_t](2\tilde{\omega}\mu l + 2\tilde{\omega}\mu g - 2\nu l - 2\mu^2 g - 2\mu^2 l - 2\nu g)$$

$$+ E[r_{t-1}](-2\omega\mu l - 2\omega\mu g)$$

$$+ \nu l^2 + 2\mu l \mu g + \nu g^2$$

$$E[r_{t+1}r_t] = E[r_t^2](\tilde{\omega} - c) - \omega E[r_t r_{t-1}] + E[r_t]c \frac{l+g}{2} \tag{115}$$

Soustavu diferenčních rovnic (113), (114), (115) zapsal autor [38] v maticovém tvaru podle (109) a získal matici M , která je pátého řádu, a pro úplnost je uveden i vektor b :

$$M = \begin{bmatrix}
\tilde{\omega} - c & -\omega & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\begin{pmatrix} 2\tilde{\omega}\mu l + 2\tilde{\omega}\mu g - 2\nu l \\ -2\mu^2 g - 2\mu^2 l - 2\nu g \end{pmatrix} & \begin{pmatrix} -2\omega\mu l \\ -2\omega\mu g \end{pmatrix} & \begin{pmatrix} \tilde{\omega}^2 - 2\tilde{\omega}\mu - 2\tilde{\omega}\mu \\ +\nu + 2\mu^2 + \nu \end{pmatrix} & (-2\tilde{\omega}\omega + 4\mu\omega) & \omega^2 \\
c(l+g)/2 & 0 & (\tilde{\omega} - c) & -\omega & 0 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}$$

$$b = \begin{bmatrix} c(l+g)/2 \\ 0 \\ \nu l^2 + 2\mu l \mu g + \nu g^2 \\ 0 \\ 0 \end{bmatrix}$$

D12 Odvození vztahu (119) pro standardní odchylku pozic trajektorie částice od jejich očekávaných hodnot

Autor [38] vychází při odvození vztahu pro standardní odchylku

$$\sigma(r_t) = \sqrt{E[r_t^2] - (E[r_t])^2}$$

z úvahy, že pro $t \rightarrow \infty$ jsou očekávané hodnoty $E[r_t], E[r_{t-1}]$ téměř nerozlišitelné od stacionárního bodu p , a proto dosadil p do rovnice (114) a následně odvodil výrazy pro hodnoty stacionárních bodů pro momenty druhého řádu:

$$\begin{aligned} E[r_{t+1}^2] &= E[r_t^2](\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) \\ &\quad + E[r_t r_{t-1}](-2\tilde{\omega}\omega + 4\mu\omega) \\ &\quad + E[r_{t-1}^2](\omega^2) \\ &\quad + 2p(\tilde{\omega}\mu l + \tilde{\omega}\mu g - vl - \mu^2 g - \mu^2 l - vg) \\ &\quad - 2p\omega\mu(l + g) \\ &\quad + vl^2 + 2\mu l \mu g + vg^2 \end{aligned}$$

Nejdříve bude samostatně upraven označený výraz:

$$\begin{aligned} 2p(\tilde{\omega}\mu l + \tilde{\omega}\mu g - vl - \mu^2 g - \mu^2 l - vg) &= 2p(\tilde{\omega}\mu(l + g) - v(l + g) - \mu^2(l + g)) = \\ &= 2p(l + g)(\tilde{\omega}\mu - v - \mu^2) = 2p(l + g)((\omega + 1)\mu - v - \mu^2) = \\ &= 2p(l + g)((\omega + 1)\mu - v - \mu^2) = 2p(l + g)\omega\mu + 2p(l + g)(\mu - v - \mu^2) \end{aligned}$$

Upravený výraz bude dosazen zpět. Dále bude použit vztah plynoucí ze (116)

$$2p(l + g) = 4p \frac{(l + g)}{2} = 4p^2$$

který bude rovněž dosazen. Výsledkem je rovnice:

$$\begin{aligned} E[r_{t+1}^2] &= E[r_t^2](\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) \\ &\quad + E[r_t r_{t-1}](-2\tilde{\omega}\omega + 4\mu\omega) \\ &\quad + E[r_{t-1}^2](\omega^2) \\ &\quad + 4p^2(\mu - v - \mu^2) \\ &\quad + vl^2 + 2\mu^2 l g + vg^2 \end{aligned} \tag{117}$$

Protože oblast stability druhého řádu leží uvnitř oblasti stability prvního řádu, autor [38] usuzuje, že pokud konverguje $E[r_t], E[r_{t-1}]$ k bodu p , bude rovněž konvergovat $E[r_t^2], E[r_t r_{t-1}]$ a $E[r_{t-1}^2]$, a sice k bodům, které označil p_{r^2} a p_{rr} tak, že:

$$z^*(t) = (p \ p \ p_{r^2} \ p_{rr} \ p_{r^2})^T$$

Rovněž tak $E[r_{t+1}^2]$ bude konvergovat k p_{r^2} . Dosadí-li autor tyto body do rovnice (117):

$$\begin{aligned} p_{r^2} &= p_{r^2}(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) \\ &\quad + p_{rr}(-2\tilde{\omega}\omega + 4\mu\omega) \\ &\quad + p_{r^2}(\omega^2) \\ &\quad + 4p^2(\mu - v - \mu^2) \\ &\quad + vl^2 + 2\mu^2l g + vg^2 \end{aligned}$$

a připojí-li rovnici (115), z níž získá obdobným dosazovacím postupem vztah pro p_{rr} :

$$\begin{aligned} E[r_{t+1}r_t] &= E[r_t^2](\tilde{\omega} - c) - \omega E[r_t r_{t-1}] + E[r_t] c \frac{l+g}{2} \\ p_{rr} &= p_{r^2}(\tilde{\omega} - c) - (\tilde{\omega} - 1)p_{rr} + p^2 c \\ p_{rr} &= p_{r^2} \left(1 - \frac{c}{\tilde{\omega}}\right) + p^2 \frac{c}{\tilde{\omega}} \end{aligned}$$

může do rovnice (117) dosadit za p_{rr} :

$$\begin{aligned} p_{r^2} &= p_{r^2}(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) \\ &\quad + \left(p_{r^2} \left(1 - \frac{c}{\tilde{\omega}}\right) + p^2 \frac{c}{\tilde{\omega}}\right)(-2\tilde{\omega}\omega + 4\mu\omega) \\ &\quad + p_{r^2}(\omega^2) \\ &\quad + 4p^2(\mu - v - \mu^2) \\ &\quad + vl^2 + 2\mu^2l g + vg^2 \end{aligned}$$

Nyní ve shodě s autorem [38] použijeme vztah $g = 2p - l$ vyplývající ze (116), který dosadíme, a pokračujeme ve výpočtu p_{r^2} :

$$\begin{aligned} p_{r^2} &- p_{r^2}(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) + p_{r^2} \left(1 - \frac{c}{\tilde{\omega}}\right)(2\omega(2\mu - \tilde{\omega})) + p_{r^2}(\omega^2) = \\ &+ p^2 \frac{c}{\tilde{\omega}}(2\omega(2\mu - \tilde{\omega})) + 4p^2(\mu - v - \mu^2) + vl^2 + 2\mu^2l g + vg^2 \end{aligned}$$

Vytkneme p_{r^2} :

$$\begin{aligned} p_{r^2} &\left(1 - \left[(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) + \left(1 - \frac{c}{\tilde{\omega}}\right)(2\omega(2\mu - \tilde{\omega})) + \omega^2\right]\right) = \\ &+ p^2 \frac{c}{\tilde{\omega}}(2\omega(2\mu - \tilde{\omega})) + 4p^2(\mu - v - \mu^2) + vl^2 + 2\mu^2l (2p - l) + v(2p - l)^2 \end{aligned}$$

Ve shodě s autorem [38] zavedeme substituci

$$\Delta = 1 - \left[(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) + \left(1 - \frac{c}{\tilde{\omega}}\right) (2\omega(2\mu - \tilde{\omega})) + \omega^2 \right] \quad (118)$$

a výraz pro stacionární bod p_{r^2} odvodíme poněkud podrobněji:

$$p_{r^2} \Delta =$$

$$p^2 \left(\frac{c}{\tilde{\omega}} (2\omega(2\mu - \tilde{\omega})) + 4(\mu - v - \mu^2) \right) + vl^2 + 4\mu^2 lp - 2\mu^2 l^2 + v4p^2 - v4pl + vl^2$$

Rovnici podělíme výrazem Δ :

$$p_{r^2} = p^2 \frac{\frac{c}{\tilde{\omega}} (2\omega(2\mu - \tilde{\omega})) + 4(\mu - v - \mu^2) + 4v}{\Delta} + p \frac{4l(\mu^2 - v)}{\Delta} + \frac{2l^2(v - \mu^2)}{\Delta}$$

Lze ukázat, že výraz

$$\frac{c}{\tilde{\omega}} (2\omega(2\mu - \tilde{\omega})) + 4(\mu - v - \mu^2) + 4v = \frac{c}{\tilde{\omega}} (2\omega(2\mu - \tilde{\omega})) + 4(\mu - \mu^2)$$

vyskytující se v prvním čitateli ve výrazu pro p_{r^2} je roven výrazu

$$\Delta - 2(\mu^2 - v)$$

a to nejlépe zpětným dosazením za Δ , vyčíslením, dosazením $\tilde{\omega} = \omega + 1$ a porovnáním obou výrazů:

$$1 - \left[(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) + \left(1 - \frac{c}{\tilde{\omega}}\right) (2\omega(2\mu - \tilde{\omega})) + \omega^2 \right] - 2(\mu^2 - v) =$$

$$1 - \tilde{\omega}^2 + 4\tilde{\omega}\mu - 2v - 2\mu^2 - 2\omega(2\mu - \tilde{\omega}) + \frac{c}{\tilde{\omega}} 2\omega(2\mu - \tilde{\omega}) - \omega^2 - 2\mu^2 + 2v =$$

$$1 - (\omega + 1)^2 + 4(\omega + 1)\mu - 2v - 2\mu^2 - 4\omega\mu + 2\omega(\omega + 1) + \frac{c}{\tilde{\omega}} 2\omega(2\mu - \tilde{\omega}) - \omega^2$$

$$-2\mu^2 + 2v =$$

$$1 - \omega^2 - 2\omega - 1 + 4\omega\mu + 4\mu - 2v - 2\mu^2 - 4\omega\mu + 2\omega^2 + 2\omega + \frac{c}{\tilde{\omega}} 2\omega(2\mu - \tilde{\omega}) - \omega^2$$

$$-2\mu^2 + 2v =$$

$$\frac{c}{\tilde{\omega}} 2\omega(2\mu - \tilde{\omega}) + 4(\mu - \mu^2)$$

což je hledaný výraz. Stacionární bod p_{r^2} je tedy dán výrazem

$$p_{r^2} = p^2 \frac{\Delta - 2(\mu^2 - v)}{\Delta} + p \frac{4l(\mu^2 - v)}{\Delta} + \frac{2l^2(v - \mu^2)}{\Delta}$$

Autor [38] následně zjišťoval, čemu se rovná rozdíl $p_{r^2} - p^2$:

$$\begin{aligned}
p_{r^2} &= p^2 \frac{\Delta}{\Delta} + p^2 \frac{-2(\mu^2 - v)}{\Delta} + p \frac{4l(\mu^2 - v)}{\Delta} - \frac{2l^2(\mu^2 - v)}{\Delta} \\
p_{r^2} - p^2 &= p^2 \frac{2(v - \mu^2)}{\Delta} - 2pl \frac{2(v - \mu^2)}{\Delta} + l^2 \frac{2(v - \mu^2)}{\Delta} \\
p_{r^2} - p^2 &= \frac{2(v - \mu^2)}{\Delta} (p^2 - 2pl + l^2) \\
\sqrt{p_{r^2} - p^2} &= \sqrt{2 \frac{(v - \mu^2)}{\Delta}} |p - l|
\end{aligned}$$

Použijeme vztah vyplývající ze (116), který dosadíme:

$$\begin{aligned}
g &= 2p - l \quad p - l = \frac{g - l}{2} \\
\sqrt{p_{r^2} - p^2} &= \sqrt{2 \frac{(v - \mu^2)}{\Delta}} |g - l| \tag{119}
\end{aligned}$$

V rovnici (119) je výraz na levé straně roven standardní odchylce pozic trajektorie systému částice r_t od jejich očekávaných hodnot:

$$\sqrt{p_{r^2} - p^2} = \sqrt{E[r_t^2] - (E[r_t])^2} = \sigma(r_t)$$

Výraz $v - \mu^2$ pod odmocninou na pravé straně rovnice (119) je roven standardní odchylce náhodné proměnné $\varphi_1 \approx \varphi_2$ vzhledem k předpokladu rovnosti koeficientů a_l, a_g :

$$\mu \approx E[\varphi_1] = a_l/2 = c/2 \approx E[\varphi_2] = a_g/2 = c/2 \tag{120}$$

$$v \approx E[\varphi_1^2] = E[\varphi_2^2] \approx \int_0^c x^2 f(x) dx = \frac{1}{c} \int_0^c x^2 dx = \frac{c^2}{3} \tag{121}$$

D13 Ověření podmínky stability 2. řádu systému částice pro dvojici koeficientů $[c, \omega] \cong [1, 48, 0, 72]$

Autor [38] vycházel při ověření podmínky stability 2. řádu ze vztahu pro standardní odchylku pozic trajektorií systému částice od jejich očekávaných hodnot (119). Mezní hodnota, při které poroste odchylka neomezeně, je pro $\Delta = 0$ (118). Autor tedy vyjádřil příslušný vztah mezi c a ω a sestavil křivku, která představuje horní hranici (*bare bone*), pod níž musí ležet dvojice $[c, \omega]$, aby systém částice popsáný rovnicemi (1) a (2) s danou dvojicí koeficientů byl stabilní.

Postup autora [38] bude nyní podán podrobněji. Nejdříve bude autorkou předkládané práce zjednodušen výraz pro Δ (118), který je zde pro přehlednost opakovaně uveden:

$$\Delta = 1 - \left[(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) + \left(1 - \frac{c}{\tilde{\omega}}\right) (2\omega(2\mu - \tilde{\omega})) + \omega^2 \right]$$

Do výrazu pro Δ dosadíme zpětně za $\tilde{\omega} = \omega + 1$ (107a) a odstraníme část závorek:

$$1 - (\omega + 1)^2 + 4(\omega + 1)\mu - 2v - 2\mu^2 - 4\omega\mu + 2\omega(\omega + 1) + \frac{c}{\omega + 1} 2\omega(2\mu - (\omega + 1)) - \omega^2$$

$$1 - \omega^2 - 2\omega - 1 + 4\omega\mu + 4\mu - 2v - 2\mu^2 - 4\omega\mu + 2\omega^2 + 2\omega + \frac{c}{\omega + 1} 2\omega(2\mu - (\omega + 1)) - \omega^2$$

$$4\mu - 2v - 2\mu^2 + \frac{c}{\omega + 1} 2\omega(2\mu - (\omega + 1))$$

Upravovaný výraz uvedeme na společný jmenovatel:

$$\frac{(4\mu - 2v - 2\mu^2)(\omega + 1) + c2\omega2\mu - c2\omega\omega - c2\omega}{\omega + 1}$$

$$\frac{4\mu\omega - 2v\omega - 2\mu^2\omega + 4\mu - 2v - 2\mu^2 + 4c\omega\mu - 2c\omega^2 - 2c\omega}{\omega + 1}$$

Do výrazu pro Δ dosadíme zpětně μ (120) a v (121) a upravíme jmenovatel:

$$\frac{4\frac{c}{2}\omega - 2\frac{c^2}{3}\omega - 2\left(\frac{c}{2}\right)^2\omega + 4\frac{c}{2} - 2\frac{c^2}{3} - 2\left(\frac{c}{2}\right)^2 + 4c\omega\frac{c}{2} - 2c\omega^2 - 2c\omega}{\omega + 1}$$

$$\frac{2c\omega - 2\frac{c^2}{3}\omega - \frac{c^2}{2}\omega + 2c - 2\frac{c^2}{3} - \frac{c^2}{2} + 2c^2\omega - 2c\omega^2 - 2c\omega}{\omega + 1}$$

$$\frac{12c\omega - 4c^2\omega - 3c^2\omega + 12c - 4c^2 - 3c^2 + 12c^2\omega - 12c\omega^2 - 12c\omega}{6(\omega + 1)}$$

$$\frac{c(12\omega - 4c\omega - 3c\omega + 12 - 4c - 3c + 12c\omega - 12\omega^2 - 12\omega)}{6(\omega + 1)}$$

$$\frac{c(12\omega + 12 - 12\omega^2 - 12\omega + c(-4\omega - 3\omega - 4 - 3 + 12\omega))}{6(\omega + 1)}$$

Upravený výraz pro Δ je následující:

$$\Delta = c * \frac{12(1 - \omega^2) + c(5\omega - 7)}{6(\omega + 1)} \quad (122)$$

Mezní hodnota, při které poroste odchylka neomezeně, je pro $\Delta = 0$:

$$12(1 - \omega^2) + c(5\omega - 7) = 0$$

Vztah mezi c a ω udává křivku, která představuje horní hranici (*bare bone*), pod níž musí ležet dvojice $[c, \omega]$, aby systém částice popsany rovnicemi (1) a (2) s danou dvojicí koeficientů byl stabilní.

$$c = -\frac{12(1 - \omega^2)}{5\omega - 7} = \frac{12(\omega^2 - 1)}{5\omega - 7} \quad (123)$$

Mezní křivka je vykreslena na obrázku Obr. 9 v kapitole 5.5.3. Dosazením do (123) lze ověřit, že bod $[c, \omega] = [1,49, 0,7213]$ leží uvnitř oblasti stability 2. řádu:

$$\omega = 0,7213$$

$$c = \frac{12(0,7213^2 - 1)}{5 * 0,7213 - 7} \cong 1,69 > 1,49$$

Vztah (119) lze upravit dosazením (120), (121) a (122):

$$\begin{aligned} \sigma(r_t) &= \sqrt{p_{r^2} - p^2} = \frac{1}{2} \sqrt{2 \frac{c^2/3 - (c/2)^2}{\Delta}} |g - l| \\ \sigma(r_t) &= \frac{1}{2} \sqrt{2 \frac{(c^2/3 - (c/2)^2) * 6(\omega + 1)}{c(12(1 - \omega^2) + c(5\omega - 7))}} |g - l| = \\ &= \frac{1}{2} \sqrt{\frac{c(\omega + 1)}{12 - 12\omega^2 + c(5\omega - 7)}} |g - l| \end{aligned} \quad (124)$$

Dosazením do (124) lze ověřit, že pro bod $[c, \omega] = [1,47876, 0,7213]$ platí $\Delta \neq 0$:

$$\begin{aligned} \sigma(r_t) &\cong \frac{1}{2} \sqrt{\frac{1,47876(0,7213 + 1)}{12 - 12 * (0,7213)^2 + 1,4786(5 * 0,7213 - 7)}} |g - l| \\ \sigma(r_t) &\cong 1,034 * |g - l| \end{aligned}$$

D14 Vymezení oblasti stability momentů prvního řádu

Při vymezení oblasti stability pro momenty prvního řádu vycházíme z charakteristické rovnice po dosazení za $1 - (1 - w)\Delta t = \omega_{\Delta t}$ podle (132):

$$\lambda^2 + (\varphi_n^i \Delta^2 t - (1 - (1 - w)\Delta t) - 1)\lambda + 1 - (1 - w)\Delta t = 0$$

$$\lambda^2 + (\varphi_n^i \Delta^2 t - \omega_{\Delta t} - 1)\lambda + \omega_{\Delta t} = 0$$

a z podmínek stability (71), (73) a (75). Rozhraní mezi reálnými a komplexními kořeny udává limitní parabola (71):

$$(\varphi_n^i \Delta^2 t - \omega_{\Delta t} - 1)^2 = 4\omega_{\Delta t}$$

Spočteme hodnoty $\varphi_n^i \Delta^2 t$ pro $(\omega_{\Delta t} = 1)$:

$$(\varphi_n^i \Delta^2 t - 2)^2 = 4$$

odtud:

$$\varphi_{n_{min}}^i = 0 \quad \varphi_{n_{max}}^i = \frac{4}{\Delta^2 t}$$

Po dosazení za $1 - (1 - w)\Delta t = \omega_{\Delta t}$ podle (132) spočteme závislost φ_n^i , w podle (73):

$$\varphi_n^i \Delta^2 t < 2(\omega_{\Delta t} + 1)$$

$$\varphi_n^i \Delta^2 t < 2(1 - (1 - w)\Delta t + 1)$$

$$\varphi_n^i < \frac{2\Delta t w - 2\Delta t + 4}{\Delta^2 t}$$

odtud pro $\varphi_{n_{min}}^i, \varphi_{n_{max}}^i$:

$$w_{min} = 1 - \frac{2}{\Delta t} \quad w_{max} = 1$$

Oblast stability momentů prvního řádu odpovídá deterministické oblasti stability:

$$S \approx S_D = \left\{ (w, E(\varphi_n^i)) : 1 - \frac{2}{\Delta t} < w < 1, 0 < E(\varphi_n^i) < \frac{1}{\Delta^2 t} (2\Delta t w - 2\Delta t + 4) \right\}$$

D15 Odvození mezní křivky pro dvojice $[a, \omega]$ a parametr α

Při odvození vychází autorka ze vztahu pro standardní odchylku pozic trajektorií systému částice od jejich očekávaných hodnot (119). Mezní hodnota, při které poroste odchylka neomezeně, je pro $\Delta = 0$ (D12:118):

$$\Delta = 1 - \left[(\tilde{\omega}^2 - 4\tilde{\omega}\mu + 2v + 2\mu^2) + \left(1 - \frac{c}{\tilde{\omega}}\right) (2\omega(2\mu - \tilde{\omega})) + \omega^2 \right]$$

Autorka tedy po vzoru autora [38] vyjádří vztah mezi a a ω představující horní hranici (bare bone), pod níž musí ležet dvojice $[a, \omega]$, aby systém částice popsany rovnicemi (1) a (2) byl stabilní.

Autorka nejdříve dosadí $\tilde{\omega} = \omega + 1$ (107a) a zjednoduší pravou stranu rovnice:

$$4\mu - 2v - 2\mu^2 + \frac{c}{\omega + 1} 2\omega(2\mu - (\omega + 1))$$

Jelikož již neplatí rovnost koeficientů $a_l = a_g = c$, bude provedena zpětná substituce k substituci (112):

$$4\mu \approx 2(E(\varphi_1) + E(\varphi_2)) = 2E(\varphi)$$

$$2v \approx E(\varphi_1^2) + E(\varphi_2^2)$$

$$2\mu^2 \approx 2 * E(\varphi_1)E(\varphi_2) = 2E(\varphi_1\varphi_2)$$

$$c = 2\frac{c}{2} \approx \frac{a_l}{2} + \frac{a_g}{2} = E(\varphi)$$

Po dosazení a dalších úpravách pravé strany rovnice, s využitím vztahů (149) a (150):

$$\begin{aligned}
& 2E(\varphi) - (E(\varphi_1^2) + E(\varphi_2^2) + 2E(\varphi_1\varphi_2)) + \frac{E(\varphi)}{\omega + 1} 2\omega(E(\varphi) - (\omega + 1)) \\
& 2E(\varphi) - E(\varphi^2) + \frac{E(\varphi)}{\omega + 1} 2\omega(E(\varphi) - (\omega + 1)) \\
& \frac{2E(\varphi)\omega + 2E(\varphi) - E(\varphi^2)\omega - E(\varphi^2) + 2(E(\varphi))^2\omega - 2E(\varphi)\omega^2 - 2E(\varphi)\omega}{\omega + 1} \\
& \frac{2E(\varphi)(\omega + 1 - \omega^2 - \omega) - E(\varphi^2)(\omega + 1) + 2(E(\varphi))^2\omega}{\omega + 1} \\
& \frac{2E(\varphi)(1 - \omega^2) - \left(\frac{a_l^2}{3} + \frac{a_g^2}{3} + \frac{a_l a_g}{2}\right)(\omega + 1) + 2(E(\varphi))^2\omega}{\omega + 1} \\
& \frac{12E(\varphi)(1 - \omega^2) - (2a_l^2 + 2a_g^2 + 3a_l a_g)(\omega + 1) + 12(E(\varphi))^2\omega}{6(\omega + 1)} \\
& \frac{12(1 - \omega^2) - E(\varphi) \left(\frac{(2a_l^2 + 2a_g^2 + 3a_l a_g)}{(E(\varphi))^2} (\omega + 1) - 12\omega \right)}{6(\omega + 1)E(\varphi)} \\
& \frac{12(1 - \omega^2) - E(\varphi) \left(\frac{2(a_l + a_g)^2 - a_l a_g}{(E(\varphi))^2} (\omega + 1) - 12\omega \right)}{6(\omega + 1)E(\varphi)} \\
& \frac{12(1 - \omega^2) - E(\varphi) \left(8(\omega + 1) + \frac{-a_l a_g}{(E(\varphi))^2} (\omega + 1) - 12\omega \right)}{6(\omega + 1)E(\varphi)}
\end{aligned}$$

Nyní autorka zpětným postupem dovodí, že označený výraz je roven výrazu $\alpha^2 - 2\alpha$, kde $\alpha = a_g/E(\varphi)$ podle (127a):

$$\alpha^2 - 2\alpha = \frac{a_g^2}{(E(\varphi))^2} - \frac{2a_g}{E(\varphi)} = \frac{a_g^2 - 2a_g E(\varphi)}{(E(\varphi))^2} = \frac{a_g^2 - a_g(a_l + a_g)}{(E(\varphi))^2} = \frac{-a_l a_g}{(E(\varphi))^2}$$

Upravovaný výraz je roven nule, je-li:

$$12(1 - \omega^2) - E(\varphi)(8(\omega + 1) + (\alpha^2 - 2\alpha)(\omega + 1) - 12\omega) = 0$$

a odtud:

$$E[\varphi] = \alpha = \frac{12(1 - \omega^2)}{4(2 - \omega) + (\alpha^2 - 2\alpha)(1 + \omega)}$$

což je hledaný vztah (151).

D16 Distribuční funkce a kritická hodnota pro *Dixonův* test odlehlých hodnot

Bude uveden příklad výpočtu hustoty pravděpodobnosti, distribuční funkce a kritické hodnoty pro zvolenou pravděpodobnost α_z pro testovací kritérium r_{10} .

Necht' množina náhodných proměnných $Y = \{Y^1, Y^2, \dots, Y^{m-1}, Y^m\}$ má konkrétní realizaci $\{y^1, y^2, \dots, y^{m-1}, y^m\}$. Jednotlivé hodnoty tvořící statistický vzorek lze uspořádat vzestupně takto:

$$y_1 < y_2 < \dots < y_{m-1} < y_m$$

kde $y_{(k)}$ je k -tá nejmenší hodnota statistického vzorku, zvaná k -tá pořadová statistika (*order statistic*). Existuje celkem $m!$ (faktoriál) statistických vzorků, jimž odpovídá totéž uspořádání. Předpokládejme, že proměnné Y^i mají spojitě náhodné rozložení s distribuční funkcí $F(Y)$, zavedme substituci $u = F(Y)$, kde U je náhodná proměnná rovnoměrného rozložení, u_k jsou realizace proměnné U . Pravděpodobnost, že

$$0 < u_1 < u_2 < \dots < u_{m-1} < u_m < 1$$

je rovna $(1/m!)$. K tomu, aby pořadová statistika u_k byla v intervalu $(u, u + du)$ je třeba, aby přesně $k - 1$ realizací bylo v intervalu $(0, u)$, jedna realizace je v intervalu $(u, u + du)$ a $m - k$ realizací je v intervalu $(u + du, 1)$. Počet způsobů volby takových realizací je:

$$\frac{m!}{(k-1)! 1! (m-k)!}$$

Pravděpodobnost, že realizace jsou ve zvolených intervalech, je:

$$(u-0)^{k-1} * (u+du-u) * (1-(u+du))^{m-k}$$

Hustota pravděpodobnosti pořadové statistiky u_k je:

$$\frac{m!}{(k-1)! (m-k)!} * u^{k-1} * du * (1-(u+du))^{m-k}$$

Je-li $F(Y)$ absolutně spojitá a $u = F(Y)$, potom $dF(Y) = f_Y(y)dy$, $du = f_Y(y)dy$, a tedy hustota pravděpodobnosti je

$$f(y_k) = \frac{m!}{(k-1)! (m-k)!} * (F(Y))^{k-1} * f_Y(y_k)dy * (1-(F(Y)))^{m-k}$$

Sdružená hustota pravděpodobnosti pro pořadové statistiky y_1, y_{m-1}, y_m je

$$f(\dots) = \frac{m!}{(m-3)!} * f_Y(y_1)dy_1 * \left(\int_{y_1}^{y_{m-1}} f_Y(z)dz \right)^{m-3} * f_Y(y_{m-1})dy_{m-1} * f_Y(y_m)dy_m$$

Pro určení hustoty pravděpodobnosti pro testovací kritérium

$$r_{10} = \frac{y_m - y_{m-1}}{y_m - y_1}$$

zavedeme substituci:

$$\begin{aligned}v &= y_m - y_1 \\r_{10}v &= y_m - y_{m-1} \\y &= y_m\end{aligned}$$

a tedy

$$\begin{aligned}y_1 &= y_m - v \\y_{m-1} &= y_m - r_{10}v\end{aligned}$$

a provedeme transformaci. Jakobián transformace je v .

Nyní integrujeme (transformovanou) sdruženou hustotu pravděpodobnosti pro pořadové statistiky y_1, y_{m-1}, y_m dle y přes $(-\infty, +\infty)$ a dle v přes $(0, +\infty)$, abychom získali hustotu pravděpodobnosti pro testovací kritérium r_{10} :

$$f(r) = \frac{m!}{(m-3)!} \int_{-\infty}^{\infty} \int_0^{\infty} \left(\int_{y-v}^{y-r_{10}v} f_Y(z) dz \right)^{m-3} f_Y(y-v) f_Y(y-r_{10}v) f_Y(y) v \, dv \, dx$$

Distribuční funkce $G(R)$ a kritická hodnota pro R je

$$G(R) = \int_0^R f(r) dr$$

Rozsah r je $0 \leq r \leq 1$, a tedy $G(0) = 0$ a $G(1) = 1$. Kritické hodnoty R jsou kořeny rovnice

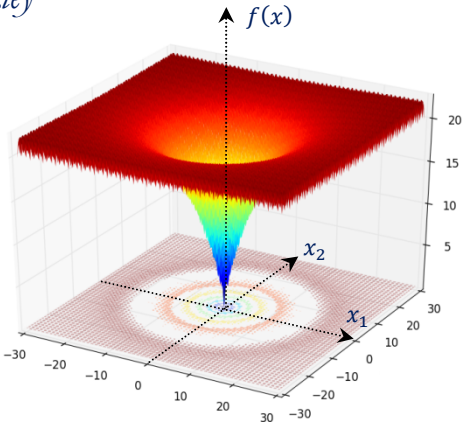
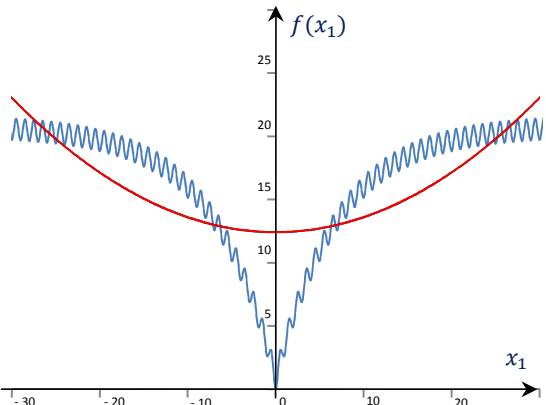
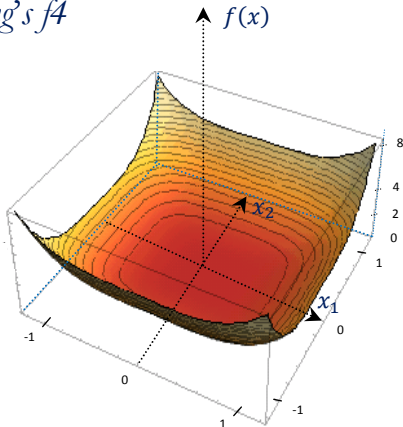
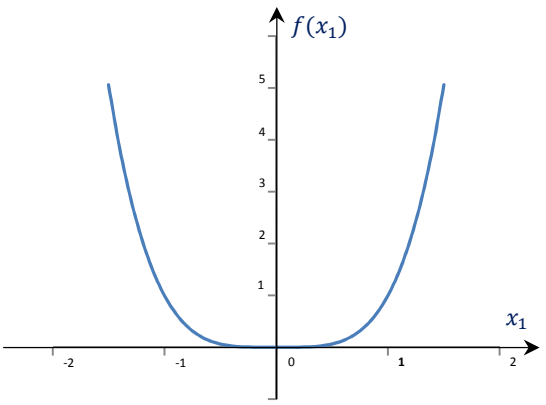
$$(1 - \alpha_z) - G(R) = 0$$

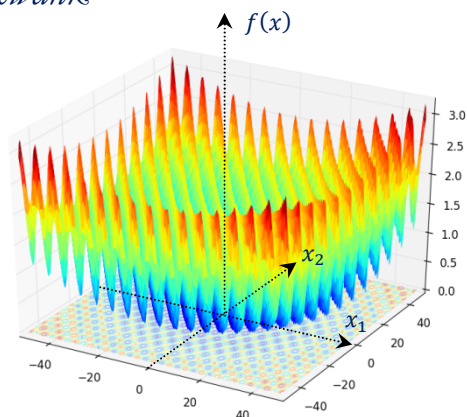
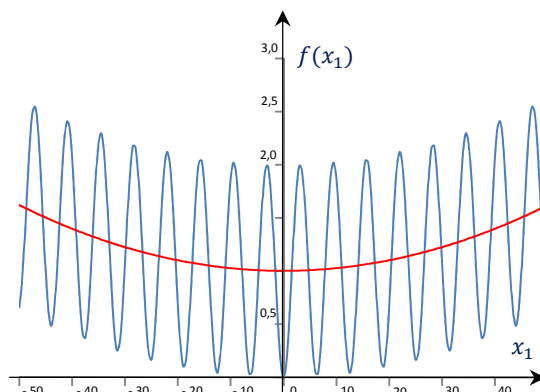
kde α_z je (zvolená) pravděpodobnost, pro niž počítáme kritickou hodnotu R .

$G(R)$ vzrůstá monotónně od 0 do 1, a tedy existuje pouze jediná kritická hodnota pro každou pravděpodobnost α_z .

Příloha A

Příloha A obsahuje vzorce a grafy účelových funkcí *Ackley*, *De Jong's f4*, *Griewank*, *Rastrigin*, *Rosenbrock*, *Schaffer's f6*, *Schwefel 26* a *Sphere*, seřazené v abecedním pořadí.

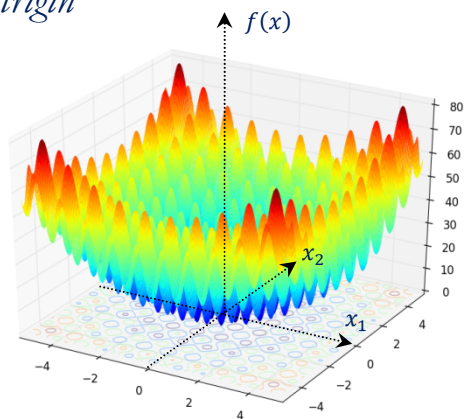
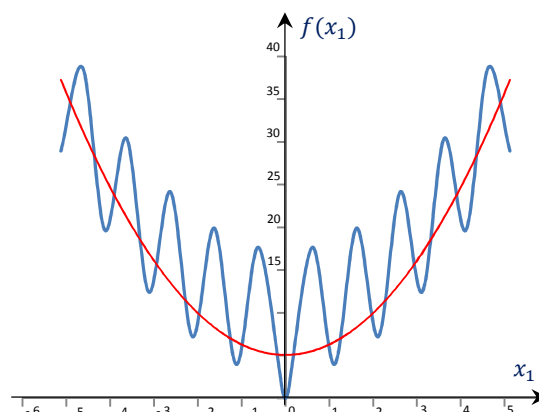
<i>Ackley</i>  Graf dvou-dimenzionálního problému $x = (x_1, x_2)$	 Jedno-dimenzionální profil a spojnice trendu (<i>fitness landscape</i>) pro $n = 2$, $x_2 = 0$
$f(x) = -20 \exp \left(-0,2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right) + 20 + e$	$-32 \leq x_i \leq 32$ $f_{\min}(x) \approx f(0, \dots, 0) = 0$
<i>De Jong's f4</i>  Graf dvou-dimenzionálního problému $x = (x_1, x_2)$	 Jedno-dimenzionální profil (<i>fitness landscape</i>) pro $n = 1$
$f(x) = \sum_{i=1}^n i x_i^4$	$-1,28 \leq x_i \leq 1,28$ $f_{\min}(x) \approx f(0, \dots, 0) = 0$

GriewankGraf dvou-dimenzionálního problému $x = (x_1, x_2)$ Jedno-dimenzionální profil a spojnice trendu
(*fitness landscape*) pro $n = 1$

$$f(x) = 1 + \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right)$$

$$-512 \leq x_i \leq 512$$

$$f_{\min}(x) \approx f(0, \dots, 0) = 0$$

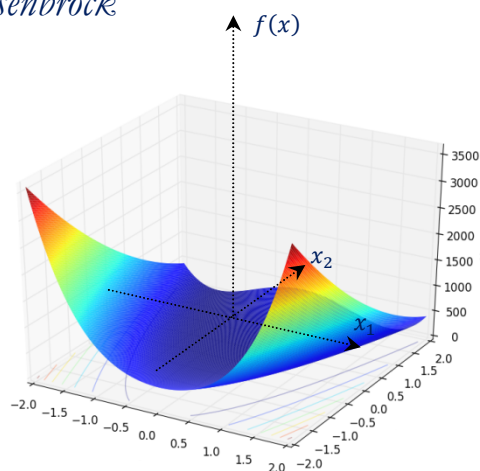
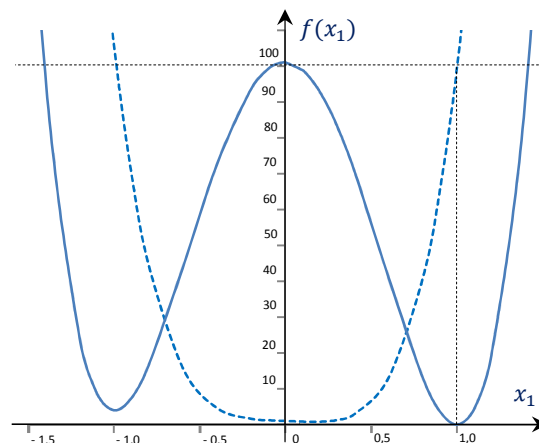
RastriginGraf dvou-dimenzionálního problému $x = (x_1, x_2)$ Jedno-dimenzionální profil a spojnice trendu
(*fitness landscape*) pro $n = 2, x_2 = 0$

$$f(x) = 10n + \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i)]$$

$$-5,12 \leq x_i \leq 5,12$$

$$f_{\min}(x) \approx f(0, \dots, 0) = 0$$

Rosenbrock

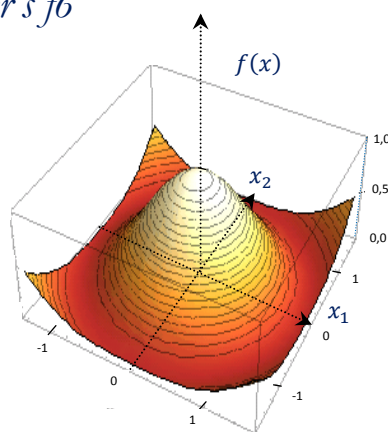
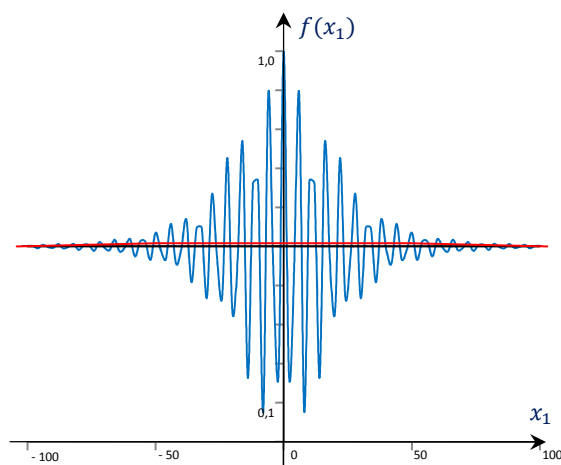
Graf dvou-dimenzionálního problému $x = (x_1, x_2)$ Jedno-dimenzionální profil (*fitness landscape*)
pro $n = 2$, $x_2 = 1$ (plná čára), $x_2 = 0$ (přerušované)

$$f(x) = \sum_{i=1}^{n-1} [100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2]$$

$$-10 \leq x_i \leq 10$$

$$f_{\min}(x) \approx f(1, \dots, 1) = 0$$

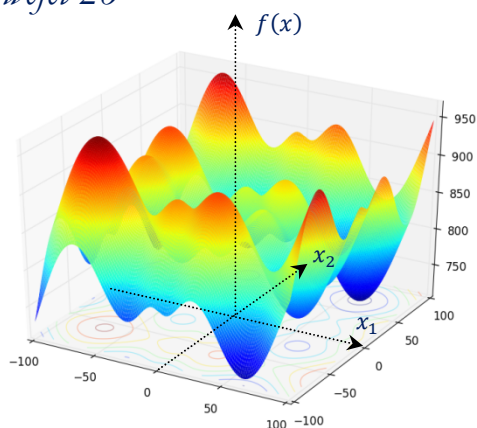
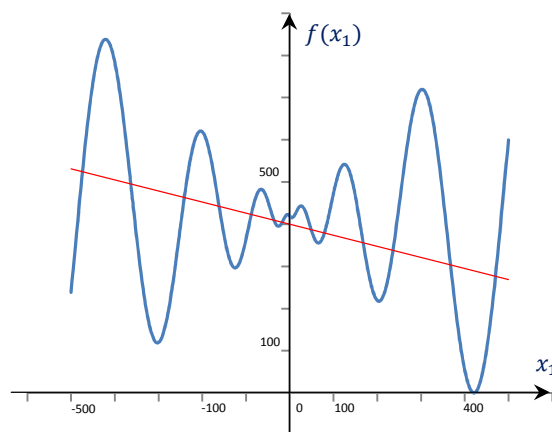
Schaffer's f6

Graf dvou-dimenzionálního problému $x = (x_1, x_2)$
Detail v okolí bodu $x = (0,0)$ Jedno-dimenzionální profil a spojnice trendu
(*fitness landscape*) pro $n = 2$, $x_2 = 0$

$$f(x) = 0,5 - \frac{(\sin \sqrt{x_1^2 + x_2^2})^2 - 0,5}{[1 + 0,001(x_1^2 + x_2^2)]^2}$$

$$-100 \leq x_i \leq 100$$

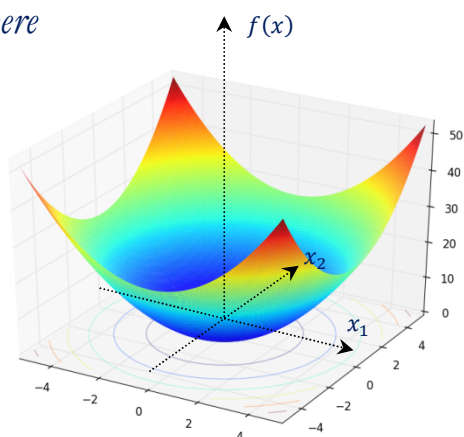
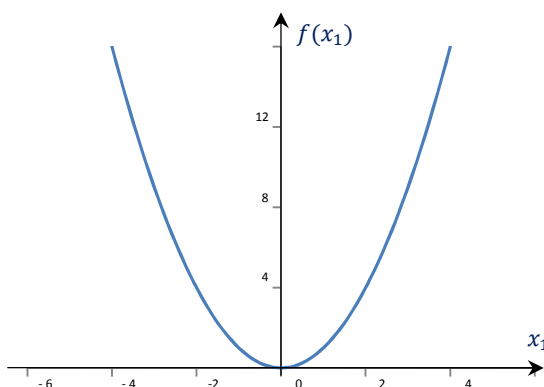
$$f_{\max}(x) \approx f(0, \dots, 0) = 1$$

Schwefel 26Graf dvou-dimenzionálního problému $x = (x_1, x_2)$ Jedno-dimenzionální profil a spojnice trendu
(*fitness landscape*) pro $n = 1$

$$f(x) = 418,982887n - \sum_{i=1}^n x_i \sin \sqrt{|x_i|}$$

$$-512 \leq x_i \leq 512$$

$$f_{\min}(x) \approx f(420.968746, \dots, 420.968746) = 0$$

SphereGraf dvou-dimenzionálního problému $x = (x_1, x_2)$ Jedno-dimenzionální profil
(*fitness landscape*) $n = 1$

$$f(x) = \sum_{i=1}^n x_i^2$$

$$-1 \leq x_i \leq 1$$

$$f_{\min}(x) \approx f(0, \dots, 0) = 0$$

Příloha B

Příloha B obsahuje popis tříd objektů, jež tvoří základ navrhovaného, rozšířeného algoritmu *PSO* vytvořeného v jazyce *Java* v prostředí *NetBeans IDE 6.9.1*. Jsou uvedeny třídy definující datové typy a třídy implementující metody výpočtů pravděpodobnosti rovnoměrného rozmístění pozic, pravděpodobnosti výlučnosti hodnoty fitness a hodnoty účelové funkce v dané pozici. Podrobná dokumentace všech zdrojových souborů (*source packages*) je k dispozici v elektronické podobě (ve formátu *HTML*) na přiloženém CD.

Datové typy a objekty

Základními datovými objekty v navrhovaném algoritmu *PSO* jsou částice, roj, množina lokálních nejlepších *fitness* a globální nejlepší *fitness*, pořadové statistiky a histogramy.

Částice je definována jako pole pozic a pole rychlostí, plus hodnota účelové funkce (*fitness*) spočtená v odpovídající pozici hyperprostoru:

```
public class Particle {  
    private Position position;  
    private Velocity velocity;  
    private double fitness;  
    :  
}  
public class Position {  
    private List<Double> x;  
    :  
}  
public class Velocity {  
    private List<Double> v;  
    :  
}
```

Pozice a rychlost částice jsou definovány jako vektory, jejichž složky jsou reálná čísla:

```
List<Double> x = new ArrayList<Double>(PARTICLE_SIZE);  
List<Double> v = new ArrayList<Double>(PARTICLE_SIZE);
```

Roj je nadefinován jako (uspořádaný) seznam částic:

```
List<Particle> swarm = new ArrayList<Particle>(SWARM_SIZE);
```

Všechna pole vázaná na velikost roje jsou definována jako *ArrayList*, včetně histogramů. Pokud by se velikost roje měnila tak, že by částice vznikaly a zanikaly, jednou stanovené intervaly pro histogram už by se neměnily. Adaptivní změna rozměru částice *particle.size()* se nepředpokládá. Předpokládá, že *swarm.size() = SWARM_SIZE*. Pokud by se počet částic dynamicky měnil, pak by tuto situaci bylo nutno ošetřit.

Lokální nejlepší pozice a odpovídající *fitness*:

```
List<Position> particlesBestPositions = new ArrayList<Position>(SWARM_SIZE);
```

```
List<Double> particlesBestFitness = new ArrayList<Double>(SWARM_SIZE);
```

Globální nejlepší pozice a odpovídající *fitness*:

```
Position globalBestPosition;
```

```
double globalBestFitness;
```

Z lokálních nejlepších *fitness* se stanovují pořadové statistiky:

```
List<Double> orderStatistics = new ArrayList<Double>(SWARM_SIZE);
```

Histogram pro stanovení rovnoměrnosti rozmístění částic je nadefinován pro každý rozměr částice. Počet intervalů v histogramu je (minimálně) roven *SWARM_SIZE*. Intervaly jsou definovány v histogramu jako pole pozorovaných (*observed*) hodnot. Očekávaná (*expected*) hodnota v každém intervalu je aktualizována v pomocné proměnné, která je pro každý interval a pro každý rozměr částice v dané iteraci táž, a v každé iteraci je hodnota této proměnné inkrementována o 1.

```
public class Histogram {
    private Observed observed;
    private double chiSquare;
    private double probabilityOfUniform;
    :
}
```

```
public class Observed {
    private List<Long> o;
    :
}
```

```
List<Long> o = new ArrayList<Long>(SWARM_SIZE);
```

```
long expected;
```

```
List<Histogram> histograms = new ArrayList<Histogram>(PARTICLE_SIZE);
```

Koeficienty ω , a_l , a_g (1) a a , α (151) a režim prohledávání *mood* (170) jsou proměnné a počítají se pro každý rozměr částice:

```
double [] c1 = new double [PARTICLE_SIZE]; //koeficient  $a_l$ 
double [] c2 = new double [PARTICLE_SIZE]; //koeficient  $a_g$ 
double [] c = new double [PARTICLE_SIZE]; //střední hodnota  $a$ 
double [] w = new double [PARTICLE_SIZE]; //koeficient  $\omega$ 
double [] alpha = new double [PARTICLE_SIZE]; //  $\alpha = 2a_g/(a_l+a_g)$  (127a)
int [] modus = new int [PARTICLE_SIZE]; // režim prohledávání (0, 1, 2, 3) mood (170)
```

Metody

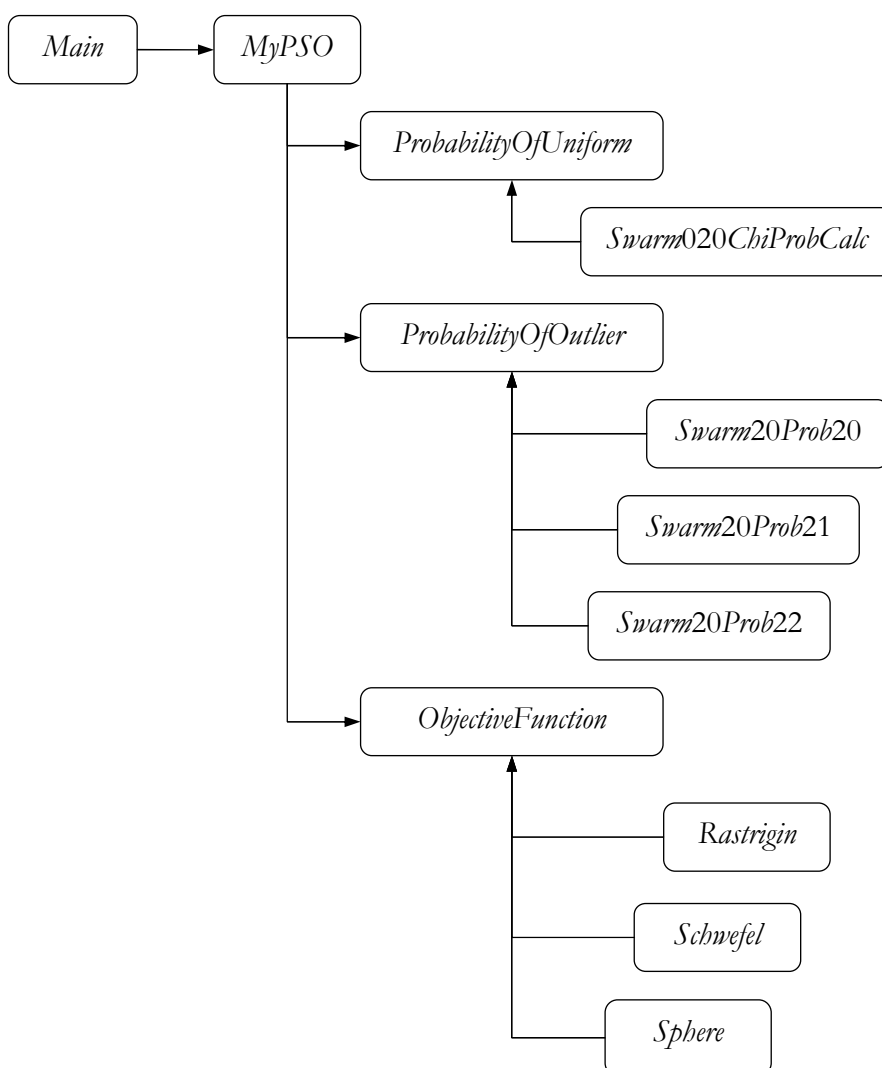
Vazby mezi třídami *Main*, *MyPso* a třídami metod výpočtů pravděpodobnosti rovnoměrného rozmístění pozic (*ProbabilityOfUniform*), pravděpodobnosti výlučnosti hodnoty *fitness* (*ProbabilityOfOutlier*) a hodnoty účelové funkce (*ObjectiveFunction*) v dané pozici vyhledávacího prostoru znázorňuje diagram na obrázku Obr. 28. Pro definice tříd je použito způsobu přetypování směrem nahoru, aby při změně testovacího kritéria nebo při změně účelové funkce nebylo nutno provádět zásah do *MyPso*.

Příklad

Objekt typu *Schwefel* je též objektem typu *ObjectiveFunction* a účelovou funkci lze zadat jako parametr přes *Main* bez nutnosti zásahu do *MyPso*. Podobně lze zadat testovací kritérium r_{22} (161). Algoritmus *PSO* je realizován jako objekt *myPso* typu *MyPso*:

```
public class Main {
    public static void main(String[] args) {
        ObjectiveFunction fitnessFunction = new Schwefel();
        ProbabilityOfOutlier outlierProbability = new Swarm20Prob22();
        MyPso myPso = new MyPso(fitnessFunction, outlierProbability);
        myPso.execute();
    }
}
```

(Konec příkladu)



Obr. 28: Vazby mezi třídami objektů navrhovaného algoritmu *PSO*.