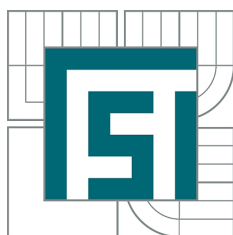


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV MATEMATIKY
FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF MATHEMATICS

KOMPRIMOVANÉ SNÍMÁNÍ V PERFUZNÍM ZOBRAZOVÁNÍ POMOCÍ MAGNETICKÉ REZONANCE

COMPRESSED SENSING IN MAGNETIC RESONANCE PERFUSION IMAGING.

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. MARIE DAŇKOVÁ

VEDOUCÍ PRÁCE
SUPERVISOR

Mgr. PAVEL RAJMIC, Ph.D.

BRNO 2014

Vysoké učení technické v Brně, Fakulta strojního inženýrství

Ústav matematiky

Akademický rok: 2013/2014

ZADÁNÍ DIPLOMOVÉ PRÁCE

student(ka): Bc. Marie Daňková

který/která studuje v **magisterském navazujícím studijním programu**

obor: **Matematické inženýrství (3901T021)**

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma diplomové práce:

Komprimované snímání v perfuzním zobrazování pomocí magnetické rezonance

v anglickém jazyce:

Compressed sensing in magnetic resonance perfusion imaging.

Stručná charakteristika problematiky úkolu:

Poslední léta jsou bohatá na nové přístupy ke zpracování vícedimenzionálních dat. Apriorní informace o signálu (např. jeho řídkost či nízkohodnostní struktura) umožňují vysokou kompresi, snadnější interpretaci či efektivnější zpracování.

Smyslem práce by bylo nastudovat tyto aktuální algoritmy a aplikovat je v tzv. perfuzní analýze magnetickou rezonancí, a to ve smyslu zrychlení snímacího procesu za pomoci principu tzv. komprimovaného snímání.

Cíle diplomové práce:

Nastudovat principy a algoritmy hledání tzv. řídkých reprezentací a komprimovaného snímání. Seznámit se s numerickými metodami pro řešení těchto úloh.

Aplikovat vybranou metodu v simulaci komprimovaného snímání perfuzního MRI signálu na simulovaných i reálných datech. Shrnout a interpretovat výsledky.

Seznam odborné literatury:

ELAD, M., Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer, 2010.

HRBÁČEK, R.; RAJMIC, P.; VESELÝ, V.; ŠPIŘÍK, J. Řídké reprezentace signálů: Úvod do problematiky. Elektrevue - Internetový časopis, 2011, roč. 2011, č. 50, s. 1-10. ISSN: 1213-1539.

MAIRAL, J., ELAD, M., SAPIRO, G., Sparse learned representations for image restoration, IASC2008.

GOOSENS, B., et al., Efficient design of a low redundant Discrete Shearlet Transform Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on, 2009, 112-124.

NEGI, P., LABATE, D., 3-D Discrete Shearlet Transform and Video Processing Image Processing, IEEE Transactions on, 2012, 21, 2944-2954.

Vedoucí diplomové práce: Mgr. Pavel Rajmic, Ph.D.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2013/2014.

V Brně, dne 22.11.2013

L.S.

prof. RNDr. Josef Šlapal, CSc.
Ředitel ústavu

prof. RNDr. Miroslav Doupovec, CSc., dr. h. c.
Děkan fakulty

Abstrakt

Perfúzní zobrazování pomocí magnetické rezonance je lékařská diagnostická metoda, která se zdá být v dnešní době velmi slibnou. Tato práce se zabývá řídkými reprezentacemi, rekonstrukcí matic s nízkou hodnotí a komprimovaným snímáním, které umožňuje překonání současných fyzikálních možností perfúzního zobrazování pomocí magnetické rezonance. Dále představuje několik modelů pro rekonstrukci naměřených perfúzních dat a uvádí numerické metody pro jejich softwarovou implementaci, která je důležitou součástí práce. Navržené modely jsou ověřeny na simulovaných i reálných perfúzních datech z magnetické rezonance.

Summary

Magnetic resonance perfusion imaging is a today's very promising method for medicine diagnosis. This thesis deals with a sparse representation of signals, low-rank matrix recovery and compressed sensing, which allows overcoming present physical limitations of magnetic resonance perfusion imaging. Several models for reconstruction of measured perfusion data is introduced and numerical methods for their software implementation, which is an important part of the thesis, is mentioned. Proposed models are verified on simulated and real perfusion data from magnetic resonance.

Klíčová slova

komprimované snímání, perfúzní zobrazování, magnetická rezonance, řídká reprezentace signálů, rekonstrukce matic s nízkou hodnotí, proximální gradientní metoda

Keywords

compressed sensing, perfusion imaging, magnetic resonance, sparse representation of signals, low-rank matrix recovery, proximal gradient method

DAŇKOVÁ, M. *Komprimované snímání v perfúzním zobrazování pomocí magnetické rezonance*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2014. 56 s. Vedoucí Mgr. Pavel Rajmic, Ph.D.

Prohlašuji, že svou diplomovou práci na téma *Komprimované snímání v perfuzním zobrazování pomocí magnetické rezonance* jsem vypracovala samostatně pod vedením Mgr. Pavla Rajmice, Ph.D. s použitím materiálů uvedených v seznamu literatury.

Bc. Marie Daňková

Děkuji vedoucímu diplomové práce Mgr. Pavlu Rajmicovi, Ph.D. za odbornou pomoc, věnovaný čas a cenné rady při psaní práce. Dále bych chtěla poděkovat Ing. Radovanu Jiříkovi, Ph.D. za poskytnutí perfúzních dat myši z magnetické rezonance a odborné rady v oblasti perfúzního zobrazování v magnetické rezonanci.

Bc. Marie Daňková

PODĚKOVÁNÍ

Výzkum popsáný v této diplomové práci byl realizován v laboratořích podpořených z projektu **SIX**; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Místo

.....
(podpis autora)

Obsah

1	Úvod	3
2	Fyzikální podstata problému	4
2.1	Perfúzní zobrazování	4
2.1.1	Modelování perfúzního zobrazování	5
2.2	Princip zobrazování magnetickou rezonancí (MRI)	7
2.2.1	Polarizace protonů	7
2.2.2	Excitace	7
2.2.3	FID a relaxace	9
2.2.4	Metody snímání dat	10
2.2.5	Prostorové zakódování	11
3	Řídká reprezentace signálů	13
3.1	Postačující podmínky pro jednoznačnost řešení	14
3.1.1	Spark matice	14
3.1.2	Vzájemná koherence	15
3.2	Metody řešení	15
3.2.1	Hladové algoritmy	15
3.2.2	Relaxace	17
4	Komprimované snímání	23
4.1	Komprimované snímání využívající řídkosti	23
4.2	Komprimované snímání založené na nízké maticové hodnotě	25
4.3	Komprimované snímání v MRI	26
5	Proximální gradientní metoda	27
5.1	Proximální operátor	27
5.1.1	Proximální operátor pro ℓ_1 -normu	30
5.1.2	Proximální operátor pro 1D totální variaci	31
5.2	Proximální gradientní metoda (dopředné-zpětné dělení)	31
6	Aplikace komprimovaného snímání na perfúzním zobrazování v magnetické rezonanci	34
6.1	Fantom perfúzního zobrazování	34
6.2	Měřicí matice (masky)	36
6.3	$\mathbf{L} + \mathbf{S}$ metoda pro rekonstrukci	38
6.4	$\mathbf{L} + \mathbf{S}$ metoda s 1D totální variací	43
6.5	$\mathbf{M}$ model	45

6.6	Vliv šumu na rekonstrukci perfúzních dat	47
6.7	Reálná data	48
7	Závěr	50
	Literatura	51
	Seznam použitých zkratek a symbolů	54
	Použité symboly	54
	Použité zkratky	55
A	Obsah CD	56

1. Úvod

Perfúzní zobrazování je experimentální lékařská diagnostická metoda. V současné době je snaha nalézt jeho využití v magnetické rezonanci, která nenese žádná rizika způsobená ionizujícím zářením. Pomocí perfúzního zobrazování v magnetické rezonanci mohou být diagnostikována onkologická a kardiovaskulární onemocnění a umožněna jejich efektivní léčba.

Pro perfúzní analýzu je nezbytně nutné mít vysoké časové a prostorové rozlišení současně, jinak by mohlo dojít ke zkreslení odhadů parametrů popisujících vlastnosti tkáně, pomocí kterých se určuje diagnóza. Využitím klasického měření v magnetické rezonanci je ale nemožné dosáhnout obou vysokých rozlišení zároveň.

Proto se přímo nabízí možnost použít komprimované snímání pro perfúzní zobrazování pomocí magnetické rezonance, které nám umožní zrychlení snímacího procesu. Pro úspěšnou následnou rekonstrukci perfúzních dat je důležité znát nějaké apriorní informace o měřených datech (např. řídkost dat v nějaké bázi nebo nízkohodnostní struktura dat). Tím je možné dosáhnout požadovaného rozlišení.

Tato práce se zabývá využitím komprimovaného snímání v perfúzní analýze magnetickou rezonancí. Úvodní kapitola 2 představuje perfúzní zobrazování spolu s magnetickou rezonancí, a to jak jejich fyzikální princip, tak i matematický popis. Další kapitola 3 se zabývá principy a algoritmy hledání tzv. řídkých reprezentací signálů. Následující kapitola 4 pojednává o komprimovaném snímání a kapitola 5 popisuje numerickou metodu pro minimalizační problémy používané pro rekonstrukci naměřených perfúzních dat. Závěrečná kapitola 6 je věnována simulaci komprimovaného snímání perfúzního zobrazování v magnetické rezonanci, která je implementována v programovém prostředí MatLab.

2. Fyzikální podstata problému

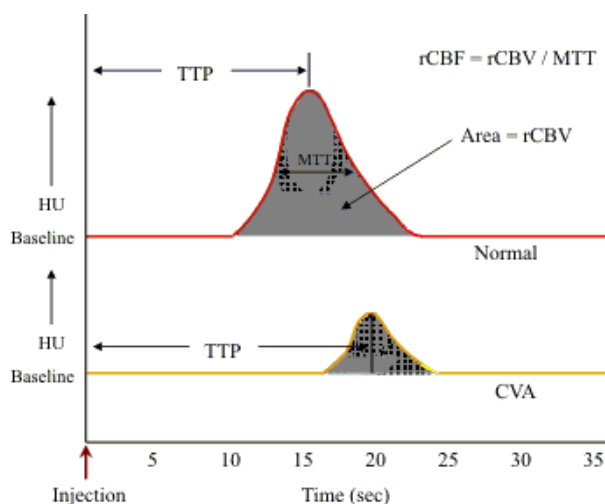
Perfúzní zobrazování je důležitý druh zobrazovací techniky používaný v medicíně pro diagnózu a hodnocení úspěšnosti léčby. Využití perfúzního zobrazování v magnetické rezonanci je jedním ze slibných možností použití této diagnostické metody. Zatím je ale perfúzní analýza v magnetické rezonanci pouze na experimentální úrovni, protože je limitována svými fyzikálními možnostmi.

2.1. Perfúzní zobrazování

Perfúzní zobrazování je důležitým nástrojem pro diagnózu kardiovaskulárních a onkologických onemocnění. V dnešní době se může tato metoda využívat v nejrozličnějších zobrazovacích modalitách jako je magnetická rezonance, počítačová tomografie, pozitronová emisní tomografie či ultrazvuk.

Základní postup perfúzního zobrazování je podobný pro všechny modality. Pacientovi je nitrožilně podána vhodná kontrastní látka (bolus) ve formě injekce nebo infúze, díky které mohou být naměřena data v oblasti zájmu. Touto oblastí může být jednak nějaký orgán, tak i část orgánu nebo jenom samostatný voxel.

Díky kardiovaskulárnímu systému je kontrastní látka rozvedena dále do těla a její časové a prostorové rozložení může být sledováno. Časové průběhy koncentrace této látky v oblasti zájmu nazýváme perfúzní křivky. Analýzou perfúzních křivek získáme odhady perfúzních parametrů nezbytných pro diagnózu. Nejdůležitějšími parametry jsou střední doba průchodu ($MTT = \text{mean transit time}$), čas dosáhnutí maximální intenzity ($TTP = \text{time to peak}$), objem krve ($BV = \text{blood volume}$, $rCBV = \text{relative cerebral blood volume}$) a krevní tok ($BF = \text{blood flow}$, $rCBF = \text{relative cerebral blood flow}$). Tyto parametry jsou schématicky znázorněny na obr. 2.1. Horní perfúzní křivka představuje zdravou oblast a dolní oblast postiženou ischemickou chorobou. Je vidět, že se tyto dvě perfúzní křivky liší ve svých parametrech [19].



Obrázek 2.1: Perfúzní křivky a jejich parametry – nahoře zdravá tkáň, dole postižená ischemickou chorobou [24]

2.1.1. Modelování perfúzního zobrazování

Pro perfúzní křivky se využívají tři základní typy modelů – matematické, prostorové a fyzikální. Pomocí těchto modelů se potom lépe odhadují perfúzní parametry. Matematické modely jsou založeny na tvarové podobnosti mezi perfúzními křivkami a nějakou matematickou funkcí, kterou se snažíme data proložit. Příkladem tohoto typu může být využití hustoty lognormálního rozdělení nebo model se zpožděním, který je založen na konvoluci hustoty normálního rozdělení (představující náhodný rozptyl způsobený turbulencemi) s exponenciální křivkou (míchání krve v srdečních komorách).

Prostorové modely jsou založeny na popisu prokrvení tkání v prostoru. Modelované křivky popisují vztah mezi umístěním aplikace kontrastní látky a vybrané oblasti zájmu. Do této kategorie spadá Erlangův model a model vyjádřený pomocí gama rozdělení pravděpodobnosti. Poslední kategorie modelů je založena na fyzikálním popisu částic po podání bolusu. Tyto modely se dají popsat pomocí parciálních diferenciálních rovnic [11].

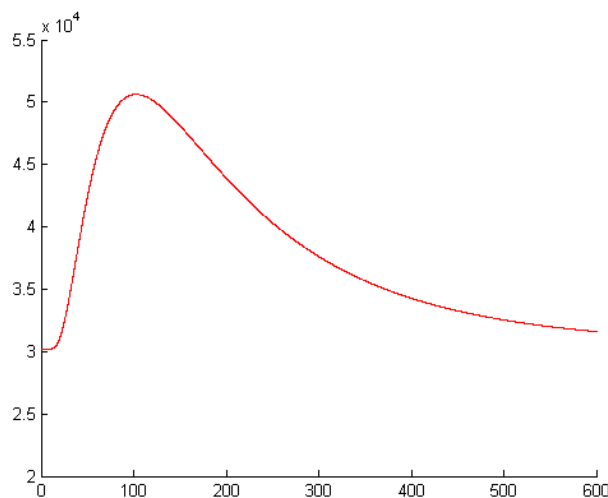
Lognormální model

V této práci jsou perfúzní křivky matematicky modelovány pomocí křivky hustoty lognormálního rozdělení. Tento model představující závislost koncentrace kontrastní látky na čase v oblasti zájmu je definován takto:

$$f(t) = \begin{cases} c & t \leq t_0 \\ c + \frac{S}{(t-t_0)\sigma\sqrt{2\pi}} e^{-\frac{(\ln(t-t_0)-\mu)^2}{2\sigma^2}} & t > t_0 \end{cases},$$

kde c představuje konstantní hodnotu bez použití kontrastní látky, t_0 zpoždění bolusu mezi místem podání a oblastí zájmu, S plochu mezi konstantní hodnotou c a křivkou a μ a σ jsou obvyklé parametry lognormálního rozdělení. Na obr. 2.2 je znázorněn lognormální model perfúzní křivky pro konkrétní volbu parametrů $t_0 = 7$, $c = 3,02 \cdot 10^4$, $S = 5,25 \cdot 10^6$, $\mu = 5,18$ a $\sigma = 0,62$ (jednotky parametrů perfúze v práci neuvádíme, pouze musí být vzájemně si odpovídající, konkrétní jednotky nás nezajímají).

Ačkoli je tento model založen na podobnosti reálných a modelových křivek, může být odvozen z distribuce průtoku krve a kinetiky kontrastní látky pomocí stochastického modelu s využitím fraktálních sítí. Každá céva se větví do dvou dceřiných cév. Je-li podíl

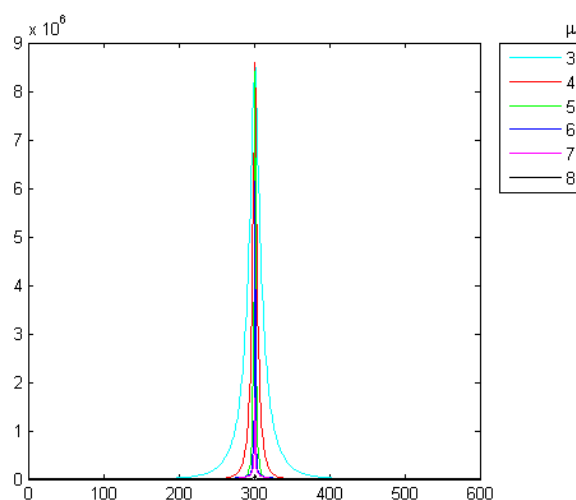


Obrázek 2.2: Perfúzní křivka s konkrétně navolenými parametry

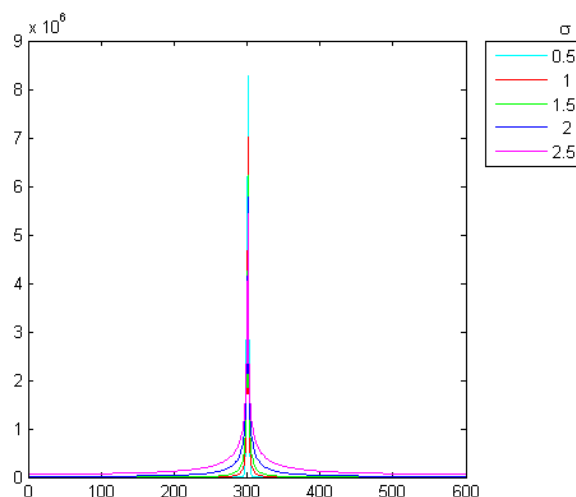
průtoku v dceřiných cévách náhodná proměnná v intervalu $(0, 1)$ s libovolnou funkcí hustoty pravděpodobnosti a síť cév má velký počet generací, distribuce průtoku je právě lognormální funkce [21].

Pro odhadnutí parametrů modelu nejprve odhadneme posunutí c jako průměr z několika prvních vzorků. Poté všechny vzorky snížíme o tuto hodnotu. Zpoždění t_0 potom můžeme dostat jako první vzorek větší než 0,1 maximální hodnoty. Dále plochu pod křivkou S spočítáme pomocí numerické integrace (např. využitím lichoběžníkové formule). Na odhad parametrů μ a σ lze použít metodu maximální věrohodnosti.

Spektrum perfúzních křivek můžeme považovat za řídké, protože velké množství jeho koeficientů se blíží k nule. Tento fakt budeme nadále v práci využívat. Na obr. 2.3 je znázorněno amplitudové spektrum pro různou volbu parametru μ a na obr. 2.4 pro různou volbu σ . Při výpočtu spektra perfúzních křivek byly ostatní parametry zvoleny takto: $t_0 = 1$, $S = 8,6 \cdot 10^6$, $c = 0$. Na výpočet spektra bylo použito 600 měření v čase [11].



Obrázek 2.3: Amplitudové spektrum perfúzní křivky pro různé μ při $\sigma = 0,7$



Obrázek 2.4: Amplitudové spektrum perfúzní křivky pro různé σ při $\mu = 5,5$

2.2. Princip zobrazování magnetickou rezonancí (MRI)

Magnetická rezonance (MR) je zobrazovací technika používaná ve zdravotnictví k zobrazení požadovaných oblastí lidského těla. V magnetické rezonanci se využívá interakce silného magnetického pole s jádry vodíku, který je obsažen ve vodě. Voda tvoří dvě třetiny lidského těla, proto je MRI vhodné prakticky pro zobrazení všech měkkých tkání s výjimkou kostí, ve kterých není téměř žádná voda. Po aplikaci silného magnetického pole je potřeba jádra „zviditelnit“ tím, že je excitujeme radiofrekvenčními pulzy (RF), a poté můžeme naměřit jejich odezvu. Abychom byli schopni rozlišit místa, ze kterých daný naměřený signál pochází, musíme přidat další slabší magnetické pole – tzv. gradientní pole [27].

2.2.1. Polarizace protonů

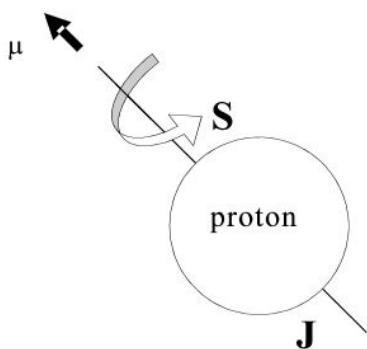
Všechny protony mají magnetický moment (μ) a spin (S) – viz obr. 2.5. Za běžných podmínek jsou osy rotace protonů (spiny) nahodile orientovány v prostoru, ale když přidáme silné vnější magnetické pole s indukcí B_0 , dojde ke srovnání těchto os ve směru siločar tohoto pole. Většina protonů se orientuje souhlasně (energeticky méně náročný stav), ale některé jsou orientovány nesouhlasně.

Navíc protony začnou vykonávat precesní pohyb (obr. 2.6). Tato frekvence precese f_0 (zvaná Larmorova) závisí na intenzitě magnetického pole a typu atomu (jeho gyromagnetickém poměru γ):

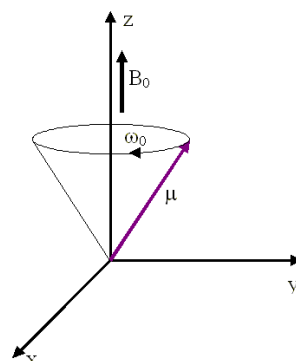
$$f_0 = \frac{1}{2\pi} \gamma B_0.$$

Díky tomu se dají zobrazovat nezávisle různé typy atomových jader, čehož se využívá v tzv. MR spektroskopii [23].

Využívané magnety jsou většinou supravodivé a běžně mají 0,5 až 2 T, ale mohou být v rozsahu 0,1 až 12 T. Pro lepší představu, jak silné jsou tyto magnety, si můžeme uvědomit, že magnetické pole Země je silné přibližně $5 \cdot 10^{-5}$ T [25, 27].



Obrázek 2.5: Znáznornění magnetického momentu μ a spinu S protonu [25]



Obrázek 2.6: Precesní pohyb protonu v magnetickém poli s indukcí B_0 [25]

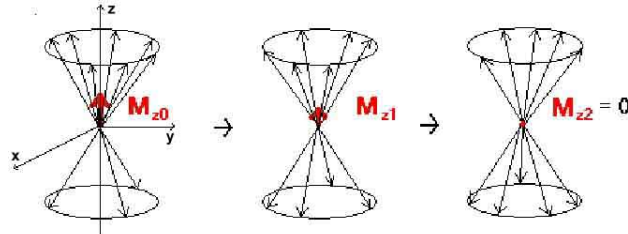
2.2.2. Excitace

Protože jsou precese jednotlivých protonů v různých fázích, dochází k vzájemnému vyrušení jejich vlivu na výsledný vektor magnetizace tkáně v rovině xy . Tento vektor je tedy rovnoběžný se směrem vnějšího magnetického pole, proto není pozorovatelný. Abychom ho

mohli naměřit, potřebujeme jej sklopit do roviny xy , ve které je umístěna přijímací cívka. Proto se přidává radiofrekvenční pulz, který musí mít frekvenci rovnou Larmorově, aby došlo k excitaci protonů. Odsud je odvozen název magnetická rezonance, protože protony rezonují na Larmorově frekvenci s RF pulzem.

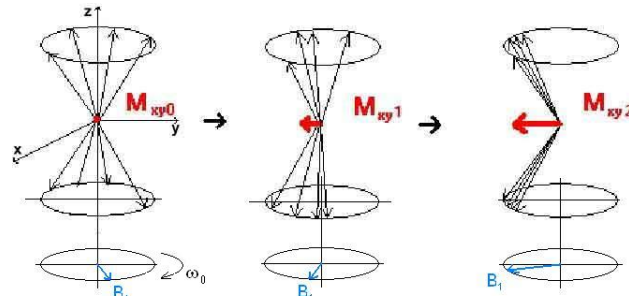
Radiofrekvenční pulz má dvojí účinek na tkáňové protony:

- Více protonů je orientováno antiparalelně k magnetickému poli, čímž dochází ke změně z -ové složky vektoru magnetizace tkáně.



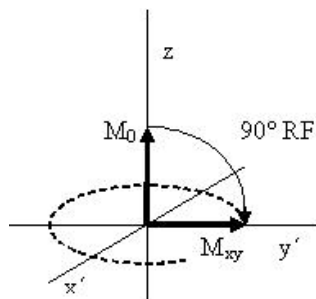
Obrázek 2.7: Po přidání RF pulzu může být více protonů orientováno antiparalelně [25]

- Fáze precesí se sjednotí, čímž vznikne příčná složka vektoru magnetizace.



Obrázek 2.8: RF pulz sjednotí fáze precesí jednotlivých protonů [25]

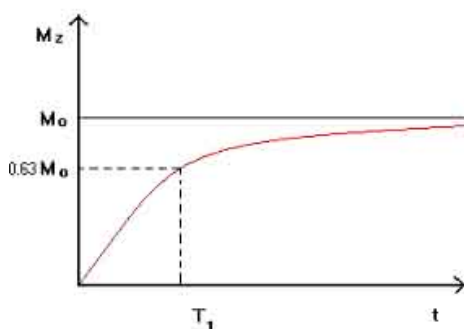
Tyto dva děje probíhají současně. Pokud použijeme nový souřadný systém, jehož osa z se shoduje s původní a jehož osy x', y' rotují s Larmorovou frekvencí kolem osy z , bude se pohyb vektoru tkáňové magnetizace jevit jako pouhé „sklápění“ do roviny xy (obr. 2.9), přičemž úhel sklopení závisí na integrálu dodané energie (tedy na velikosti RF pulzu a délce jeho trvání). Téměř vždy se používá RF pulz, který sklápí vektor o 90° , ale často bývá následován dalšími, které využívají různých vlastností tkáně pro dosažení požadovaného kontrastu (tzv. pulzní sekvence) [25].



Obrázek 2.9: „Sklápění“ vektoru magnetizace do roviny xy přidáním RF pulzu [25]

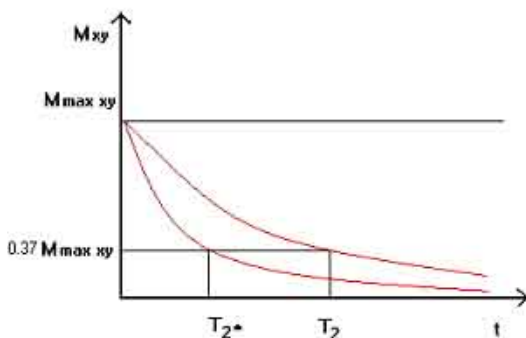
2.2.3. FID a relaxace

Po dodání energie 90° radiofrekvenčním pulzem začnou vektory magnetizace rotovat v rovině xy s Larmorovou frekvencí, a tím indukují elektrický proud v cílce umístěné v této rovině. Takto získaný signál se označuje jako tzv. FID (free induction decay), který má harmonický průběh s exponenciálně klesající amplitudou. Jakmile přestane působit RF pulz, dojde k tzv. relaxaci, kdy se protony snaží dostat do rovnovážného stavu (ve směru siločar magnetického pole). Vektor tkáňové magnetizace M_z ve směru osy z nabývá opět svou velikost (tzv. podélná longitudinální neboli spin-mřížková relaxace). Časový průběh nárůstu je exponenciální a můžeme ho znázornit tzv. T_1 křivkou (obr. 2.10), kde konstanta T_1 udává čas, za jaký dojde k obnovení velikosti M_z na 63 % své původní velikosti. Také přestane působit synchronizační efekt radiofrekvenčního pulzu. Vlivem magnetic-



Obrázek 2.10: T_1 relaxace [25]

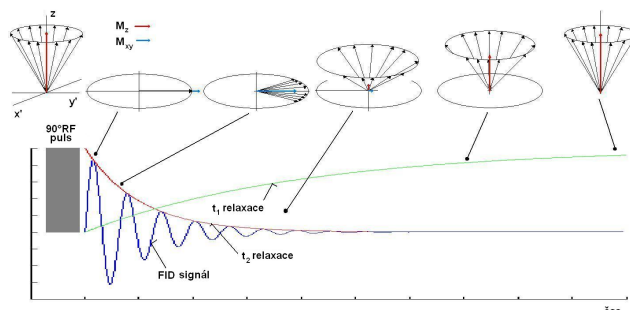
kých polí jednotlivých částic, které způsobují drobné lokální nehomogenity magnetického pole, budou jednotlivé protony precedovat s nepatrně rozdílnými kmitočty a dojde tak k postupné ztrátě fázové jednotnosti precedujících protonů (spin-spinová relaxace), a tím také k zániku příčné složky vektoru tkáňové magnetizace M_{xy} . Změnu velikosti v čase popisuje T_2 křivka (obr. 2.11), která má také charakter exponenciály. T_2 relaxační konstanta pak udává čas, za který dojde k poklesu velikosti M_{xy} na 37 % svého maxima. Ve skutečnosti je pokles příčné složky tkáňové magnetizace ovlivněn ještě drobnými změnami v nehomogenitě vnějšího magnetického pole. Pokles je tak podstatně strmější a příslušnou relaxační konstantu označujeme jako T_2^* .



Obrázek 2.11: T_2 relaxace [25]

Na obr. 2.12 je zobrazeno rozložení vektorů magnetických momentů v různých časových okamžicích po excitaci 90° RF pulzem. Těsně po odeznění RF pulzu jsou jednotlivé vektory ve fázi a výsledný vektor magnetizace je skloněn do roviny xy . Navenek tedy můžeme pozorovat vektor magnetizace v rovině xy . V přijímací cílce se začne indukovat FID

signál. Jelikož je vždy $T_2 < T_1$, rychleji se uplatňuje T_2 relaxace a amplituda FID signálu klesá exponenciálně s konstantou T_2 (resp. T_2^*). Zároveň, ale pomaleji, se uplatňuje taky relaxace T_1 , což způsobí růst magnetizace ve směru osy z . Celý systém konverguje k rovnovážnému stavu, který trval před excitací [14, 25].



Obrázek 2.12: Měření FID signálu [25]

2.2.4. Metody snímání dat

Pokud využijeme v MRI pouze sekvenci 90° RF pulzu, zobrazíme tak hustotu protonů v jednotlivých tkáních. Takto ovšem nemusíme zobrazit všechny požadované rozdíly v tkáních. K dosažení potřebného kontrastu se proto využívají další pulzní sekvence, které berou v potaz i relaxační časy T_1 a T_2 (resp. T_2^*).

Metoda saturation recovery

Tato zobrazovací metoda se skládá z několika po sobě jdoucích rovnoměrně časově rozložených RF pulzů, z nichž každý otáčí vektor magnetizace o 90° . Po prvním RF pulzu se vektor magnetizace sklopí do roviny xy a poté se na něm začne uplatňovat relaxace. Protože různé tkáně mají různou hodnotu T_1 , bude rychlost podélné relaxace v různých tkáních rozdílná. Poté jsou přidány další 90° RF pulzy, po kterých se ihned měří FID. Interval mezi po sobě jdoucími pulzy se označuje jako doba opakování pulsu TR (pulse repetition time). V oblastech s krátkým relaxačním časem T_1 se vektor po prvním pulzu vrátí do polohy blízko osy z a po aplikaci druhého pulzu se sklopí mírně pod rovinu xy . Naopak v oblastech s dlouhým relaxačním časem T_1 se vektor vrací pouze těsně nad rovinu xy , takže se po druhém RF pulzu dostane hodně pod rovinu xy . Díky tomu se bude FID signál lišit v tkáních s různou relaxační dobou T_1 (obdobně po dalších pulzech). Vhodným nastavením času TR lze docílit optimálního kontrastu zobrazovaných tkání.

Technika inversion recovery

Zobrazování pomocí inversion recovery je tvořeno sekvencí dvou pulzů. Jako první je použit 180° RF pulz, který překlápí vektor magnetizace z kladného směru osy z do záporného. Poté se začne uplatňovat relaxace T_1 a magnetizace se vrací do rovnovážného stavu. Potom následuje 90° RF pulz, který překlápí vektor magnetizace do roviny xy , a po něm se měří FID. Vektor magnetizace, který se rychle vrátí do kladného směru osy z kvůli krátkému T_1 , se po excitaci druhým RF pulzem dostane do kladného směru osy x (y) – dostaneme světlou část obrazu, zatímco vektor v tkáních s dlouhou T_1 do záporného směru (tmavá část obrazu). Ve srovnání s metodou saturation recovery příkládá metoda

inversion recovery větší váhu na T_1 , takže je vhodná, pokud potřebujeme dostat vysoký kontrast.

Spin-echo sekvence

Skládá se z 90° RF pulzu a jednoho nebo několika následujících 180° RF pulzů. Po 90° RF pulzu je vektor magnetizace sklopený do roviny xy a začíná se projevovat T_2 relaxace, kdy dochází k rozfázování. Následuje-li ale další 180° RF pulz, který překlopí jednotlivé spiny v rovině xy o 180° , spiny se opět sfázuji a v přijímací cílce je detekován echo signál, jehož amplituda je závislá na T_2 tkáně (T_2^* se neuplatní – vliv nehomogenity B_0 při rozfázování se vyruší při sfázování). Kontrast v obrázku lze nastavit pomocí času TR (time repeat) a TE (time echo). Bude-li $TR \gg T_1$, pak bude obrázek T_2 -váhovaný. Bude-li TR srovnatelný s T_1 , bude obrázek při malých TE T_1 -váhovaný, při větších TE T_2 -váhovaný.

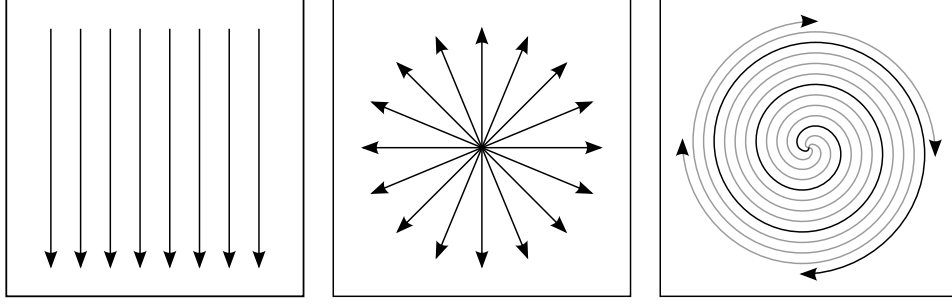
Technika gradient-echo

Tato sekvence začíná 90° RF pulzem, který sklopí vektor magnetizace do roviny xy . K vyvolání echa je zde ale narozdíl od předešlé techniky použit gradient magnetického pole místo dalšího pulzu. Je-li k magnetickému poli B_0 přidáno gradientní magnetické pole, budou sousedící protony precesovat s mírně odlišnou Larmorovou frekvencí. To způsobí rozfázování jednotlivých spinů. Následuje gradient s opačným znaménkem, který znovu sfázuje jednotlivé spiny, a tím vyvolá echo. Pokles amplitudy FID signálu je tady závislý na relaxačním čase T_2^* a obrázek tedy bude T_2^* -váhovaný. Tato technika pracuje s menšími časy TE než předchozí a pro excitaci se využívá menších úhlů, což vede k možnosti menších časů TR, proto se jedná o velmi rychlou zobrazovací techniku, která slouží jako základ zobrazovacích technik používaných v současnosti [12, 25].

2.2.5. Prostorové zakódování

Při použití výše uvedených technik chybí v měřených datech jakákoliv prostorová informace. Proto se přidává další slabší magnetické pole, tzv. gradientní pole, které způsobí, že se kmitočet precese bude lineárně lišit v prostoru. To si můžeme názorně představit na analogii s klavírem. Každá klávesa na klavíru se rozeznívá jinou frekvencí. Když klavírista zahráje nějaký akord, slyšíme všechny tyto frekvence zároveň, ale zkušený muzikant dokáže rozeznat, které noty a jak silně byly zahrány.

FID signál vytvořený v přítomnosti gradientního pole je také polyfonní. Prostorové souřadnice uvnitř těla pacienta se dají přirovnat ke klávesám klavíru. Zaznamenaný FID signál je jako „tóny“ s frekvencí lineárně se měnící podle polohy. Pokud budeme gradientní pole měnit, získáme různé trajektorie v k -prostoru. Nejčastěji používané typy trajektorií jsou zobrazeny na obr. 2.13. Z těchto trajektorií již můžeme snadno rekonstruovat 2D nebo 3D obrazy pacientova těla pomocí inverzní Fourierovy transformace [18].



Obrázek 2.13: Nejběžnější typy trajektorií v k -prostoru při snímání MRI dat

Matematický popis MRI

Gradientní pole $X(\mathbf{z})$ na pozici $\mathbf{z} \in \mathbb{R}^3$ můžeme zapsat ve tvaru $X(\mathbf{z}) = |X(\mathbf{z})|e^{-i\phi(\mathbf{z})}$, kde $|X(\mathbf{z})|$ je amplituda a $\phi(\mathbf{z})$ fáze. Označíme-li $\mathbf{G} \in \mathbb{R}^3$ gradient tohoto magnetického pole, frekvence precese (Larmorova) může být popsána jako

$$\omega(\mathbf{z}) = \kappa(B_0 + \langle \mathbf{G}, \mathbf{z} \rangle), \quad \mathbf{z} \in \mathbb{R}^3,$$

kde B_0 je intenzita vnějšího statického magnetického pole a κ fyzikální konstanta. Pokud se gradient $\mathbf{G}$ mění v čase, $\mathbf{G} : \langle 0, T \rangle \rightarrow \mathbb{R}^3$, fáze magnetizace $\phi(\mathbf{z}) = \phi(\mathbf{z}, t)$ je integrál

$$\phi(\mathbf{z}, t) = 2\pi\kappa \int_0^t \langle \mathbf{G}(\tau), \mathbf{z} \rangle d\tau,$$

kde $t = 0$ odpovídá času excitace radiofrekvenčním pulzem. Nyní definujeme funkci $\mathbf{k}$, $\mathbf{k} : \langle 0, T \rangle \rightarrow \mathbb{R}^3$ jako

$$\mathbf{k}(t) = \kappa \int_0^t \mathbf{G}(\tau) d\tau.$$

Přijímací cívka integruje přes celý objem a naměří signál

$$f(t) = \int_{\mathbb{R}^3} |X(\mathbf{z})| e^{-2\pi i \langle \mathbf{k}(t), \mathbf{z} \rangle} d\mathbf{z} = \mathcal{F}(|X|)(\mathbf{k}(t)),$$

kde $\mathcal{F}(|X|)(\xi)$ značí trojdimenzionální Fourierovu transformaci amplitudy $|X|$ magnetizace [10].

V MRI můžeme také měřit jednotlivé řezy těla. Nejprve nastavíme gradientní pole na z -ové souřadnici, které způsobí, že se rezonanční frekvence budou lineárně lišit podél této souřadnice. Když potom excitujeme protony radiofrekvenčním pulzem s úzkým rozsahem frekvencí, vybereme tím jeden řez, ve kterém Larmorovy frekvence odpovídají frekvencím RF pulzu. V tomto případě je trojrozměrná Fourierova transformace nahrazena dvojrozměrnou [12].

Naměřený MRI signál je tedy Fourierova transformace prostorově závislé amplitudy $|X|$ magnetizace (obraz), která je podvzorkována na křivce $\{\mathbf{k}(t) : t \in [0, T]\} \subset \mathbb{R}^3$. Opakováním několika různých radiofrekvenčních excitací dostaneme vzorky (diskrétní) Fourierovy transformace funkce $|X|$ podél různých křivek $\mathbf{k}_1, \dots, \mathbf{k}_L$ v $\mathbb{R}^3$. Potřebný čas na měření je úměrný počtu L takových křivek, proto bychom chtěli jejich počet minimalizovat [10].

3. Řídká reprezentace signálů

Problematikou řídkých reprezentací signálů jsem se zabývala již ve své bakalářské práci [7]. Signál $\mathbf{y}$ (např. zvuk, obraz či video) můžeme reprezentovat pomocí aditivního modelu, ve kterém vyjádříme $\mathbf{y}$ jako lineární kombinaci základních „stavebních bloků“ $\mathbf{a}_i$:

$$\mathbf{y} = \sum_i x_i \mathbf{a}_i,$$

kde x_i jsou váhy neboli souřadnice $\mathbf{y}$ v systému $\{\mathbf{a}_i\}$. Jednotlivým $\mathbf{a}_i$ říkáme atomy, celý systém $\{\mathbf{a}_i\}$ nazýváme slovník.

Tento model můžeme také zapsat v maticovém tvaru jako soustavu lineárních rovnic

$$\mathbf{A}\mathbf{x} = \mathbf{y},$$

kde $\mathbf{x}$ je vektor neznámých a $\mathbf{y}$ známý vektor (může to být pozorování, měření, signál). Tento systém je nedourčený, tj. máme více neznámých než kolik je rovnic. Předpokládáme, že matice $\mathbf{A}$ je plné hodnosti, a proto existuje nekonečně mnoho řešení.

Před samotným zavedením řídké reprezentace signálů je třeba vysvětlit několik pojmů.

Definice 3.1. [16] Nechť $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = (x_1, \dots, x_n)^\top$. Pak ℓ_p norma vektoru $\mathbf{x}$ je definována takto:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad \text{pro } 1 \leq p < \infty,$$

$$\|\mathbf{x}\|_p = \sum_{i=1}^n |x_i|^p \quad \text{pro } 0 < p < 1,$$

$$\|\mathbf{x}\|_\infty = \max_i |x_i|,$$

$$\|\mathbf{x}\|_0 = |\text{supp}(\mathbf{x})| = |\{i : x_i \neq 0, i = 1, \dots, n\}|.$$

Poznámka. O normu se ve skutečnosti jedná pouze, když $1 \leq p \leq \infty$. V ostatních případech pro $0 \leq p < 1$ se o normu nejedná, protože neplatí pozitivní homogenita, je to pouze metrika. Pro zjednodušení se ale v práci dále používá označení ℓ_p norma. Stejným způsobem lze zavést i ℓ_p normu pro matice.

Definice 3.2. [6] Nechť $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = (x_1, \dots, x_n)^\top$. Pak *diskrétní 1D totální variace* $\|\mathbf{x}\|_{\text{tv}}$ vektoru $\mathbf{x}$ je definována jako

$$\|\mathbf{x}\|_{\text{tv}} = \sum_{i=1}^{n-1} |x_{i+1} - x_i|.$$

Diskrétní 1D totální variace je seminorma, protože nerozlišuje body, ale je pozitivně homogenní a splňuje trojúhelníkovou nerovnost.

Dalším důležitým pojmem je řídkost vektoru.

Definice 3.3. [16] Vektor $\mathbf{x}$ nazveme *k-řídkým*, pokud platí:

$$\|\mathbf{x}\|_0 \leq k.$$

Tedy k -řádký vektor má nejvýše k nenulových složek. Nás budou zajímat pouze řádká řešení – řešení s co nejvíce nulovými složkami. Experimentálně bylo totiž zjištěno, že reálné signály můžeme přibližně reprezentovat jako lineární kombinaci pomocí malého počtu základních „stavebních bloků“. Matematicky můžeme tento problém popsat jako řešení tohoto minimalizačního problému:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ vzhledem k } \mathbf{Ax} = \mathbf{y}. \quad (\text{P0})$$

V tomto optimalizačním problému hledáme vektor $\mathbf{x}$ jako nejřidší řešení soustavy $\mathbf{Ax} = \mathbf{y}$, přičemž známe vektor $\mathbf{y} \in \mathbb{R}^m$ a matici $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m < n$.

Signál také může být zašuměný a přesné řešení neexistuje. Proto v praxi často uvažujeme odchylku od přesného řešení neboli:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ vzhledem k } \|\mathbf{Ax} - \mathbf{y}\|_2 \leq \epsilon,$$

kde ϵ je povolená odchylka od přesného řešení [16].

3.1. Postačující podmínky pro jednoznačnost řešení

3.1.1. Spark matice

Pro jednoznačnost řešení je nutné zavést pojem spark (doslova „jiskra“).

Definice 3.4. [16] Číslo $\text{spark}(\mathbf{A})$ definujeme jako nejmenší počet sloupců matice $\mathbf{A}$, které jsou lineárně závislé. Formálně:

$$\text{spark}(\mathbf{A}) = \min_{\substack{\mathbf{z} \in \ker \mathbf{A}, \\ \mathbf{z} \neq \mathbf{0}}} \|\mathbf{z}\|_0,$$

kde $\ker \mathbf{A}$ značí jádro lineárního zobrazení určeného maticí $\mathbf{A}$.

Pro nenulovou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$, kde $m < n$, platí, že spark může nabývat hodnot $\text{spark}(\mathbf{A}) \in \{2, \dots, m+1\}$, přičemž hodnoty 2 je dosaženo, když je jeden sloupec přímo násobkem jiného.

Nyní následuje tvrzení, které poskytuje postačující podmínku pro jednoznačnost řešení.

Tvrzení 3.1. [16] *Pokud má soustava $\mathbf{Ax} = \mathbf{y}$ řešení $\mathbf{x}$ splňující*

$$\|\mathbf{x}\|_0 < \frac{\text{spark}(\mathbf{A})}{2},$$

pak $\mathbf{x}$ je nutně nejřidší možné řešení a žádné jiné řešení se stejnou řádkostí neexistuje.

Nalezení $\text{spark}(\mathbf{A})$ je bohužel výpočtově velmi náročné, proto se snažíme nalézt jiné podmínky zaručující jednoznačnost řešení.

3.1.2. Vzájemná koherence

Definice 3.5. [16] *Vzájemná koherence (mutual coherence) matice $\mathbf{A}$ je definována jako největší absolutní normovaný skalární součin dvou různých sloupců matice $\mathbf{A}$:*

$$\mu(\mathbf{A}) = \max_{\substack{1 \leq j, k \leq n, \\ j \neq k}} \frac{|\mathbf{a}_j^\top \mathbf{a}_k|}{\|\mathbf{a}_j\|_2 \cdot \|\mathbf{a}_k\|_2},$$

kde $\mathbf{a}_j$ označuje j -tý sloupec matice $\mathbf{A}$.

Jedině ortogonální matice mají nulovou koherenci, protože její sloupce jsou po dvou ortogonální (tedy čítec je roven nule), zatímco nedourčená soustava bude mít vždy nenulovou koherenci.

Tvrzení 3.2. [16] *Pro libovolnou matici $\mathbf{A}$ platí:*

$$\text{spark}(\mathbf{A}) \geq 1 + \frac{1}{\mu(\mathbf{A})}.$$

Tato podmínka nám umožňuje zdola ohraničit $\text{spark}(\mathbf{A})$ a přitom není zdaleka tak výpočtově náročná. Z toho plyne:

Tvrzení 3.3. [16] *Pokud má soustava $\mathbf{A}\mathbf{x} = \mathbf{y}$ řešení $\mathbf{x}$ splňující*

$$\|\mathbf{x}\|_0 < \frac{1}{2} \left(1 + \frac{1}{\mu(\mathbf{A})} \right),$$

pak $\mathbf{x}$ je nutně nejřidší možné a jediné takové. Navíc tohoto řešení lze dosáhnout pomocí ℓ_1 -minimalizace (vysvětleno dále viz část 3.2.2).

Naší snahou je tedy používání maximálně nekoherentních slovníků, které se nejvíce přibližují ortogonální matici [16, 13].

3.2. Metody řešení

Předpokládejme, že $\text{spark}(\mathbf{A}) > 2k_0$ a existuje k_0 -řídke řešení soustavy. Pak je toto řešení nejřidší možné a jednoznačné. Pro nalezení přesného řešení bychom museli projít všech $\binom{m}{k_0}$ kombinací podmnožin atomů matice, což je NP-těžký problém (nedeterministicky polynomiální problém). Proto se vyvinulo několik aproximačních metod řešení tohoto problému, které nejsou tolik přesné, ale zato jsou rychlejší. Tyto metody se dělí do dvou hlavních kategorií – hladové algoritmy a relaxační algoritmy [16].

3.2.1. Hladové algoritmy

Hlavní myšlenka

Předpokládejme, že slovník $\mathbf{A}$ má $\text{spark}(\mathbf{A}) > 2$ a existuje řídke řešení soustavy lineárních rovnic $\mathbf{A}\mathbf{x} = \mathbf{y}$ s řidkostí $k = 1$. To znamená, že $\mathbf{y}$ je skalárním násobkem některého sloupce matice $\mathbf{A}$. Tento atom můžeme nalézt pomocí m kroků. V j -tém kroku minimalizujeme chybu $\epsilon(j) = \|\mathbf{a}_j z_j - \mathbf{y}\|_2$. Takto dostaneme optimální volbu konstanty $z_j^* = \mathbf{a}_j^\top \mathbf{y} / \|\mathbf{a}_j\|_2^2$. Pokud se chyba rovná nule, našli jsme hledané řešení.

Ve většině případů ovšem nemáme jenom řešení s řídkostí 1, ale s řídkostí $k = k_0$. Předpokládejme tedy, že slovík $\mathbf{A}$ má $\text{spark}(\mathbf{A}) > 2k_0$. Pak hledané řešení je nejřidší možné a je zaručena jeho jednoznačnost. Jak bylo uvedeno výše, hledat přesné řešení je výpočtově velmi náročné, proto se hladové algoritmy pokouší iterativně nalézt v každém kroku jeden atom, který má největší podíl na řešení (tj. má nejmenší reziduální ℓ_2 chybu při aproximaci $\mathbf{y}$). Na začátku položíme $\mathbf{x}^0 = 0$ a v každém kroku aproximujeme řídké řešení tak, aby nenulové prvky $\mathbf{x}$ příslušely množině aktivních atomů (na začátku je tato množina prázdná). V každé iteraci přidáme do této množiny vybraný atom. Tento postup není přesný, proto musíme nastavit maximální odchylku od přesného řešení, se kterou jsme spokojeni [9].

Ortogonalní sdružovací metoda (OMP)

Ortogonalní sdružovací metoda (Orthogonal Matching Pursuit) je jednou z nejpoužívanějších hladových algoritmů, proto ji uvedeme jako první. Jak bylo předesláno výše, v tomto algoritmu položíme prvotní aproximaci řídkého řešení $\mathbf{x}^0 = 0$, spočítáme počáteční reziduum $\mathbf{r}^0 = \mathbf{y} - \mathbf{A}\mathbf{x}^0 = \mathbf{y}$. Samozřejmě na začátku také množina všech nenulových koeficientů S^0 je prázdná (této množině budeme říkat nosič). Potom můžeme přistoupit k jednotlivým krokům metody. Nejprve se vyjádří reziduální chyba při aproximaci pro všechna $j = 1, 2, \dots, n$:

$$\begin{aligned}\epsilon(j) &= \min_{z_j} \|\mathbf{a}_j z_j - \mathbf{r}^{k-1}\|_2^2 = \left\| \frac{\mathbf{a}_j^\top \mathbf{r}^{k-1}}{\|\mathbf{a}_j\|_2^2} \mathbf{a}_j - \mathbf{r}^{k-1} \right\|_2^2 = \\ &= \|\mathbf{r}^{k-1}\|_2^2 - 2 \frac{(\mathbf{a}_j^\top \mathbf{r}^{k-1})^2}{\|\mathbf{a}_j\|_2^2} + \frac{(\mathbf{a}_j^\top \mathbf{r}^{k-1})^2}{\|\mathbf{a}_j\|_2^2} = \\ &= \|\mathbf{r}^{k-1}\|_2^2 - \frac{(\mathbf{a}_j^\top \mathbf{r}^{k-1})^2}{\|\mathbf{a}_j\|_2^2}.\end{aligned}$$

Z těchto vypočítaných chyb vybereme nejmenší z nich. Označme index nejmenší chyby j_0 , formálně zapsáno: $\forall j \notin S^{k-1} : \epsilon(j_0) \leq \epsilon(j)$. Poté přidáme do nosiče S^k nalezený index j_0 . Dále spočítáme $\mathbf{x}^k$ pomocí minimalizace $\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2$, ale nesmíme zapomenout, že $\mathbf{x}^k$ je nenulové pouze na pozicích S^k . To nás vede k tomu, že $\mathbf{x}_{S^k} = \mathbf{A}_{S^k}^+ \mathbf{y}$, kde $\mathbf{x}_{S^k}$ je nenulová část $\mathbf{x}$ a $\mathbf{A}_{S^k}$ je matice, která obsahuje pouze sloupce matice $\mathbf{A}$ odpovídající nosiči S^k (znak $+$ nad maticí označuje pseudoinverzi [16]). Nakonec spočítáme nové reziduum – pokud je menší než požadovaná odchylka, máme hledané řešení při dané míře nepřesnosti, jinak provedeme další iteraci [9].

Sdružovací metoda (MP)

Další metoda zvaná sdružovací metoda (Matching Pursuit) je velice podobná OMP, ale je o něco rychlejší za cenu menší přesnosti. Tato metoda také nejprve spočítá chyby $\epsilon(j)$ a nalezne index j_0 . Řídké řešení $\mathbf{x}$ ale nepočítá pomocí metody nejmenších čtverců, jako tomu bylo u OMP, ale nechává nezměněné všechny jeho nenulové složky z předchozí iterace, pouze přidá nový koeficient do nově nalezené nenulové složky vektoru. Za tento koeficient volíme optimální konstantu $z_{j_0}^*$ [9].

Normování

Oba již dříve uvedené hladové algoritmy (OMP a MP) byly popsány pro obecnou matici $\mathbf{A}$, která nemá normované sloupce. Můžeme ale normovat sloupce pomocí $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{W}$, kde $\mathbf{W}$ je diagonální matice, která má na diagonále $1/\|\mathbf{a}_i\|_2$. Potom můžeme v předchozích algoritmech použít $\tilde{\mathbf{A}}$ místo $\mathbf{A}$. Je to výhodné zejména proto, že místo hledání nejmenší chyby $\epsilon(j)$ v každé iteraci stačí nalézt největší (v absolutní hodnotě) skalární součin rezidua $\mathbf{r}^{k-1}$ a normovaného vektoru matice $\tilde{\mathbf{A}}$. To značně urychlí výpočet. Když tedy použijeme k výpočtu $\tilde{\mathbf{A}}$, nalezneme řádkový vektor $\tilde{\mathbf{x}}$. Nás obvykle ale zajímá vektor původní matice – k získání původního řádkového vektoru $\mathbf{x}$ musíme $\tilde{\mathbf{x}}$ zleva vynásobit maticí $\mathbf{W}$ [9]:

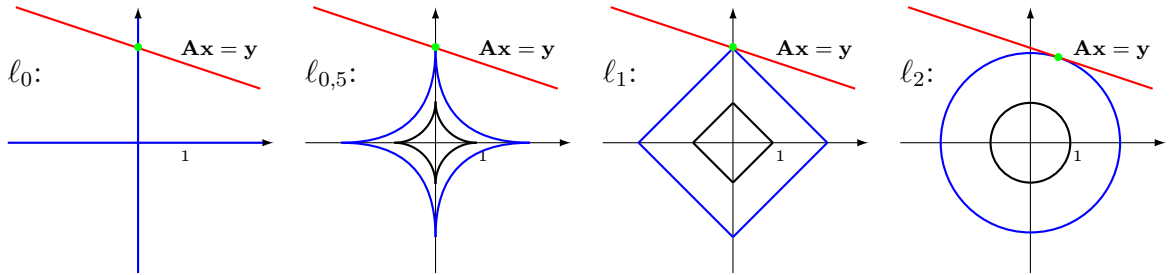
$$\mathbf{y} = \tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{A}\mathbf{W}\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x} \quad \Rightarrow \quad \mathbf{x} = \mathbf{W}\tilde{\mathbf{x}}.$$

3.2.2. Relaxace

Norma ℓ_0 není konvexní funkce, tudíž není možné na náš problém použít žádnou z řady metod a algoritmů konvexní optimalizace, které jsou dnes k dispozici. Protože ale ℓ_p normy jsou konvexní pro $p \geq 1$, nabízí se nám otázka, zda by se nedala ℓ_0 norma aproximovat nejbližší konvexní normou – tedy ℓ_1 . Tím by se náš problém převedl na úlohu:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k } \mathbf{A}\mathbf{x} = \mathbf{y}. \quad (\text{P1})$$

Jak ukážeme za chvíli, za určitých podmínek lze použít ℓ_1 normu místo ℓ_0 . Navíc se často řešení obou úloh shodují. Obr. 3.1 ilustruje řešení úlohy $\min_{\mathbf{x}} \|\mathbf{x}\|_p$ vzhledem k $\mathbf{A}\mathbf{x} = \mathbf{y}$ v $\mathbb{R}^2$, postupně pro $p = 0; 0,5; 1; 2$. Prostor všech přípustných řešení reprezentuje červená přímka. Z obrázku je patrné, že řešení problémů (P0) a (P1) je shodné. Řešení v případě, kdy použijeme eukleidovskou normu, $p = 2$, je však odlišné (jedná se o řešení s minimální energií) [14, 16].



Obrázek 3.1: Vrstevnice norem $\ell_0, \ell_{0,5}, \ell_1$ a ℓ_2 a jejich dotyk s nadrovinou určenou soustavou $\mathbf{A}\mathbf{x} = \mathbf{y}$

Podmínky ekvivalence řešení pomocí ℓ_0 - a ℓ_1 -minimalizace

Vlastnost nulového prostoru

Pro $T \subset \{1, \dots, n\}$ označíme $\mathbf{x}_T \in \mathbb{R}^n$ vektor odvozený z $\mathbf{x} \in \mathbb{R}^n$ tak, že prvky na pozicích patřících do množiny T zachováme a ostatní vynulujeme. Komplement T označíme $T^c = \{1, \dots, n\} \setminus T$. Nyní následuje definice vlastnosti nulového prostoru:

Definice 3.6. [16] Řekneme, že matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ má *vlastnost nulového prostoru řádu k* s konstantou $\gamma \in (0, 1)$, pokud platí:

$$\|\mathbf{u}_T\|_1 \leq \gamma \|\mathbf{u}_{T^c}\|_1$$

pro všechny množiny $T \subset \{1, \dots, n\}$, $|T| \leq k$ a pro všechny vektory $\mathbf{u} \in \ker \mathbf{A}$.

Vlastnost nulového prostoru umožňuje nalezení k -řádkého řešení pomocí ℓ_1 minimalizace a zaručuje jeho jednoznačnost. Ověření této podmínky ale není úplně jednoduché. Protože reálné signály nejsou řídke v pravém slova smyslu, ale mají místo nulových složek malé nenulové hodnoty, budeme potřebovat definovat chybu aproximace řídkým vektorem:

Definice 3.7. [16] *Chyba nejlepší aproximace* vektoru $\mathbf{x} \in \mathbb{R}^n$ k -řádkým vektorem $\mathbf{z}$ v normě ℓ_p je definována jako

$$\sigma_k(\mathbf{x})_p = \inf_{\mathbf{z} \in \Sigma_k^n} \|\mathbf{x} - \mathbf{z}\|_p,$$

kde $\Sigma_k^n = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_0 \leq k\}$.

Následující tvrzení udává horní odhad chyby i v ostatních případech:

Tvrzení 3.4. [16] *Nechť matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ má vlastnost nulového prostoru řádu k s konstantou $\gamma \in (0, 1)$. Nechť $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} = \mathbf{A}\mathbf{x}$ a $\mathbf{x}^* \in \mathbb{R}^n$ je řešení ℓ_1 minimalizace. Potom*

$$\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \frac{2(1 + \gamma)}{1 - \gamma} \sigma_k(\mathbf{x})_1.$$

Pokud tedy existuje nějaké nejvýše k -řádké řešení soustavy $\mathbf{y} = \mathbf{A}\mathbf{x}$ za splnění předpokladů tvrzení, pak $\sigma_k(\mathbf{x})_1 = 0$ a nulovost pravé strany vynutí $\mathbf{x} = \mathbf{x}^*$ (díky tomu, že $\|\mathbf{x} - \mathbf{x}^*\|_1 = 0$). To znamená, že hledané řídké řešení nalezneme i ℓ_1 optimalizací.

Zároveň platí obrácená implikace – pokud je možné ze soustavy $\mathbf{y} = \mathbf{A}\mathbf{x}$ rekonstruovat všechny k -řádké vektory $\mathbf{x}$ pomocí ℓ_1 -minimalizace, pak matice $\mathbf{A}$ má vlastnost nulového prostoru řádu k s nějakou konstantou $\gamma \in (0, 1)$ [16].

Vlastnost zeslabené izometrie

Další možností, jak zajistit ekvivalenci řešení pomocí ℓ_0 - a ℓ_1 -minimalizace, je definovat vlastnost zeslabené izometrie, která je oproti vlastnosti nulového prostoru výpočtově přijatelnější a navíc je stabilní i pod vlivem šumu.

Definice 3.8. [16] *Konstanta zeslabené izometrie δ_k matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je nejmenší číslo takové, že platí:*

$$(1 - \delta_k) \leq \frac{\|\mathbf{A}\mathbf{z}\|_2^2}{\|\mathbf{z}\|_2^2} \leq (1 + \delta_k)$$

pro všechny vektory $\mathbf{z} \in \Sigma_k^n$. Řekneme, že matice $\mathbf{A}$ má vlastnost zeslabené izometrie řádu k s konstantou δ_k , pokud $\delta_k \in (0, 1)$.

Zeslabení izometrie (lineární zobrazení zachovávající délku vektorů) spočívá v omezení se jen na všechny podmatice $\mathbf{A}$ o k sloupcích a dále nepožaduje přesnou izometrii,

ale povoluje malou odchylku δ_k . To znamená, že všechny matice musejí být přibližně ortonormální.

Do definice vlastnosti zeslabené izometrie je nutné obecně zahrnout všechny podmatice o maximálně k sloupcích, protože dopředu nevíme, které prvky vektoru $\mathbf{x}$ budou nenulové, a proto také nevíme, které atomy matice $\mathbf{A}$ se budou podílet na reprezentaci signálu $\mathbf{y}$.

Konstantu zeslabené izometrie můžeme spočítat přímo:

$$\delta_k = \max_{\substack{T \subset \{1, \dots, n\}, \\ |T| \leq k}} \|\mathbf{A}_T^\top \mathbf{A}_T - \mathbf{I}\|_{2 \rightarrow 2},$$

kde norma matice $\|\cdot\|_{2 \rightarrow 2}$ je definována jako:

$$\|\mathbf{B}\|_{2 \rightarrow 2} = \max_{\mathbf{x}} \frac{\|\mathbf{B}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}.$$

Podobně jako u vlastnosti nulového prostoru lze i u vlastnosti zeslabené izometrie určit chybu aproximace ℓ_1 -minimalizací. Také lze nalézt vzájemný vztah mezi vlastností zeslabené izometrie a vlastností nulového prostoru.

Tvrzení 3.5. [16] *Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$ má vlastnost zeslabené izometrie řádu $K = k + h$ s konstantou $\delta_K \in (0, 1)$. Pak $\mathbf{A}$ má vlastnost nulového prostoru řádu k s konstantou*

$$\gamma = \sqrt{\frac{k(1 + \delta_K)}{h(1 - \delta_K)}}.$$

Dále můžeme pomocí vlastnosti zeslabené izometrie omezit chybu aproximace řešení ℓ_1 -minimalizací.

Tvrzení 3.6. [16] *Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$ má vlastnost zeslabené izometrie řádu $3k$ s konstantou $\delta_{3k} < \frac{1}{3}$. Nechť $\mathbf{A}\mathbf{x} = \mathbf{y}$ pro $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{x}^* \in \mathbb{R}^n$ je řešení ℓ_1 -minimalizace. Pak*

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C \frac{\sigma_k(\mathbf{x})_1}{\sqrt{k}},$$

kde C je konstanta závisající pouze na δ_{3k} .

Tvrzení 3.7. [16] *Nechť $\mathbf{A}$ má vlastnost zeslabené izometrie řádu $2k$ s konstantou δ_{2k} , $\delta_{2k} < \sqrt{2} - 1 \approx 0,4142$. Nechť $\mathbf{A}\mathbf{x} = \mathbf{y}$ pro $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{x}^* \in \mathbb{R}^n$ je řešení ℓ_1 -minimalizace. Pak*

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C \frac{\sigma_k(\mathbf{x})_1}{\sqrt{k}}$$

a

$$\|\mathbf{x} - \mathbf{x}^*\|_1 \leq C \sigma_k(\mathbf{x})_1$$

pro nějakou konstantu C závisající pouze na δ_{2k} .

Další tvrzení specifikuje vlastnosti v případě bílého šumu (gaussovského).

Tvrzení 3.8. [16] Předpokládejme, že matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ má vlastnost zeslabené izometrie řádu $2k$ s konstantou

$$\delta_{2k} < \frac{2}{3 + \sqrt{\frac{7}{4}}} \approx 0,4627.$$

Následující vztahy platí pro všechna $\mathbf{x} \in \mathbb{R}^n$. Nechť signál je naměřen se šumem $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, $\|\mathbf{e}\|_2 \leq \epsilon$, $\mathbf{x}^*$ je řešením úlohy:

$$\min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{vzhledem k} \quad \|\mathbf{A}\mathbf{z} - \mathbf{y}\|_2 \leq \epsilon.$$

Potom

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_1 \frac{\sigma_k(\mathbf{x})_1}{\sqrt{k}} + C_2 \epsilon$$

pro kladné konstanty C_1 a C_2 závisující pouze na δ_{2k} .

Z uvedeného tvrzení vyplývá, že celková chyba má 2 části: část, která závisí pouze na řídkosti, a část, která závisí na velikosti šumu.

Matice, které by měly vlastnost zeslabené izometrie s předem danými parametry, se zatím nepodařilo deterministicky sestavit. Byly nalezeny pouze matice, které tuto vlastnost splňují jen s vysokou pravděpodobností. Uvedená tvrzení však neříkají nic o tom, jestli je ℓ_1 -minimalizace ekvivalentní s ℓ_0 -minimalizací v případě, kdy matice $\mathbf{A}$ nemá vlastnost zeslabené izometrie. Experimentálně bylo zjištěno, že lze použít i některé matice, které tuto vlastnost nemají [16].

Iterativně převažovaná metoda nejmenších čtverců (IRLS)

Jak již bylo řečeno, další možností, jak nalézt řídké řešení soustavy $\mathbf{A}\mathbf{x} = \mathbf{y}$, je použít relaxaci ℓ_0 normy. Speciálně pro tyto účely byla vyvinuta iterativně převažovaná metoda nejmenších čtverců (Iterative Reweighted Least Squares), která reprezentuje ℓ_p normu (pro pevně dané $p \in (0, 1)$) pomocí vážené ℓ_2 normy. Předpokládejme, že máme určené přibližné řešení $\mathbf{x}_{k-1}$ a položíme $\mathbf{X}_{k-1} = \text{diag}(|\mathbf{x}_{k-1}|^q)$. Uvažujme, že tato matice je invertibilní a $\|\mathbf{X}_{k-1}^{-1}\mathbf{x}\|_2^2$ je rovno $\|\mathbf{x}\|_{2-2q}^{2-2q}$. Jestliže zvolíme $q = 1 - p/2$, uvedený výraz nahrazuje ℓ_p normu, $\|\mathbf{x}\|_p^p$. Není-li matice $\mathbf{X}_{k-1}$ invertibilní, použijeme místo inverze $\mathbf{X}_{k-1}$ pseudo-inverzi, $\|\mathbf{X}_{k-1}^+\mathbf{x}\|_2^2$. Pokusíme se nyní řešit tento problém:

$$\min_{\mathbf{x}} \|\mathbf{X}_{k-1}^+\mathbf{x}\|_2^2 \quad \text{vzhledem k} \quad \mathbf{A}\mathbf{x} = \mathbf{y}.$$

Tento problém má tu výhodu, že jej lze řešit pomocí lineární algebry a Lagrangeových multiplikátorů. Jako účelovou funkci využijeme klasickou ℓ_2 normu. Tedy máme

$$L(\mathbf{x}) = \|\mathbf{X}_{k-1}^+\mathbf{x}\|_2^2 + \lambda^\top (\mathbf{y} - \mathbf{A}\mathbf{x})$$

$$\Rightarrow \frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{0} = 2(\mathbf{X}_{k-1}^+)^2 \mathbf{x} - \mathbf{A}^\top \lambda.$$

Pokud budeme uvažovat, že $\mathbf{X}_{k-1}^+$ je invertibilní, znamená to, že má nenulové prvky na hlavní diagonále [9]. Můžeme proto položit $(\mathbf{X}_{k-1}^+)^{-1} = \mathbf{X}_{k-1}$. V tomto případě můžeme psát řešení ve tvaru:

$$\mathbf{x}_k = \frac{1}{2} \mathbf{X}_{k-1}^2 \mathbf{A}^\top \lambda.$$

V obecném případě, kdy jsou některé prvky na hlavní diagonále $\mathbf{X}_{k-1}$ nulové, řešení dané těmito nulovými prvky odpovídá složkám v $\mathbf{x}_k$. Za předpokladu, že nulové složky $\mathbf{x}_k$ mohou zůstat v dalších iteracích (kvůli řídkosti řešení), tento vzorec na výpočet $\mathbf{x}_k$ je smysluplný a dává správný výsledek. Nyní se vrátíme k rovnici $\mathbf{A}\mathbf{x} = \mathbf{y}$, do které dosadíme právě spočítané $\mathbf{x}_k$. Dostaneme:

$$\frac{1}{2} \mathbf{A} \mathbf{X}_{k-1}^2 \mathbf{A}^\top \lambda = \mathbf{y} \quad \Rightarrow \quad \lambda = 2 (\mathbf{A} \mathbf{X}_{k-1}^2 \mathbf{A}^\top)^{-1} \mathbf{y}.$$

Protože inverze obecně nemusí existovat, raději použijeme pseudoinverzi. Celkově tedy máme:

$$\mathbf{x}_k = \mathbf{X}_{k-1}^2 \mathbf{A}^\top (\mathbf{A} \mathbf{X}_{k-1}^2 \mathbf{A}^\top)^+ \mathbf{y}.$$

Přibližné řešení $\mathbf{x}_k$ použijeme pro vytvoření nové diagonální matice $\mathbf{X}_k$ a můžeme začít novou iteraci.

V praxi se ukázalo, že jakmile je nějaká složka vektoru $\mathbf{x}_k$ nulová, v algoritmu nikdy nedojde k tomu, že by byla nahrazena nenulovou složkou. Proto se za startovací vektor nebere vektor nulový, ale vektor, jehož složky jsou rovny jedné. Na rozdíl od předchozích algoritmů výpočet řešení končí, jakmile $\|\mathbf{x}_k - \mathbf{x}_{k-1}\|_2$ je menší než zadaná požadovaná přesnost. Nejčastěji volíme $p = 1$.

Nesmíme ovšem zapomenout, že zatímco ℓ_0 norma nerozlišuje velikost nenulových složek $\mathbf{x}_k$, ℓ_p norma na velikosti složek přímo závisí. Proto algoritmus vybírá za nenulové složky vektoru $\mathbf{x}_k$ ty, které násobí v ℓ_p normě největší atomy matice $\mathbf{A}$, a tím je přesné řešení ovlivněno. Z tohoto důvodu se matice $\mathbf{A}$ před vlastním výpočtem normuje. Aby bylo nalezené řídké řešení normované matice $\mathbf{A}$ ekvivalentní s řešením původní matice, na konci výpočtu (stejně jako tomu bylo u hladových algoritmů) se musí po složkách vydělit příslušnými normami matice $\mathbf{A}$ [9].

Modifikované IRLS – s uvažováním odchylky od přesného řešení

Jak již bylo řečeno dříve, signál může být zašuměný a přesné řešení nemusí existovat. V případě relaxace se tedy pokoušíme řešit tento problém:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon.$$

Využitím příslušných Lagrangeových multiplikátorů dostaneme:

$$\min_{\mathbf{x}} \lambda \|\mathbf{x}\|_1 + \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2,$$

kde Lagrangeův multiplikátor λ je funkcí $\mathbf{A}$, $\mathbf{y}$ a ϵ . V případě IRLS položíme $\mathbf{X} = \text{diag}(|\mathbf{x}|)$ a dostaneme $\|\mathbf{x}\|_1 = \mathbf{x}^\top \mathbf{X}^{-1} \mathbf{x}$. Jestliže máme danou aproximaci $\mathbf{x}_{k-1}$, nastavíme $\mathbf{X}_{k-1} = \text{diag}(|\mathbf{x}_{k-1}|)$ a pokusíme se řešit:

$$\min_{\mathbf{x}} \lambda \mathbf{x}^\top \mathbf{X}^{-1} \mathbf{x} + \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2,$$

což je kvadratický optimalizační problém. Pokud účelovou funkci zderivujeme a položíme rovnou nule, dostaneme soustavu lineárních rovnic, kterou lze řešit pomocí klasické lineární algebry.

Největším problémem je zvolení λ tak, aby nalezené řešení co nejvíce odpovídalo přesnému řešení. Volba λ závisí na směrodatné odchylce i řídкости řešení. Někdy se také λ aproximuje a postupně se iterativně zpřesňuje. Vzniklý systém lineárních rovnic se často řeší pomocí metody sdružených gradientů [9].

4. Komprimované snímání

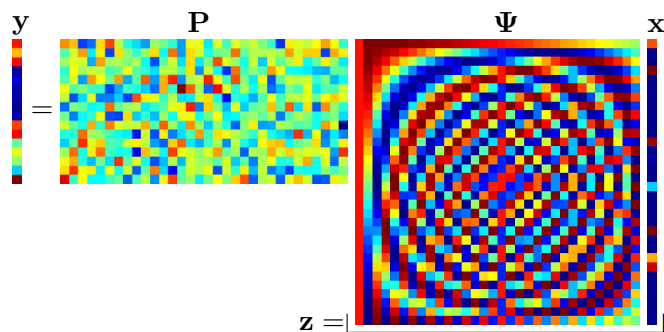
Běžný způsob získání a komprese dat je adaptivní vzhledem k signálu. Nejprve naměříme všechna data, potom provedeme nějakou transformaci a vyhodnotíme získané koeficienty. Protože chceme signál komprimovat, aby nám nezabral moc místa v paměti, většinu koeficientů dále neuložíme (úplně je „zahodíme“). Můžeme si to dovolit, protože z nějakého důvodu nesou málo informace. Typickým příkladem tohoto přístupu je kompresní formát JPEG, u kterého nejprve získáme hodnoty všech pixelů fotografie, potom provedeme dvojrozměrnou diskretní kosinovou transformaci (DCT) a kvantujeme vzniklé koeficienty. Nakonec ponecháme pouze některé koeficienty (nejčastěji ty s největší velikostí) a zakódujeme jenom jejich pozice/indexy a velikost.

Komprimované snímání (compressed sensing, compressive sampling, CS) přichází s jinou strategií: snímat signál lineárně a neadaptivně, a to pouze tolikrát, kolik je skutečně třeba! Tento přístup je asymptoticky stejně dobrý jako adaptivní přístup. Komprimované snímání je možné provádět jenom díky nějakým apriorním informacím o měřeném signálu. Umožní nám urychlit měřicí proces, ale cenou za to je delší doba rekonstrukce, protože není lineární [3, 15].

4.1. Komprimované snímání využívající řídkosti

Nejčastěji využívaným předpokladem v komprimovaném snímání je řídkost signálu v nějaké vhodné reprezentaci (omezíme se pouze na ortonormální báze). V komprimovaném snímání jde o stejný problém jako (P0), ale matice $\mathbf{A}$ má speciální konstrukci.

Označíme bázi Ψ a tedy signál $\mathbf{z}$ má vyjádření $\mathbf{z} = \Psi\mathbf{x}$, kde $\mathbf{x}$ je k -řídký. Cílem je provést „malý počet“ neadaptivních měření, které budou mít charakter skalárních součinů se signálem, což lze zapsat jako $\mathbf{y} = \mathbf{P}\mathbf{z} = \mathbf{P}\Psi\mathbf{x}$. Zde $\mathbf{P}$ je tzv. měřicí matice rozměru $m \times n$ a jednotlivé složky vektoru $\mathbf{y}$ jsou výsledky měření, které vznikají jako lineární kombinace



Obrázek 4.1: Ilustrační schéma komprimovaného snímání pro řídký signál: naměřený vektor $\mathbf{y}$ je roven součinu měřicí matice $\mathbf{P}$, unitární matice Ψ a řídkého vektoru $\mathbf{x}$. Do procesu snímání vstupuje vektor $\mathbf{z}$, který je pozorovatelný a není sám řídký, ale je řídký v nějaké bázi Ψ , zde v ortonormální bázi zpětné DCT. Obrázek je proveden v pseudobarevné škále, tzn. nejnižší hodnotě v maticích je přiřazena modrá barva, nejvyšší červená a hodnoty mezi nimi jsou znázorněny barvami přecházejícími postupně od modré k červené podle zvolené palety barev.

vzorků signálu. Počet měření je $m \ll n$. Je vidět, že v komprimovaném snímání matice $\mathbf{A} = \mathbf{P}\Psi$, celkově máme tento problém:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ vzhledem k } \mathbf{y} = \mathbf{P}\Psi\mathbf{x}. \quad (4.1)$$

Pokud označíme $\tilde{\mathbf{x}}$ bod, ve kterém uvedená účelová funkce nabývá svého minima, můžeme pak po jeho nalezení jednoduše získat samotný signál $\mathbf{z}$ jako $\mathbf{z} = \Psi\tilde{\mathbf{x}}$. Ilustrační schéma komprimovaného snímání je znázorněné na obr. 4.1.

Měřicí matice se většinou uvažují ve tvaru $\mathbf{P} = \mathbf{R}\Phi$. Zde Φ je (zatím nespecifikovaná) matice $n \times n$ a $\mathbf{R}$ je matice, která vznikne z jednotkové matice $n \times n$ ponecháním pouze některých (náhodně) vybraných řádků, a tedy funguje jako výběr řádků z Φ . Náhodný výběr v $\mathbf{R}$ se přitom řídí nejčastěji rovnoměrným rozdělením pravděpodobnosti. Tedy celkově roli matice $\mathbf{A}$ hraje nyní matice $\mathbf{A} = \mathbf{R}\Phi\Psi$. Celkové schéma komprimovaného snímání je zobrazené na obr. 4.2.

Nyní zbývá zodpovědět otázku, kolik je třeba provést měření tak (tj. jaký musí být počet řádků m matice $\mathbf{P}$), abychom byli schopni z nich úspěšně rekonstruovat signál pomocí ℓ_1 -technik. V případě náhodných měřicích matic ($\mathbf{R}$) s rovnoměrným rozdělením pravděpodobnosti závisí tento počet na vzájemné koherenci μ . Ve speciálním případě, kdy je matice složena ze dvou ortonormálních bází Φ a Ψ , tedy $[\Phi, \Psi]$, máme

$$\mu([\Phi, \Psi]) = \max_{1 \leq i, j \leq n} |\Psi_i^\top \Phi_j|.$$

Hodnota $\mu([\Phi, \Psi])$ se pohybuje mezi $\frac{1}{\sqrt{n}}$ a 1. Následující tvrzení uvádí podmínku, za jaké je zaručena přesná rekonstrukce z měření.

Tvrzení 4.1. [3] *Nechť je dán signál $\mathbf{z}$, který má v bázi Ψ k -řádkou reprezentaci $\mathbf{x}$. Pak řešení ℓ_1 -minimalizace*

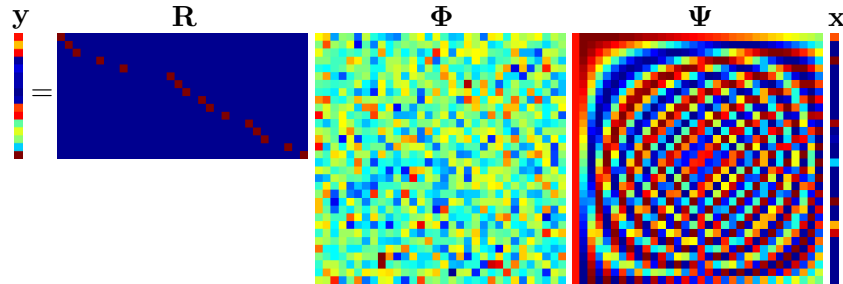
$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ vzhledem k } \mathbf{y} = \mathbf{R}\Phi\Psi\mathbf{x},$$

kde $\mathbf{y}$ jsou měření, je současně s vysokou pravděpodobností nejřidší možné, pokud je zvolen počet řádků m matice $\mathbf{R}$ takto:

$$m \geq C \cdot \mu^2([\Phi, \Psi]) \cdot k \cdot n \cdot \ln n$$

pro nějakou kladnou konstantu C .

Z uvedeného tvrzení je patrné, že počet měření závisí na řídkosti pouze lineárně. Koherence kvadraticky ovlivňuje nutný počet měření. Proto je snaha hledat takové dvojice, jejichž koherence je minimální – u párů $[\Phi, \Psi]$ s koherencí $\frac{1}{\sqrt{n}}$ stačí řádově $k \cdot \ln n$ měření. Naopak, jak koherence roste, přestává být měření dle této podmínky výhodné, až v určitém momentu přestává mít smysl, neboť přeroste počet vzorků signálu n [3, 14, 15].



Obrázek 4.2: Celkové schéma komprimovaného snímání: součin matic $\mathbf{R}$ a Φ tvoří měřicí matici $\mathbf{P}$

4.2. Komprimované snímání založené na nízké maticové hodnotnosti

Komprimované snímání můžeme použít také na matice $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}(\mathbb{C}^{n_1 \times n_2})$. Řídkost je zde nahrazena předpokladem nízké hodnotnosti $\mathbf{X}$. Měření provádíme pomocí nějakého lineárního zobrazení $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}^m$, kde $m < n_1 n_2$, tedy naměřená data $\mathbf{y}$ dostaneme jako

$$\mathbf{y} = \mathcal{A}(\mathbf{X}) \in \mathbb{R}^m.$$

Abychom byli schopni rekonstruovat $\mathbf{X}$ z $\mathbf{y}$, budeme předpokládat, že $\mathbf{X}$ má hodnotnost nejvýše $h(\mathbf{X}) = r \ll \min\{n_1, n_2\}$. Naivní přístup řešení tohoto optimalizačního problému

$$\min h(\mathbf{X}) \text{ vzhledem k } \mathbf{y} = \mathcal{A}(\mathbf{X})$$

je NP-těžký problém, ale můžeme ho relaxovat podobně jako v předchozím případě. Uvažujme singulární rozklad (SVD) matice $\mathbf{X}$, tj.

$$\mathbf{X} = \sum_{l=1}^n \sigma_l \mathbf{u}_l \mathbf{v}_l^*,$$

kde $n = \min\{n_1, n_2\}$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ jsou singulární čísla $\mathbf{X}$ a $\mathbf{u}_l \in \mathbb{R}^{n_1}$, $\mathbf{v}_l \in \mathbb{R}^{n_2}$ jsou popořadě levé a pravé singulární vektory. Matice $\mathbf{X}$ má hodnotnost r právě tehdy, když vektor $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{X})$ singulárních čísel je r -řídský, tj. $h(\mathbf{X}) = \|\boldsymbol{\sigma}(\mathbf{X})\|_0$. Podobně jako tomu bylo předtím, můžeme tento problém také relaxovat. Definujeme proto tzv. nukleární normu jako ℓ_1 -normu singulárních čísel, tj.

$$\|\mathbf{X}\|_* = \|\boldsymbol{\sigma}(\mathbf{X})\|_1 = \sum_{l=1}^n \sigma_l(\mathbf{X}). \quad (4.2)$$

Nyní můžeme uvažovat relaxovaný minimalizační problém s nukleární normou

$$\min \|\mathbf{X}\|_* \text{ vzhledem k } \mathbf{y} = \mathcal{A}(\mathbf{X}). \quad (4.3)$$

Tento problém je už konvexní a je tedy efektivně řešitelný. Konkrétní metoda na řešení tohoto problému a jemu podobných je uvedena v další kapitole.

Na řešitelnost a jednoznačnost tohoto problému byla vyvinuta podobná teorie jako na řídké signály. Uvažují se při tom náhodná zobrazení $\mathcal{A}$. Ukazuje se, že matice $\mathbf{X}$ s hodnotností nejvýše r mohou být s vysokou pravděpodobností rekonstruovány z m měření [10], jestliže

$$m \geq Cr \max\{n_1, n_2\}.$$

Tato hranice je optimální, jestliže pravá strana uvedené nerovnice odpovídá počtu stupňů volnosti popisujících matici typu $n_1 \times n_2$ s hodnotností r . Existují také jiné podmínky využívající např. koherenci matic, ale ty zde neuvádíme. Všechny podmínky ale pouze zaručují vysokou pravděpodobnost rekonstrukce. Například kdybychom měřili jednotlivé prvky matice, která má pouze jeden nenulový prvek, pak bychom ho s vysokou pravděpodobností naměřili, a proto by jeho určení bylo skoro nemožné [2, 10].

4.3. Komprimované snímání v MRI

Měření v MRI je velmi pomalé. Například při perfúzním zobrazování v MRI lze nyní snímat jeden řez v maximálním rozlišení přibližně 100×100 px asi za 4 s. Vyšší rozlišení není možné, protože bychom nebyli schopni díky délce trvání zachytit průběh koncentrace kontrastní látky. Proto se zde přímo nabízí použití komprimovaného snímání na zrychlení procesu měření. Jedinou nevýhodou je, že rekonstrukce obrazu bude trvat déle.

Abychom mohli využít komprimované snímání, je nutno využít nějakých apriorních informací o měřených datech. Například můžeme využít faktu, že perfúzní křivka je řídká ve spektru, video (obraz měnící své parametry v čase díky perfúzi) má nízkou maticovou hodnotu (perfúzní křivky mají podobný průběh v rámci jednoho typu tkáně) nebo malé totální variace perfúzních křivek.

Připomeňme, že v MRI je signál měřen přímo ve frekvenční oblasti. Mění se gradientní pole určuje trajektorii v k -prostoru, proto matice Φ je daná – je to přímo 2D (3D) Fourierova transformace. Díky tomu každé měření představuje jeden Fourierův koeficient. Protože je matice Φ daná, je důležité se zaměřit na konstrukci měřících matic pomocí výběru řádků z Φ . Ale ani ty nemohou být libovolné, protože v k -prostoru nemohou být jakékoliv trajektorie kvůli fyzikálním a fyziologickým omezením. V dnešních systémech jsou gradienty limitovány svojí maximální amplitudou ($|G| < G_{\max}$) a maximální rychlostí jejich náběhu ($|\frac{dG}{dt}| < S_{\max}$). Navíc velké amplitudy gradientu a jejich rychlé změny mohou způsobovat stimulaci periferní nervové soustavy [18].

5. Proximální gradientní metoda

V předchozí kapitole jsem viděli, že při používání komprimovaného snímání lze rekonstruovaná data vyjádřit jako řešení vhodné optimalizační úlohy (např. viz (4.1) nebo (4.3)). Pokud využijeme CS pouze s předpokladem řidkosti (úloha (4.1)), můžeme k vyřešení této optimalizační úlohy využít algoritmy pro získání řídké reprezentace uvedené v části 3.2.

Pokud použijeme další předpoklady, musíme se potýkat s otázkou, jakou metodu použít k vyřešení sestavené optimalizační úlohy pro rekonstrukci naměřených dat. Jednou z možností je využití proximální gradientní metody, která je dostatečně flexibilní na tento druh problémů. Dříve než si popíšeme samotný algoritmus, představíme si proximální operátory, které proximální gradientní metoda využívá.

5.1. Proximální operátor

Proximální operátor konvexní funkce je přirozeným rozšířením pojmu projekčního operátoru na konvexní množinu. Tento operátor hraje důležitou roli v analýze a numerickém řešení konvexních optimalizačních problémů.

Dále budeme potřebovat definovat několik pojmů.

Definice 5.1. [28] Řekneme, že funkce $f(\mathbf{x})$ je *zdola polospojita* na podmnožině U nějakého metrického prostoru, jestliže pro libovolné $\mathbf{x}_0 \in U$ platí:

$$\liminf_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \geq f(\mathbf{x}_0).$$

Definice 5.2. [4] Nechť f je zdola polospojita konvexní funkce s neprázdným definičním oborem, $f : \mathbb{R}^n \rightarrow (-\infty, \infty)$. *Subdiferenciál* funkce f je operátor $\partial f : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$, který definujeme jako

$$\mathbf{x} \mapsto \left\{ \mathbf{u} \in \mathbb{R}^n : (\forall \tilde{\mathbf{x}} \in \mathbb{R}^n) \quad (\tilde{\mathbf{x}} - \mathbf{x})^\top \mathbf{u} + f(\mathbf{x}) \leq f(\tilde{\mathbf{x}}) \right\}. \quad (5.1)$$

Jednotlivé prvky $\mathbf{u}$ subdiferenciálu se nazývají *subgradients*.

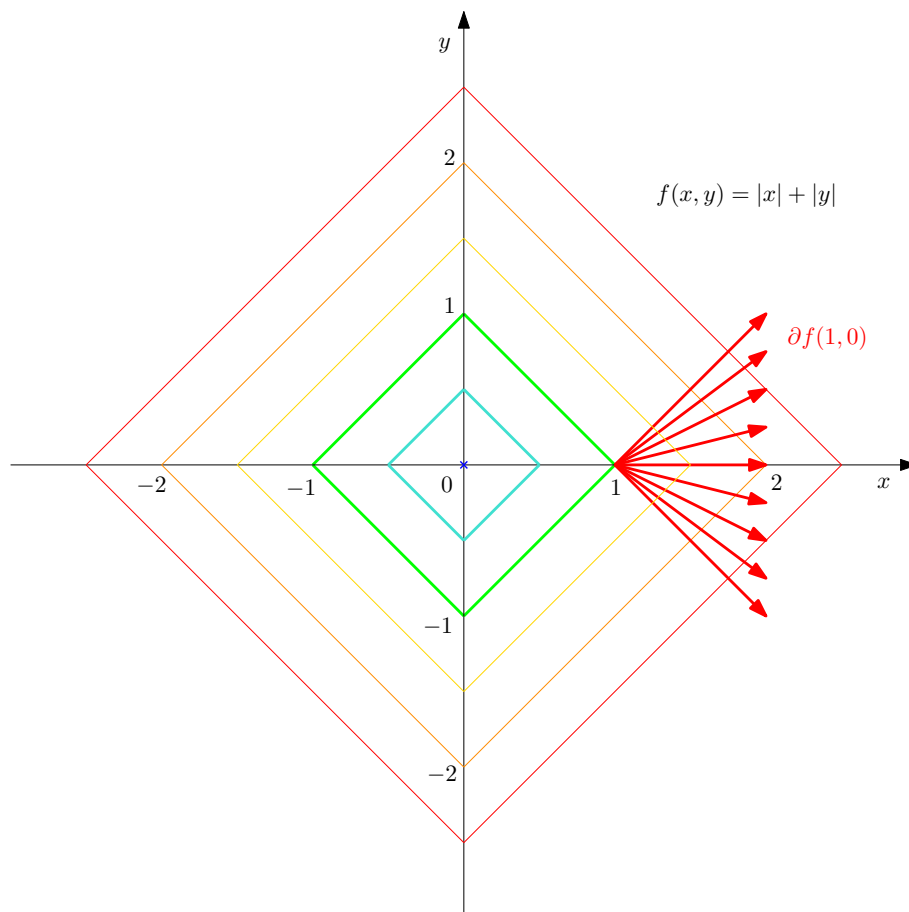
Subgradient ∂f představuje zobecněný pojem gradientu $\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$. V bodech $\mathbf{x}$, kde je funkce f diferencovatelná, se subgradient rovná gradientu.

Příklad: Nalezněte subdiferenciál funkce $f(x, y) = |x| + |y|$ v bodě $[x, y] = [1, 0]$.

Řešení: Napíšeme podmínku (5.1) pro subdiferenciál:

$$\begin{aligned} (\tilde{\mathbf{x}} - \mathbf{x})^\top \mathbf{u} + f(\mathbf{x}) &\leq f(\tilde{\mathbf{x}}), \\ (\tilde{x} - x, \tilde{y} - y) \begin{pmatrix} u \\ v \end{pmatrix} + |x| + |y| &\leq |\tilde{x}| + |\tilde{y}|, \\ (\tilde{x} - 1, \tilde{y}) \begin{pmatrix} u \\ v \end{pmatrix} + 1 &\leq |\tilde{x}| + |\tilde{y}|, \\ (\tilde{x} - 1)u + \tilde{y}v + 1 &\leq |\tilde{x}| + |\tilde{y}|. \end{aligned}$$

Protože uvedená nerovnost musí platit pro všechna u, v , lze odsud jednoduše určit (rozložením jednotlivých možností kombinací znamének $u, \tilde{x}, \tilde{y}$), že subgradients jsou vektory tvaru $(u, v) = (1, v)$, přičemž $-1 \leq v \leq 1$. Subdiferenciál funkce $f(x, y) = |x| + |y|$ v bodě $[1, 0]$ je znázorněn na obr. 5.1.



Obrázek 5.1: Subdiferenciál funkce $f(x, y) = |x| + |y|$ v bodě $[1, 0]$ – barevně jsou vykresleny vrstevnice f , červené vektory představují vybrané subgradienty

Definice 5.3. [4] Nechť C je neprázdná podmnožina $\mathbb{R}^n$. Potom *indikátorová (charakteristická) funkce množiny C* je

$$\iota_C : \mathbf{x} \mapsto \begin{cases} 0, & \mathbf{x} \in C, \\ \infty, & \mathbf{x} \notin C. \end{cases}$$

Nyní zavedeme proximální operátor.

Definice 5.4. [4] Nechť $f : \mathbb{R}^n \rightarrow (-\infty, \infty)$ je zdola polospojité konvexní funkce s neprázdným definičním oborem. Pro každé $\mathbf{x} \in \mathbb{R}^n$, minimalizační problém

$$\min_{\mathbf{y} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + f(\mathbf{y})$$

má jediné řešení. Bod, ve kterém uvedená účelová funkce nabývá minima, označíme $\text{prox}_f(\mathbf{x})$. Operátor $\text{prox}_f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ takto definovaný nazveme *proximální operátor f* .

Je zajímavé si povšimnout, že proximální operátor je obecnějším pojmem než projekce P_C bodu $\mathbf{x} \in \mathbb{R}^n$ na neprázdnou uzavřenou konvexní množinu $C \subset \mathbb{R}^n$. Tato projekce je řešením problému

$$\min_{\mathbf{y} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \iota_C(\mathbf{y}).$$

Navíc také funkce ι_C splňuje předpoklady v definici proximálního operátoru.

Proximální operátor f můžeme charakterizovat touto inkluzí:

$$(\forall (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^n \times \mathbb{R}^n) \quad \mathbf{p} = \text{prox}_f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \mathbf{p} \in \partial f(\mathbf{p}), \quad (5.2)$$

která vychází z podmínek pro minimum: $\mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) \Rightarrow \mathbf{0} \in \partial f$. Inkluzi je možno redukovat na

$$(\forall (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^n \times \mathbb{R}^n) \quad \mathbf{p} = \text{prox}_f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \mathbf{p} = \nabla f(\mathbf{p}),$$

když je f diferencovatelná.

Věta 5.1 (Vlastnosti proximálního operátoru). [5] *Nechť φ je zdola polospojité konvervní funkce s neprázdným definičním oborem, $\varphi : \mathbb{R}^n \rightarrow (-\infty, \infty)$. Pak platí následující:*

(i) *Posunutí: Nechť $\psi = \varphi(\cdot - \mathbf{z})$, kde $\mathbf{z} \in \mathbb{R}^n$. Potom $\text{prox}_\psi(\mathbf{x}) = \mathbf{z} + \text{prox}_\varphi(\mathbf{x} - \mathbf{z})$.*

(ii) *Změna měřítka: Nechť $\psi = \varphi\left(\frac{\cdot}{\rho}\right)$, kde $\rho \in \mathbb{R} \setminus \{0\}$. Potom $\text{prox}_\psi(\mathbf{x}) = \rho \text{prox}_{\frac{\varphi}{\rho^2}}\left(\frac{\mathbf{x}}{\rho}\right)$.*

Důkaz. Položíme $\mathbf{p} = \text{prox}_\psi(\mathbf{x})$. Podle (5.2) je to ekvivalentní s $\mathbf{x} - \mathbf{p} \in \partial \psi(\mathbf{p})$. Postupně můžeme psát:

(i)

$$\begin{aligned} \mathbf{x} - \mathbf{p} &\in \partial \psi(\mathbf{p}), \\ \mathbf{x} - \mathbf{p} &\in \partial \varphi(\mathbf{p} - \mathbf{z}), \\ (\mathbf{x} - \mathbf{z}) - (\mathbf{p} - \mathbf{z}) &\in \partial \varphi(\mathbf{p} - \mathbf{z}), \\ \mathbf{p} - \mathbf{z} &= \text{prox}_\varphi(\mathbf{x} - \mathbf{z}), \\ \mathbf{p} = \text{prox}_\psi(\mathbf{x}) &= \mathbf{z} + \text{prox}_\varphi(\mathbf{x} - \mathbf{z}). \end{aligned}$$

(ii)

$$\begin{aligned} \mathbf{x} - \mathbf{p} &\in \partial \psi(\mathbf{p}), \\ \mathbf{x} - \mathbf{p} &\in \frac{1}{\rho} \partial \varphi\left(\frac{\mathbf{p}}{\rho}\right), \\ \frac{\mathbf{x}}{\rho} - \frac{\mathbf{p}}{\rho} &\in \partial \left(\frac{\varphi}{\rho^2}\right)\left(\frac{\mathbf{p}}{\rho}\right), \\ \mathbf{p} = \text{prox}_\psi(\mathbf{x}) &= \rho \text{prox}_{\frac{\varphi}{\rho^2}}\left(\frac{\mathbf{x}}{\rho}\right). \end{aligned}$$

□

Z hlediska zpracování signálů mají proximální operátory jednoduchou interpretaci jako odstraňování šumu (tzv. denoising). Předpokládejme standardní problém odstraňování šumu, kdy chceme rekonstruovat signál $\tilde{\mathbf{x}} \in \mathbb{R}^n$ z pozorování

$$\mathbf{y} = \tilde{\mathbf{x}} + \mathbf{w},$$

kde $\mathbf{w} \in \mathbb{R}^n$ je šum modelu. Tento problém můžeme zformulovat do podoby proximálního operátoru, kde $\|\cdot - \mathbf{y}\|^2/2$ zaručuje malou odchylku od dat se šumem a f představuje apriorní informace o $\tilde{\mathbf{x}}$ [4].

5.1.1. Proximální operátor pro ℓ_1 -normu

Nyní si ukážeme, jak odvodit proximální operátor pro ℓ_1 -normu. Podle definice můžeme napsat:

$$\text{prox}_{\lambda \|\cdot\|_1}(\mathbf{x}) = \arg \min_{\mathbf{y}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{y}\|_1,$$

což rozepíšeme jako:

$$\min_{\mathbf{y}} \underbrace{\frac{1}{2} \sum_{i=1}^n (x_i - y_i)^2 + \lambda \sum_{i=1}^n |y_i|}_{f(\mathbf{y})}.$$

Budeme hledat toto minimum postupně na jednotlivých intervalech, kde je absolutní hodnota diferencovatelná:

1) $y_i > 0$

$$\frac{\partial f}{\partial y_i} = y_i - x_i + \lambda = 0,$$

$$y_i = x_i - \lambda > 0 \quad \Rightarrow \quad x_i > \lambda,$$

2) $y_i < 0$

$$\frac{\partial f}{\partial y_i} = y_i - x_i - \lambda = 0,$$

$$y_i = x_i + \lambda < 0 \quad \Rightarrow \quad x_i < -\lambda.$$

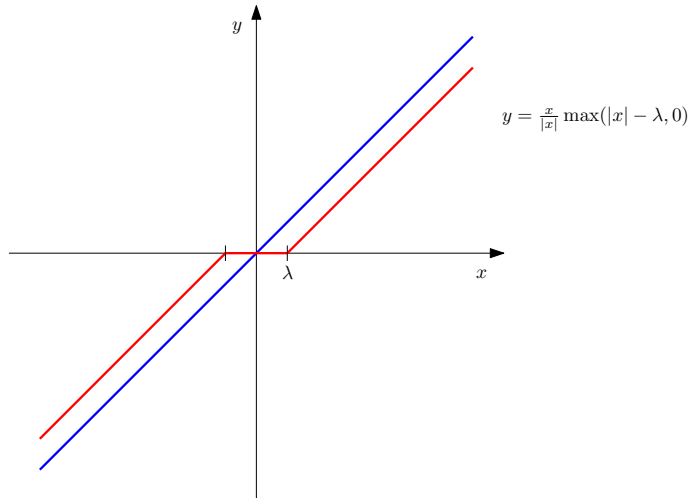
3) Jediným případem, který jsme neuvažovali, je $y_i = 0$, které je hledaným minimem pro zbylé případy x_i , tedy $-\lambda \leq x_i \leq \lambda$.

Celkově tedy můžeme proximální operátor ℓ_1 -normy napsat jako

$$y_i = \frac{x_i}{|x_i|} \max(|x_i| - \lambda, 0).$$

Tato funkce se taky někdy nazývá měkké prahování, které je znázorněné na obr. 5.2.

Protože nukleární norma je definovaná jako ℓ_1 -norma singulárních čísel (viz (4.2)), je tedy proximální operátor nukleární normy měkké prahování singulárních čísel.



Obrázek 5.2: Červeně znázorněna identita a modře její měkké prahování

5.1.2. Proximální operátor pro 1D totální variaci

Proximální operátor 1D totální variace napíšeme podle definice takto:

$$\text{prox}_{\lambda\|\cdot\|_{\text{tv}}}(\mathbf{x}) = \arg \min_{\mathbf{y}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{y}\|_{\text{tv}},$$

což rozepíšeme jako:

$$\min_{\mathbf{y}} \frac{1}{2} \sum_{i=1}^n (x_i - y_i)^2 + \lambda \sum_{i=1}^{n-1} |y_{i+1} - y_i|. \quad (5.3)$$

Tento proximální operátor se nedá analyticky vyjádřit jako tomu bylo u ℓ_1 -normy. Na jeho spočítání existují různé algoritmy, které jsou většinou iterativní. Nedávno se objevil i přímý algoritmus [6] na jeho výpočet. Tento algoritmus využívá duální problém k (5.3):

$$\min_{\mathbf{u} \in \mathbb{R}^{n+1}} \sum_{i=1}^n (x_i - u_i + u_{i-1})^2$$

$$\text{vzhledem k } |u_i| \leq \lambda, \forall i = 1, \dots, n-1, \text{ a } u_0 = u_n = 0.$$

Pokud najdeme řešení $\mathbf{u}$ duální úlohy, lze jednoduše dostat řešení $\mathbf{y}$ primární úlohy jako

$$y_i = x_i - u_i + u_{i-1}, \quad \forall i = 1, \dots, n. \quad (5.4)$$

Řešení $\mathbf{y}$ a $\mathbf{u}$ je jednoznačně určeno Karush-Kuhn-Tuckerovými podmínkami [1], které vedou na

$$\begin{aligned} u_0 = u_n = 0 \text{ a } \forall i = 1, \dots, n-1 \\ \left\{ \begin{array}{ll} u_i \in \langle -\lambda, \lambda \rangle & \text{pokud } y_i = y_{i+1}, \\ u_i = -\lambda, & \text{pokud } y_i < y_{i+1}, \\ u_i = \lambda, & \text{pokud } y_i > y_{i+1}. \end{array} \right. \end{aligned} \quad (5.5)$$

Algoritmus prochází postupně jednotlivé prvky y_i . Na pozici i se snaží prodloužit stávající segment $\mathbf{y}$ položením $y_{i+1} = y_i$. Pokud to není možné bez porušení (5.4) a (5.5), algoritmus jde zpět na poslední pozici, kde se objevil skok v $\mathbf{y}$, ověří stávající segment do této pozice, začne s novým segmentem a pokračuje stejným způsobem dál, dokud nedojde do konce.

5.2. Proximální gradientní metoda (dopředné-zpětné dělení)

Máme danou konvexní úlohu

$$\min_{\mathbf{x}} g(\mathbf{x}) + h(\mathbf{x}), \quad (5.6)$$

přičemž $h : \mathbb{R}^n \rightarrow \mathbb{R}$ je zdola polospojité konvexní funkce a $g : \mathbb{R}^n \rightarrow \mathbb{R}$ je konvexní funkce, která je diferencovatelná s β -lipschitzovským spojitým gradientem ∇g , tj.

$$(\forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n) \quad \|\nabla g(\mathbf{x}) - \nabla g(\mathbf{y})\| \leq \beta \|\mathbf{x} - \mathbf{y}\|, \quad (5.7)$$

kde $\beta \in (0, \infty)$. Tato úloha má alespoň jedno řešení, pokud $g + h$ je koercivní, tj.

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} g(\mathbf{x}) + h(\mathbf{x}) = \infty.$$

Pokud navíc $g + h$ je ryze konvexní (tedy g nebo h je ryze konvexní), pak je toto řešení jediné.

Tvrzení 5.1. *Nechť $\mathbf{x} \in \mathbb{R}^n$ a $t \in (0, \infty)$. Pak*

$$\mathbf{x} \text{ je řešením (5.6)} \Leftrightarrow \mathbf{x} = \text{prox}_{th}(\mathbf{x} - t\nabla g(\mathbf{x})).$$

Tato charakterizace $\mathbf{x}$ jakožto pevného bodu naznačuje možnost řešení problému (5.6) iterativně takto:

$$\mathbf{x}^{k+1} = \text{prox}_{th}(\mathbf{x}^k - t\nabla g(\mathbf{x}^k)) \quad (5.8)$$

pro vhodnou hodnotu parametru t^k . Tato iterativní metoda se nazývá proximální gradientní metoda nebo také dopředné-zpětné dělení (forward-backward splitting). Skládá se ze dvou kroků – dopředného (explicitního) gradientního kroku, ve kterém využíváme pouze g na výpočet $\mathbf{x}^{k+\frac{1}{2}} = \mathbf{x}^k - t^k \nabla g(\mathbf{x}^k)$, a zpětného (obecně implicitního) proximálního kroku využívající pouze h na výpočet $\mathbf{x}^{k+1} = \text{prox}_{th}(\mathbf{x}^{k+\frac{1}{2}})$. Pokud $t^k < \frac{2}{\beta}$, pak posloupnost $\{\mathbf{x}^k\}$ generovaná algoritmem (5.8) konverguje k řešení problému (5.6) [4, 5, 20].

Ukážeme si, jak funguje proximální gradientní metoda na tomto jednoduchém příkladu:

Příklad: Nalezněte řešení problému

$$\min_x \frac{1}{2}(x-2)^2 + \frac{1}{2}|2x+1|.$$

Řešení: Označíme si $g(x) = \frac{1}{2}(x-2)^2$ a $h(x) = \frac{1}{2}|2x+1|$. Vidíme, že obě funkce jsou konvexní a spojitě a navíc g je diferencovatelná. Spočítáme lipschitzovskou konstantu β tak, aby platilo (5.7), tedy konkrétně:

$$|g'(x) - g'(y)| \leq \beta|x - y|.$$

Dosadíme-li do uvedeného vztahu gradient g , což je $g'(x) = x - 2$, dostaneme:

$$|x - 2 - (y - 2)| = |x - y| \leq 1 \cdot |x - y|,$$

z čehož ihned vyplývá, že $\beta = 1$. Funkce $g + h$ je také koercivní, protože

$$\lim_{|x| \rightarrow \infty} g(x) + h(x) = \lim_{|x| \rightarrow \infty} \frac{1}{2}(x-2)^2 + \frac{1}{2}|2x+1| = \infty$$

a funkce g je navíc ryze konvexní. Můžeme tedy použít proximální gradientní metodu k nalezení jednoznačného řešení.

Bod $x_0 = 3$ zvolíme za startovací bod. Dále zvolíme konstantní krok $t = 0,3 < \frac{2}{\beta} = 2$ (také bychom mohli v každé iteraci tento krok měnit). V každé iteraci se nejprve posuneme v t -násobku gradientu g :

$$x^{k+\frac{1}{2}} = x^k - t \cdot g'(x^k) = x^k - 0,3 \cdot (x^k - 2) = 0,7x^k + 0,6.$$

V dalším kroku použijeme na tento bod proximální operátor th :

$$x^{k+1} = \text{prox}_{th}(x^{k+\frac{1}{2}}).$$

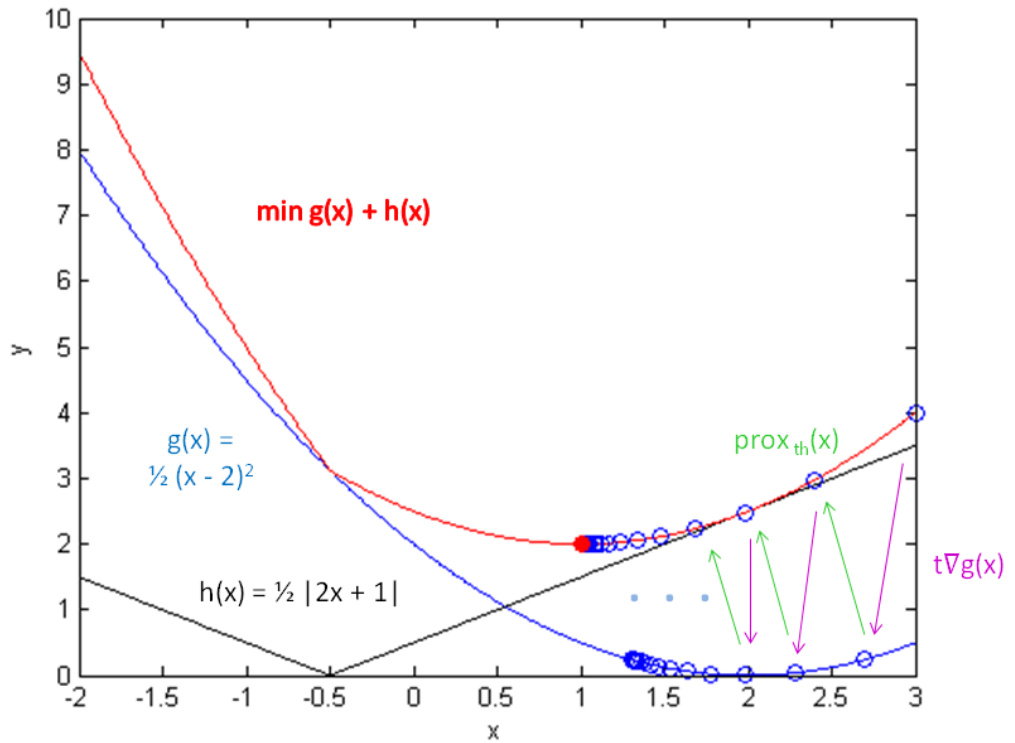
Připomeňme, že proximální operátor ℓ_1 -normy je měkké prahování, tj.

$$\text{soft}(x, \lambda) = \frac{x}{|x|} \max(|x| - \lambda, 0).$$

Využijeme-li vlastnosti pro posunutí a změnu měřítka proximálního operátoru (viz věta 5.1), dostaneme:

$$x^{k+1} = \frac{1}{2} \text{soft} \left(2x^{k+\frac{1}{2}} + 1; 0,3 \cdot 2^2 \cdot \frac{1}{2} \right) - \frac{1}{2} = \frac{1}{2} \text{soft} \left(2x^{k+\frac{1}{2}} + 1; 0,6 \right) - \frac{1}{2}.$$

V iteracích pokračujeme tak dlouho, dokud se dva po sobě jdoucí body „moc neliší“, konkrétně tedy zvolíme konec, pokud $|x^{k+1} - x^k| < 10^{-5} \cdot |x^k|$. Tento algoritmus zastaví v nalezeném minimu v bodě $x = 1$ po 32 iteracích. Hodnota účelové funkce v tomto bodě je 2. Celý průběh výpočtu můžeme vidět na obr. 5.3.



Obrázek 5.3: Průběh proximální gradientní metody při minimalizaci funkce $g+h$. Nalezené minimum je vyznačeno červeným bodem.

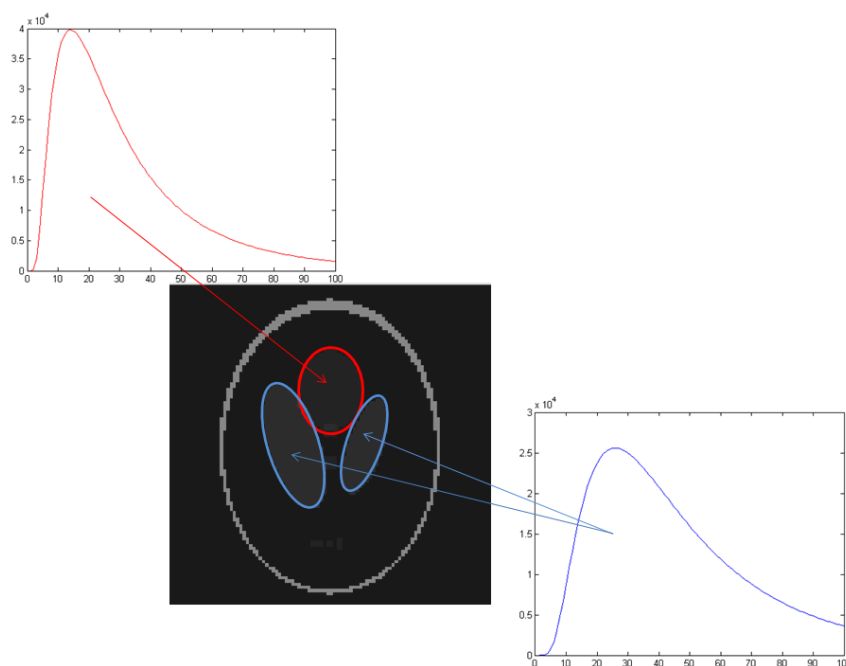
6. Aplikace komprimovaného snímání na perfúzním zobrazování v magnetické rezonanci

Předchozí kapitoly nastínily problematiku perfúzního zobrazování v magnetické rezonanci a možnost jeho zrychlení pomocí komprimovaného snímání. Nyní si ukážeme, jak prakticky využít komprimované snímání v magnetické rezonanci. K tomuto účelu využijeme fantom perfúzního zobrazování, na kterém nasimulujeme měření v magnetické rezonanci a následnou rekonstrukci pomocí komprimovaného snímání. Poté tyto metody ověříme na reálných perfúzních datech myši. Celá simulace je provedena v programovém prostředí MatLab (2010b).

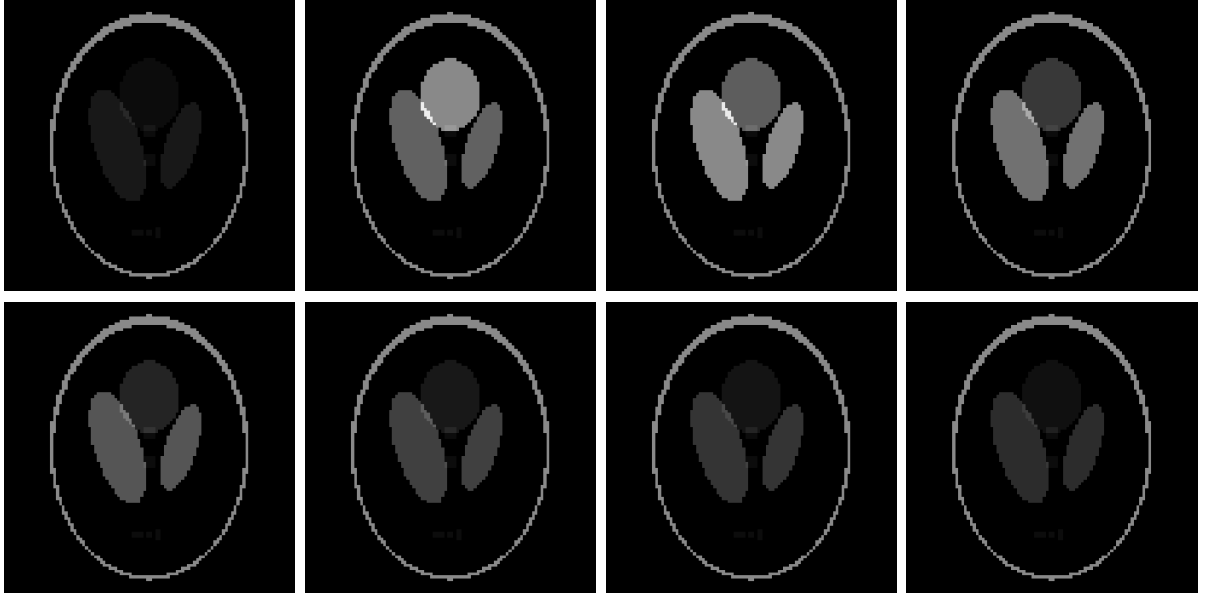
6.1. Fantom perfúzního zobrazování

Abychom mohli dobře nasimulovat měření perfúzního zobrazování v magnetické rezonanci, potřebujeme nejprve vytvořit jeho kvalitní model – tedy fantom perfúzního zobrazování. Základ fantomu je tvořen modifikovaným Shepp-Loganovým fantomem v MatLabu [22], který simuluje řez mozku. Je tvořen deseti elipsami, z nichž se některé překrývají. K tomuto fantomu jsou v čase postupně přidávány hodnoty perfúzních křivek vytvořených pomocí lognormálního modelu (viz část 2.1.1). Fantom perfúzního zobrazování obsahuje dva druhy lognormálních křivek, které se liší svými parametry. Na obr. 6.1 je schématicky znázorněno, ke kterým elipsám se tyto perfúzní křivky přidávají. Parametry křivek se dají libovolně volit.

Pro účely simulace jsem využila pro vytvoření fantomu tyto parametry perfúzních křivek (odhadnuté z reálných dat): konstantní hodnota bez použití kontrastní látky c



Obrázek 6.1: Schématické znázornění vytvoření fantomu perfúzního zobrazování

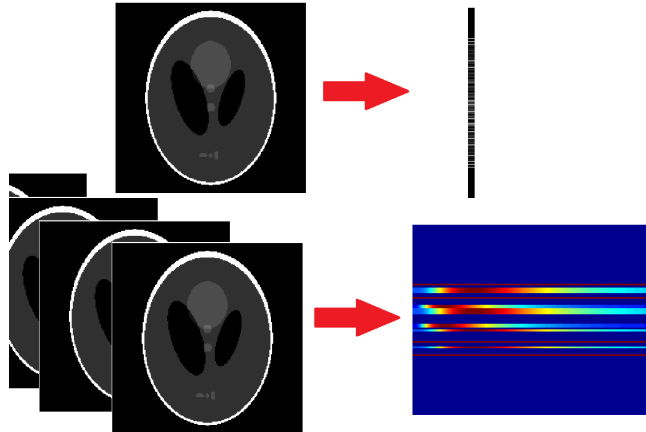


Obrázek 6.2: Průběh fantomu perfúzního zobrazování ve vybraných časech

byla vzata ze samotných parametrů modifikovaného Shepp-Loganova fantomu, zpoždění bolusu $t_0 = 1$ a plocha pod křivkou $S = 8,6 \cdot 10^6$ u obou druhů křivek. Červená křivka (viz obr. 6.1) má další parametry zvolené jako $\mu = 5,5$ a $\sigma = 0,7$; modrá křivka $\mu = 5$ a $\sigma = 0,8$. Dále bylo potřeba zvolit celkovou dobu pozorování $T = 600$ a frekvenci snímání měření $f = \frac{1}{6}$ (předpokládá se soudělnost T a f). Protože samotný modifikovaný Shepp-Loganův fantom nabývá hodnot z intervalu $\langle 0, 1 \rangle$, byly perfúzní křivky před samotným přidáním normovány tak, aby jejich maximální hodnota po aplikaci byla 1 (ale v místech překryvu oblastí s perfúzí se jejich hodnoty sčítají, takže v nich mohou přesáhnout hodnotu 1).

Nejčastěji jsem využila pro simulace fantom perfúzního zobrazování o velikosti 100×100 px. Některé snímky tohoto fantomu v různých časech můžete vidět na obr. 6.2 (zobrazen 1. a potom každý další 14. snímek).

Fantom perfúzního zobrazování je ve skutečnosti trojrozměrná matice dat (video), se kterou se při výpočtech hůře pracuje. Proto se před používáním transformuje do dvojrozměrné matice [20]. Každý snímek se přeskládá do vektoru (sloupec po sloupci) a jednotlivé snímky se vloží do sloupců matice, se kterou se následně pracuje. Jednotlivé řádky takto vytvořené matice potom obsahují přímo perfúzní křivky. Schéma této transformace je znázorněné na obr. 6.3.



Obrázek 6.3: Přeskládání video dat do matice (použita dvojí pseudobarevná škála)

6.2. Měřicí matice (masky)

Pro měření perfúzních dat v MRI se používají tzv. měřicí matice (masky), které je potřeba vhodně navrhnout, abychom byli schopni dobře rekonstruovat data pomocí komprimovaného snímání. Měřicí masky se skládají z jedniček a nul (zobrazeno barevně – bílá barva představuje jedničku a černá nulu). Tyto masky se aplikují na jednotlivé snímky přímo v k -prostoru. Na pozici, kde máme jedničku, naměříme Fourierův koeficient, jinde ne.

V magnetické rezonanci se nemohou používat libovolné gradientní sekvence, proto by měly tyto masky obsahovat spojitě trajektorie. Pokud bychom ale uvažovali, že tyto snímky jsou součástí 3D měření objemu, pak mohou trajektorie být spojitě ve třetím rozměru, tj. v příčných řezech se jevit jako diskrétní body.

Pro simulaci měření jsem použila několik druhů měřicích masek:

Náhodná maska

Náhodná měřicí matice využívá rovnoměrné rozdělení pravděpodobnosti v intervalu $(0, 1)$ (ozn. $Ro(0, 1)$) pro výběr pozic, na kterých je umístěná jednička. Zvolíme parametr θ udávající pravděpodobnost poměru počtu jedniček ku počtu prvků v celé matici. Potom vygenerujeme matici $\mathbf{M} = (m_{ij})_{i,j=1}^n$ nezávislých náhodných veličin $m_{ij} \sim Ro(0, 1)$. Prvek a_{ij} měřicí matice $\mathbf{A}$ bude roven jedné, jestliže $m_{ij} \leq \theta$. Dále je tu možnost měřit vždy jenom jeden koeficient z páru komplexně sdružených Fourierových koeficientů (náhodně vybíráme prvky pouze z jedné „poloviny“ k -prostoru, popsáno podrobněji dále viz věta 6.1).

Náhodná nerovnoměrná maska

Nerovnoměrná náhodná maska je generována obdobným způsobem jako náhodná, ale pro každý prvek a_{ij} definujeme pravděpodobnost zvolení jedničky h_{ij} . Dále opět zvolíme parametr θ udávající pravděpodobnost poměru počtu jedniček ku počtu prvků v celé matici. Určíme hodnotu $\mu = \frac{\sum_{i=1}^n \sum_{j=1}^n P(\mathbb{X}_{ij} \leq h_{ij})}{n^2}$, kde $\mathbb{X}_{ij} \sim Ro(0, 1)$. Pokud $\theta \leq \mu$, vezmeme $\tilde{h}_{ij} = k \cdot h_{ij}$, přičemž k určíme tak, aby platilo $\theta = k \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n P(\mathbb{X}_{ij} \leq h_{ij})}{n^2}$. Naopak pokud $\theta > \mu$, můžeme vzít $\tilde{h}_{ij} = \min\{h_{ij} + q, 1\}$, kde q určíme tak, aby platilo $\theta = \frac{\sum_{i=1}^n \sum_{j=1}^n P(\mathbb{X}_{ij} \leq \tilde{h}_{ij})}{n^2}$. Nakonec vygenerujeme matici $\mathbf{M} = (m_{ij})_{i,j=1}^n$ nezávislých náhodných veličin $m_{ij} \sim Ro(0, 1)$. Prvek a_{ij} měřicí matice $\mathbf{A}$ bude roven jedné, jestliže $m_{ij} \leq \tilde{h}_{ij}$.

Při simulacích jsem využila pravděpodobnost

$$h_{ij} = \min \left\{ \left(\frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} + c \right)^6, 1 \right\},$$

kde $x = \frac{6i-3n-3}{n-1}$, $y = \frac{6j-3n-3}{n-1}$ a konstanta c byla zvolena tak, aby $\mu \doteq 0,5$ (experimentálně upravená hustota dvourozměrného normálního rozdělení).

Radiální maska

Radiální maska obsahuje přímky vedoucí počátkem k -prostoru, jejichž úhel svíraný s osou x je generován rovnoměrným rozdělením. Počet těchto přímek můžeme měnit. Další variantu této masky obsahuje pouze polopřímky začínající na nulových souřadnicích (dále označeno jako radiální paprsky).

Spirální maska

Spirální maska obsahuje logaritmické spirály, které můžeme v polární soustavě souřadnic (ρ, φ) zapsat pomocí rovnice [26]:

$$\rho = ae^{b\varphi}.$$

Můžeme zvolit počet těchto spirál. Každá další spirála bude otočená kolem počátku souřadnic o stejný úhel oproti předchozí spirále tak, aby tento úhel byl i mezi první a poslední. Parametry a, b spirály jsou generovány náhodně, $a \sim Ro(0,5; 1,5)$ a $b \sim Ro(0,1; 1,1)$.

Výše uvedené typy masek jsou zobrazeny na obr. 6.4.

Při simulacích jsem po aplikaci měřicí masky na perfúzní data využila této vlastnosti Fourierovy transformace:

Věta 6.1. *Nechť $\mathcal{F}(\mathbf{f}) = (F_{ij})_{i,j=0}^{n-1}$ je 2D diskretní Fourierova transformace $\mathbf{f} = (f_{ij})_{i,j=0}^{n-1} \in \mathbb{R}^{n \times n}$, tj.*

$$F_{uv} = \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} f_{kl} \cdot e^{-2\pi i \frac{ku+lv}{n}}.$$

Potom platí

$$F_{u,v} = \overline{F_{n-u,n-v}} \quad \forall u, v = 0, \dots, n-1,$$

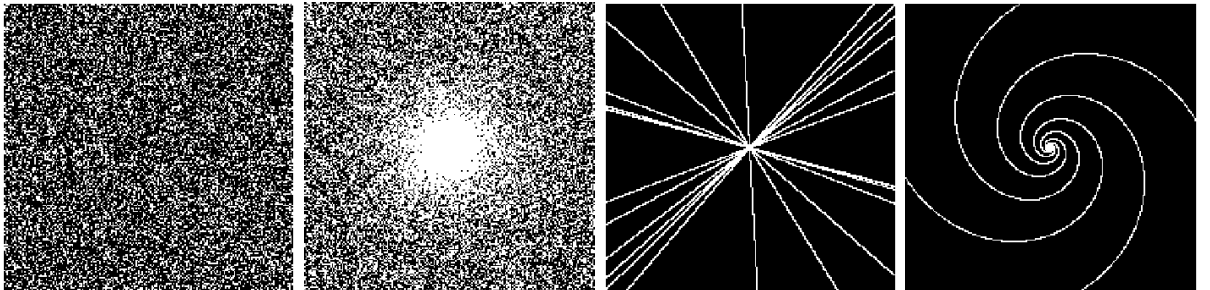
kde pokládáme $F_{n,v} = F_{0,v}$ a $F_{u,n} = F_{u,0}$.

Důkaz.

$$\begin{aligned} \overline{F_{n-u,n-v}} &= \overline{\sum_{k=0}^{n-1} \sum_{l=0}^{n-1} f_{kl} \cdot e^{-2\pi i \frac{(n-ku)+(n-lv)}{n}}} = \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} f_{kl} \cdot e^{2\pi i \frac{(n-ku)+(n-lv)}{n}} = \\ &= \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} f_{kl} \cdot \underbrace{e^{4\pi i}}_1 \cdot e^{-2\pi i \frac{ku+lv}{n}} = F_{u,v}. \end{aligned}$$

□

Důsledkem této věty je, že Fourierův koeficient $F_{0,0}$ je reálný, v případě sudého n také $F_{\frac{n}{2}, \frac{n}{2}}, F_{0, \frac{n}{2}}$ a $F_{\frac{n}{2}, 0}$. Ostatní koeficienty tvoří páry vzájemně komplexně sdružených koeficientů. Když tedy naměříme pomocí měřicí masky v nějakém čase pouze jeden koeficient z tohoto páru, použijeme jeho hodnotu i pro vytvoření druhého koeficientu. V přítomnosti šumu při naměření obou koeficientů je navíc můžeme „zprůměrovat“, tj. místo $F_{u,v}$ vzít $F_{u,v}^* = \frac{1}{2} (F_{u,v} + \overline{F_{n-u,n-v}})$. Tím potlačíme šum při rekonstrukci obrazu a vyhneme se řešení, která by nebyla reálná.



Obrázek 6.4: Měřicí matice (zleva doprava): náhodná, nerovnoměrně náhodná, radiální a spirální

6.3. L + S metoda pro rekonstrukci

Protože měření v magnetické rezonanci je velice pomalé, využijeme při simulaci perfúzního zobrazování komprimované snímání, které nám umožní zrychlit snímání dat. Abychom byli schopni následně rekonstruovat původní signál bez velkých chyb, je nutné znát nějaké apriorní informace o snímaných datech a s jejich pomocí sestavit vhodný model pro získání rekonstrukce (viz část 4.3).

Jako první jsem využila tzv. **L + S** model pro perfúzní data, který byl představen v článku [20]. Tento model využívá toho, že perfúzní křivky mají podobný průběh v jednom typu tkáně a že perfúzní křivky mají řídké spektrum. Proto se **L + S** model snaží rozložit hledanou rekonstrukci videa **M** na součet dvou matic, a to matice s nízkou hodnotí **L** (s malým počtem nenulových singulárních čísel) a řídké matice ve spektru po řádcích **S** (s málo nenulovými prvky řádkového spektra). Celý problém se dá zformulovat jako tato konvexní optimalizační úloha:

$$\min_{\mathbf{L}, \mathbf{S}} \frac{1}{2} \| \underbrace{E(\mathbf{L} + \mathbf{S})}_{\mathbf{M}} - \mathbf{d} \|_2^2 + \lambda_L \|\mathbf{L}\|_* + \lambda_S \|T\mathbf{S}\|_1, \quad (6.1)$$

kde T je 1D Fourierova transformace po řádcích, E představuje podvzorkovanou prostorovou 2D Fourierovu transformaci pro každý snímek (podvzorkování provedeno pomocí měřicí matice), d jsou podvzorkovaná (naměřená) data v k -prostoru a λ_L, λ_S jsou vhodné regularizační konstanty. První člen v účelové funkci s eukleidovskou normou vynucuje, aby se hledaná rekonstrukce **M** moc neodchylovala od naměřených dat, druhý člen s nukleární normou zajišťuje nízkou hodnotu **L** (podobnost perfúzních křivek v rámci tkáně) a třetí člen s ℓ_1 normou řídkost perfúzních křivek.

Konvexní úlohu (6.1) můžeme řešit pomocí proximální gradientní metody (viz část 5.2). Označíme $g = \frac{1}{2} \|E(\mathbf{L} + \mathbf{S}) - \mathbf{d}\|_2^2$ a $h = \lambda_L \|\mathbf{L}\|_* + \lambda_S \|T\mathbf{S}\|_1$. Obě funkce jsou konvexní a spojitě a navíc g je diferencovatelná. Funkce $g + h$ je také koercivní, protože

$$\lim_{\|\mathbf{L}\|, \|\mathbf{S}\| \rightarrow \infty} g(x) + h(x) = \lim_{\|\mathbf{L}\|, \|\mathbf{S}\| \rightarrow \infty} \frac{1}{2} \|E(\mathbf{L} + \mathbf{S}) - \mathbf{d}\|_2^2 + \lambda_L \|\mathbf{L}\|_* + \lambda_S \|T\mathbf{S}\|_1 = \infty.$$

Ověřili jsme tedy, že existuje řešení problému (6.1). Pokud máme E podvzorkovanou Fourierovu transformaci, nemáme ani jednu funkci ryze konvexní (protože eukleidovská norma nebude v některých směrech ohraničená). Pokud ale použijeme měřicí masky s náhodnými parametry, bude řešení s vysokou pravděpodobností jediné [20].

Pro určení délky kroku t nejprve spočítáme lipschitzovskou konstantu β tak, aby platilo (5.7):

$$\|\nabla g(\mathbf{x}) - \nabla g(\mathbf{y})\| \leq \beta \|\mathbf{x} - \mathbf{y}\|.$$

Do tohoto vztahu dosadíme $\nabla g = E^*(E\mathbf{x} - d)$:

$$\|E^*(E\mathbf{x} - d) - E^*(E\mathbf{y} - d)\| = \|E^*E(\mathbf{x} - \mathbf{y})\|.$$

Když E představuje Fourierovu transformaci, vynásobením E^*E dostaneme identický operátor, a tedy $\beta = 1$.

Při obecné volbě E platí, že β je rovno největšímu singulárnímu číslu E^*E [20]. V případě podvzorkované Fourierovy transformace by toto největší singulární číslo bylo rovněž jedna (mimo podvzorkování úplného), proto $\beta = 1$ představuje optimální volbu lipschitzovské konstanty. Z toho vyplývá, že pro konvergenci proximální gradientní metody je

potřeba zvolit krok $t < \frac{2}{\beta} = 2$. V práci jsem nadále vždy použila konstantní krok $t = 1$, který této podmínce vyhovuje.

Za počáteční bod pro proximální gradientní metodu bereme bod $\mathbf{M} = d$. V každé iteraci počítáme postupně:

$$\begin{aligned}\mathbf{L}^k &= \text{svt}(\mathbf{M}^{k-1} - \mathbf{S}^{k-1}, \lambda_L), \\ \mathbf{S}^k &= T^{-1}(\text{soft}(T(\mathbf{M}^{k-1} - \mathbf{L}^k), \lambda_S)), \\ \mathbf{M}^k &= \mathbf{L}^k + \mathbf{S}^k - E^*(E(\mathbf{L}^k + \mathbf{S}^k) - d),\end{aligned}$$

přičemž $\text{svt}(\mathbf{L}, \lambda) = \mathbf{U}\text{soft}(\mathbf{\Sigma}, \lambda)\mathbf{V}^*$, kde $\mathbf{L} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ je singulární rozklad matice $\mathbf{L}$. Iterace ukončíme, pokud je relativní změna $\mathbf{M}$ menší než zvolená tolerance ϵ :

$$\|\mathbf{M}^k - \mathbf{M}^{k-1}\|_2 < \epsilon \cdot \|\mathbf{M}^{k-1}\|_2.$$

V této práci je zvolena dále tolerance vždy $\epsilon = 10^{-5}$.

Konkrétní výsledek rekonstrukce se zvolenými parametry lze vidět na obr. 6.5. Porovnání úspěšnosti rekonstrukce pro různé procento naměřených koeficientů je zobrazeno na obr. 6.6 pro náhodnou masku a na obr. 6.7 pro masku typu radiální paprsky.

Jednotlivá měření s následnou rekonstrukcí pomocí $\mathbf{L} + \mathbf{S}$ modelu byla opakována při naměření jiného procenta koeficientů (pro různé parametry masek). Úspěšnost rekonstrukce jsem následně měřila pomocí SNR:

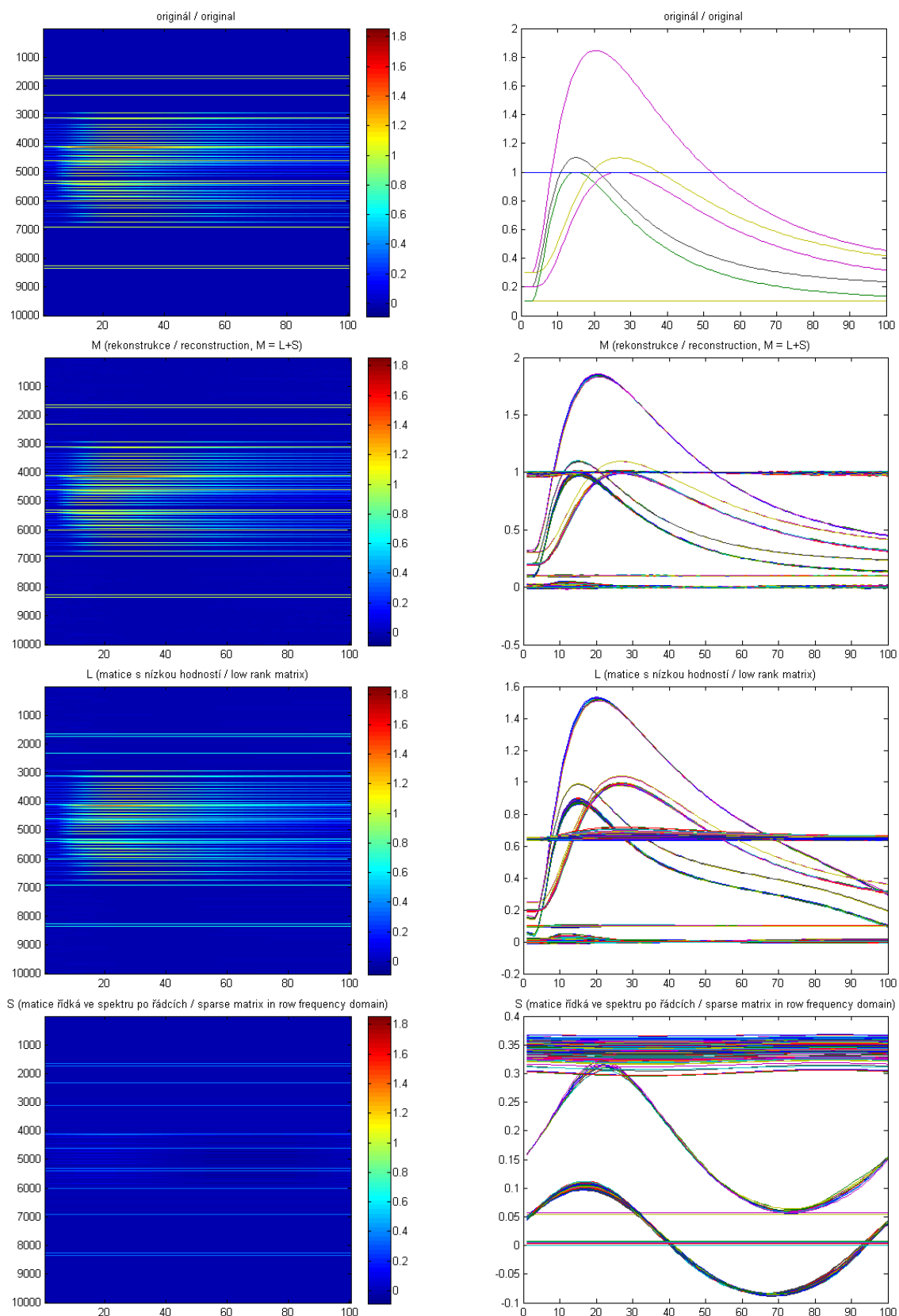
$$\text{SNR} = 20 \cdot \log_{10} \frac{\|\mathbf{D} - \overline{\mathbf{D}}\|_2}{\|\mathbf{M} - \mathbf{D}\|_2} \text{ [dB]},$$

kde $\mathbf{D}$ je perfúzní fantom a $\overline{\mathbf{D}}$ jeho průměr. Závislost SNR na počtu naměřených koeficientů je možné vidět na obr. 6.8.

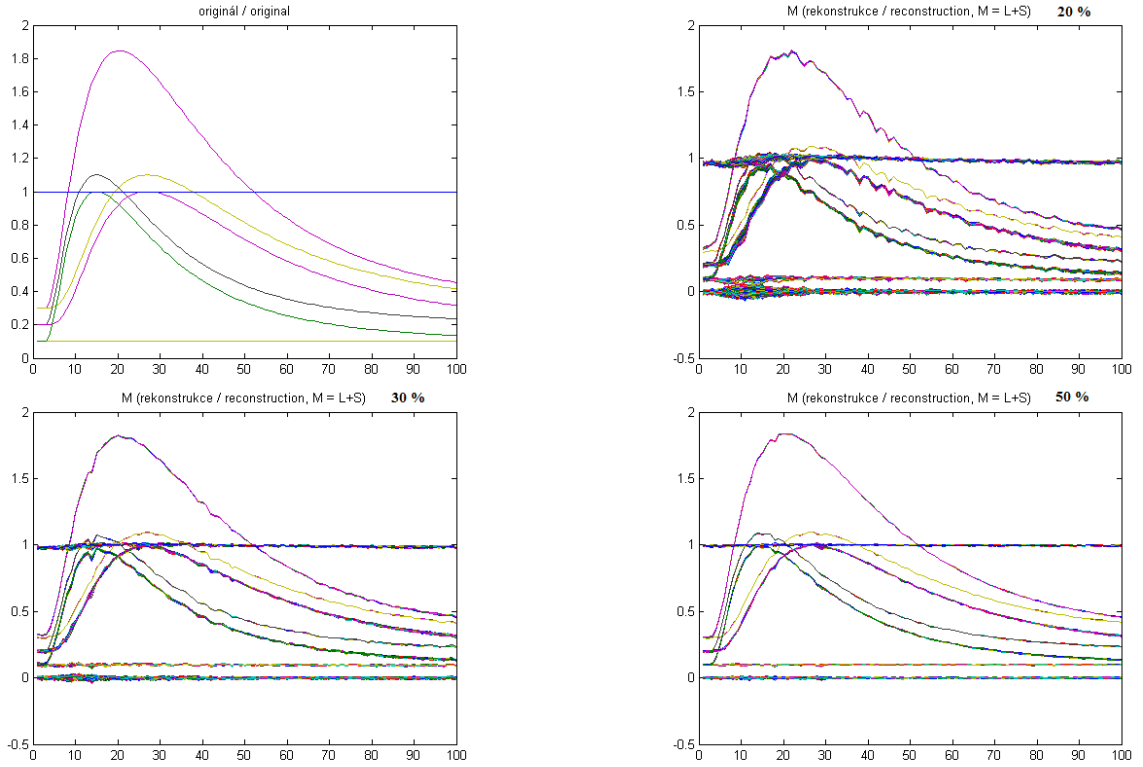
Z grafu je patrné, že nejlepších výsledků dosahuje při nižším počtu naměřených koeficientů (přibližně do 35 %) maska typu radiální paprsky a při vyšším počtu naměřených dat nerovnoměrná náhodná maska. Jen o něco horší je radiální maska a při malém počtu měřených koeficientů také spirální maska. Pro větší fantom dosahuje rekonstrukce lepších výsledků. Je to nejspíše způsobené tím, že celkově máme k dispozici větší množství naměřených koeficientů než u menšího fantomu.

Nejhorší výsledky ze všech použitých druhů masek má náhodné podvzorkování. Je to způsobeno tím, že náhodné vzorkování vybírá měřené hodnoty rovnoměrně v celém k -prostoru, zatímco ostatní druhy masek mají větší koncentraci měřených vzorků v počátku k -prostoru, kde má spektrum větší absolutní hodnoty (při jejich naměření se proto dopustíme menší chyby). Na grafu lze také pozorovat, že při zvyšování množství koeficientů u spirálního podvzorkování nedochází k tak výraznému zvyšování SNR jako u náhodného a radiálního, což je dáno tím, že spirály nevybírají vzorky, které se nacházejí v „rozech“ k -prostoru. Proto jsem implementovala u spirální masky navíc ještě možnost použití „čtvercovaté“ spirály, která by se mohla do „rohů“ dostat lépe.

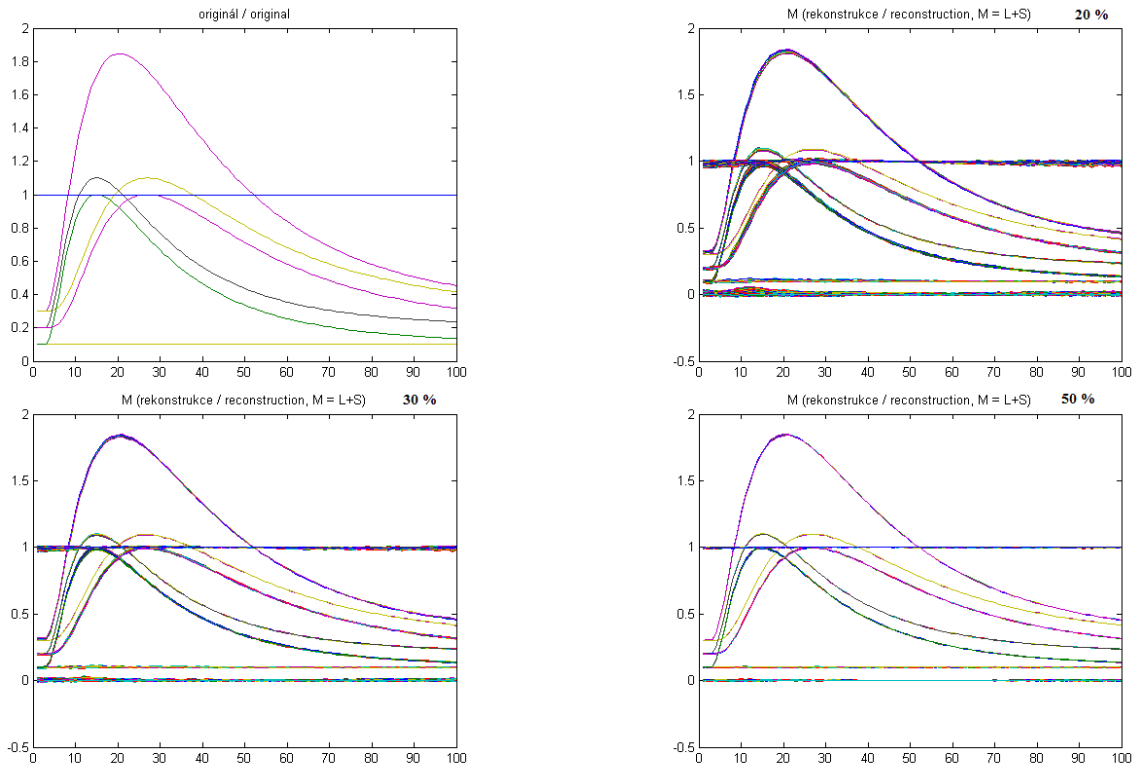
Vzhledově se zdají vyhovující rekonstrukce zhruba od SNR 36 dB. U masky typu radiální paprsky by tedy stačilo naměřit přibližně 10 % vzorků pro fantom o velikosti 150×150 px a 12 % pro 100×100 px. Naopak u náhodné masky by bylo třeba alespoň 25 % pro menší fantom a 30 % vzorků pro větší.



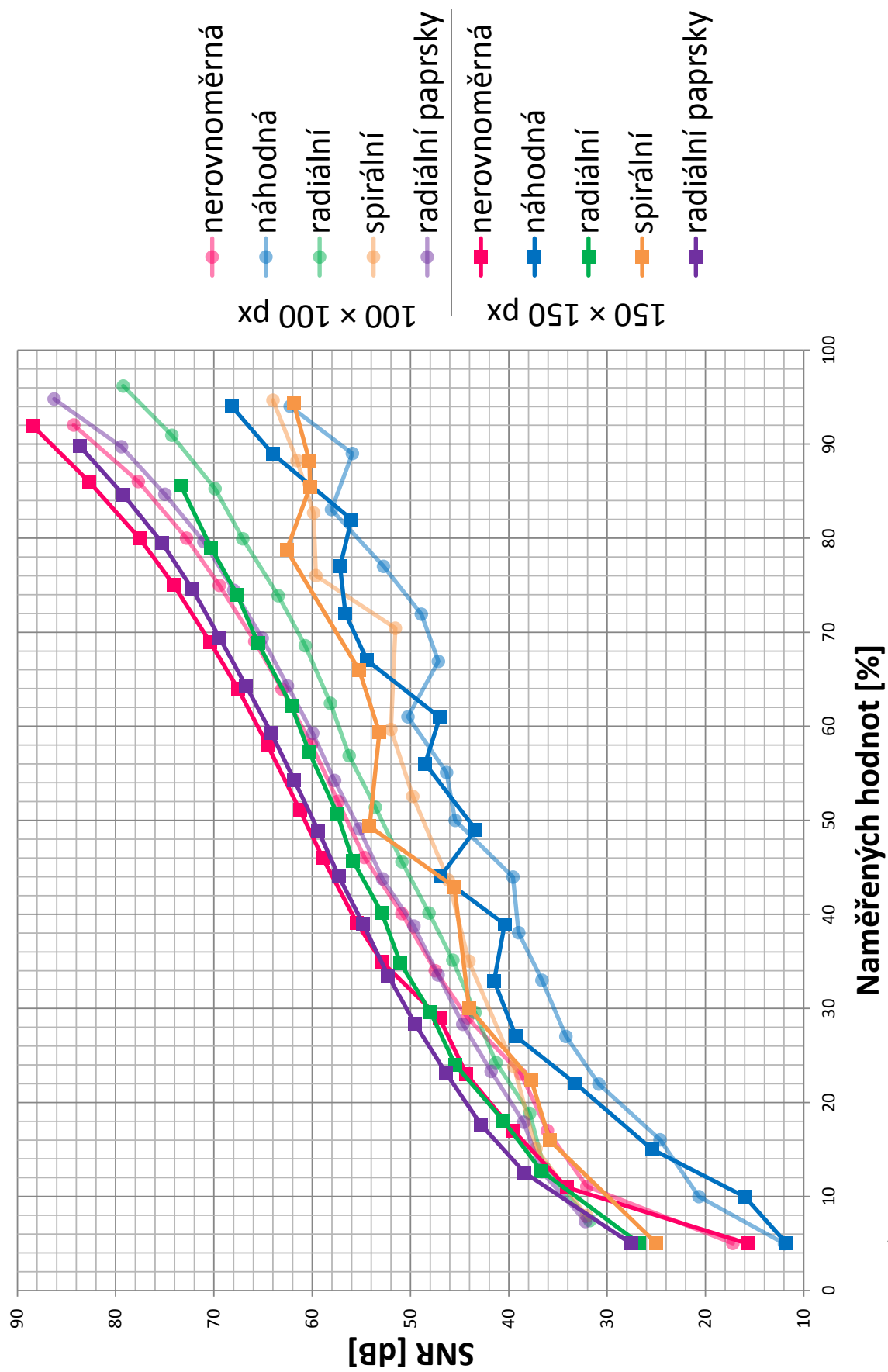
Obrázek 6.5: Nahoře originální matice fantomu o velikosti 100×100 px a perfúzní křivky fantomu (řádky této matice). Pod ní jednotlivé matice i perfúzní křivky rekonstrukce pomocí radiální masky z 20 přímek (16,75 %), v účelové funkci nastaveno $\lambda_L = 0,5$ a $\lambda_S = 0,1$. Algoritmus spočítal M za 595 s a rekonstrukce dosáhla SNR 46,69 dB.



Obrázek 6.6: Originální perfúzní křivky fantomu o velikosti 100×100 px a po rekonstrukci při různých procentech naměřených koeficientů s využitím náhodné masky (použito $\lambda_L = 2$ a $\lambda_S = 0,5$).



Obrázek 6.7: Originální perfúzní křivky fantomu o velikosti 100×100 px a po rekonstrukci při různých procentech naměřených koeficientů s využitím masky typu radiální paprsky (použito $\lambda_L = 2$ a $\lambda_S = 0,5$).



Obrázek 6.8: Úspěšnost rekonstrukce měřená podle SNR v závislosti na % naměřených hodnot pro různé druhy měřících masek a dvě velikosti fantomu ($\lambda_L = 2$ a $\lambda_S = 0,5$)

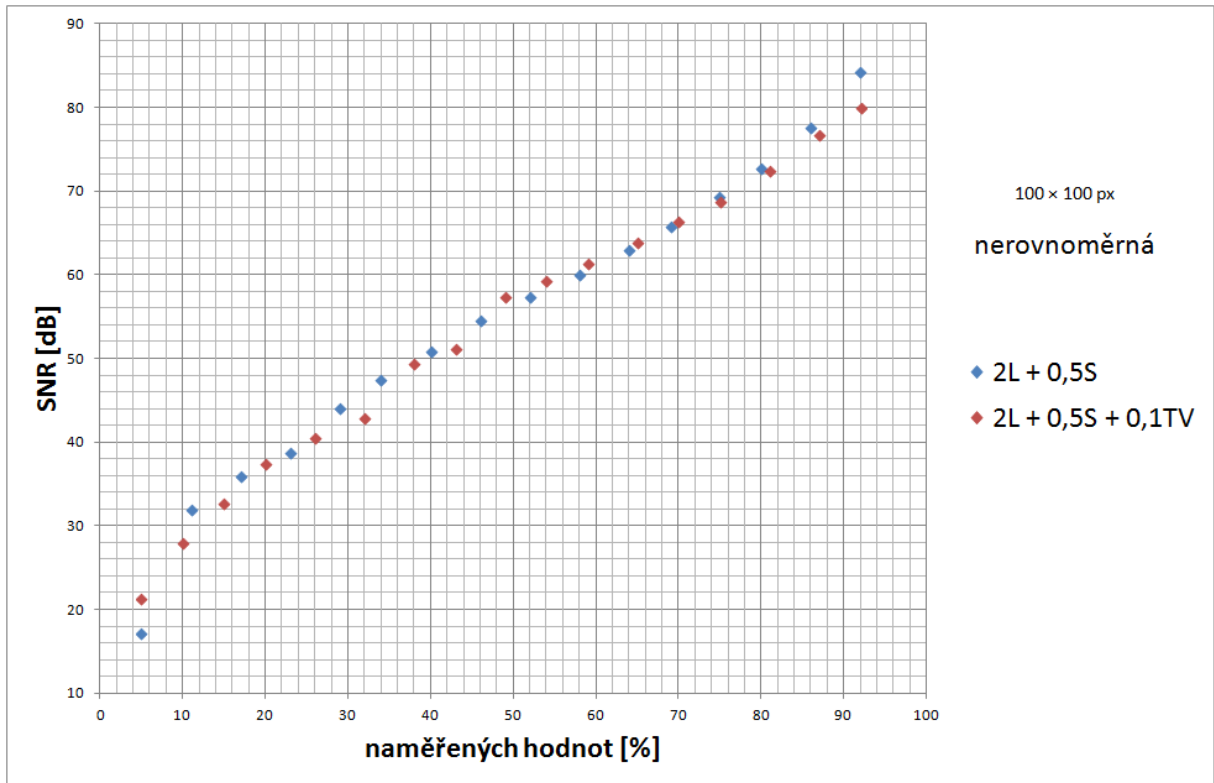
6.4. L + S metoda s 1D totální variací

U $\mathbf{L} + \mathbf{S}$ modelu někdy při nízkém počtu naměřených koeficientů dochází k tomu, že rekonstruované perfúzní křivky nejsou hladké. Proto jsem k tomuto modelu přidala další člen s 1D totální variací perfúzních křivek (do časové osy), která nedovolí, aby vznikaly velké skoky v perfúzních křivkách. Tento model obohacený o totální variaci zapíšeme jako tento problém:

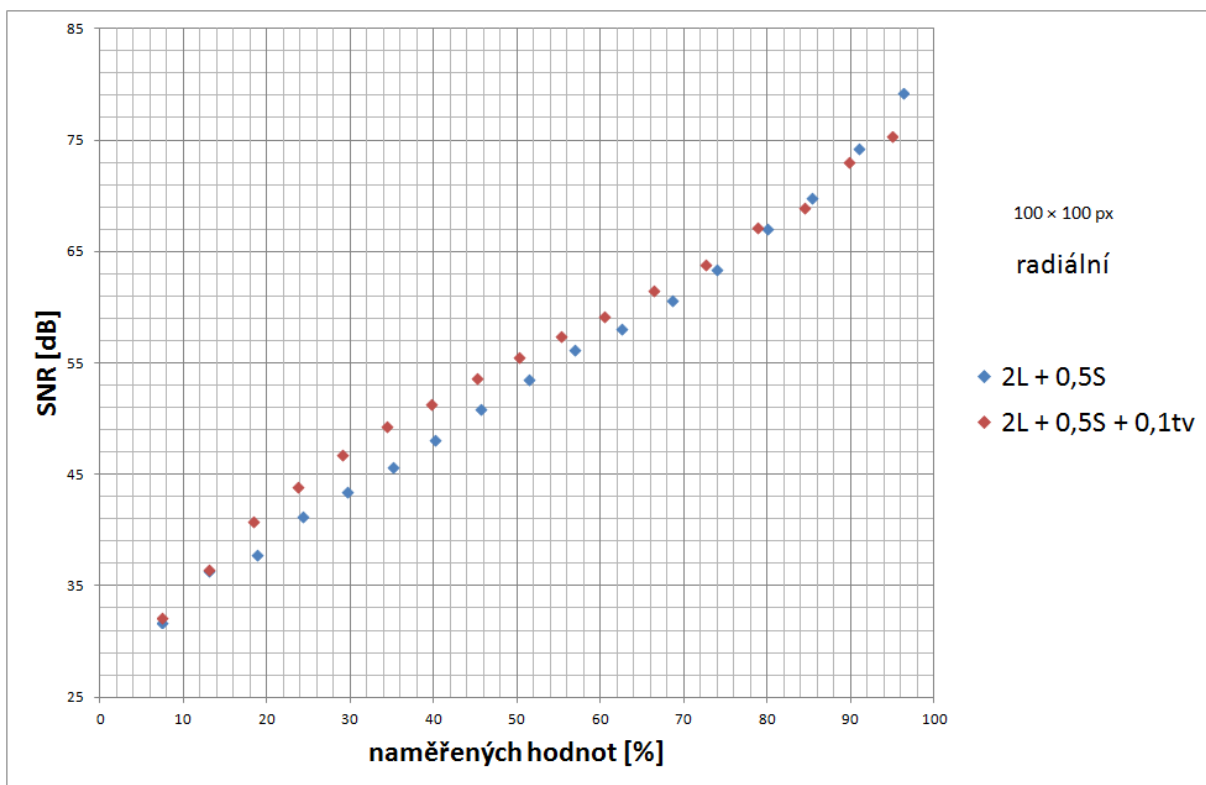
$$\min_{\mathbf{L}, \mathbf{S}} \frac{1}{2} \left\| E \underbrace{(\mathbf{L} + \mathbf{S})}_{\mathbf{M}} - \mathbf{d} \right\|_2^2 + \lambda_L \|\mathbf{L}\|_* + \lambda_S \|\mathbf{TS}\|_1 + \lambda_{tv} \sum_i \|\mathbf{M}_{i,:}\|_{tv}.$$

Pro nalezení řešení tohoto minimalizačního problému použijeme opět jako u $\mathbf{L} + \mathbf{S}$ modelu proximální gradientní metodu.

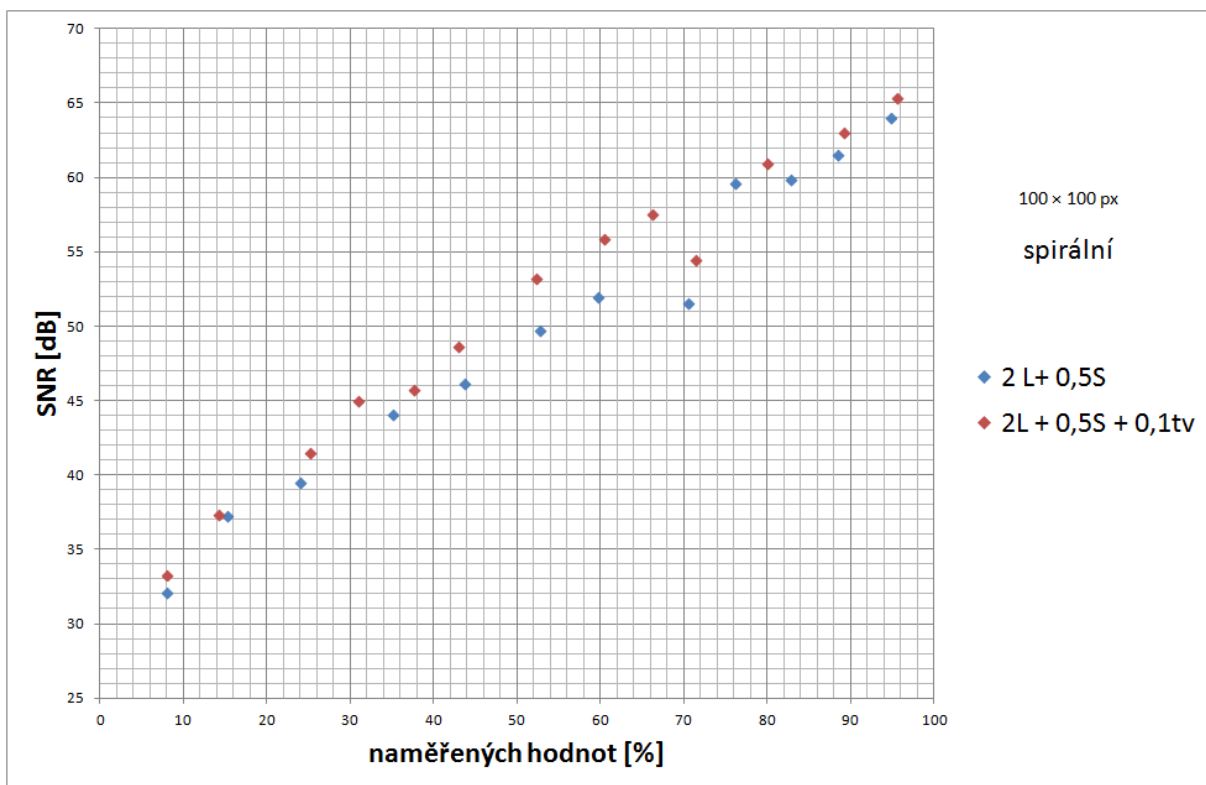
Porovnála jsem tento model s $\mathbf{L} + \mathbf{S}$ metodou. Výsledky je možné vidět na grafech 6.9, 6.10, 6.11 a 6.12. Mírné zlepšení oproti předchozímu modelu nastalo u radiální masky a masky typu radiální paprsky při naměření 15–50 % hodnot a také u spirální masky při 20–70 %, což splnilo očekávaný efekt. Naopak u náhodné nerovnoměrné masky totální variace nepřinesla zlepšení.



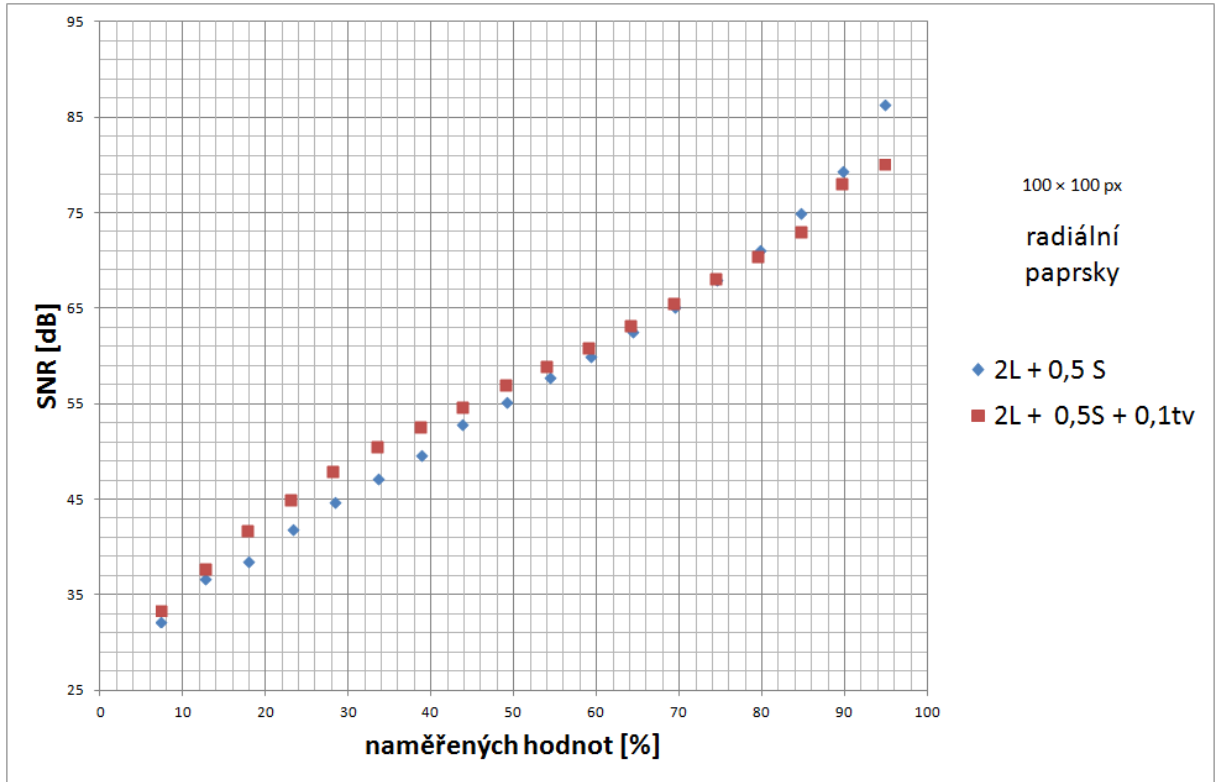
Obrázek 6.9: Porovnání úspěšnosti rekonstrukce $\mathbf{L} + \mathbf{S}$ metody a s přidáním 1D diskretní totální variace pro nerovnoměrnou náhodnou masku, parametry byly zvoleny $\lambda_L = 2$, $\lambda_S = 0,5$ a $\lambda_{tv} = 0,1$



Obrázek 6.10: Porovnání úspěšnosti rekonstrukce $\mathbf{L} + \mathbf{S}$ metody a s přidáním 1D diskretní totální variace pro radiální masku, parametry byly zvoleny $\lambda_L = 2$, $\lambda_S = 0,5$ a $\lambda_{tv} = 0,1$



Obrázek 6.11: Porovnání úspěšnosti rekonstrukce $\mathbf{L} + \mathbf{S}$ metody a s přidáním 1D diskretní totální variace pro spirální masku, parametry byly zvoleny $\lambda_L = 2$, $\lambda_S = 0,5$ a $\lambda_{tv} = 0,1$



Obrázek 6.12: Porovnání úspěšnosti rekonstrukce $\mathbf{L} + \mathbf{S}$ metody a s přidáním 1D diskrétní totální variace pro masku typu radiální paprsky, parametry byly zvoleny $\lambda_L = 2$, $\lambda_S = 0,5$ a $\lambda_{tv} = 0,1$

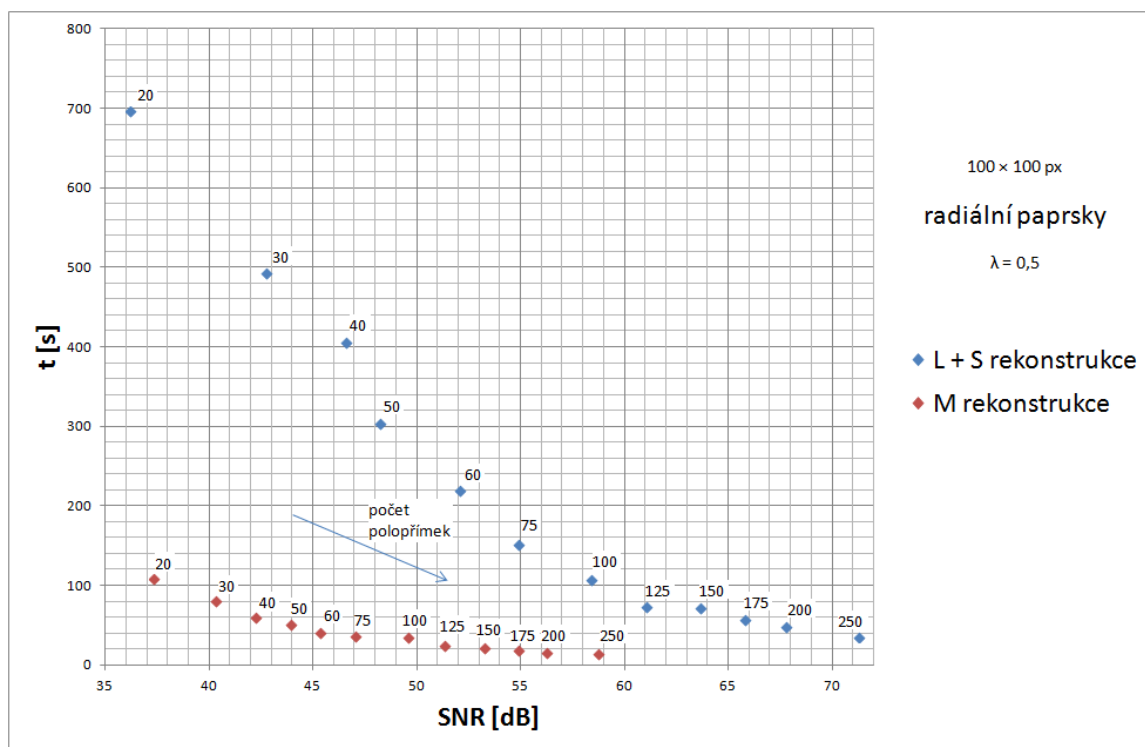
6.5. M model

Jako další model jsem aplikovala zároveň podmínku nízké hodnoty i řídkosti spektra na celou rekonstrukci $\mathbf{M}$. Tento model můžeme popsat pomocí tohoto optimalizačního modelu [17, 20]

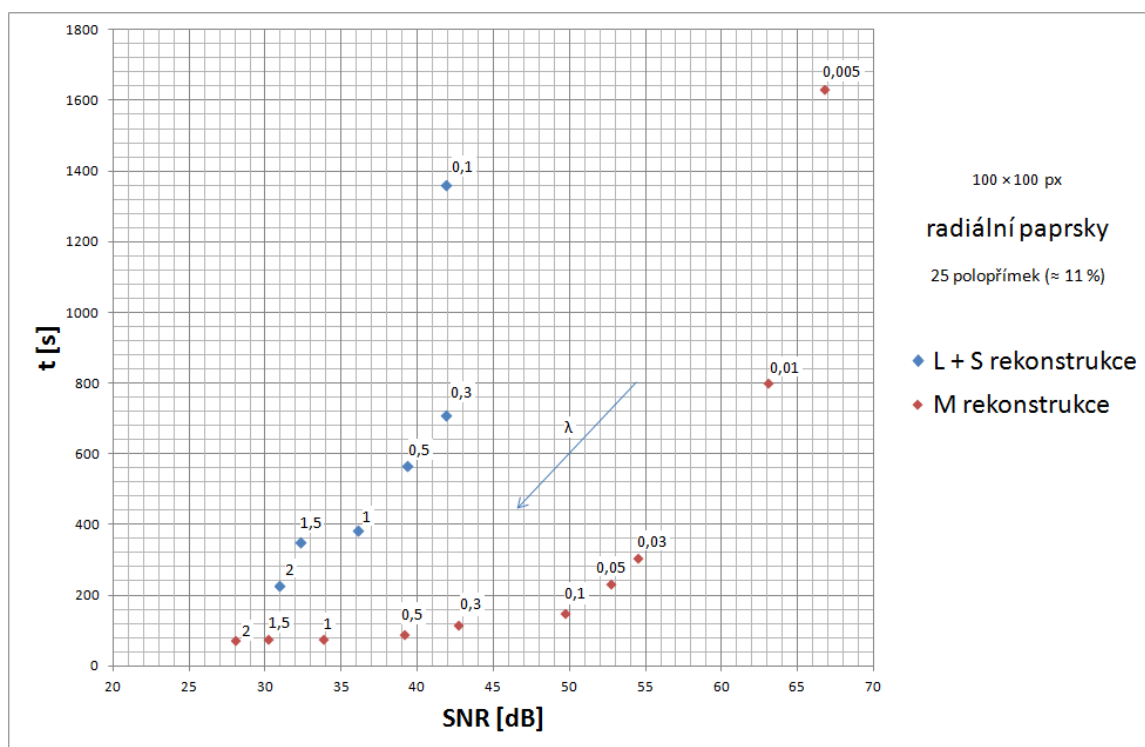
$$\min_{\mathbf{M}} \frac{1}{2} \|\mathbf{E}\mathbf{M} - \mathbf{d}\|_2^2 + \lambda_L \|\mathbf{M}\|_* + \lambda_S \|\mathbf{T}\mathbf{M}\|_1.$$

Opět jej řešíme pomocí proximální gradientní metody.

Tento model jsem porovnávala s původním $\mathbf{L} + \mathbf{S}$ modelem. U obou modelů jsem sledovala závislost výpočetního času na dosaženém SNR při změně počtu naměřených dat (viz graf 6.13) a při změně parametrů λ v účelové funkci, kdy jsem volila oba parametry vždy stejné (viz graf 6.14). Je vidět, že při stejné nastavených parametrech dosahuje $\mathbf{M}$ model nižšího SNR než $\mathbf{L} + \mathbf{S}$ model, ale pokud bychom parametry u obou modelů nastavili tak, že by výpočet trval stejnou dobu, dosahuje $\mathbf{M}$ mnohem lepšího SNR než $\mathbf{L} + \mathbf{S}$.



Obrázek 6.13: Porovnání **L + S** a **M** modelu pro rekonstrukci dat. Graf ukazuje časovou náročnost výpočtu na dosaženém SNR při změně počtu naměřených polopřímek u masky typu radiální paprsky. Parametry byly zvoleny $\lambda_L = \lambda_S = 0,5$.



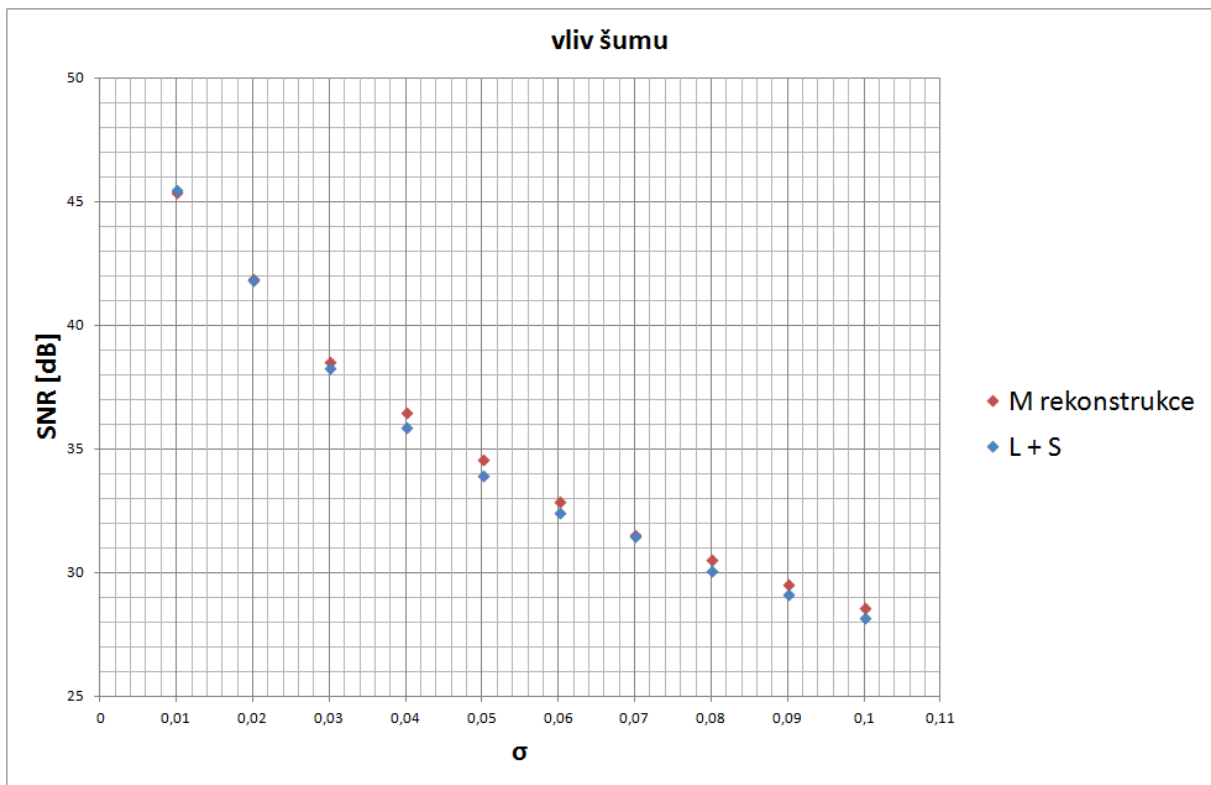
Obrázek 6.14: Porovnání **L + S** a **M** modelu pro rekonstrukci dat. Graf ukazuje časovou náročnost výpočtu na dosaženém SNR při změně parametrů λ_L a λ_S , přičemž bylo položeno $\lambda_L = \lambda_S$. K měření byla použita maska typu radiální paprsky a bylo naměřeno 25 polopřímek ($\approx 11\%$).

6.6. Vliv šumu na rekonstrukci perfúzních dat

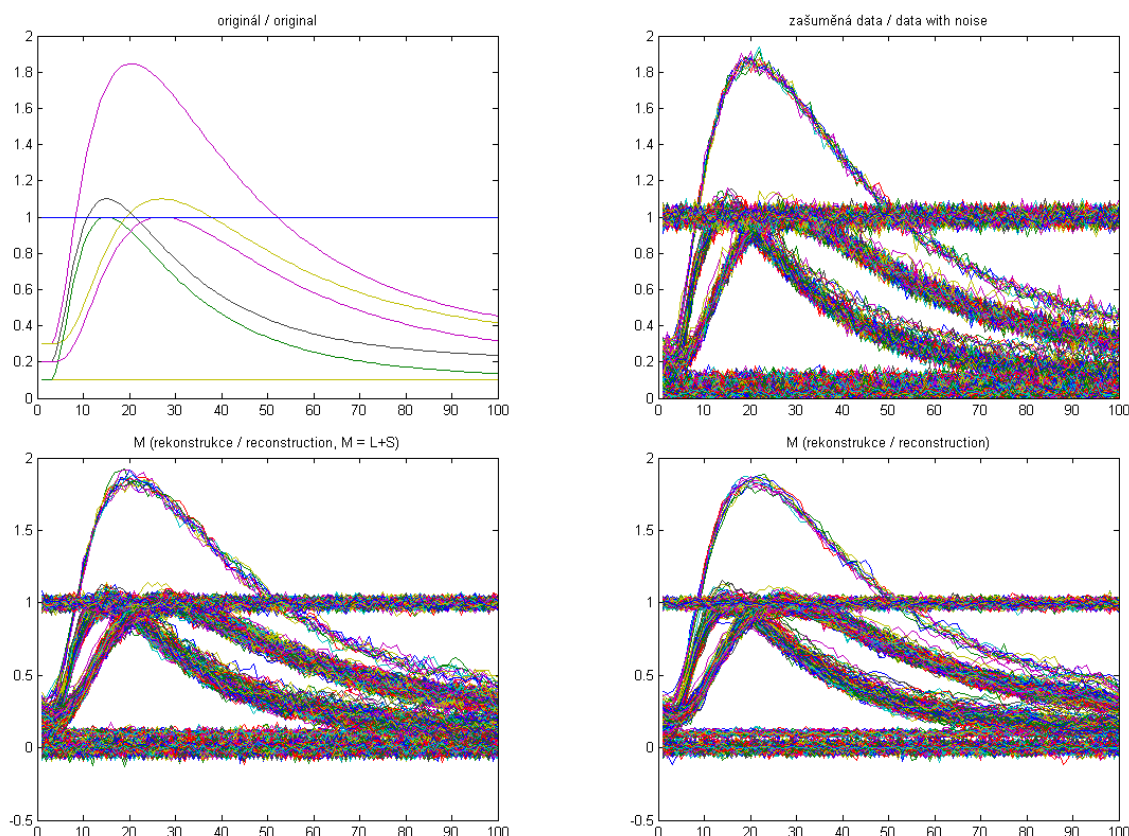
Protože v reálných datech se vyskytuje velké množství šumu, je důležité vědět, jak moc ovlivní rekonstrukci měřených perfúzních dat. Šum se v magnetické rezonanci modeluje jako aditivní šum v k -prostoru, tj. měřená data v každém snímku k -prostoru $\mathbf{D} = \tilde{\mathbf{D}} + \xi$, kde $\tilde{\mathbf{D}}$ jsou nezašuměná data a ξ šum modelu. Předpokládáme, že šum $\xi = \mathbb{X} + i\mathbb{Y}$, kde $\mathbb{X}$ a $\mathbb{Y}$ jsou nezávislé náhodné veličiny se stejným rozdělením $N(0, \sigma^2)$.

Porovnávala jsem závislost úspěšnosti rekonstrukce v závislosti na směrodatné odchylce šumu pro $\mathbf{L} + \mathbf{S}$ model a $\mathbf{M}$ model. Protože úspěšnost rekonstrukce závisí na volbě regularizačních konstant λ , nejprve jsem zjistila, pro jakou volbu parametrů je SNR (téměř) nejvyšší při dané směrodatné odchylce šumu. Na obr. 6.15 je zobrazeno nejvyšší dosažené SNR v závislosti na směrodatné odchylce šumu σ pro oba modely. Z grafu je vidět, že oba modely mají podobnou úspěšnost rekonstrukce pod vlivem šumu.

Konkrétní výsledek je možno vidět na obr. 6.16. Samotná zašuměná data se směrodatnou odchylkou šumu $\sigma = 0,04$ mají SNR 16,15 dB. Poté bylo provedeno měření z 50 polopřímek s využitím masky typu radiální paprsky. Následná rekonstrukce pomocí $\mathbf{L} + \mathbf{S}$ modelu dosáhla SNR 35,77 dB nastavením $\lambda_L = 1,4$ a $\lambda_S = 0,1$ a pomocí $\mathbf{M}$ modelu SNR 36,03 dB nastavením $\lambda_L = 0,6$ a $\lambda_S = 0,6$. Data se nepodařilo při rekonstrukci zbavit šumu, ale podařilo se zvýšit SNR.



Obrázek 6.15: Závislost úspěšnosti rekonstrukce na směrodatné odchylce šumu σ při volbě regularizačních konstant tak, aby bylo dosaženo co nejlepšího SNR. Byla použita maska typu radiální paprsky a rekonstrukce byla provedena z 50 polopřímek ($\doteq 20\%$)



Obrázek 6.16: Vlevo nahoře originální perfúzní data a vedle s přidáním aditivním šumem se směrodatnou odchylkou 0,04 (zobrazena absolutní hodnota). Dole rekonstruovaná data z 50 polopřímek užitím masky typu radiální paprsky pomocí $L + S$ modelu vlevo a M modelu vpravo.

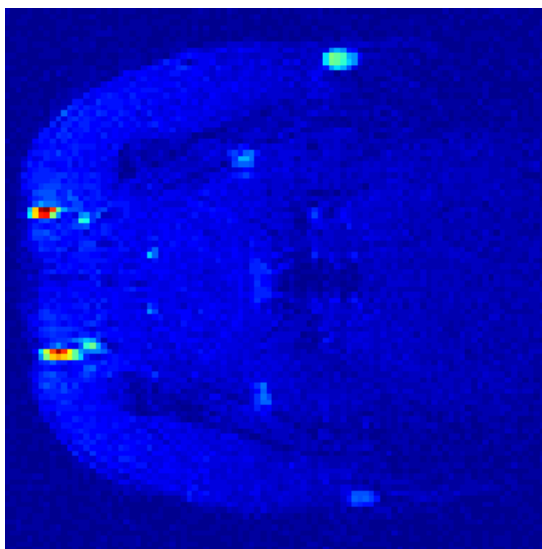
6.7. Reálná data

Jako poslední jsem výše uvedené metody testovala na reálných datech perfúzního zobrazování pomocí magnetické rezonance. Pro tyto účely mi byla poskytnuta perfúzní data myši od Ústavu přístrojové techniky Akademie věd ČR. Jednotlivé snímky těchto dat jsou rozměru 96×128 px a obsahují celkem 600 snímků pořízených v čase (jeden snímek je pořízený přibližně za 4 s). Protože se ale v magnetické rezonanci častěji používají snímky čtvercového rozměru, napsaný program je určený pouze pro čtvercové snímky, a proto pro účely simulace jsem využila pouze necelou část (oříznutou v prostorové oblasti zprava a zleva) každého snímku o velikosti 96×96 px. První snímek těchto (oříznutých) dat lze vidět na obr. 6.17 a perfúzní křivky myši na obr. 6.18. Odsud je patrné, že všechny perfúzní křivky už nemají ideální modelový průběh.

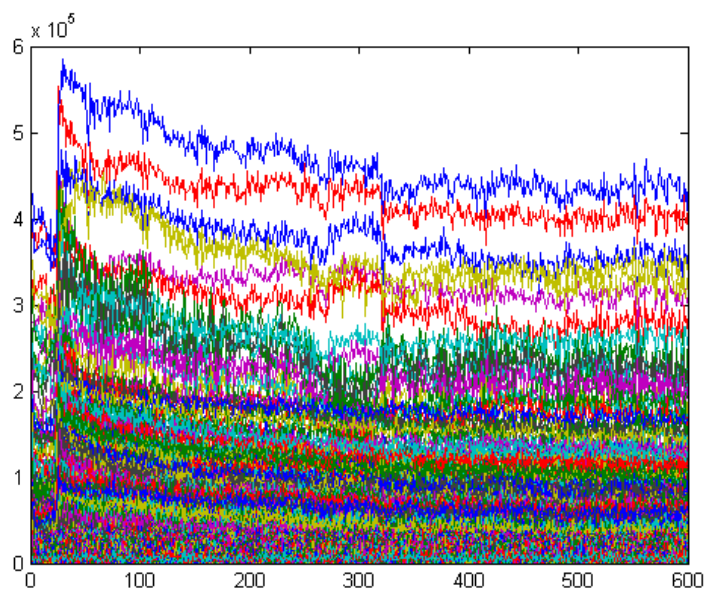
Poskytnutá data jsou už v prostorové oblasti, tzn. inverzní Fourierova transformace naměřených dat. Protože se v magnetické rezonanci vyskytuje šum, spektrum měřených dat není komplexně sdružené (viz věta 6.1). Proto po provedení inverzní Fourierovy transformace naměřených dat nedostaneme reálné hodnoty. Dále se pak pracuje jenom s absolutní hodnotou těchto dat. (Poskytnutá perfúzní data myši tedy obsahují pouze absolutní velikost naměřených hodnot po provedení zpětné Fourierovy transformace.) Proto při simulaci, kdy měříme data přímo v k -prostoru, dostaneme trochu jiné koeficienty než původní. Také rekonstrukci nelze porovnat k datům bez šumu jako u fantomu.

Při simulaci komprimovaného snímání na těchto reálných datech jsem pomocí masky typu radiální paprsky naměřila 50 polopřímek, což je přibližně 36 %. Pomocí **L+S** modelu se podařilo dosáhnout SNR 39,17 dB a pomocí **M** modelu SNR 40,42 dB, což je méně než u simulovaných dat. Regularizační parametry byly přitom nastaveny $\lambda_L = 100\,000$ a $\lambda_S = 10\,000$. Jejich hodnoty byly zvoleny tak, aby rekonstrukce měla co nejvyšší SNR. Jsou výrazně větší než u fantomu, což je dáno tím, že reálná data myši obsahují také podstatně vyšší hodnoty koeficientů.

Podařilo se tedy ověřit, že pomocí komprimovaného snímání lze naměřit výrazně méně koeficientů při zachování určité přesnosti. Díky tomu lze měření zrychlit, a tím naměřit data s vyšším prostorovým a časovým rozlišením.



Obrázek 6.17: První snímek z reálných perfúzních dat myši



Obrázek 6.18: Perfúzní křivky myši

7. Závěr

Perfúzní zobrazování pomocí magnetické rezonance je experimentální zobrazovací metoda určená pro lékařskou diagnózu. Avšak pro perfúzní analýzu je nezbytně nutné mít vysoké časové a prostorové rozlišení současně, což je ale nemožné s využitím klasického měření v magnetické rezonanci.

Tato diplomová práce nejprve představuje problematiku perfúzního zobrazování a magnetické rezonance. Dále se zabývá tzv. řídkými reprezentacemi a komprimovaným snímáním, které umožňuje zrychlení snímacího procesu. Následně jsou uvedeny numerické metody pro minimalizační problémy používané pro rekonstrukci naměřených perfúzních dat.

Hlavním cílem této diplomové práce bylo využití komprimovaného snímání pro simulaci perfúzního zobrazování pomocí magnetické rezonance v programovém prostředí MatLabu. Pro tyto účely byl nejprve vytvořen fantom perfúzního zobrazování a navrženy měřicí matice pro komprimované snímání. K sestavení několika modelů pro následnou rekonstrukci perfúzních dat bylo využito několika apriorních informací o měřených datech.

Jako první byl představen $\mathbf{L} + \mathbf{S}$ model využívající nízké maticové hodnoty dat a řídkosti perfúzních křivek. Následně byl tento model vylepšen přidáním diskrétní 1D totální variace. Jako poslední byl uveden $\mathbf{M}$ model pro rekonstrukci. Porovnáním s $\mathbf{L} + \mathbf{S}$ modelem se ukázalo, že $\mathbf{M}$ model je při stejné kvalitě rekonstrukce rychlejší pro výpočet. Dále byl zkoumán vliv aditivního šumu v měřených datech na úspěšnost rekonstrukce perfúzních dat. Bylo zjištěno, že $\mathbf{L} + \mathbf{S}$ i $\mathbf{M}$ model mají podobnou stabilitu pod vlivem šumu.

Nakonec byly oba modely ověřeny na reálných perfúzních datech myši. U rekonstrukce bylo dosaženo nižšího SNR než u simulovaných dat, ale pořád je možno provést mnohem méně měření, čímž se sníží doba snímání, při zachování vysokého SNR nutného k perfúzní analýze.

Některé výsledky této práce byly publikovány v článku [8] ve sborníku workshopu Nové směry v biomedicínském inženýrství.

Literatura

- [1] Bazaraa, M. S.; Sherali, H. D.; Shetty, C. M.: Nonlinear programming: theory and algorithms. 3rd ed. Hoboken: John Wiley, 2006, xv, 853 s. ISBN 04-714-8600-0.
- [2] Candès, E. J.; Recht, B.: Exact Matrix Completion via Convex Optimization. *Foundations of Computational Mathematics*. 2009, vol. 9, issue 6, s. 717-772. DOI: 10.1007/s10208-009-9045-5.
- [3] Candès, E. J.; Wakin, M. B.: An introduction to compressive sampling. *IEEE Signal Processing Magazine*, ročník 25, č. 2, 2008: s. 21–30, ISSN 1053-5888.
- [4] Combettes, P. L.; Pesquet, J.-C.: Proximal Splitting Methods in Signal Processing. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer, 2011.
- [5] Combettes, P. L.; Wajs, V. R.: Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* 4, 1168–1200 (2005)
- [6] Condat, L.: A Direct Algorithm for 1D Total Variation Denoising. *IEEE Signal Processing Letters*. 2013, vol. 20, issue 11, s. 1054-1057. DOI: 10.1109/LSP.2013.2278339.
- [7] Daňková, M.: *Aplikace lineární algebry a optimalizace ve zpracování obrazů*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2014. 38 s. Vedoucí Mgr. Pavel Rajmic, Ph.D.
- [8] Daňková, M.; Rajmic, P.; Jiřík, R.: *Komprimované snímání perfúzního zobrazování v MRI*. Nové směry v biomedicinském inženýrství. Vyd. 1. V Brně: Vysoké učení technické, 2013, 106 s. ISBN 978-80-214-4814-8.
- [9] Elad, M.: *Sparse and redundant representations: from theory to applications in signal and image processing*. New York: Springer, 2010, 376 s. ISBN 978-1-4419-7010-7.
- [10] Fourcart, S.; Rauhut, H.: A mathematical introduction to compressive sensing. Dordrecht: Birkhäuser, xviii, 625 s. Applied and numerical harmonic analysis. ISBN 978-0-8176-4947-0.
- [11] Harabiš, V.; Kolář, R.; Mézl, M.; Jiřík, R.: Comparison and evaluation of indicator dilution models for bolus of ultrasound contrast agents. *Physiological Measurement*. 2013-02-01, vol. 34, issue 2, s. 151-162. DOI: 10.1088/0967-3334/34/2/151.
- [12] Hendee, W.; Morgan, C.: Magnetic resonance imaging Part I — Physical principles. *West J. Med.* 141(4), 491–500 (1984)
- [13] Hrbáček, R.: *Compressive sampling a simulace one-pixel camera*: bakalářská práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2011. 48 s. Vedoucí práce byl Mgr. Pavel Rajmic, Ph.D.

- [14] Hrbáček, R.: *Využití řídké reprezentace signálu při snímání a rekonstrukci v nukleární magnetické rezonanci*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2013. 108 s. Vedoucí práce byl Mgr. Pavel Rajmic, Ph.D.
- [15] Hrbáček, R.; Rajmic, P.; Veselý, V.; Špiřík, J.: Řídké reprezentace signálů: komprimované snímání. *Elektrorevue - Internetový časopis*, 2011, roč. 2011, č. 67, s. 1–8. ISSN: 1213- 1539. Dostupné z: <http://elektrorevue.cz/cz/download/ridke-reprezentace-signalu--komprimovane-snimani/>
- [16] Hrbáček, R.; Rajmic, P.; Veselý, V.; Špiřík, J.: Řídké reprezentace signálů: úvod do problematiky. *Elektrorevue - Internetový časopis*, 2011, roč. 2011, č. 50, s. 1–10. ISSN: 1213- 1539. Dostupné z: www.elektrorevue.cz/files/200000751-638ac6484b
- [17] Lingala, S. G.; Yue Hu, Dibella, E.; Jacob, M.: Accelerated Dynamic MRI Exploiting Sparsity and Low-Rank Structure: k-t SLR. *IEEE Transactions on Medical Imaging*. 2011, vol. 30, issue 5, s. 1042-1054. DOI: 10.1109/TMI.2010.2100850. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5705578>
- [18] Lustig, M.; Donoho, D. L.; Santos, J. M.; Pauly, J. M.: Compressed Sensing MRI. *IEEE Signal Processing Magazine*. 2008, vol. 25, issue 2, s. 72-82. DOI: 10.1109/MSP.2007.914728. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4472246>
- [19] Mézl, M.; Jiřík, R.; Harabiš, V.: *Akvizice a zpracování dat v ultrazvukové perfuzní analýze. Nové směry v biomedicínském inženýrství*. Vyd. 1. V Brně: Vysoké učení technické. ISBN 978-80-214-4814-8.
- [20] Otazo, R.; Candès, E. J.; Sodickson, D.: Low-rank and sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. Submitted to *Magnetic Resonance in Medicine*, Sept. 2013.
- [21] Qian, H.; Bassingthwaite, J. B.: A Class of Flow Bifurcation Models with Lognormal Distribution and Fractal Dispersion. *Journal of Theoretical Biology*. 2000, vol. 205, issue 2, s. 261-268. DOI: 10.1006/jtbi.2000.2060. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0022519300920605>
- [22] Shepp, L. A.; Logan, B. F.: The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science*. 1974, vol. 21, issue 3, s. 21-43. DOI: 10.1109/TNS.1974.6499235. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6499235>
- [23] Starčuk, Z.; Krupa, P.; Starčuk jr., Z.; Horký, J.: ^1H in vivo MR spektroskopie v klinické neurologii. *Neurologie pro praxi*, ročník 6, č. 3, 2005: s. 140–148, ISSN 13359592. Dostupné z: http://www.solen.sk/index.php?page=pdf_view&pdf_id=675&magazine_id=3
- [24] Difusion & perfusion imaging [online]. 2006 [cit. 2014-03-10]. Dostupné z: <http://spinwarp.ucsd.edu/neuroweb/Text/br-710epi.htm>
- [25] fMRI [online]. 2013 [cit. 2014-03-22]. Dostupné z: <http://fmri.mchmi.com>

- [26] *Wikipedie – internetová encyklopedie*: Logaritmická spirála [online]. 2014 [cit. 2014-04-21]. Dostupné z: http://cs.wikipedia.org/wiki/Logaritmická_spirála
- [27] Lustig, M.: Principles of MRI [online]. 2013 [cit. 2014-03-22]. Dostupné z: http://inst.eecs.berkeley.edu/~ee225e/sp13/notes/Lecture01_011712.pdf
- [28] *Wikipedie – internetová encyklopedie*: Polospojitosť [online]. 2014 [cit. 2014-04-11]. Dostupné z: <http://cs.wikipedia.org/wiki/Polospojitosť>

Seznam použitých zkratek a symbolů

Použité symboly

c	konstantní hodnota bez použití kontrastní látky (lognormální model)
t_0	zpoždění bolusu mezi místem podání a oblastí zájmu (lognormální model)
S	plocha mezi konstantní hodnotou c a křivkou (lognormální model)
μ, σ	parametry lognormálního rozdělení (lognormální model)
μ	magnetický moment
S	spin (moment hybnosti částice)
B_0	indukce magnetického pole
f_0	Larmorova frekvence
γ	gyromagnetický poměr
M	magnetizace
T_1, T_2, T_2^*	relaxační časy
$\mathbf{G}$	gradient magnetického pole
$\mathbf{I}$	jednotková matice
k	řidkost vektoru
$\ \cdot\ _p$	p norma vektoru $\mathbf{x}$
$\ \cdot\ _{\text{tv}}$	diskrétní 1D totální variace
$\ \cdot\ _*$	nukleární norma
$\text{spark}(\mathbf{A})$	spark matice $\mathbf{A}$
$\ker \mathbf{A}$	jádro lineárního zobrazení určeného maticí $\mathbf{A}$
$\mu(\mathbf{A})$	vzájemná koherence
γ	konstanta nulového prostoru
$\sigma_k(\mathbf{x})_p$	chyba nejlepší aproximace vektoru $\mathbf{x}$ k -řidkým vektorem
δ_k	konstanta zeslabené izometrie
∂f	subdiferenciál funkce f
ι_C	indikátorová funkce množiny C
prox_f	proximální operátor funkce f
soft	měkké prahování

Použité zkratky

BF	krvní tok (blood flow)
BV	objem krve (blood volume)
CS	komprimované snímání (compressed sensing)
FID	volné doznívání indukce (free induction decay)
IRLS	iterativně převažovaná metoda nejmenších čtverců (iterative reweighted least squares)
MP	sdužovací metoda (matching pursuit)
MR	magnetická rezonance
MRI	zobrazování magnetickou rezonancí (magnetic resonance imaging)
MTT	střední doba průchodu (mean transit time)
OMP	ortogonální sdužovací metoda (orthogonal matching pursuit)
rCBF	relativní mozkový krevní tok (relative cerebral blood flow)
rCBV	relativní objem krve mozku (relative cerebral blood volume)
RF	radiofrekvenční (pulz)
SNR	poměr signálu k šumu (signal to noise ratio)
SVD	rozklad na singulární čísla (singular value decomposition)
TE	čas ozvěny (time echo)
TR	čas opakování (time repeat)
TTP	čas dosáhnutí maximální intenzity (time to peak)

A. Obsah CD

Příložené CD obsahuje tyto zdrojové kódy v programovém prostředí MatLab (2010b):

<i>demo_LS_decomposition_MRI.m</i>	hlavní program, který simuluje měření videa perfúzního zobrazování v MRI pomocí komprimovaného snímání a následně provádí rekonstrukci dat, která mohou být jak simulovaná, tak i reálná
<i>conj_coef.m</i>	dopočítávání komplexně sdružených Fourierových koeficientů
<i>export_avi.m</i>	export 3D dat do formátu .avi
<i>generate_mask.m</i>	vytvoření měřicí masky pro celé video
<i>lognormal_model.m</i>	vytvoření perfúzních křivek pomocí lognormálního modelu
<i>LS_decomposition.m</i>	rekonstrukce dat pomocí L + S modelu
<i>M_decomposition.m</i>	rekonstrukce dat pomocí M modelu
<i>mask.m</i>	vytvoření jedné měřicí matice
<i>nonuniform_random_mask.m</i>	vytvoření nerovnoměrné náhodné masky
<i>phantom3d.m</i>	Shepp-Loganův fantom ve 3D (autor Matthias Christian Schabel)
<i>phantom_video.m</i>	vytvoření fantomu perfúzních dat
<i>prox_tv1d.m</i>	proximální operátor diskrétní 1D totální variace
<i>MGEmultiFA_sec01_inp_130328.mat</i> ...	perfúzní data myši