



STŘEDOŠKOLSKÁ ODBORNÁ ČINNOST

Řídce vzorkovaná data a jejich zpracování ve
spektroskopii nukleární magnetické rezonance

AUTOR Benedikt Peřko
ŠKOLA Gymnázium Matyáše Lercha, Brno
KRAJ Jihomoravský
ŠKOLITEL Mgr. Jiří Nováček, Ph.D.
OBOR 2. Fyzika

Brno 2014

Prohlášení

Prohlašuji, že svou práci na téma Řídce vzorkovaná data a jejich zpracování ve spektroskopii nukleární magnetické rezonance jsem vypracoval samostatně pod vedením Mgr. Jiřího Nováčka, Ph.D. a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Dále prohlašuji, že tištěná i elektronická verze práce SOČ jsou shodné a nemám závažný důvod proti zpřístupňování této práce v souladu se zákonem č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a změně některých zákonů (autorský zákon) v platném znění.

V Brně dne 16. 3. 2014

Podpis:



Jihomoravský kraj

**MASARYKOVA
UNIVERZITA**

**Přírodovědecká
fakulta**

**Národní centrum pro
výzkum biomolekul**

**Středoevropský
technický institut**

Poděkování:

Tímto bych chtěl poděkovat svému školiteli Mgr. Jířímu Nováčkovi, Ph.D. za ochotu, přátelský přístup a mnoho času, který této práci věnoval. Také děkuji za hodiny konzultací, které pro mne vždy byly příjemně stráveným časem.

Dále děkuji Jihomoravskému kraji a JCM za finanční podporu této práce.

ABSTRAKT:

Práce se zabývá zpracováním signálu měřeného ve spektroskopii nukleární magnetické rezonance (NMR) a rekonstrukcí NMR spekter z řídkce navzorkovaných dat. Je provedena analýza nejpoužívanějších algoritmů generování schémat pro měření řídkce vzorkovaného NMR signálu. Rekonstruovaná spektra jsou následně porovnávána na základě parametrů určujících kvalitu NMR dat (citlivost, rozlišitelnost signálů). Účelem práce je nalezení optimálního rozložení řídkce vzorkovaných dat poskytující maximální citlivost a rozlišení. Jednotlivá schémata jsou podrobena simulaci a výstupní data jsou graficky porovnávána. Následně je posuzována náhodnost schematických rozdělení a zkoumána podstata spektrálního šumu.

Klíčová slova: NMR, zpracování dat, řídké vzorkování, citlivost, rozlišení, spektrální šum

ABSTRACT:

This work deals with processing of signal in nuclear magnetic resonance spectroscopy (NMR) and reconstruction of NMR spectra from sparsely sampled data. An analysis of algorithms for generating sparsely sampled NMR signal is conducted. Reconstructed spectra are compared by parameters determining the quality of NMR data (sensitivity and resolution of signals). The goal of this work is to find an optimal distribution of sparsely sampled data points yielding the highest sensitivity and resolution. Each scheme is subjected to simulation and the outputs are compared graphically. Afterwards, the randomness of the schematical distributions is examined and the nature of spectral noise is investigated.

Key words: NMR, data-processing, sparse sampling, sensitivity, resolution, spectral noise

Obsah

ÚVOD	5
1. TEORIE	6
1.1. NMR EXPERIMENT	6
1.1.1. <i>Signál</i>	7
1.1.2. <i>Vícerozměrový experiment</i>	8
1.1.3. <i>Termální šum</i>	8
1.1.4. <i>Citlivost a rozlišení</i>	9
1.2. ZPRACOVÁNÍ DAT	10
1.2.1. <i>Fourierova transformace</i>	10
1.2.2. <i>Vícerozměrová Fourierova transformace</i>	12
1.2.3. <i>Nyquistův teorém</i>	13
1.2.4. <i>Optimální délka měření</i>	14
2. ŘÍDKÉ VZORKOVÁNÍ	16
2.1. NÁHODNÉ VZORKOVÁNÍ	16
2.2. SCHÉMATICKÉ VZORKOVÁNÍ	16
2.2.1. <i>Zachování rozlišení</i>	19
2.2.2. <i>Porovnání citlivosti</i>	20
2.2.3. <i>Dva signály</i>	23
2.2.4. <i>Rozložení šumu</i>	25
2.2.5. <i>Ukazatel kvality</i>	27
ZÁVĚR	30
LITERATURA	31
SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK	32
SEZNAM PŘÍLOH	33

Úvod

Spektroskopie nukleární magnetické resonance (NMR) je jedinečnou metodou, která umožňuje studovat biomakromolekuly (proteiny, sacharidy, nukleové kyseliny a jejich komplexy) v roztoku na atomární úrovni. Analýzou NMR dat lze získat informaci o vzájemném rozmístění jednotlivých atomů ve studované molekule a tím určit její trojrozměrný strukturní model nebo studovat interakce, jichž se biomakromolekuly účastní. Díky tomu nalézá NMR uplatnění například ve farmaceutickém průmyslu.

Princip NMR je založen na existenci nenulového vnitřního momentu hybnosti některých jader - jaderného spinu. Při interakci spinu s magnetickým polem dochází k precesi magnetického momentu jader okolo vektoru určujícího směr magnetického pole. Při umístění zkoumané látky do magnetického pole se pak tato látka vyznačuje nenulovou magnetizací. Typický NMR experiment je iniciován působením radiofrekvenčního pulsu na zkoumanou látku umístěnou v magnetickém poli. Výstupem NMR měření je časová závislost celkové magnetizace při návratu do rovnovážného stavu (tzv. FID – Free Induction Decay).

Jedním z několika úskalí této metody je její nízká citlivost, která se projevuje výrazným zastoupením elektronického šumu v měřených datech a vlastní NMR signál je tak ve spektrech často nerozpoznatelný. K odlišení signálu od šumu je proto nutné akumulovat data z opakovaných NMR měření, což je hlavním důvodem časové náročnosti NMR experimentů. Nicméně zkoumané látky se často vyznačují nízkou stabilitou, která vyžaduje, aby všechna měření byla provedena v co nejkratší možné době. Z tohoto důvodu a také z důvodu finanční náročnosti NMR měření se vyvíjejí metody, které umožňují kompletní rekonstrukci signálu s nízkými nároky na množství experimentálně získaných dat.

Tato práce se zaměřuje na řídké vzorkování NMR dat, které umožňuje výrazně snížit časovou náročnost měření. Hlavní motivací je provést kompletní rekonstrukci NMR signálu z nekompletní sady experimentálních dat při zachování důležitých charakteristik signálu (rezonanční frekvence, rozlišitelnost signálů) a docílit tak radikálního snížení doby experimentu. Pro řídké vzorkování signálu byla v literatuře navržena různá schémata a tato práce posuzuje jejich kvalitu a vhodnost použití pomocí počítačové simulace signálu.

1. Teorie

1.1. NMR EXPERIMENT

NMR se zabývá studiem biomakromolekul v kapalné nebo pevné fázi. Při umístění jádra atomu s nenulovým spinem do magnetického pole dochází k periodickému pohybu (rotaci) vektoru magnetického momentu jádra. Frekvence této rotace je charakteristická pro každé jádro zkoumané látky a je označována jako rezonanční frekvence daného jádra. Cílem NMR experimentu je identifikace jednotlivých jader ve vzorku na základě jejich zjištění jejich rezonančních frekvencí.¹

Průběh a princip experimentu:

Společně se spinem jádra existuje jeho magnetický moment, který je souhlasně či protikladně orientovaný:

$$\hat{\mu} = \gamma \hat{S} \quad (1.1)$$

Jeho směr i velikost tedy závisí na gyromagnetickém poměru (γ), který může být kladný či záporný, a spinu (\hat{S}).

Na počátku jsou magnetické momenty jader v rovnováze a zkoumaná látka nevykazuje žádnou vnější magnetizaci. Tato rovnováha je porušena působením magnetického pole (konvenčně v ose z) o síle 5-20 Tesla, což důsledkem přítomnosti jaderného spinu způsobuje rotaci magnetického momentu jader po trajektorii opisující plášť kužele. Vlivem perturbace magnetického pole způsobené okolními jádry a elektrony se tato trajektorie mění a celkový magnetický moment směřuje ve výsledku souhlasně s aplikovaným magnetickým polem. Tím se vzorek začne navenek vyznačovat nenulovou celkovou magnetizací M :

$$M(t) = M_{eq}(1 - e^{-t/T_1}) \quad (1.2)$$

kde M značí celkovou magnetizaci vzorku a T_1 konstantu longitudinální magnetizace. Tento jev je důsledkem energetické výhodnosti souhlasné orientace magnetického momentu jader se směrem vnějšího magnetického pole, neboť platí:

$$E = -\mu \cdot B \quad (1.3)$$

Protože signál ve směru vnějšího magnetického pole je obtížně detekovatelný, sklápí se získaná magnetizace vzorku pomocí radiofrekvenčního pulsu do roviny xy kolmé k hlavnímu magnetickému poli. Následnou rotaci celkového magnetického momentu lze zaznamenat cívkami položenými v této rovině. Periodická oscilace magnetického momentu způsobuje vznik střídavého elektrického proudu v detekčních cívkách, jehož frekvence je shodná s frekvencí rotace vektoru magnetizace. Důsledkem desynchronizace jednotlivých nukleárních magnetických momentů dochází ke ztrátě signálu v rovině xy

$$M(t) = M_{eq} e^{-t/T_2}, \quad (1.4)$$

kde T_2 značí konstantu transverzální magnetizace.

1.1.1. Signál

Cílem NMR spektroskopie je získat informaci o rezonanční frekvenci jader při precesi kolem hlavního magnetického pole. Úhlová rychlost precese se nazývá Larmorova frekvence a je dána vztahem¹

$$\boldsymbol{\omega} = -\gamma \mathbf{B}_0 \quad (1.5)$$

Protože je pozorovaná frekvence úměrná velikosti aplikovaného magnetického pole \mathbf{B}_0 , vztahuje se v praxi pozorovaná rezonanční frekvence k rezonanční frekvenci referenčního jádra, aby mohla být jednoduše porovnána data měřená při různých magnetických polích \mathbf{B}_0 :

$$\delta = \frac{\omega - \omega_{ref}}{\omega_{ref}} \quad (1.6)$$

kde δ značí chemický posun a ω_{ref} rezonanční frekvenci referenčního jádra.

Aby bylo možné určit směr precese, musí být signál detekován ze dvou směrů. Tento postup se nazývá kvadrurní detekce a jeho výstupem je komplexní signál. Cívkami spektrometru prochází v průběhu experimentu střídavý proud, jenž je formou signálu: $x_X(t) = Ae^{-R_2 t} \cos \omega t + \varepsilon$ a $x_Y = Ae^{-R_2 t} \sin \omega t + \varepsilon$

Tyto výrazy jsou ekvivalentní k rovnici

$$x(t) = Ae^{i\omega t - R_2 t} + \varepsilon \quad (1.7)$$

1.1.2. Vícerozměrový experiment

V praxi se často měří korelace mezi více jádry. Takový signál se získává NMR experimenty sestávajícími se z většího množství radiofrekvenční pulzů. Získaný signál se potom dělí na jedinou přímo a několik nepřímo detekovaných dimenzí.

Měřením získáváme časovou doménu signálu, která má D dimensí, přitom první je měřena přímo a zbylé nepřímo, na základě opakování NMR experimentu při současně změně délky prodlev mezi radiofrekvenčními pulzy. Tímto dochází k amplitudové modulaci výsledného přímo detekovaného signálu. Každá dimenze musí obsahovat reálnou a imaginární složku signálu, které jsou získávány kvadraturní detekcí, proto pro získání jedné naměřené hodnoty v nepřímých dimenzích je nutné NMR experiment opakovat 2^D krát.

Uvažujme dvoudimensionální experiment. NMR experiment začne aplikací radiofrekvenčního pulzu. Po něm následuje prodleva τ_2 , která se mění po diskrétních hodnotách při jednotlivých opakováních NMR experimentu, a za ní série radiofrekvenčních pulzů, po které je signál ponechán volnému zániku a je detekován v po dobu τ_1 . Tento postup je mnohokrát opakován se změnou času τ_2 , aby byla vytvořena matice dat $\tau_1 \times \tau_2$.

Obecně v D-dimenzionálním experimentu pokrývají pulsy v různých časech $\tau_2[n_2\Delta t_2]$ až $\tau_D[n_D\Delta t_D]$ tensor D-tého řádu (D-dimenzionální kvádr) s hodnotami signálu jako (D+1)ou dimensí.

1.1.3. Termální šum

Metoda NMR je zatížena dvěma typy šumu: termální, pocházející z nedokonalosti spektrometru, a spektrální, jež je způsoben například omezením šířky frekvenční domény (tzv. aliasing) nebo diskontinuitou v časové doméně.

Termální šum má několik vlastností plynoucích z jeho náhodnosti.²

Je Gaussovsky náhodný, proto vychyluje výsledek měření v kladném i záporném smyslu:

$$\sum_i \varepsilon_i = 0 \quad (1.8)$$

Je nekorelovaný:

$$\sum_i \varepsilon_i \varepsilon_j = 0 \quad (1.9)$$

Má nenulovou energii:

$$\sum_i \varepsilon_i^2 = \sigma^2 \quad (1.10)$$

kde σ značí střední kvadratickou hodnotu šumu.

Potom při součtu dvou různých náhodných sekvencí šumu platí:

$$\begin{aligned} \sum_i (\varepsilon_i + \varepsilon'_i)^2 &= \sum_i (\varepsilon_i^2 + 2\varepsilon_i \varepsilon'_i + \varepsilon_i'^2) = \sum_i \varepsilon_i^2 + \sum_i \varepsilon_i'^2 = 2\sigma^2 = \sigma_0^2 \\ \sigma_0 &= \sqrt{2}\sigma \end{aligned} \quad (1.11)$$

Z toho lze indukovat obecný závěr:

$$\sum_i (\sum_j \varepsilon_{ij})^2 = \sum_i (\sum_j \varepsilon_{ij}^2 + \sum_{k \neq j} \sum_j \varepsilon_{ij} \varepsilon_{ik}) = \sum_i \sum_j \varepsilon_{ij}^2 = n \sum_i \varepsilon_{ij}^2 = n\sigma^2 = \sigma_0^2$$

kde $n = \sum_j 1$. Potom

$$\sigma_0 = \sqrt{n}\sigma \quad (1.12)$$

Energie i střední hodnota šumu tedy stoupá s odmocninou z počtu měřených bodů.

1.1.4. Citlivost a rozlišení

Jedním z ukazatelů kvality spektra je dosažitelné rozlišení, jež je určeno kvantitou šířky signálu v polovině jeho výšky. Čím lepší rozlišení, tedy čím užší signály, tím je možné ve spektrech rozlišit signály o méně rozdílných rezonančních frekvencích. Naopak v případě nízkého rozlišení se signály o blízkých rezonančních frekvencích vzájemně překrývají bez možnosti přesné identifikace.

Signál v reálné frekvenční doméně je reprezentován absorpčním lorentziánem (viz 1.2.1):

$$X(\omega) = \frac{R_2}{R_2^2 + (\omega - \omega_0)^2} \quad (1.13)$$

kde $R_2 = \frac{1}{T_2}$ značí prudkost zániku signálu. Šířka absorpčního lorentziánu v polovině maxima je dána vztahem $L = \frac{R_2}{\pi}$ v Hz^2 . Z toho plyne, že pomaleji zanikající signál poskytuje lepší rozlišení.

Čím déle jsme takový signál schopni měřit, tím lepší rozlišení získáme (viz 1.2.1 rovnice (1. 26)). Na druhou stranu příliš dlouhá doba měření způsobí, že šum začne převládat nad vlastním signálem, což snižuje citlivost experimentu.

Citlivost je mírou rozeznatelnosti signálu od šumu a často je vyjádřen jako poměr maxima signálu ku střední nebo maximální hodnotě šumu, tedy jako

$$S/N = \frac{S_{max}}{N_{rms}} \quad (1.14)$$

nebo

$$S/N = \frac{S_{max}}{N_{max}} \quad (1.15)$$

V této práci byl pro ukazatele citlivosti zvolen vztah (1. 15).

1.2. Zpracování dat

1.2.1. Fourierova transformace

Jelikož požadovaným výstupem experimentu je frekvence jádra, aplikuje se na časově proměnný signál Fourierova transformace:

$$X(\nu) = \int_{-\infty}^{+\infty} x(t)e^{-i\nu t} dt \quad (1.16)$$

Signál snímaný detekčními cívkami je z důvodu zpracování a uložení na výpočetní technice digitalizován, tedy převeden do diskrétní podoby:

$$x[n] = \int_{-\infty}^{\infty} x(t)\delta(t - n\Delta t) dt \quad (1.17)$$

kde Δt je převrácenou hodnotou snímací frekvence. Funkce signálu v diskretní podobě je

$$x[n] = Ae^{i2\pi f n - R_2 n} + \varepsilon \quad (1.18)$$

Pro diskretní body přechází integrál v sumu a Fourierovou transformací v diskretním tvaru lze přejít do frekvenční domény:

$$x[n] \xrightarrow{\mathcal{F}} X[k]$$

kde n jsou diskretní body v časové doméně (nebo také $n\Delta t = t$) a $\frac{k}{\Delta t}$ je frekvence v hertzích. Samotná rovnice analýzy diskretní Fourierovy transformace má tvar

$$X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{i2\pi kn}{N}\right) \quad (1.19)$$

kde N je počet měřených bodů.

V maticové podobě nabývá Fourierova transformace tvaru²

$$\mathbf{F}_{kn} = \begin{bmatrix} e^{-\frac{i2\pi n}{N}} & \dots & e^{-i2\pi n} \\ \vdots & e^{-\frac{i2\pi kn}{N}} & \vdots \\ e^{-i2\pi n} & \dots & e^{-i2\pi nN} \end{bmatrix} \quad (1.20)$$

který se aplikuje na získaná data časové domény tvořící N -dimensionální vektor \mathbf{x} . Jak lze vidět z rovnice analýzy, sumace probíhá přes index n , proto podle pravidla násobení matic zastupuje pořadí sloupců. Fourierova transformace znamená promítnutí signálového vektoru na bázi inverzní Fourierovy transformace, jež je hermitovským konjugátem matice samotné Fourierovy transformace, jak plyne z rovnice syntézy

$$x[n] = \sum_{k=0}^{N-1} X[k] \exp\left(\frac{i2\pi kn}{N}\right) \quad (1.21)$$

v maticovém tvaru potom

$$\mathbf{F}^{-1} = \mathbf{F}_{nk}^* = \begin{bmatrix} e^{\frac{i2\pi}{N}} & \dots & e^{i2\pi} \\ \vdots & e^{\frac{i2\pi kn}{N}} & \vdots \\ e^{i2\pi} & \dots & e^{i2\pi N} \end{bmatrix} \quad (1.22)$$

Vlastnosti Fourierovy transformace vysvětlují rozlišení spektra. Dle konvolučního teorému je Fourierova transformace součinu dvou funkcí rovna konvoluci Fourierových transformací každé funkce. Potom pro signál definovaný v sekci 1.1.1 platí:

$$F(e^{i\omega t} e^{-R_2 t}) = F(e^{i\omega t}) * F(e^{-R_2 t}) \quad (1.23)$$

Fourierovou transformací exponenciály počínající v čase $t_0 = 0$ je:

$$F(e^{-R_2 t}) = \int_0^{+\infty} e^{-(i\omega + R_2)t} dt = \frac{R_2 - j\omega}{R_2^2 + \omega^2} \quad (1.24)$$

Tato rovnice vysvětluje tvar spektrálního vrcholu, který znázorňuje frekvenci signálu.

Měření probíhá po omezený čas, proto je signál násoben jednotkovými skoky. Fourierova transformace těchto skoků je:

$$\begin{aligned} F(u(t) - u(t_{max})) &= \int_{-\infty}^{+\infty} (u(t) - u(t_{max})) e^{-j\omega t} dt \\ &= \frac{\sin \omega t_{max} + i(\cos \omega t_{max} - 1)}{\omega} \end{aligned} \quad (1.25)$$

Důsledkem pro reálnou část spektra je zhoršení rozlišení, protože spektrum je dle (1.22) dáno konvolucí s funkcí

$$g(t) = \frac{\sin \omega t_{max}}{\omega} \quad (1.26)$$

1.2.2. Vícerozměrová Fourierova transformace

Vícerozměrový signál má tvar¹

$$x(t_1, \dots, t_D) = \prod_j (A_j e^{i2\pi f_j t_j - R_{j2} t_j} + \varepsilon_j) \quad (1.27)$$

Po proběhnutí digitizace

$$x[n_1, \dots, n_D] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x(t_1, \dots, t_D) \prod_j \delta(t_j - n_j \Delta t_j) dt_j \quad (1.28)$$

je získán tvar diskrétního signálu

$$x[n_1, \dots, n_D] = \prod_j (A_j e^{i2\pi f_j n_j - R_{j2} n_j} + \varepsilon_j), \quad (1.29)$$

kde D je počet dimenzí a ε je funkcí šumu. Časová doména signálu tak tvoří tensor D -tého řádu. Každá dimenze přechází do frekvenční domény nezávisle na ostatních², tzn.:

$$x[n_1, \dots, n_D] \rightarrow x_1[k_1, n_2, \dots, n_D] \rightarrow x_2[k_1, k_2, n_3, \dots, n_D] \rightarrow x_D[k_1, \dots, k_D] \equiv X$$

Fourierova transformace tak v D dimensích nabývá tvaru

$$X[k_1, \dots, k_D] = \sum_j \sum_{n_j=0}^{N-1} x[n_1, \dots, n_D] \prod_j \exp\left(-\frac{i2\pi k_j n_j}{N}\right) \quad (1.30)$$

V maticové podobě pro $D = 2$ dostáváme tvar Fourierovy transformace

$$X_{k_1 k_2} = (\mathbf{F}_{k_2 n_2} (\mathbf{F}_{k_1 n_1} \mathbf{x}_{n_1 n_2})^T)^T \quad (1.31)$$

1.2.3. Nyquistův teorém

Negativním důsledkem diskretizace je spektrální artefakt zvaný aliasing, který plyne z periodičnosti funkce e^{ix} . Spektrum se periodicky opakuje ve vyšších frekvencích, signály s vysokou frekvencí tak mohou zkreslovat spektrum tím, že se promítají na místech nižších frekvencí, naproti tomu ideální kontinuální tvar ($\Delta t \rightarrow 0$) je nezkreslený. Tento jev popisuje Nyquistův teorém.

Z tvaru diskrétní Fourierovy transformace vyplývá vlastnost limitující maximální velikost frekvence spektra bez aliasingu²:

$$\begin{aligned} \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi n}{N}(k+N/2)} &= \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi n k}{N}} e^{-i\pi n} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi n k}{N}} e^{i\pi n} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi n}{N}(k-\frac{N}{2})} \end{aligned}$$

$$X\left[\frac{N}{2}\right] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi n}{N}(0+N/2)} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi n}{N}(0-\frac{N}{2})} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i2\pi n\Delta t \frac{-N/2}{N\Delta t}}$$

$$= X\left[-\frac{N}{2}\right]$$

$$X\left[\frac{N}{2}\right] = X\left[-\frac{N}{2}\right] = X[k_{max}] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i(n\Delta t)\frac{-2\pi}{2\Delta t}}$$

$$|v_{max}| = \frac{1}{2\Delta t} \quad (1.32)$$

Nyquistův teorém tak říká, že aby bylo dosaženo spektra bez aliasingu, nesmí signál obsahovat frekvence vyšší než $\frac{1}{2\Delta t}$ a nižší než $-\frac{1}{2\Delta t}$, kde Δt značí časovou vzdálenost mezi body v časové doméně. Toto platí i naopak, totiž snímací frekvence ($\frac{1}{\Delta t}$) musí být dvakrát vyšší než nejvyšší frekvence ve spektru ($\frac{1}{2\Delta t}$).

1.2.4. Optimální délka měření

Čím déle je signál měřen, tím více šumu se projeví ve spektru. Bylo zjištěno⁶, že z hlediska citlivosti je nejvýhodnější délkou měření čas odpovídající $t_{max} = 1,26R_2^{-1}$, na místě je proto srovnání řídkého vzorkování s uniformním o stejné časové délce.

Důkaz⁶:

Během měření signál zaniká, a tak se zmenšuje jeho obsah ve frekvenční doméně. Výška signálu ve spektru bude úměrná jeho množství detekovaného v časové doméně:

$$X[k_0] \propto \int_0^{t_0} e^{-tR_2} dt = \frac{1 - e^{-tR_2}}{R_2} \quad (1.33)$$

$$X \equiv S$$

Ze sekce 1.1.3 z rovnice (1.12) vyplývá úměra mezi šumem a počtem měřených bodů (tím pádem i času):

$$\sigma_0 = \sigma\sqrt{n} = \sigma\sqrt{\frac{t_{max}}{\Delta t}} \quad (1.34)$$

Protože Δt značí převrácenou hodnotu snímací frekvence, a je tudíž konstantní, můžeme psát

$$\sigma_0 \propto \sqrt{t_{max}} \quad (1.35)$$

$$\sigma_0 \equiv N$$

Vložením obou časových úměr (1. 26) a (1. 28) do poměru získáváme vztah

$$S/N \propto \frac{1 - e^{-t_{max}R_2}}{R_2\sqrt{t_{max}}} \quad (1.36)$$

Maximum této úměry leží ve stacionárním bodě:

$$\frac{d}{dt_{max}} \frac{1 - e^{-t_{max}R_2}}{R_2\sqrt{t_{max}}} = 0 \quad (1.37)$$

$$\frac{(2R_2t_{max} + 1)e^{-t_{max}R_2} - 1}{2R_2t_{max}^{3/2}} = 0 \quad (1.38)$$

Jako řešení vyplývá:

$$t_{max} = \frac{\log(2R_2t_{max} + 1)}{R_2} \quad (1.39)$$

Jehož hodnotou je právě $t_{max} \approx 1,26R_2^{-1}$.

2. Řídké vzorkování

2.1. Náhodné vzorkování

Aby bylo dosaženo časové efektivity, je vhodné minimalizovat množství měřených dat a analyzovat, která data jsou esenciální pro rekonstrukci signálu.

Jednou z možností je vynechat měření dat na konci časové domény. Opodstatnění takového postupu je v tom, že tato data obsahují kvůli exponenciálnímu zániku méně signálu, zato však stále stejně šumu. Větší diskontinuita v časové doméně znamená ale nižší rozlišení (rovnice (1. 26) sekce 1.2.1).

Další možností je měření náhodných bodů z celé časové domény. Náhodné vzorkování zachovává rozlišení, neboť to je nejvíce určeno koncovými body časové domény. Negativem tohoto postupu je nižší citlivost, neboť některé body s vysokým obsahem signálu na počátku domény nejsou měřeny.

Mimo zcela náhodné vzorkování lze usměrnit koncentraci náhodně generovaných bodů v časové doméně pomocí schémat definovaných funkcemi, které znázorňují hustotu bodů v závislosti na čase. Dle literatury³⁴⁵ určuje různý výběr bodů časové domény podobu finálního spektra.

2.2. Schématické vzorkování

Tato kapitola rozebírá vlastnosti konkrétních schémat na základě statistiky získané počítačovou simulací signálu. Bylo simulováno 100 spekter každého rozdělení při hustotách 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 65%, 75%, 90%, 100%. Z těchto dat byly získány obecné charakteristiky signálu (viz 1.1.4). V dalších sekcích jsou znázorněny některé speciální případy spekter.

Pro simulaci signálu byl použit program Octave (www.gnu.org/software/octave/). Bylo naměřeno 512 bodů signálu s rezonanční frekvencí 350 Hz doprovázeným Gaussovsky rozloženým šumem a konstantou zániku $R_2 = 1,26/512$, která je zdůvodněna v sekci

1.2.4. Měření proběhla za použití šesti různých schémat generování náhodných bodů, které jsou vyjádřeny funkcemi hustoty rozložení bodů.

Uniformní rozdělení:

$$g(t) = u(t) - u(t - 1,26R_2^{-1}) \quad (2. 1)$$

Exponenciální rozdělení:

$$g(t) = \{u(t) - u(t - 1,26R_2^{-1})\}e^{-R_2 t} \quad (2. 2)$$

Gaussovské rozdělení:

$$g(t) = \{u(t) - u(t - 1,26R_2^{-1})\}e^{-\frac{R_2^2 t^2}{1,26}} \quad (2. 3)$$

Cosinové ve druhé mocnině:

$$g(t) = \{u(t) - u(t - 1,26R_2^{-1})\} \cos^2\left(\frac{\pi t}{2N\Delta t}\right) \quad (2. 4)$$

Poissonovské:

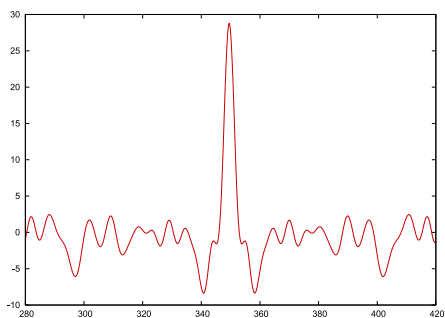
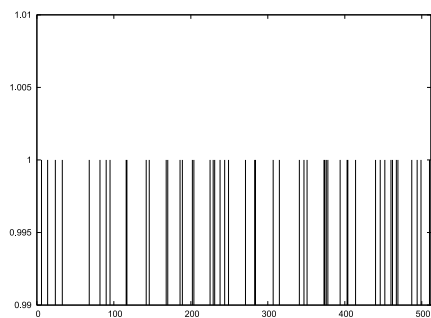
$$g(t) = \{u(t) - u(t - 1,26R_2^{-1})\}e^{-\lambda} \quad (2. 5)$$

kde $\lambda_1 = \sin\left(\frac{t\pi}{N\Delta t}\right) = \sin(\Phi_1(t))$ a $\lambda_2 = \sin\left(\frac{t\pi}{N\Delta t}\right) = \sin(\Phi_2(t))$

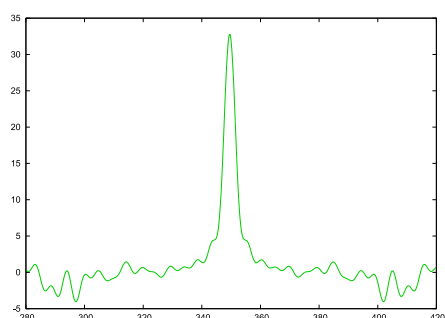
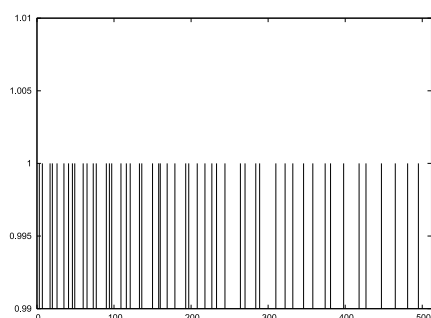
Rozložení bodů dle funkcí je dosaženo následujícím algoritmem. Časová doména signálu je před měřením rozdělena do n intervalů $\langle t_{i-1}; t_i \rangle$, kde n je počet měřených bodů a i značí i -tý interval. Krajní body těchto intervalů jsou vázány rekurzivním vztahem

$$\int_{t_{i-1}}^{t_i} g(t) dt = \frac{\int_0^{t_{max}} g(t) dt}{n} \quad (2. 6)$$

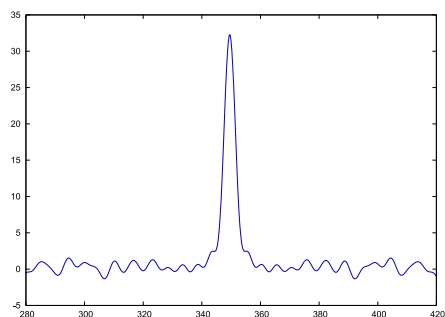
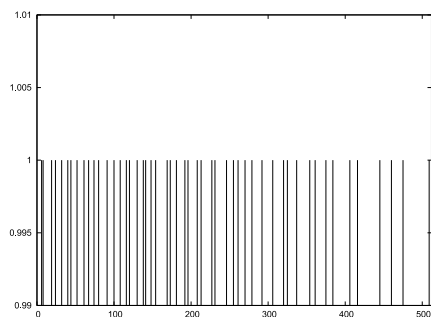
Do každého z těchto intervalů je náhodně generován jeden bod. Na grafech 1 až 6 jsou znázorněna konkrétní rozdělení bodů vytvořená na základě schémat (2. 1) až (2. 6) při 10 procentech hustoty.



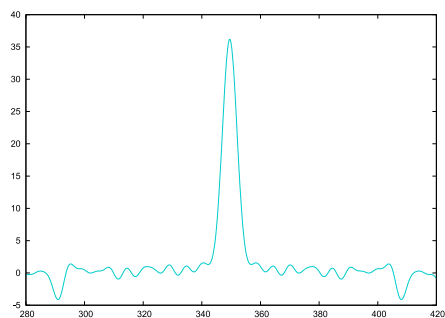
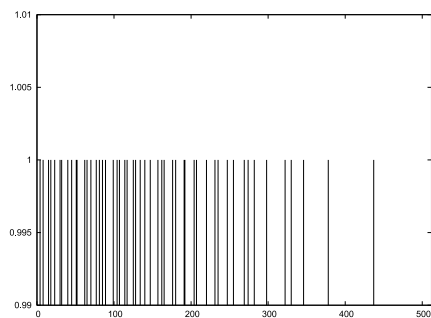
Graf 1: Uniformní rozdělení a jeho spektrum



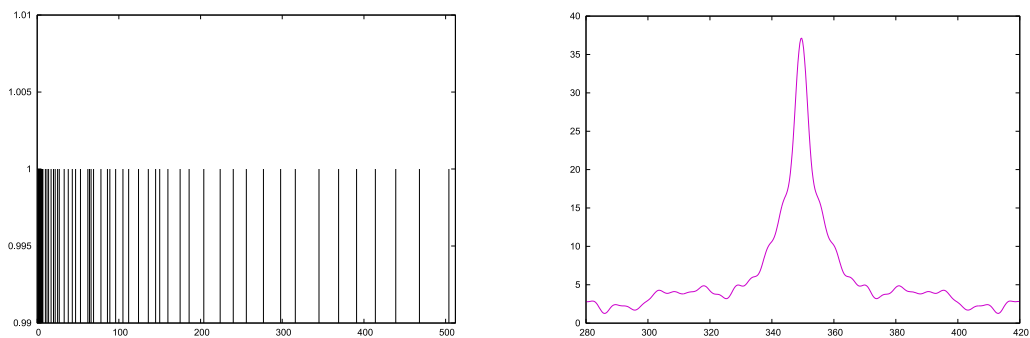
Graf 2: Exponenciální rozdělení a jeho spektrum



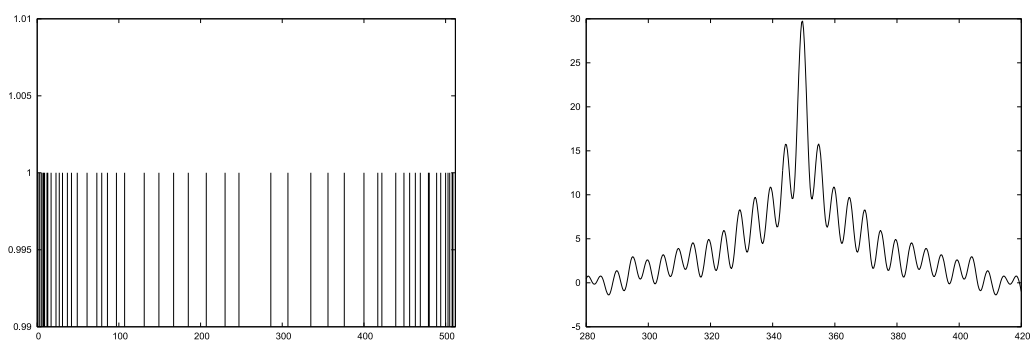
Graf 3: Gaussovské rozdělení a jeho spektrum



Graf 4: Cosinové rozdělení a jeho spektrum



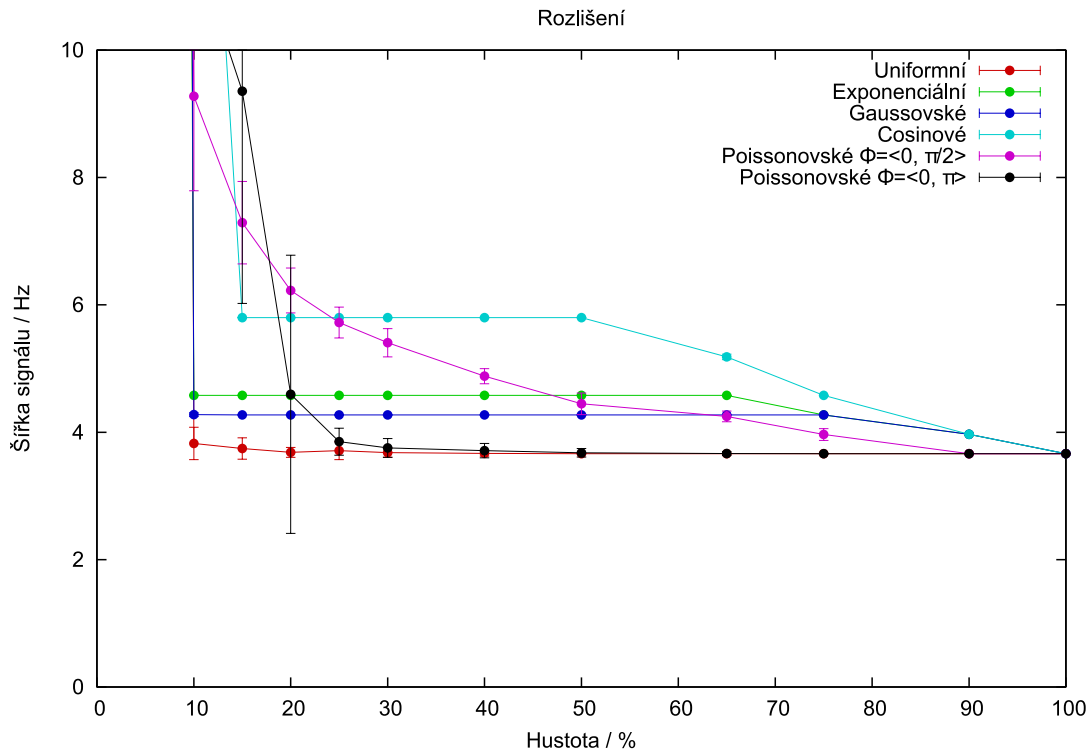
Graf 5: Poissonovské rozdělení pro $\lambda_1 = \sin\left(\frac{t\pi}{2}\right)$ a jeho spektrum



Graf 6: Poissonovské rozdělení pro $\lambda_2 = \sin\left(\frac{t\pi}{N\Delta t}\right)$ a jeho spektrum

2.2.1. Zachování rozlišení

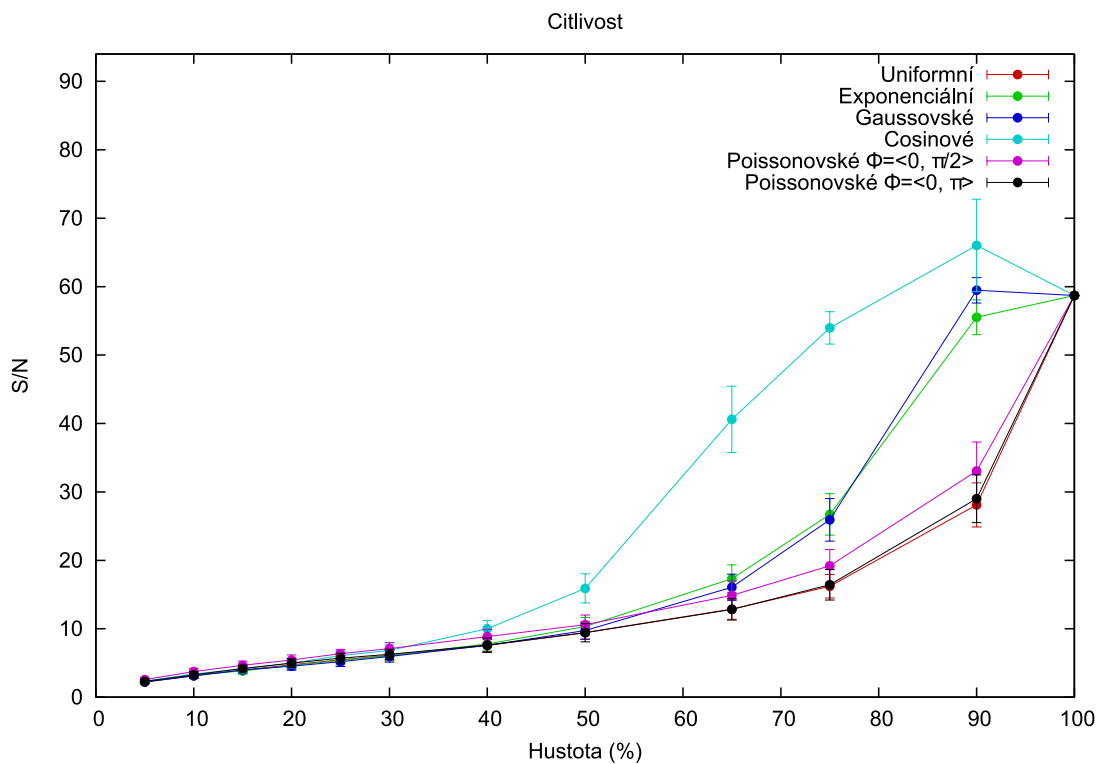
Každé schéma je charakterizováno množstvím bodů, z něhož je ještě možno rekonstruovat signál za přijatelného rozlišení spektra. Šírky signálů všech spekter pod 10 procent hustoty bodů prudce divergují. Při hustotách nad 20 procent zůstává rozlišení v přijatelných mezích u všech rozdělení (graf 7). Dále je pozorovatelné, že nejlepšího rozlišení je dosaženo použitím uniformní rozdělení či při hustotách nad 30 procent poissonovského rozdělení pro $\lambda_2 = \sin\left(\frac{t\pi}{N\Delta t}\right)$. Nejméně vhodnými se z hlediska rozlišení jeví cosinové a poissonovské rozdělení pro $\lambda_1 = \sin\left(\frac{t\pi}{2}\right)$, přičemž rozlišení cosinového rozdělení je stabilnější.



Graf 7: Šířka testovacího signálu v závislosti na hustotě a schématu rozdělení

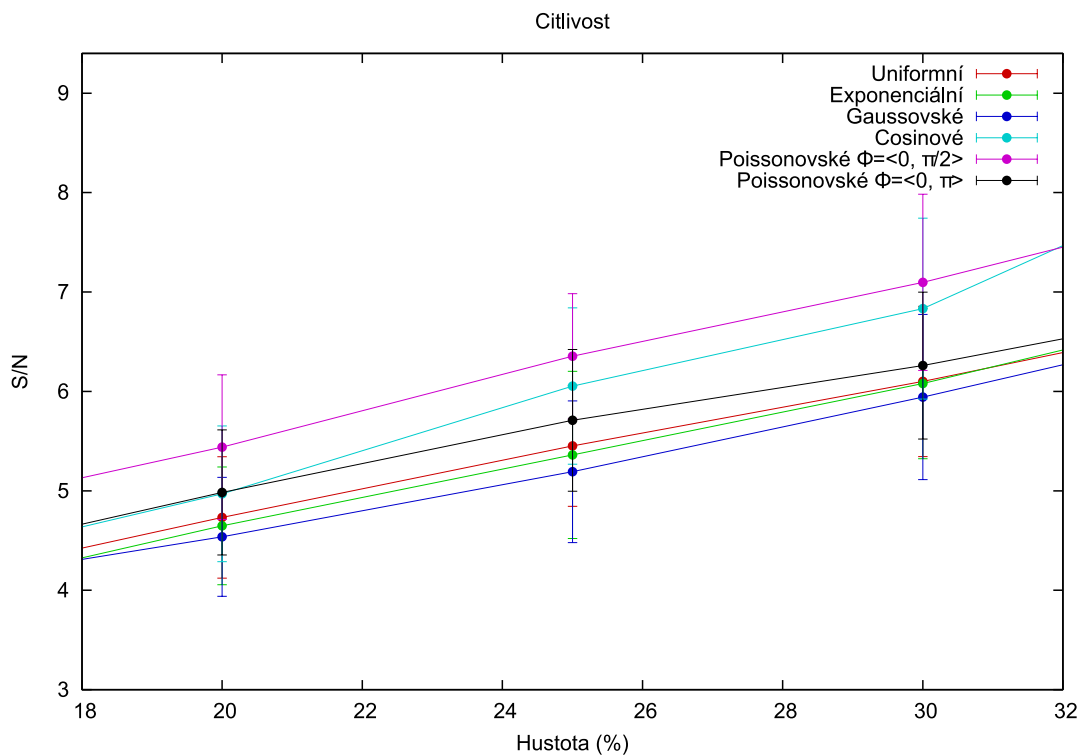
2.2.2. Porovnání citlivosti

Na grafu 8 je vykreslen poměr signálu ku šumu v závislosti na použitém schématu a hustotě bodů. Se zvyšující se densitou se zvyšuje rozdíl mezi jednotlivými rozděleními, zatímco od hustoty 90 procent konvergují všechna schémata k totožné citlivosti, protože při vysokých hustotách kvůli velkému množství bodů nelze sestavit rozdělení, které by odpovídalo funkcím (2. 2) až (2. 5).



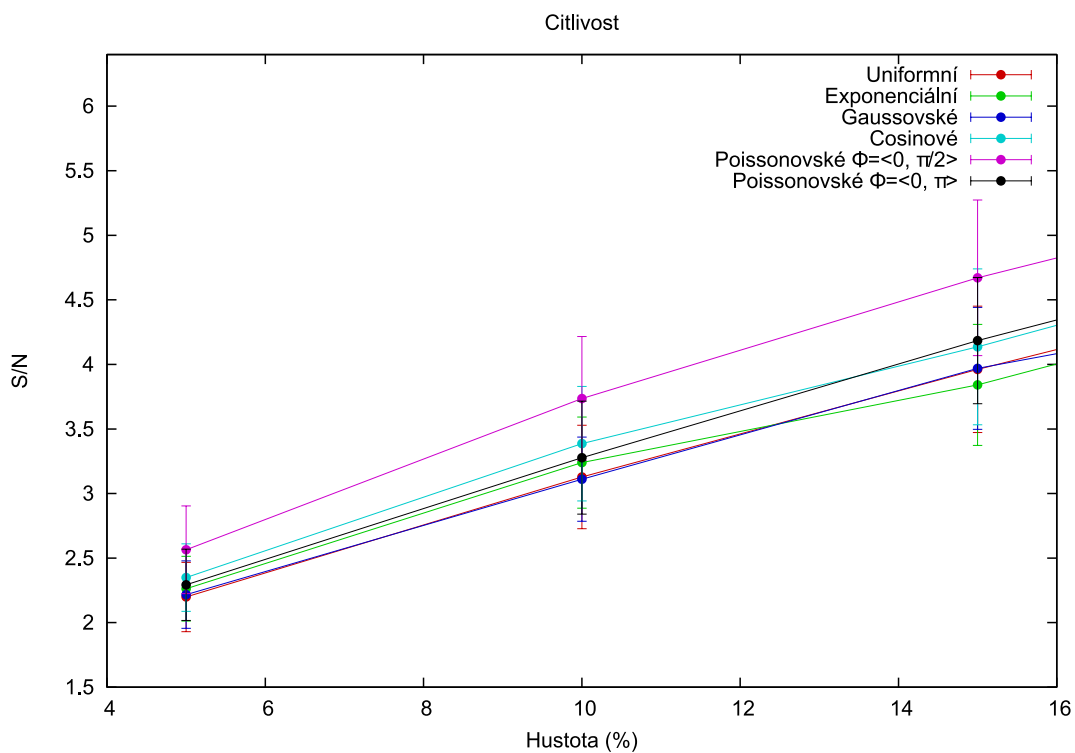
Graf 8: Poměr signálu ku šumu v závislosti na hustotě bodů

Při vysokých hustotách bodů je optimálním z hlediska citlivosti cosinové rozdělení, které klade velký důraz na počáteční body, proto v porovnání s ostatními schémata disponuje vysokou citlivostí v této oblasti hustot.



Graf 9: Citlivost spekter při středních hodnotách hustoty bodů

Pro střední hustoty bodů je nejcitlivějším schématem Poissonovské ($\pi/2$) rozdělení.

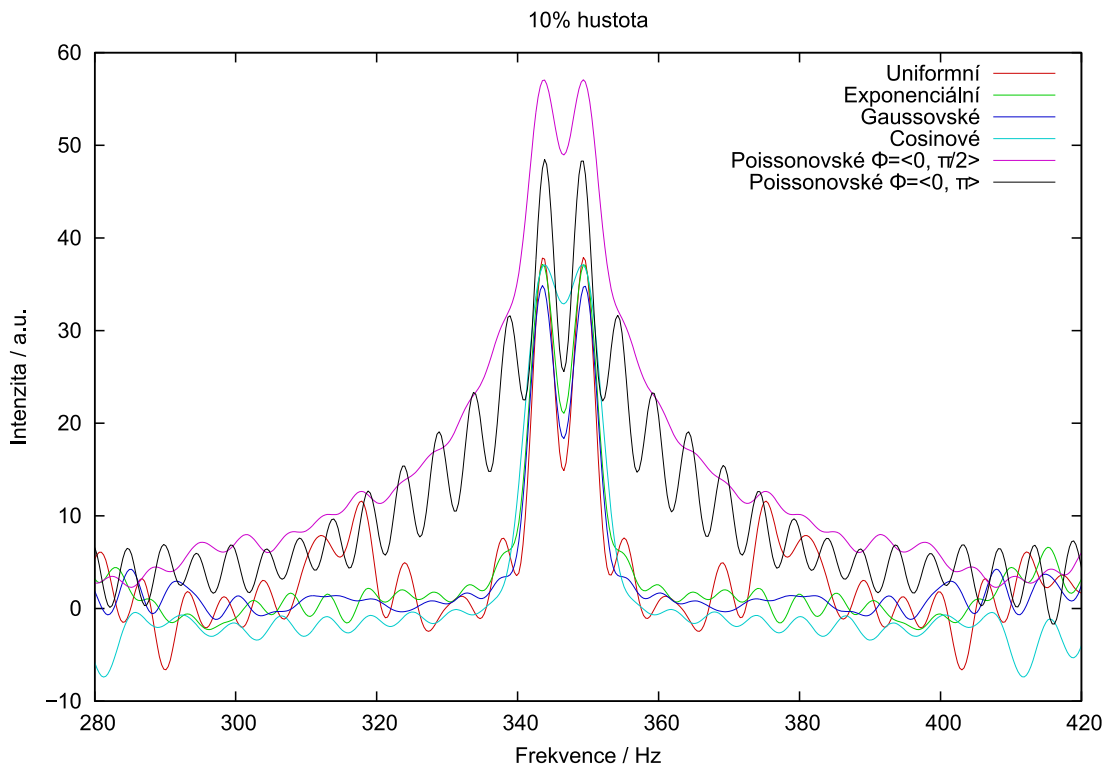


Graf 10: Nízké hustoty a jejich poměr signálu ku šumu v závislosti na densitě

Při nízkých hustotách se zdá být z hlediska citlivosti nejvhodnější Poissonovského ($\pi/2$), jak se ale ukazuje v sekci 2.2.5, je vysoká citlivost kompenzována velkou šířkou signálu.

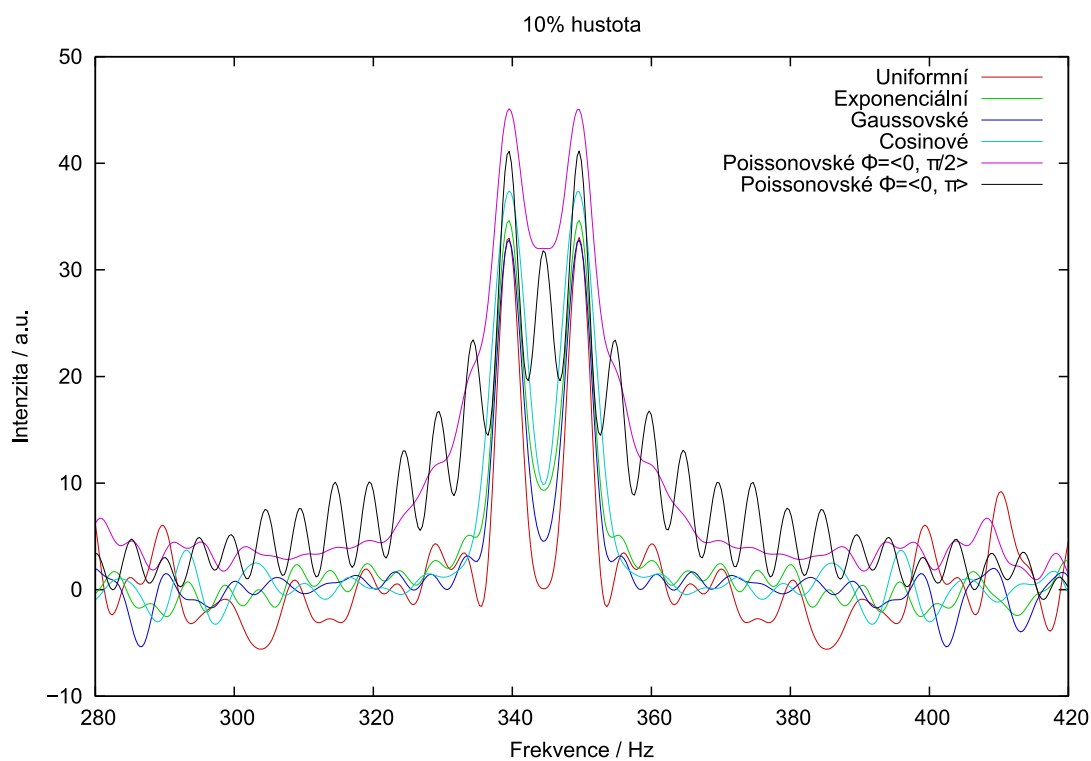
2.2.3. Dva signály

Při určité hustotě, která je charakteristická pro každé schéma (sekce 2.2.1), začne prudce klesat rozlišení. Tento jev lze pozorovat na grafu, kdy dva signály s 10% hustotou vzorkování a rozdílem frekvencí 6 Hz splývají v závislosti na schématu.



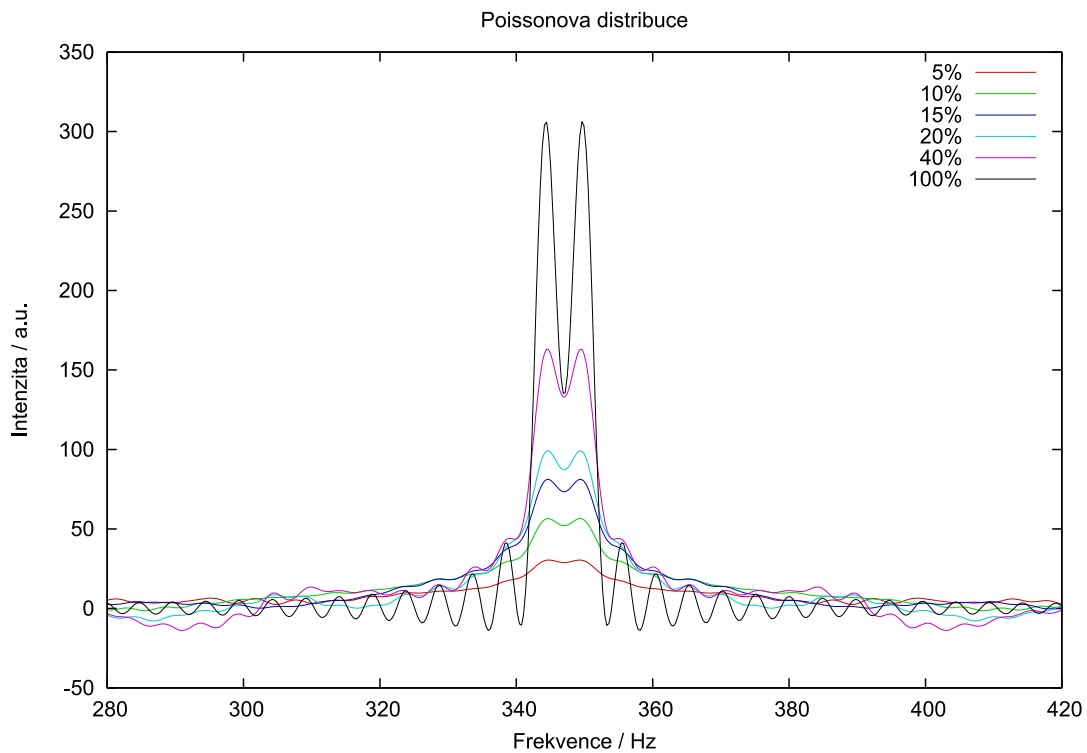
Graf 11: Simulace dvou blízko položených signálů za použití jednotlivých rozdělí

Graf 11 ukazuje, že signály poissonovského ($\pi/2$) a cosinového rozdělí hustoty experimentálně měřených dat nejsou rozeznatelné s dostatečnou přesností, zatímco nejlepší rozlišení signálů poskytují uniformní a Gaussovské. Tento výsledek je v souladu se statistickými výsledky získanými v sekci 2.2.1. U některých schémat s vysokou citlivostí se objevují falešné signály, jako je tomu například na grafu 12. Poissonovské (π) rozdělí zde má zdánlivě vysoké rozlišení, nicméně vysoký obsah šumu způsobil vznik falešného signálu mezi dvěma zkoumanými signály.



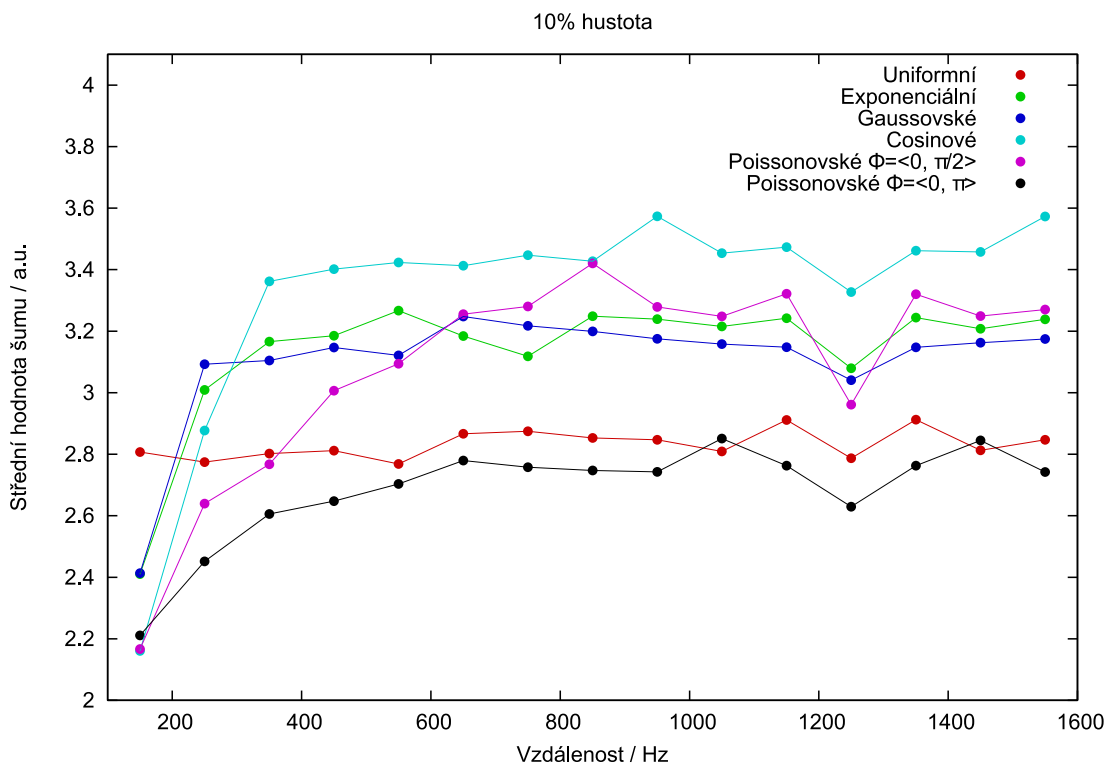
Graf 12: Dva signály se vzdáleností 10 Hz. Uprostřed spektra Poissonovského (π) rozdělení se nachází falešný signál.

Rozlišení je určeno hustotou bodů. Při nízkých hustotách se tak snižuje rozeznatelnost více signálů, jak lze vidět například v případě Poissonovského ($\pi/2$) rozložení zobrazeného na grafu 12, jehož dva signály se liší o 5 Hz.



Graf 13: Spektrum dvou signálů při použití Poissonovského rozdělení za různých hustot bodů

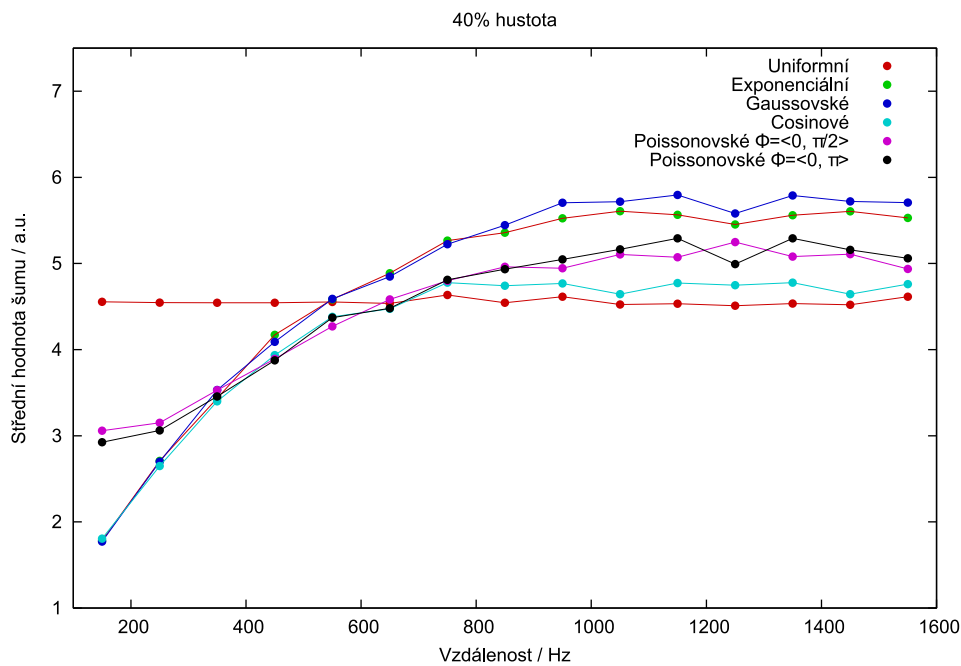
2.2.4. Rozložení šumu



Graf 14: Distribuce šumu ve frekvenční doméně v závislosti na vzdálenosti od signálu

Rozložení šumu za použití schémata distribuce ukazuje, že schémata disponují vlastností posunu artefaktů od signálu, a tak vzniká v jeho blízkosti oblast s nízkým obsahem šumu. Graf 14 ukazuje průměrné střední kvadratické hodnoty šumu v pásmech o šířce 100 Hz v závislosti na vzdálenosti od signálu. Dle očekávání je šum při uniformním rozložení experimentálně měřených hodnot rovnoměrně rozprostřen po celém spektru, u ostatních schémat pozorujeme různé hladiny šumu s různou mírou posunutí šumu. K nejrazantnějšímu posunu šumu dochází při použití cosinového rozdělení, naproti tomu nejmírněji stoupá šum u poissonovského schématu pro $\lambda_2 = \sin\left(\frac{t \pi}{N\Delta t}\right)$, které se v tomto ohledu blíží uniformnímu rozdělení.

S vyšší hustotou bodů (graf 15) se posunutí šumu propaguje méně razantně, protože se méně projevuje odlišnost od uniformního rozložení.



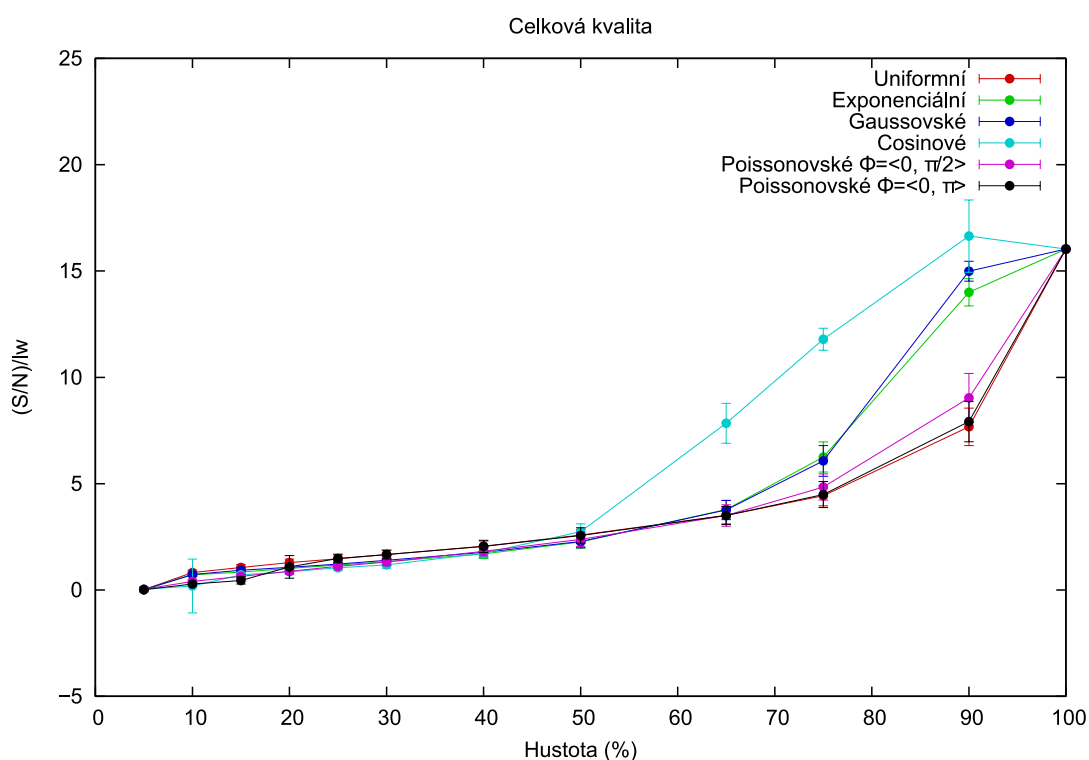
Graf 15: Rozložení šumu při vyšší hustotě bodů

Díky posunu šumu je dosaženo vysoké citlivosti v blízkosti signálu, nicméně pokud se ve spektru nachází více signálů, může posunutý šum znemožnit jejich detekci. Toto negativum se v případě simulovaného signálu projevuje v závislosti na rozdělení bodů. Například v případě spekter znázorněných na grafu 14 by rozdíl frekvencí dvou signálů neměl být vyšší než 200 Hz u neuniformních rozdělení.

2.2.5. Ukazatel kvality

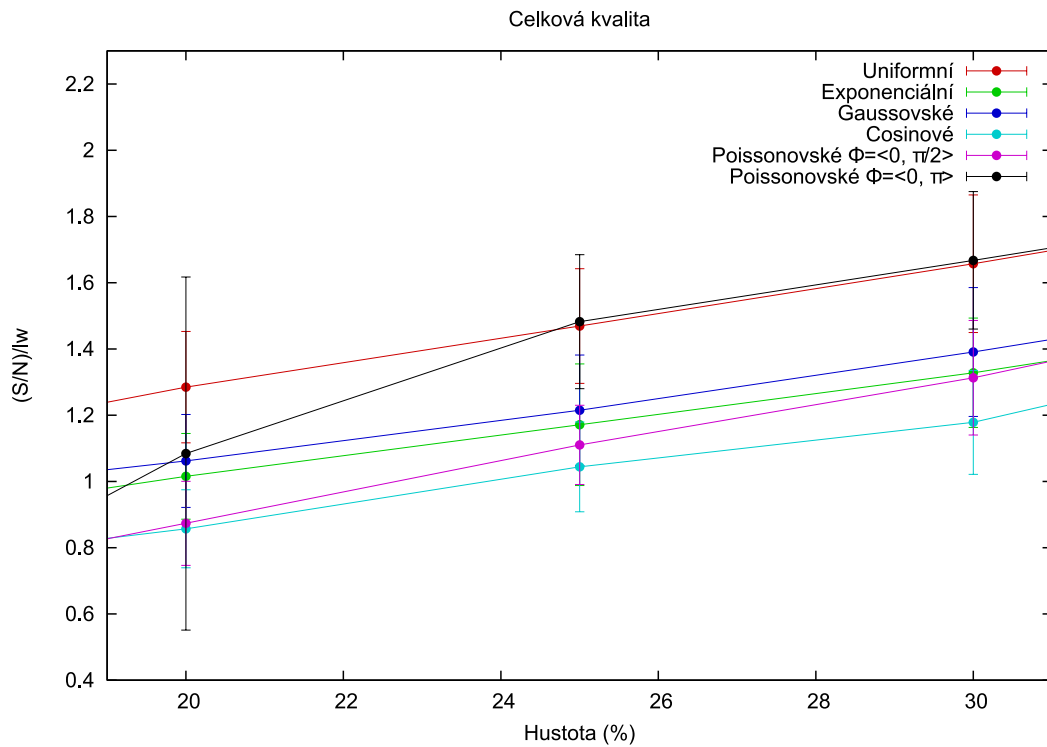
V sekcích 2.2.1 a 2.2.2 bylo znázorněno, že charakteristikami kvality spektra je citlivost i rozlišení. Z tohoto důvodu byl v této práci jako míra kvality spektra zvolen poměr citlivosti a rozlišení, jak byly definovány v sekci 1.1.4. Potom je kvalita spektra dána vztahem:

$$q = \frac{S/N}{L} \quad (2.7)$$



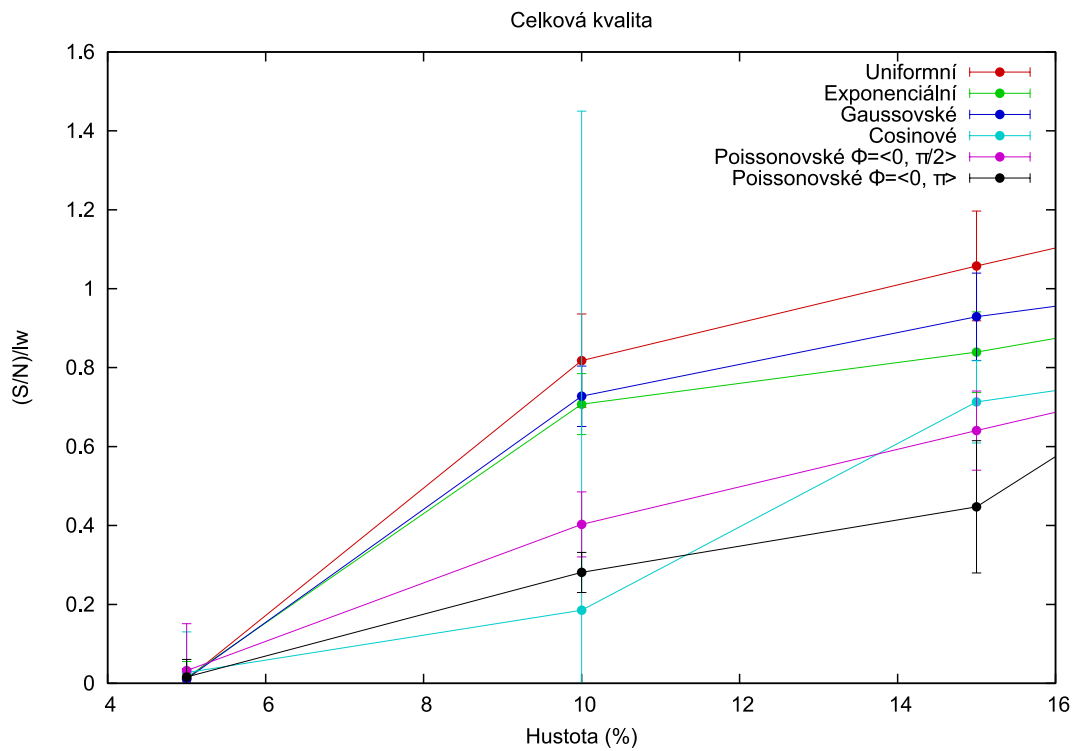
Graf 16: Kvalita spektra v závislosti na hustotě a rozdělení

Kvalita spektra je při vysokých hustotách dána převážně citlivostí, která doprovází jednotlivá schémata, neboť, jak plyne ze sekce 2.2.1, je rozlišení v této oblasti stabilní pro všechna schémata. Z grafu 16 proto plyne výhodnost cosinového rozdělení při vysokých hustotách bodů. Příčinou je vysoká citlivost, kterou toto schéma disponuje, jak je zmíněno v sekci 2.2.2.



Graf 17: Kvalita spekter pro střední hustoty bodů

Pro střední hustoty (20-30%) jsou uniformní a poissonovské (π) optimálními schémata.



Graf 18: Kvalita spekter při malých hustotách bodů

Při hustotách nižších než 10 procent konverguje kvalita spekter prudce k nule (graf 18) kvůli nízkému rozlišení způsobeném nedostatkem bodů, jak lze pozorovat na grafu 7. Přijatelnými schémata v této oblasti hustot mohou být uniformní, gaussovské či exponenciální rozdělení.

Závěr

Tato práce dokázala, že je možné rekonstruovat spektrum ze zlomku experimentálně měřených dat a zachovat přitom rozlišení srovnatelné s celistvými daty. Pro spektra rekonstruovaná z malé hustoty bodů se jako podstatné pro posouzení kvality projevilo rozlišení a bylo statisticky zachyceno.

Získané výsledky mohou sloužit při výběru vhodného schématu. Volba schématu je obecně závislá na důrazu kladeném buď na rozlišení, nebo citlivost. Pokud je upřednostňována citlivost, je na místě použití cosinového rozdělení, naopak pro rozlišení je nejvhodnější schéma s uniformní hustotou měřených dat. Při analýze rozložení šumu se projevila důležitá vlastnost schémat s neuniformní hustotou generovaných bodů, a to vznik pásma s nízkým obsahem šumu v blízkém okolí signálu. To může ovlivnit volbu schématu při aplikaci, neboť v některých případech komplikuje detekci více signálů.

Vzhledem ke kvalitě, jak je definována v této práci, se jako nejvhodnější schéma jeví cosinové při vysokých hustotách, poissonovské pro $\lambda_2 = \sin\left(\frac{t\pi}{N\Delta t}\right)$ při hustotách 25-50 procent, nakonec pro hustoty pod 25% shledáváme jako optimální rozdělení s uniformní hustotou měřených bodů.

Literatura

- [1] LEVITT, Malcolm H. *Spin dynamics: basics of nuclear magnetic resonance*. 2nd ed. Hoboken, NJ: John Wiley, c2008, xxv, 714 p., [7] p. of plates. ISBN 04-705-1117-6.
- [2] HOCH, Jeffrey C a Alan S STERN. *NMR data processing*. New York: Wiley-Liss, c1996, xi, 196 p. ISBN 04-710-3900-4.
- [3] HYBERTS, Sven G., Scott A. ROBSON a Gerhard WAGNER. Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra. *Journal of Biomolecular NMR*. 2013, vol. 55, issue 2. DOI: 10.1007/s10858-012-9698-2. Dostupné z: <http://link.springer.com/10.1007/s10858-012-9698-2>
- [4] KAZIMIERCZUK, Krzysztof, Anna ZAWADZKA a Wiktor KOŹMIŃSKI. *Journal of Magnetic Resonance*. 2008, vol. 192, issue 1, s. 123-130. DOI: 10.1016/j.jmr.2008.02.003. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S1090780708000530>
- [5] KAZIMIERCZUK, Krzysztof, Anna ZAWADZKA a Wiktor KOŹMIŃSKI. Narrow peaks and high dimensionalities: Exploiting the advantages of random sampling. *Journal of Magnetic Resonance*. 2009, vol. 197, issue 2, s. 219-228. DOI: 10.1016/j.jmr.2009.01.003. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S109078070900007X>
- [6] ROVNYAK, D., J.C. HOCH, A.S. STERN a G. WAGNER. Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy. *Journal of Biomolecular NMR*. 2004, vol. 30, issue 1, s. 1-10. DOI: 10.1023/B:JNMR.0000042946.04002.19. Dostupné z: <http://link.springer.com/10.1023/B:JNMR.0000042946.04002.19>

Seznam symbolů, veličin a zkratk

*	konvoluce
δ	chemický posun, Diracova delta funkce
F	Fourierova transformace
NMR	nukleární magnetická rezonance

Seznam příloh

Příloha 1. Zdrojové kódy