Atelier BEST/NUS
mercredi 14 octobre 2015

B. Brutscher / A. Favier / E. Lescop / N. Morellet
Programme de la journée

9h30-10h30: cours introduction
10h30-12h30: TP (groupe 1) / TD (groupe 2)

12h30-14h00: déjeuner

14h00-15h00: TP (groupe 1) / TD (groupe 2)

15h00-15h15: pause

15h15-18h15: TP (groupe 2) / TD (groupe 1)

18h15-19h00: Discussion générale
Multidimensional NMR

Real world time scale

Recycling delay

~ 0.1 s  1 s
Data acquisition

Digital acquisition (discrete numbers)
Signal collected for given regularly spaced increments DW (associated to spectral width SW)

\[ AQ = DW \times TD \]
\[ DW = \frac{1}{2 \times SW} \]
\[ AQ = \frac{DW}{2 \times SW} \]
Resolution and acquisition time (aq)

Resolution increases when AQ is longer \( AQ = \frac{TD}{2SW} \)
(TD is not the only relevant parameter)

If AQ is too long, we collect noise and no signal:
maximal resolution is achieved but we decrease the SNR
Rule of thumb: \( AQ \sim 1.2 \ T_2 \)
Optimal recycling delay

\[
\left( \frac{S}{N} \right)_t = \frac{1 - e^{-\frac{t_{\text{rec}}}{T_1}}}{\sqrt(T_{\text{scan}})}
\]

Optimal at 1.25 \( T_1 \)
The sampling problem

Sampling indirect dimension takes time to get resolution
The sampling problem

<table>
<thead>
<tr>
<th>1D</th>
<th>2D</th>
<th>3D</th>
<th>4D</th>
<th>5D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s-1mn</td>
<td>10mn-1h</td>
<td>5h-72h</td>
<td>week-month</td>
<td>year?</td>
</tr>
</tbody>
</table>

Duration of nD spectrum follows an exponential growth

Problems:
- Spectrometer time usage
- Sample stability
- Fast processes not accessible
How to collect nD spectra faster?

- Reduce interscan delay (fast pulsing)
  - wait...
  - approx. 0.1 s
  - 1 - 2 s

- Reduce number of repetitions of pulse sequence
  - Hadamard spectroscopy (frequency domain)
  - Spatial frequency encoding
  - Reduced dimensionality / projection
  - Spectral compression (extensive aliasing, ASCOM)
  - Non-linear sampling methods (time domain)
How to collect nD spectra faster?

- **Reduce interscan delay (fast pulsing)**
  - BEST/ SOFAST-HMQC

- **Reduce number of repetitions of pulse sequence**
  - Hadamard spectroscopy (frequency domain)
  - Spatial frequency encoding
  - Reduced dimensionality / projection
  - Spectral compression (extensive aliasing, ASCOM)
  - Non-linear sampling methods (NUS)

![Diagram showing different time delays and signal acquisition](attachment:Diagram.png)
ASCOM: Automated Spectral Compression

A simple tool to reduce the number of repetitions

\[
\text{Spectral resolution} \quad \alpha \quad \frac{\text{Number of repetitions}}{\text{Spectral width}}
\]

Reduced spectral width \(\Rightarrow\) Reduced number of repetitions

* Lescop, Schanda, Rasia, Brutscher, JACS, 2007
Aliasing in the case of complex data acquisition

(States, States-TPPI or Echo-Antiecho mode)

The peaks fold on the other side of the spectrum
This is a translation by 30 ppm
The ASCOM software does the calculation within seconds ([http://www.icsn.cnrs-gif.fr/download/nmr](http://www.icsn.cnrs-gif.fr/download/nmr)) Integrated in Topspin (see example in a few minutes)
Advantages:
- Same content of information (resolution, dispersion)
- But faster (25/5.7 = 4.4 times)
Advantages:
- Same content of information (resolution, dispersion)
- But faster (25/5.7 = 4.4 times)
ASCOM application

Use the ASCOM-optimized SW($^{15}$N) to all other $^1$H-$^{15}$N planes:
-3D/4D triple resonance experiments
-$^{15}$N relaxation experiments
-RDC measurements
-Kinetic experiment (H/D exchange, …)

Compatible with any pulsesequence (BEST, …) but also with other sampling modes (cf NUS)

But: (1) any change in relative $^1$H-$^{15}$N peak position (interaction, temperature, pH, …) requires a new optimization!

(2) Mostly useful for spectra with good signal-to-noise ratio initially because:
-the signal-to-noise also reduces (we accumulate less)
-useful only if NS=2 (otherwise keep NS=2 and use increased SW)
Approach limited to small-medium sized proteins

Ubiquitin (76 a.a.)

Hyl-1 (103 a.a.)

Wzb (147 a.a.)
Application to 3D triple resonance experiments

Number of resolved peaks

SW(13C)

SW(15N)

4th dimension

13CO
1H
15N

13CO
1H
15N

13CO
1H
15N
Application to 3D triple resonance experiments

Combined optimization of:

SW(N)/SW(C)

Unstructured

\(\gamma\)-synuclein

Hyl-1
(103 aa)

MSG
(~700 aa)

Black area:
Combination of \(^{15}\text{N}/^{13}\text{C}\) spectral widths for which no additional overlapping occurs compared with the reference spectrum

\(\text{SW}^{(15}\text{N})=6.6\text{ppm} / \text{SW}^{(13}\text{C})=2.6\text{ppm}\)

Compared with \(\text{SW}^{(15}\text{N})=30\text{ppm} / \text{SW}^{(13}\text{C})=15\text{ppm}\), the 3D experiment can be recorded \textbf{30 times faster}\n
Application to Hyl-1

Folded part (~90 aa)

15N HSQC

HN-CO plane
Application to Hyl-1

Significant help in backbone resonance assignment
Non Uniform Sampling (NUS): 2D

Instead of collecting all $N$ points from $t_1=0$ to $t_{1\text{max}}$,
We collect a subset $N_{\text{keep}} < N$

NUS percentage defined as $N_{\text{keep}} / N$
Non-linear Uniform Sampling (NUS): 3D
Non-linear Uniform Sampling (NUS) : 3D

Only a fraction of ($t_1$/$t_2$) pairs are collected
Non-linear Uniform Sampling (NUS) : 3D

Only a fraction of \((t_1/t_2)\) pairs are collected
Non-linear Uniform Sampling (NUS)

Fourier Transform (FT) requires a complete matrix.
How to deal with incomplete sampling?

Fourier Transform (FT) requires a complete matrix.

Solution (1): replace missing points by 0.

Satisfies prerequisite but provides artefacts because the original signal is strongly perturbed.
How to deal with incomplete sampling?

Simulation: superposition of two peaks / **No noise**

Blue: A fraction of experimental data are set to 0.
How to deal with incomplete sampling?

Blue: A fraction of experimental data are set to 0.
How to deal with incomplete sampling?

Blue: A fraction of experimental data are set to 0.
How to deal with incomplete sampling?

Filling with zeroes introduce severe artefacts / structure noise.
How to deal with incomplete sampling?

Fourier Transform (FT) requires a complete matrix

Solution (2): predict the missing points.
Several methods (MaxEnt, MDD, CS, …) developed in the groups of Koźmiński W, Billeter M, Marion D, Hoch J, Orekhov V….

Reviews:


Billeter M, Orekhov V, Novel Sampling Approaches in Higher Dimensional NMR (book)
Practical implementation: qMDD / Topspin

For real usage, any method should be available as a software, easy-to-use, robust, with minimal parameters to set, and with rapid computation.

qMDD (free): https://groups.google.com/forum/#!forum/mddnmr
(1) FT (zeroed-spectra)
(2) MDD
(3) Compressed Sensing

Topspin (Licence)
Included with xfb/tf3d
(available at ICSN for IR-RMN users)
Multiway Decomposition Approach (MDD)

Principle:
Any nD spectra can be considered as the sum of various nD peaks. Each peak $k$ is defined as the direct product (‘⊗’) of 1D shapes $F_{1}^{k}, \ldots, F_{n}^{k}$.

$$S(\omega_{1} \ldots \omega_{N}) = \sum_{k} F_{1}^{k}(\omega_{1}) \otimes F_{2}^{k}(\omega_{2}) \ldots \otimes F_{N}^{k}(\omega_{N})$$

MDD consists in finding the shapes $F_{1}^{k}, \ldots, F_{n}^{k}$, that satisfy the experimental spectrum. The fit can be done in the time domain where missing data are simply removed from the analysis. A critical parameter (NCOMP) is the number of component (peaks) expected. Too low, weak peaks will not be modelled.

The shapes can then be used to reconstruct missing data prior to regular FT.
Compressed Sensing

Principle: Providing the frequency spectrum $S$ is sparse (i.e. zero values at most locations), $S$ can be reconstructed a few experimental time data $b$, by solving

$$\min_S ||S||^1 \text{ subject to } FT^{-1}(S) = b$$

With $$||S||_1 = \sum_i |S_i|^1$$

Many NMR spectra are sparse in the frequency domain, with sparsity increasing as the number of dimensions increases. The 0-norm should be used in theory but is practically not convenient. The 1-norm is easier.
Compressed Sensing

Many NMR spectra are sparse in the frequency domain, with sparsity increasing as the number of dimensions increases. The 0-norm should be used in theory but is practically not convenient.

Different algorithms to minimize $||S||_1$
- Iterative re-weighted least squares (IRLS)
- Iterative soft thresholding (IST)

Iterative Soft Threshold (hmsIST) outlined

NUS synthetic two line spectrum

FFT produces spectrum with many artifacts (point spread function, PSF)

Copy top 2% of spectrum and store in different location

FFT\(^{-1}\) and zero skipped time domain data points - subtract

FFT and iterate

After a sufficient number of iteration all PSF artifacts are eliminated

Very fast procedure

Hyberts et al. J. Biomol. NMR 52(4) 2012
Compressed Sensing: robust to high dynamic range (NOESY)

Accuracy of peak intensities in the reconstructed 2D NOESY spectrum of azurin.

Weak peaks may not be recognized and disappear below noise level but otherwise, peak intensity seems to be reliable.

Fidelity of the lineshape reconstruction. The I13 HN-Ha peak from the IRLS-reconstructed DQF COSY spectrum of ubiquitin is shown.
Sampling schedule: uniform distribution

Exponentially damped sinusoid

Good SNR
Poor resolution

Poor SNR
Good resolution

Sampling uniformly or non-uniformly impacts peak intensity and resolution (similar effect as window function)
In absence of knowledge about signal -> uniform sampling
Perfect for constant-time evolution ($^{15}$N dimension of HNC exp)
Sampling schedule: matched acquisition (exponential distribution)

Exponentially damped sinusoid

- Good SNR
- Poor resolution
- Poor SNR
- Good resolution

Collect more increments at the beginning and less at the end of the fid (or pseudo-fid)
Good for real time evolution ($^{15}$N/$^{13}$C dim of HSQC or $^{13}$C dim of HNC)
Sampling schedule: exponential/sin distribution

Exponentially damped sinusoid + J coupling

Good SNR
Poor resolution

Poor SNR
Good resolution

Concentrate more on regions with good SNR and less near 0.

Good for (1) $^{13}$C dimension of HNCA/HNCOCA/HNCACB/HNOCOCAB (not for the CT version) where $J_{C\alpha C\beta} = 35$ Hz is active when high resolution is needed or (2) J-coupled HSQC (IPAP, …)
Sampling schedule:

Pure Random

Exponential weight
Sampling schedule: TOPSPIN

NUS (Non Uniform Sampling) parameters

- **NUS Help**: Show NUS help
- **NusAMOUNT [%]**: 100
- **NusPOINTS**: 1500
- **NusJSP [Hz]**: 0 0 0
- **NusT2 [sec]**: 1 1
- **NusSEED**: 54321
- **NUSLIST**: automatic
- **Calculate**: Calculate point spread function
- **Show**: Display NUS point spread

NUS (Non Uniform Sampling) parameters

- **Mdd_mod**: cs
- **MddCEXP**: FALSE
- **MddCT_SP**: FALSE
- **MddF180**: FALSE
- **MddNCOMP**: 0
- **MddPHASE**: 0 0
- **MddSRSIZE [ppm]**: 0
- **MDD mode**: MDD mode
- **RMDD/MDD flag**: RMDD/MDD flag
- **Constant time**: Constant time
- **Delayed sampling flag**: Delayed sampling flag
- **Number of components**: Number of components
- **Phase**: Phase
- **Sub region size**: Sub region size
Sampling schedule: TOPSPIN schedule

2D

3D
Speeding up multidimensional NMR

Standard NMR techniques

1D → 2D → 3D → 4D → 5D
seconds → minutes → hours → weeks → years

Fast-pulsing NMR

Sparse non-uniform data sampling (NUS)
Longitudinal relaxation, inter-scan delay & sensitivity

Longitudinal relaxation

Scan time [sec]

S/N  \sim \sqrt{\text{number of scans}}
\sim \frac{1}{\sqrt{\text{scan time}}}

Sensitivity

Scan time / seconds

z-magnetization

Longitudinal relaxation

Scan time [sec]
\[ \begin{align*}
1^H \text{ longitudinal relaxation mechanisms} \\
\text{Solomon or Bloch-McConnell equations:} \\
- \frac{d}{dt} \begin{pmatrix} W_{1z} \\ H_{1z} \\ \vdots \\ H_{nz} \end{pmatrix} &= \begin{pmatrix} \rho_W & 0 & 0 & 0 & 0 \\ 0 & \sigma_{12} & \cdots & \sigma_{1n} & \sigma_{11} \\ 0 & \sigma_{21} & \sum_j \rho_{2j} & \cdots & \sigma_{2n} \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & \sigma_{n1} & \sigma_{n2} & \cdots & \sum_j \rho_{nj} \end{pmatrix} \begin{pmatrix} W_{1z} - W_{1z}^0 \\ H_{1z} - H_{1z}^0 \\ \vdots \\ H_{nz} - H_{nz}^0 \end{pmatrix} + \\
& \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -k_{ex,1} & k_{ex,1} & 0 & 0 & 0 \\ -k_{ex,2} & 0 & k_{ex,2} & 0 & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ -k_{ex,n} & 0 & 0 & 0 & k_{ex,n} \end{pmatrix} \begin{pmatrix} W_{1z} \\ H_{1z} \\ \vdots \\ H_{nz} \end{pmatrix} \\
\end{align*} \]

- $^1H-^1H$ dipolar interactions ($\sigma_{ij}$)
- Chemical exchange with water $^1H$ ($k_{ex}$)

$^1H$ polarization recovery depends on the spin state of all other protons in the protein and the bulk water.
Amide $^1$H longitudinal relaxation enhancement (LRE)

Non-selective excitation

- **Aliphatic $^1$H saturated ($\bullet$)**
  - $M_z = 0$
  - $M_z = 0$
  - $M_z = 0$

$T_1 \approx 0.8 - 4.0\text{ s}$

Selective $H^N$ excitation

- **Aliphatic $^1$H in equilibrium ($\blacksquare$)**
  - $M_z = M_{eq}$
  - $M_z = 0$
  - $M_z = M_{eq}$

$T_1 \approx 0.1 - 0.4\text{ s}$
Amide $^1$H longitudinal relaxation enhancement (LRE)

**Structured part**: $T_1 \approx 0.9$ s; $T_1 \approx 200$ ms

**Disordered part**: $T_1 \approx 2$ s; $T_1 \approx 60$ ms
Computed sensitivity curves for different $T_1$ time constants

Gain in Sensitivity

... and speed
BEST & SOFAST experiments

**high-speed and optimal-sensitivity regime**

- Sensitivity gain of a factor of 2-10
- ~ 30-200% more sensitive

**Sensitivity vs Recovery time [sec]**

- BEST
- Standard

Sensitivity [a.u.] vs Recovery time [sec]

0 1 2 3 4

0 1 2 3 4
BEST: Band-Selective Excitation Short-Transient

Band-selective pulses
BIP pulse pairs (360° rotation)

INEPT

BEST-HSQC

BEST-TROSY

E: Excitation pulse (EBURP2)
FB: Flip-back pulse (time-reversed EBURP2)
R: Refocusing pulse (REBURP)
Additional tricks for enhancing sensitivity

1. **1H polarization enhancement**
   - Using Ernst-angle excitation (**SOFAST**)

2. **Heteronuclear polarization enhancement**
   - **BEST-TROSY**

**1H pathway**

**15N pathway**

contributes to **15N pathway** in next scan
The SOFAST HMQC experiment

**SOFAST:** Band-Selective Optimized-Flip-Angle Short-Transient

Combining the advantages of Longitudinal-relaxation enhancement and Ernst-angle excitation!
The SOFAST HMQC experiment

**standard SOFAST**

- highest sensitivity

**FTA-SOFAST**

- for use with a fast mixing device
- small flip angle

**ST-SOFAST**

- no $^{15}$N decoupling, $\sqrt{2}$ sensitivity loss
- gradient coherence selection

**J-SOFAST**

- J-coupling-based selection
- natural abundance samples
SOFAST HMQC - 2D spectra in a few seconds

Ubiquitin (8.6 kDa, 0.2 mM)

TET-2 (486 kDa, 80 \( \mu \text{M} \), ~ 1 mM protomer concentration)

tRNA\textsuperscript{Val} (26 kDa kDa, 0.9 mM)
BEST-TROSY – Enhancing $^{15}$N steady-state polarization

**1H pathway**

$$H_z \xrightarrow{\text{INEPT}} \pm 2H_z N_x = \pm (H_\alpha N_x - H_\beta N_x) \xrightarrow{\text{ST2-PT}} N_\beta H_x$$

**$^{15}$N pathway**

$$N_z \xrightarrow{\text{INEPT}} N_x = (H_\alpha N_x + H_\beta N_x) \xrightarrow{\text{ST2PT}} N_\beta H_x$$

**3rd pathway**

$$\ldots \xrightarrow{\text{relaxation } (\Delta)} H_z \xrightarrow{\text{ST2-PT}} -N_z \xrightarrow{180N} +N_z$$

- Benefits from **high B$_0$ fields** (increased CSA-DD CCR & long $^{15}$N $T_1$)
- **$^{15}$N polarization increases for short recycle delays** (BEST) and long relaxation delays $\Delta$

BEST-HSQC and BEST-TROSY Experiments

**BEST $^1$H-$^{15}$N HSQC**

- BEST $^1$H-$^{15}$N HSQC experiment at 600 MHz, 25°C.
- Data for 8.6 kDa, 12 kDa, and 21 kDa proteins shown.
- Intensity (arbitrary unit) vs. $T_{\text{scan}}$ (sec).

**BEST $^1$H-$^{15}$N TROSY**

- BEST $^1$H-$^{15}$N TROSY experiment at 600 MHz, 25°C.
- Data for 8.6 kDa, 12 kDa, and 21 kDa proteins shown.
- Intensity (arbitrary unit) vs. $T_{\text{scan}}$ (sec).
BEST-HSQC and BEST-TROSY Experiments

**i-HNCA**

(Δ ≈ 100ms)

**HNCO**

(Δ ≈ 60ms)

**BEST-HSQC**

**BEST-TROSY**

Average gain: ~ 110%

Average gain: ~ 70%

Correlation peak
BEST-HSQC and BEST-TROSY Experiments

Application to a 138-residue (16 kDa) protein

**BEST-TROSY** versus **BEST-HSQC**

Peak intensity [a.u.]

(iHNCA, 800 MHz, $T_{rec} = 200$ ms)

### Sensitivity enhancement of up to an order of magnitude !!

Weak signals are enhanced more !!

**BEST-TROSY** versus standard

Peak intensity [a.u.]

(iHNCA, 800 MHz, $T_{rec} = 200$ ms)
3D BEST-TROSY experiments

- High resolution in $^{15}$N dimension using semi CT editing

α-synuclein (600 MHz)

CT ($t_1^{\text{max}} = 24$ ms) vs. Semi-CT ($t_1^{\text{max}} = 80$ ms)
Fast-pulsing techniques (SOFAST, BEST) - Summary

- **SOFAST**
  - Most sensitive and fastest 2D experiment
  - but: MQ line widths, no extension to 3D!

- **BEST-HSQC**
  - $\geq$ 3D experiments possible
  - SQ line widths
  - Optimal performance at $< 800$ MHz

- **BEST-TROSY**
  - $\geq$ 3D experiments possible
  - Highest spectral resolution
  - Optimal performance at $\geq 800$ MHz

- Significant gain in speed, sensitivity, and evtl. spectral resolution
SOFAST/ BEST experiments – Applications

- **Sample quality control**
  High sensitivity 2D H-N correlation experiments
  Information on structural compactness from HET-SOFAST/BEST
  Information on oligomerization state from 1D BT-TRACT

- **NMR assignment**
  BEST HNC, BEST HNC\(^+\) and BEST HNN correlation experiments
  Pro-selective 2D H-N experiments.

- **Structural & dynamic information**
  Measurement of backbone scalar couplings and RDCs
  Solvent exchange rates from HET-SOFAST/BEST
  Conformational exchange from BT-CPMG-RD experiments

- **Kinetic information**
  Real-time 2D and 3D SOFAST/BEST experiments
NMRlib: easy use of BEST/SOFAST experiments

3D BT-HNN experiments
BT-HNC$^+$ experiments – improved performance

Example: HNCA

\[ H \xrightarrow{J_{NH}} N \xrightarrow{1J_{NCA}, 2J_{NCA}} CA \xrightarrow{t_1} \text{back transfer} \]

- \( N^i_x \)
- \( N^i_y \)
- \( N^i_{CA} \)
- \( N^i_{CA'} \)

Only 2 out of 4 pathways contribute to the NMR signal!
All 4 pathways contribute to the NMR signal!

BT-HNC$^+$ experiments – improved performance
BT-HNC⁺ experiments – improved performance

\[ 2\Delta_1 = 20 \text{ms} \]

- HNCA
- HNCA+
- CTP- II
  HN(CO)CA-like
- CTP- I
  HNCA

\[ \tau_c = 4 \text{ ns} \]

\[ \tau_c = 10 \text{ ns} \]

Intensity (a.u.)
Amino-acid-type information for sequential assignment

CA/CB chemical shifts

Peptide fragments:

1  3  7
2  5
4
6
8

Amino-acid-type anchoring

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTHLVLRGLG

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTHLVLRGLG
Amino-acid-type selection

Amino-acid-type editing

HADamard AMino-ACid-type edited
2D (HCBCACO)NH experiment

![Diagram of amino-acid-type editing](image)

- **Black**: Ala, Val, Ile
- **Red**: Asn, Asp
- **Green**: Cys, aromatic residues
- **Blue**: Gly
- **Yellow**: Ser
- **Magenta**: Thr
- **Cyan**: all others

Lescop, Rasia & Brutscher, *JACS (130) 2008*, 5014.
Intra-residue * and sequential HADAMAC

Complementary assignment tools for IDPs

Internet ressources

BEST:  

ASCOM:  http://www.icsn.cnrs-gif.fr/download/nmr

qMDD:  https://groups.google.com/forum/#!forum/mddnmr  
http://pc8.nmr.gu.se/~mdd/Downloads

NUS (Wagner):  
http://gwagner.med.harvard.edu/intranet/hmsIST/index.html