Multidimensional Cramér-Rao Lower Bound for Non-uniformly Sampled NMR Signals

Månsson, Anders; Jakobsson, Andreas; Akke, Mikael

Published in:
European Signal Processing Conference

2014

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
MULTIDIMENSIONAL CRAMÉR-RAO LOWER BOUND FOR NON-UNIFORMLY SAMPLED NMR SIGNALS

Anders Månsson*, Andreas Jakobsson†, and Mikael Akke†

*Dept. of Mathematical Statistics, Lund University, Sweden
†Dept. of Biophysical Chemistry, Lund University, Sweden

ABSTRACT
In this work, we extend recent results on the Cramér-Rao lower bound for multidimensional non-uniformly sampled Nuclear Magnetic Resonance (NMR) signals. The used signal model is more general than earlier models, allowing for the typically present variance differences between the direct and the different indirect sampling dimensions. The presented bound is verified with earlier presented 1- and R-dimensional bounds as well as with the obtainable estimation accuracy using the statistically efficient non-linear least squares estimator. Finally, the usability of the presented bound is illustrated as a measure of the obtainable accuracy using three different sampling schemes for a real 15N-HSQC NMR experiment.

Index Terms— NMR spectroscopy, Cramér-Rao lower bound, Non-uniform sampling

1. INTRODUCTION
In the last decade, powerful methods in Nuclear Magnetic Resonance (NMR) spectroscopy have been developed to study protein dynamics involved in increasingly complex biological phenomena [1–3]. Dynamic parameters, such as correlation times and amplitudes of motion, are determined by model fitting against sets of relaxation rates of specified NMR coherences. With the advent of modern spectrometers, a sufficient signal-to-noise ratio (SNR) is often reached before all evolution dimensions have been explored [5], although, regrettably, these measurements rely on very time-consuming experimental acquisition schemes, which are deemed necessary to obtain robust estimates of the relaxation rates, but also limit the applications to particularly stable biomolecular systems. Traditional NMR spectroscopy samples the signals uniformly, such that the signal power is measured at regular intervals. However, recent breakthroughs in compressive sensing and non-uniform sampling (NUS) have shown that one may achieve the same spectral resolution with a non-uniformly sampled damped multidimensional sinusoidal signal [a model that has been found to accurately describe various forms of NMR signals [9,10]. The CRLB offers a helpful tool to a priori evaluate the limitations of different experimental designs, by allowing a measure of the achievable accuracy for a particular setting and sampling scheme. The CRLB has previously been derived for a similar harmonic models, see e.g. [11–14]. These models, however, are either restricted to situations of low dimensionality such as 1- or 2-D models [12,13], contain a limited number of modes [15], or consider only undamped sinusoids [11–14]. In particular, the presented work may be viewed as a generalization of the results of [12], where the Fisher information matrix (FIM) is derived for a 2-D dampened sinusoidal model, with a somewhat stronger model structure than presented here, and [14], where the authors finds an expression for the CRLB of a uniformly sampled R-dimensional undamped sinusoidal model. In order to verify the presented bound, we compare the found CRLB expression with earlier 1- and R-dimensional bounds, as well as with the performance of the statistically efficient non-linear least squares (NLS) estimator. The usability of the bound is further illustrated by determining the achievable variance for a measured 2-D NMR signal when using three different sampling schemes.

2. DATA MODEL
Consider a non-uniformly sampled R-dimensional NMR signal containing F modes, such that (see e.g. [10,15])

$$x_{m_1m_2...m_R} = \sum_{f=1}^{F} c_f e^{i\phi_f} \prod_{r=1}^{R} e^{(\gamma_{r,f}+\omega_{r,f})t} r_{m_r}$$

where \(x_{m_1m_2...m_R}\) is an entry of the signal tensor, \(\mathbf{X} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}\), with \(c_f, \phi_f, \omega_{r,f} \in \mathbb{R}\), and \(\gamma_{r,f} \in \mathbb{R}^-\) denoting the amplitude, phase, frequency, and the damping coefficient in the \(r:\text{th}\) dimension for the \(f:\text{th}\) mode, respectively.

An implementation of the herein presented CRLB is publicly available from \(\text{http://www.github.com/monopolis/CRLB}\).

This work was supported in part by the Swedish Research Council, and the Crafoord and Carl Trygger foundations.
Without loss of generality, it is assumed that the frequency vectors \( \omega_f = (\omega_{1,f}, \ldots, \omega_{R,f})^T \) are distinct. Furthermore, let
\[
\mathbf{Y} = \mathbf{X} + \mathbf{E}
\]  
(2)
where \( \mathbf{Y} \) denotes the measurement tensor, and \( \mathbf{E} \) the noise tensor, where each fiber \( e_{m_2 \ldots m_R} \), i.e., the column vector formed by varying the first index while keeping the remaining indices fixed, is assumed to be an additive circularly symmetric Gaussian white noise, such that
\[
e_{m_2 \ldots m_R} \sim \mathcal{CN}(0, \sigma_{m_2 \ldots m_R}^2),
\] where \( \mathcal{CN} \) denotes the complex-valued Gaussian distribution.

It is worth noting that, in accordance with [12], the noise variance is allowed to vary over the indirect dimensions.

3. THE CRAMÉR-RAO LOWER BOUND

Let \( \mathcal{L}(\theta) = \ln p(\mathbf{Y}; \theta) \) denote the log-likelihood of \( \mathbf{Y} \), given the \( 2^R (R+1) \)-dimensional parameter vector \( \theta \), formed from all the unknown parameters. Let the data set \( \mathcal{M} \), with elements \( \mathbf{m} = (m_1, \ldots, m_R) \), define an \( R \)-dimensional grid over the sampling times \( t_m = (t_{1m_1}, \ldots, t_{Rm_R})^T \), and
\[
s_{kp}(\mathbf{m}) = -\frac{2}{\sigma_m} \exp \left\{ (\gamma_p + \gamma_k)^T t_m \right\}
\]  
(3)
\[
\psi_{kp}(\mathbf{m}) = \cos \left\{ (\omega_p - \omega_k)^T t_m + \phi_p - \phi_k \right\}
\]  
(4)
\[
\varphi_{kp}(\mathbf{m}) = \sin \left\{ (\omega_p - \omega_k)^T t_m + \phi_p - \phi_k \right\}
\]  
(5)
where \( \gamma_p = (\gamma_{1,p}, \ldots, \gamma_{R,p}) \). Then, the elements of the FIM may be formed as (see, e.g., [16])
\[
[F(\theta)]_{kl} = \mathbb{E} \left\{ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta_k \partial \theta_l} \right\}
\]  
(6)
where \( [F(\theta)]_{kl} \) denotes the \((k, l)\):th element of \( F(\theta) \), and
\[
\mathbb{E} \left\{ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta_k \partial \theta_l} \right\} = \sum_{m \in \mathcal{M}} s_{kp}(\mathbf{m}) \psi_{kp}(\mathbf{m})
\]

It is worth noting that, for all \( k, p = 1, \ldots, F \) and \( \mathbf{m} \in \mathcal{M} \),
\[
s_{kp}(\mathbf{m}) = s_{p,k}, \varphi_{k,p} = 0, \psi_{k,p} = 1, \varphi_{p,k} = -\varphi_{p,k}, \text{ and } \psi_{p,k} = \psi_{k,p}.
\] The CRLB may then be found as the diagonal elements of the matrix \( [F(\theta)]^{-1} \) (see also [17] for further details).

4. NON-UNIFORM SAMPLING SCHEMES

Various NUS schemes that may be used to improve both the performance and the acquisition time of NMR experiments have been examined in the recent literature (see, e.g., [8][18]).
The goal of such schemes is to reduce the number of samples needed, while maintaining most of the information (in a loose sense) of the signal. Reducing the number of samples yields a speed-up of the experiment acquisition time, which is critical in many applications, e.g. in the case of high-dimensional spectra, where the curse of dimensionality otherwise makes a high-resolution experiment intractable, or in situations where either sample or instruments imposes time constraints and thereby indirectly limits the achievable resolution (see e.g. [8]). Here, we apply the CRLB to evaluate the performance of three sampling schemes in a high-dimensional setting, namely uniform sampling, an exponential sampling scheme suggested in [12], which was evaluated for a two-dimensional case in [13] through the use of CRLBs, as well as a sinusoidal-weighted Poisson gapped sampling scheme proposed in [19]. These methods are tailored especially for signals containing exponential decay in the time-domain. For signals with other characteristics, these schemes will need tweaking in order to work as intended. The sampling scheme of Schmieder et al. [19] may be expressed as follows: given a stopping time $T$, a line-width $L$, and a sample size $J$, with $t_0 = 0$, then

$$t_{j+1} = -\frac{1}{\lambda} \ln \left( \frac{\exp\{-LT_j\} - 1}{J-1} \right) \quad (7)$$

for $j = 1 \ldots J - 1$. It should be noted that this sampling scheme will generally generate off-grid samples. Proceeding to the Poisson-based scheme, which given a uniform grid $\{0, \ldots, M - 1\}$, may be formed as follows: with $t_0 = 0$,

$$t_{j+1} = t_j + k_j + 1 \quad (8)$$

for $j = 1 \ldots J - 1$, where $k_j$ is a random variable drawn from a Poisson process defined by $p(k_j | \lambda_j) = (\lambda_j^k e^{-\lambda_j})/(k!)$ with a time-dependent average gap length $\lambda_j = \Lambda \sin \left( \pi \frac{2\gamma_0}{J+1} \right)$, where the parameter $\Lambda$ is tuned so that $t_J = M - 1$. When applying these schemes to higher dimensional signals, each of the indirect dimensions are sampled individually using the NUS-scheme. The direct dimension is throughout sampled using a uniform scheme.

5. NUMERICAL RESULTS

We proceed to verify the presented CRLB expression as compared to earlier presented 1- and $R$-dimensional bounds [13, 14], as well as with the mean square error (MSE) of the statistically efficient NLS estimator. To ensure correct convergence, we here for simplicity initialize the NLS-estimator to values close to the true parameters by perturbing the values using a random uniform distribution. The perturbations are scaled such that each parameter is perturbed at most 0.1% of its magnitude. Beginning with the 1-D case, Fig. 1 illustrates

$$\text{SNR} = \frac{\|X\|^2}{\prod_{r=1}^R M_r \sigma_r^2} \quad (11)$$

and where $\| \cdot \|_F$ denotes the Frobenious norm, and with $\sigma_r^2$ being the variance of the noise tensor. Clearly, the results coincide; the bounds and accuracy for the other parameters behave similarly. In the second and third experiment, we amend additional damping/frequency dimensions to the parameter configuration of [20] with the specifics of the experiments detailed in [17]. A 2-D signal is created analogously to [10] and is evaluated over the rectangular grid $\{0, \ldots, 127\} \times \{0, \ldots, 31\}$. Fig. 2 shows the presented bound for the frequency of the first mode as compared to the bound derived in [13] and with the MSE of the NLS-estimator, evaluated using 1000 Monte Carlo simulations. Here,

$$p_1 = (c_1, \gamma_1, f_1, \phi_1) = (0.9, -0.1, 0.1, 0.11)$$

$$p_2 = (c_2, \gamma_2, f_2, \phi_2) = (0.5, -0.1, 0.2, 0.8) \quad (9)$$

and

$$y_k = \sum_{j=1}^F s(p_j; t_k) \quad (10)$$

with $s(c, \gamma, f, \phi; t_k) = ce^{(\gamma+i2\pi f)t_k+i\phi}, t_k = 0, \ldots, 127$, and where the SNR is defined as (see also [22])

An analog scheme that generates on-grid samples is presented in [21].

Fig 3. Comparing the CRLB to the bound presented in [14] for $f_{1,1}$ shows a close agreement at small damping coefficients. With larger damping coefficients, the bounds grow apart. The bounds and accuracy for the other parameters behave similarly.

the presented bound for the frequency of the first mode as compared to the bound derived in [13] and with the MSE of the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>0.9</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.1</td>
</tr>
<tr>
<td>$f_1$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.11</td>
</tr>
<tr>
<td>$c_2$</td>
<td>0.5</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.1</td>
</tr>
<tr>
<td>$f_2$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.8</td>
</tr>
</tbody>
</table>

The bounds and accuracy for the other parameters behave similarly. In the second and third experiment, we amend additional damping/frequency dimensions to the parameter configuration of [20] with the specifics of the experiments detailed in [17]. A 2-D signal is created analogously to [10] and is evaluated over the rectangular grid $\{0, \ldots, 127\} \times \{0, \ldots, 31\}$. Fig. 2 shows the presented bound for the frequency of the first mode as compared to the bound derived in [13] and with the MSE of the NLS-estimator, evaluated using 1000 Monte Carlo simulations. Here,
two bounds are in almost complete agreement. As the dampening magnitude increases ($\gamma_{r, f} \sim 10^{-1}$), the two bounds grow further apart. It should be pointed out that, since our model contains a larger number of parameters it will invariably have a larger CRLB. The magnitude of the difference can, however, as shown, be small in comparison to the actual values of the CRLBs. Finally, using a measured NMR signal from a $^{15}$N-HSQC experiment of a Histidine sample, with pH 11.14 (25°C), acquired at 600 MHz, we examine the achievable performance of the aforementioned sampling schemes. The data set consists of 1024 x 512 samples distributed on a uniform grid. From the measured signal, the 12 most dominant modes are estimated using the sparse estimation algorithm presented in [23]. In order to maintain a correct SNR for the signal generated from the estimates, the amplitudes of the chosen modes are scaled such that the periodogram of the estimated signal is to the periodogram of the original signal. The correct noise level is subsequently chosen such that noise floor of the two periodograms are close to identical (see also [17] for further details). From the estimated parameters, the CRLBs of the various schemes are then computed. The sampling schemes are tuned for a 3x speed-up in terms of acquisition time while covering the same intervals as the original grid. For the exponential sampling scheme, we chose $L = 0.01$. It may be noted, that the herein presented CRLBs should be viewed as approximative, as the measured signal in fact also contains additional modes. The results of this experiment are presented in Table 1 where one finds the estimated damping parameters and the corresponding 3σ-confidence interval, computed via the corresponding CRB values. The total number of samples in all dimensions are for the three sampling schemes 1.048.576, 349.184, and 349.184, respectively. Observing the data, one finds that it is possible by the usage of a NUS scheme, to achieve a 3x speed-up at little to no loss in performance. For the exponential sampling scheme, we note an increase in performance as compared to the uniform sampling scheme. This is likely due to scheme producing off-grid samples, but also due to the fact that the scheme samples much sparser at larger times in the indirect dimension, and thereby maintains a higher SNR.

6. REFERENCES


<table>
<thead>
<tr>
<th>Estimand</th>
<th>Uniform</th>
<th>Exponential</th>
<th>Poisson</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\gamma}_{1,1}$</td>
<td>-0.00343 ± 3.16e-05</td>
<td>2.80e-05</td>
<td>4.67e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,2}$</td>
<td>-0.00800 ± 7.31e-05</td>
<td>8.38e-05</td>
<td>1.05e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,3}$</td>
<td>-0.00343 ± 3.51e-05</td>
<td>3.11e-05</td>
<td>5.18e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,4}$</td>
<td>-0.000800 ± 8.10e-05</td>
<td>9.29e-05</td>
<td>1.17e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,5}$</td>
<td>-0.00343 ± 5.59e-05</td>
<td>4.97e-05</td>
<td>8.26e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,6}$</td>
<td>-0.00784 ± 1.27e-04</td>
<td>1.47e-04</td>
<td>1.81e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,7}$</td>
<td>-0.00343 ± 3.78e-05</td>
<td>3.35e-05</td>
<td>5.59e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,8}$</td>
<td>-0.008000 ± 8.74e-05</td>
<td>1.01e-04</td>
<td>1.26e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,9}$</td>
<td>-0.00343 ± 3.78e-05</td>
<td>3.34e-05</td>
<td>5.58e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,10}$</td>
<td>-0.000800 ± 8.73e-05</td>
<td>1.00e-04</td>
<td>1.26e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,11}$</td>
<td>-0.00343 ± 6.00e-05</td>
<td>5.43e-05</td>
<td>8.96e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,12}$</td>
<td>-0.00713 ± 1.24e-04</td>
<td>1.52e-04</td>
<td>1.80e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,13}$</td>
<td>-0.00343 ± 5.05e-05</td>
<td>4.47e-05</td>
<td>7.47e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,14}$</td>
<td>-0.000800 ± 1.17e-04</td>
<td>1.34e-04</td>
<td>1.68e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,15}$</td>
<td>-0.00343 ± 5.31e-05</td>
<td>4.70e-05</td>
<td>7.85e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,16}$</td>
<td>-0.008000 ± 1.23e-04</td>
<td>1.41e-04</td>
<td>1.76e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,17}$</td>
<td>-0.00250 ± 4.76e-05</td>
<td>4.94e-05</td>
<td>8.26e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,18}$</td>
<td>-0.000800 ± 1.41e-04</td>
<td>1.64e-04</td>
<td>2.05e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,19}$</td>
<td>-0.00250 ± 4.81e-05</td>
<td>4.98e-05</td>
<td>8.34e-05</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,20}$</td>
<td>-0.000800 ± 1.42e-04</td>
<td>1.65e-04</td>
<td>2.07e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,21}$</td>
<td>-0.00250 ± 7.30e-05</td>
<td>7.57e-05</td>
<td>1.27e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,22}$</td>
<td>-0.000800 ± 2.16e-04</td>
<td>2.51e-04</td>
<td>3.15e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,23}$</td>
<td>-0.00250 ± 7.37e-05</td>
<td>7.64e-05</td>
<td>1.28e-04</td>
</tr>
<tr>
<td>$\hat{\gamma}_{1,24}$</td>
<td>-0.000800 ± 2.18e-04</td>
<td>2.53e-04</td>
<td>3.17e-04</td>
</tr>
</tbody>
</table>


