A highly efficient blocked Gibbs sampler reconstruction of multidimensional NMR spectra

Ji Won Yoon and Simon P. Wilson
School of Computer Science and Statistics,
Trinity College Dublin, Ireland

K. Hun Mok
School of Biochemistry and Immunology,
Trinity College Dublin, Ireland

Abstract

Projection Reconstruction Nuclear Magnetic Resonance (PR-NMR) is a new technique to generate multi-dimensional NMR spectra, which have discrete features that are relatively sparsely distributed in space. A small number of projections from lower dimensional NMR spectra are used to reconstruct the multi-dimensional NMR spectra. We propose an efficient algorithm which employs a blocked Gibbs sampler to accurately reconstruct NMR spectra. This statistical method generates samples in Bayesian scheme. Our proposed algorithm is tested on a set of six projections derived from the three-dimensional 700 MHz HNCO spectrum of HasA, a 187-residue heme binding protein.

1 Introduction

Current methods used to determine the structure of molecules at atomic level resolution include X-ray crystallography and NMR spectroscopy. Whereas determining molecular structure with X-ray diffraction requires well-ordered crystals or fibres, NMR spectroscopy has the advantage of being able to analyse solids, liquids, or even gaseous states. One of the challenges for NMR spectroscopy is chemical shift overlap, which arises quite frequently in the spectra of biomolecules (such as proteins and nucleic acids) where a large number of nuclei are present. In order to overcome this, various pulse sequences to increase spectral resolution have been in development. Such *multi-dimensional NMR spectroscopy technologies* share one drawback in that the acquisition times are extremely long, exponentially increasing with the number of experiments. For example, suppose that k time points are sampled in each evolution period for a K-dimensional NMR experiment that detects K nuclei types. In this case, the total experimental time is determined by the number of time points (k), the number of scans (N_a) averaged directly, and the duration of each scan (Simon & Sattler, 2004). Thus, when $N_a = 10$ and there are 60 time points per evolution period, the minimum experimental time will be 10 minutes ($60 \times 10s$) for a 2D (K=2), 10 hours for a 3D (K=3), 25 days for a 4D (K=4), and so on. For these reasons, several investigators have been trying to speed up these measurements by more efficient approaches (Mishkovsky & Frydman, 2009; Felli & Brutscher, 2009; Mok et al., 2007). Many approaches are based on the concept of *accordion spectroscopy* (Bodenhausen & Ernst, 1982). One recent approach to address this problem is GFT-NMR by Kim and Szyperski (Kim & Szyperski, 2003). GFT-NMR experiments detect sums and differences of chemical shifts by linking the various evolution dimensions. Another approach which is mainly considered in this paper is Projection Reconstruction NMR (PR-NMR), initially developed by Freeman and Kupce (Kupce & Freeman, 2003a; Freeman & Kupce, 2004; Kupce & Freeman, 2004b,a).

To implement the reconstruction of multi-dimensional NMR spectra, we have investigated the use of several deterministic and statistical approaches (Yoon et al., 2006; Yoon & Godsill, 2006). Although deterministic algorithms, including Back projection and Lowest algorithms, can reconstruct the spectra rapidly, the reconstructed spectra are rather different from the underlying ground truth. Meanwhile, the statistical approaches based on a Bayesian scheme are relatively slow but provide closer reconstruction to the ground truth than the deterministic approaches.

In addition, there are two types of schemes to reconstruct spectra from projections: the pixel-by-pixel approach and the peak-by-peak approach. In our previous papers (Yoon & Godsill, 2006; Yoon et al., 2006; Yoon & Fitzgerald, 2009), we demonstrated Maximum Likelihood (ML) reconstruction, Maximum a Posterior (MAP) reconstruction and Maximum Entropy (ME) (Mobli et al., 2006) reconstruction and truncated Gibbs sampling reconstruction (Yoon & Fitzgerald, 2009) for the pixel-by-pixel ap-

Appearing in Proceedings of the 13^{th} International Conference on Artificial Intelligence and Statistics (AISTATS) 2010, Chia Laguna Resort, Sardinia, Italy. Volume 9 of JMLR: W&CP 9.

proach. We also showed that Reversible Jump Markov Chain Monte Carlo (RJMCMC) reconstruction based on the peak-by-peak scheme can work better than other approaches based on the pixel-by-pixel approach in that it reduces the complexity of the problem and speeds up the calculation. However, if the shape of the peaks does not fit to given profiles such as Gaussian or Laplacian distribution, the peak-by-peak approach may separate a large peak into a myriad of small peaks to make the large peak. This may make peak-by-peak approaches less efficiently. Worse, it may result in unwanted shapes of peaks on the spectra. Therefore, we can still use the pixel-by-pixel approaches since they are not dependent on the shape of the peaks and we showed the superiority of truncated Gibbs sampling reconstruction in the pixel-by-pixel based approaches. However, it is known that the separate estimation of highly dependent parameters by Gibbs sampling requires a long time to converge to stationary distribution. Therefore, we propose an algorithm which is updating the high dimensional spectra jointly per block while our previous truncated algorithm updates every pixel in a Gibbs scheme. As known, the blocked Gibbs sampler works more efficiently if the hidden variables to be inferred are highly dependent.

Three contributions are described in this paper. Firstly, we introduce a robust mathematical model for PR-NMR, which is similar to Bayesian mapping of disease model based on intrinsic Gaussian Markov Random Field. Secondly, we suggest a heuristic algorithm to build an efficient approximation of a marginalised precision matrix. Lastly, the reconstruction of the multi-dimensional NMR spectra is done in a blocked sampling framework to consider the highly dependent spatial properties.

2 Mathematical model for PR-NMR

2.1 A hierarchical model for PR-NMR

In PR-NMR, the input dataset is a small number of projections obtained at different projection angles. Let $\mathbf{f}$ be the underlying NMR spectra of interest. Suppose $\mathbf{y}$ is a set of projections and Λ is a set of projection angles such that $\mathbf{y}^{(r)}$ is a projection of the rth angle where $r = 1, \cdots, |\Lambda|$. Here, $|\cdot|$ is the number of elements of a set and φ_r denotes a scaling factor. Since the projections are the sum of individual intensity in frequency domains, the PR-NMR data is then defined as

$$\begin{aligned} p(\mathbf{y}^{(r)}|\eta, \theta) &= \mathcal{N}(\cdot; \varphi_r \mathbf{G}_r \eta, \kappa_r^{-1}\mathbf{I}) \\ \eta &= \mathbf{z}^T\beta + \mathbf{f} + \mathbf{u} \end{aligned} \quad (1)$$

where $\mathbf{f}$ represents a random field associated with some properties of the spatial domain and $\mathbf{u}$ is an unstructured term where $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \rho^{-1}\mathbf{I})$. Here, $\mathcal{N}$ denotes Normal distribution and $\mathbf{G}_r$ is a linear projection matrix for the rth angle. With this mathematical model, to reconstruct the multi-dimensional NMR spectra is equivalent to a problem to infer $\mathbf{f}$ from the posterior $p(\mathbf{f}|\mathbf{y})$. We set $\mathbf{x} = \{\eta, \mathbf{f}\}$ and $\theta = \{\kappa_{1:|\Lambda|}, \tau, \rho, \beta, \varphi_{1:|\Lambda|}\}$. We demonstrate a Gibbs sampler based NMR reconstruction algorithm using Gaussian Markov Random Field (GMRF) prior (Besag, 1974; Rue & Held, 2005). However, this algorithm is only for a small size spectra since the projection matrix $\mathbf{G}_r$ is extremely large and it is practically infeasible to calculate the inverse of the matrix $\mathbf{G}_r^{-1}$, which is often used in reconstruction. Alternatively, we introduce blocked Gibbs sampler approaches drawing samples from its blocked conditionals in order to reduce the size of projection matrix and to maintain the Markovian dependency in spatial property.

2.2 Gaussian Markov Random Fields

Gaussian Markov Random fields (GMRFs) are defined as discrete Gaussian fields with a Markov property of conditional independence of a component with all other given its neighbours (Mardia, 1988). They have seen widespread application in statistical modelling, for example in spatio-temporal models (Besag et al., 1991) and dynamic linear models (West & Harrison, 1997). The GMRF is also one of the most popular schemes for images (Weir, 1997; Hanson & Wecksung, 1983; Hunt, 1977; Therrien, 1993).

A random variable $\mathbf{f}$, which denotes spectra, is used to specify the GMRF. More formally, let $\mathbf{f} = (f_1, \cdots, f_N)^T$ have a normal distribution with mean μ and covariance Σ when there are N pixels in the image where T denotes the transpose operation. We define $\Delta_{i,j}$ as differenced values of $\mathbf{f}$ at site (i, j) on the lattice. In this paper, a second order random walk is defined through letting $\Delta_{i,j} = \mathbf{f}_{i+1,j} + \mathbf{f}_{i-1,j} + \mathbf{f}_{i,j+1} + \mathbf{f}_{i,j-1} - 4\mathbf{f}_{i,j}$ be independent Gaussians, for $i, j = 2, \ldots, n-1$. The distribution on $\mathbf{f}$ is of Gaussian form but the precision matrix is not of full rank; this is known as an intrinsic GMRF and they are widely used as prior distributions in Bayesian latent models (Rue & Held, 2005). The prior distribution for $\mathbf{f}$ is then

$$\begin{aligned} p(\mathbf{f}|\theta) &\propto \exp\left\{-\frac{\tau}{2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\Delta_{i,j}^2\right\} \\ &= \exp\left\{-\frac{\tau}{2}(\mathbf{Df})^T(\mathbf{Df})\right\} \\ &= \exp\left\{-\frac{\tau}{2}\mathbf{f}^T\mathbf{Q}_f\mathbf{f}\right\} \end{aligned} \quad (2)$$

where $\mathbf{Q}_f = \mathbf{D}^T\mathbf{D}$ and τ is a scale parameter for $\tau \in \theta$.

2.3 Posterior distribution

Our interest is inference of $p(\mathbf{x}, \theta|\mathbf{y})$ and its conditional distribution of $\mathbf{x}$ given θ is defined by

$$p(\mathbf{x}|\mathbf{y}, \theta) \propto \exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \sum_{r=1}^{|\Lambda|} g_r(\mathbf{x})\right\} \quad (3)$$

where $g_r(\mathbf{x})$ is $\log p(\mathbf{y}^{(r)}|\mathbf{x},\theta_r)$. Let $\mathbf{x} = (\mathbf{x}_A, \mathbf{x}_{-A})$ where A represents a set of indexes of the vector $\mathbf{x}$ and $-A$ denotes the others. For a large size data, we need the conditional distribution for a blocked area A of $\mathbf{x}$, which is notated by $\mathbf{x}_A$, rather than full joint distribution. Since our interest turns to $p(\mathbf{x}_A|\mathbf{x}_{-A},\mathbf{y},\theta)$, we need conditional prior for the blocked data, $p(\mathbf{x}_A|\mathbf{x}_{-A},\theta)$. Using Eq. (1), the conditional prior distribution is defined by $p(\mathbf{x}_A|\mathbf{x}_{-A},\theta) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}_A-\mu_A)^T\mathbf{Q}_A(\mathbf{x}_A-\mu_A)\right\}$ where $\mu_A = \begin{bmatrix} \mathbf{0} \\ -\frac{\mathbf{z}_A^T\beta}{\rho} \end{bmatrix}$ and

$$\mathbf{Q}_A = \begin{bmatrix} \rho\mathbf{I}_{K\times K} & -\rho\mathbf{I}_{K\times K} \\ -\rho\mathbf{I}_{K\times K} & \rho\mathbf{I}_{K\times K} + \tau\mathbf{Q}_{f_A} \end{bmatrix} \quad (4)$$

and the details for derivation of Eq. (4) are described in Appendix A. Here $\mathbf{Q}_{f_A}$ is obtained by an approximation in Section 2.3.2. Therefore, the conditional posterior of our main interest is

$$\begin{aligned} & p(\mathbf{x}_A|\mathbf{x}_{-A},\mathbf{y},\theta) \\ \propto\ & \exp\left\{-\frac{1}{2}(\mathbf{x}_A-\mu_A)^T\mathbf{Q}_A(\mathbf{x}_A-\mu_A) + \sum_{r=1}^{|\Lambda|} g_r(\mathbf{x})\right\} \\ \propto\ & \exp\left\{-\frac{1}{2}(\mathbf{x}_A-\mu_A^*)^T\mathbf{Q}_A^*(\mathbf{x}_A-\mu_A^*)\right\}. \end{aligned} \quad (5)$$

Then we can draw samples $\mathbf{x}_A$ on the conditional posterior distribution of Eq. (5): $\mathbf{x}_A \sim q(\mathbf{x}_A|\mathbf{x}_{-A},\mathbf{y},\theta) = p(\mathbf{x}_A|\mathbf{x}_{-A},\mathbf{y},\theta)\mathbf{I}_{\mathbf{x}_A\geq 0}(\mathbf{x}_A)$ where $\mathbf{I}$ is an indicator function for truncation. It is known that projection reconstruction is an inverse problem and has a myriad of optima. Under the assumption that the spectra have non-negative intensity, we adapt truncation with the indicator function to remove such unrealistic local optima. We describe only equations for the blocked update for $\mathbf{x}_A$ in main sections and the estimation for the other parameters θ is illustrated in Appendix B.

2.3.1 Approximating $p(\mathbf{x}_A|\mathbf{x}_{-A},\mathbf{y},\theta)$

The next stage is to obtain μ_A^* and $\mathbf{Q}_A^*$ of Eq. (5). Let η and η' denote the current and proposed values of η and partition η into η_A and η_{-A} for an index subset A. After rearranging the order of elements, we now have $\eta = \begin{bmatrix} \eta_A \\ \eta_{-A} \end{bmatrix}$ and $\eta' = \begin{bmatrix} \eta_A' \\ \eta_{-A} \end{bmatrix}$ when η_A is only updated to η_A'. Thus we have $\eta' = \eta - \begin{bmatrix} \eta_A \\ \eta_{-A}=\mathbf{0} \end{bmatrix} + \begin{bmatrix} \eta_A' \\ \eta_{-A}=\mathbf{0} \end{bmatrix}$. Therefore the second part of the right-hand side of Eq. (3), which means log likelihood (L), is

$$\begin{aligned} L &= \sum_{r=1}^{|\Lambda|} g_r(\mathbf{x}) = \log p(\mathbf{y}|\mathbf{x},\theta) \\ &= -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\mathbf{y}^{(r)} - \varphi_r\mathbf{G}_r\eta'\right)^T\kappa_r\mathbf{I}\left(\mathbf{y}^{(r)} - \varphi_r\mathbf{G}_r\eta'\right) \\ &= -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\mathbf{y}^{(r)} - \varphi_r\mathbf{G}_r\eta + \varphi_r\mathbf{G}_r\begin{bmatrix}\eta_A \\ \eta_{-A}=\mathbf{0}\end{bmatrix}\right. \\ &\quad \left. -\varphi_r\mathbf{G}_r\begin{bmatrix}\eta_A' \\ \eta_{-A}=\mathbf{0}\end{bmatrix}\right)^T\kappa_r\mathbf{I} \\ &\quad \times\left(\mathbf{y}^{(r)} - \varphi_r\mathbf{G}_r\eta + \varphi_r\mathbf{G}_r\begin{bmatrix}\eta_A \\ \eta_{-A}=\mathbf{0}\end{bmatrix}\right. \\ &\quad \left. -\varphi_r\mathbf{G}_r\begin{bmatrix}\eta_A' \\ \eta_{-A}=\mathbf{0}\end{bmatrix}\right). \end{aligned} \quad (6)$$

Let $\bar{\mathbf{y}}^{(r)} = \mathbf{y}^{(r)} - \varphi_r\mathbf{G}_r\eta$ and $\mathbf{G}_r = (\bar{\mathbf{G}}_r, \neg\bar{\mathbf{G}}_r)$ where $\bar{\mathbf{G}}_r$ is a set of columns which are associated with η_A and $\neg\bar{\mathbf{G}}_r$ is its complement, then we have the following equation:

$$\begin{aligned} L &= -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\bar{\mathbf{y}}^{(r)} + \varphi_r\bar{\mathbf{G}}_r\eta_A - \varphi_r\bar{\mathbf{G}}_r\eta_A'\right)^T \\ &\quad \times\kappa_l\mathbf{I}\left(\bar{\mathbf{y}}^{(r)} + \varphi_r\bar{\mathbf{G}}_r\eta_A - \varphi_r\bar{\mathbf{G}}_r\eta_A'\right) \\ &= -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\eta_A' - \frac{\bar{\mathbf{G}}_r^{-1}\bar{\mathbf{y}}^{(r)}}{\varphi_r} - \eta_A\right)^T \\ &\quad \times\kappa_r\varphi_r^2\bar{\mathbf{G}}_r^T\bar{\mathbf{G}}_r\left(\eta_A' - \frac{\bar{\mathbf{G}}_r^{-1}\bar{\mathbf{y}}^{(r)}}{\varphi_r} - \eta_A\right). \end{aligned} \quad (7)$$

Note the comparison between $\mathbf{G}_r$ and $\bar{\mathbf{G}}_r$. When $\mathbf{y}^{(r)}$ is a $K\times 1$ vector and the η is a $N\times 1$ vector, $\mathbf{G}_r$ is $K\times N$. In case of blocked area, $\bar{\mathbf{y}}^{(r)}$ is represented by a $n\times 1$ vector so the size of $\bar{\mathbf{G}}_r$ is $K\times n$ where $N >> n$. Finally, we have the following posterior distribution:

$$\begin{aligned} & \log p(\mathbf{x}_A|\mathbf{x}_{-A},\mathbf{y},\theta) \\ &= -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\eta_A' - \frac{\bar{\mathbf{G}}_r^{-1}\bar{\mathbf{y}}^{(r)}}{\varphi_r} - \eta_A\right)^T \\ &\quad \times\kappa_r\varphi_r^2\bar{\mathbf{G}}_r^T\bar{\mathbf{G}}_r\left(\eta_A' - \frac{\bar{\mathbf{G}}_r^{-1}\bar{\mathbf{y}}^{(r)}}{\varphi_r} - \eta_A\right) \\ &\quad -\frac{1}{2}(\mathbf{x}_A-\mu_A)^T\mathbf{Q}_A(\mathbf{x}-\mu_A) + \text{const} \\ &= -\frac{1}{2}\left(\mathbf{x}_A-\mu_A^*\right)^T\mathbf{Q}_A^*\left(\mathbf{x}_A-\mu_A^*\right) + \text{const} \end{aligned} \quad (8)$$

where $\mathbf{Q}_A^* = \begin{bmatrix} \sum_{r=1}^{|\Lambda|}\kappa_r\varphi_r^2\bar{\mathbf{G}}_r^T\bar{\mathbf{G}}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \mathbf{Q}_A$ and $\mu_A^* = \mathbf{Q}_A^{*-1}\Psi$ where

$$\Psi = \left\{\begin{bmatrix} \sum_{r=1}^{|\Lambda|}\kappa_r\varphi_r^2\bar{\mathbf{G}}_r^T\bar{\mathbf{G}}_r(\frac{\bar{\mathbf{G}}_r^{-1}\bar{\mathbf{y}}^{(r)}}{\varphi_r} + \eta_A) \\ \mathbf{0} \end{bmatrix} + \mathbf{Q}_A\mu_A\right\}$$

2.3.2 Approximating $\mathbf{Q}_{f_A}$

In Eq. (4), we find that $\mathbf{Q}_{f_A}$ is not equal to $\mathbf{Q}_f$ which is a precision matrix of the intrinsic GMRF of Eq. (2) since $\mathbf{Q}_{f_A}$ is a marginalized precision matrix on $\mathbf{f}_A$. Thus, we

cannot easily build an intrinsic GMRF for $\mathbf{f}_A$ to build $\mathbf{Q}_{f_A}$ by using Eq. (2). However, it is rather difficult to obtain exact values of the marginalized precision due to following reasons. Firstly, it is intractable to calculate the marginalized precision matrix given a large precision matrix $\mathbf{Q}_f$ while it is straightforward to find marginalized covariance Σ_{f_A} given a covariance Σ_f where $\Sigma_f = \begin{bmatrix} \Sigma_{f_A} & \Sigma_2 \\ \Sigma_2^T & \Sigma_3 \end{bmatrix}$. Worse, in a high dimensional application domain such as a 512×512 image, the size of the full precision matrix is $512^2 \times 512^2$, which is impractical computationally. Therefore, heuristic approximation algorithms to build the marginalized precision matrix without such huge computational load are required. Fig. 1 exemplifies the basic idea of the heuristic approximation which we propose in this paper. We name this algorithm *Proportional increase approximation*. Blue and bright grey squares represent a chopped (blocked) area of interest and a whole area (spectra). The heuristic algorithm for the efficient selection makes an orange area which is larger than the chopped area but still much smaller than the whole area. After linking each corner between blue and bright grey squares, the orange area is determined by varying the parameter t for $0 \leq t \leq 1$. Now we can simply obtain the marginalized precision by inverting the precision matrix of the orange area, which is much cheaper than inverting the whole area.

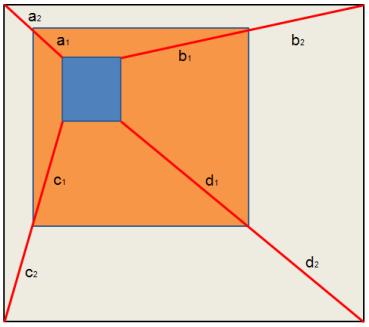


Figure 1: The size of approximated area in *Proportional increase approximation* is determined with varying t: $(t : 1-t) = (a_1 : a_2) = (b_1 : b_2) = (c_1 : c_2) = (d_1 : d_2)$ for $0 \leq t \leq 1$

In Fig. 2, three heuristic algorithms are described: regular increase (a, d), infinite increase (b, e) and proportional increase (c, f) models. Graphs (a, b, c) and (d, e, f) of this figure explain the possible expansion of the area in smaller and larger areas respectively. Black arrows shows the possibility of the expansion for each method.

We highlighted the performance of this heuristic approach with other naive heuristics as shown in Figs. 3 and 4. Fig. 3 shows the comparison of standard deviations of differences of elements between a true precision and an approximated precision by three different approaches. In order

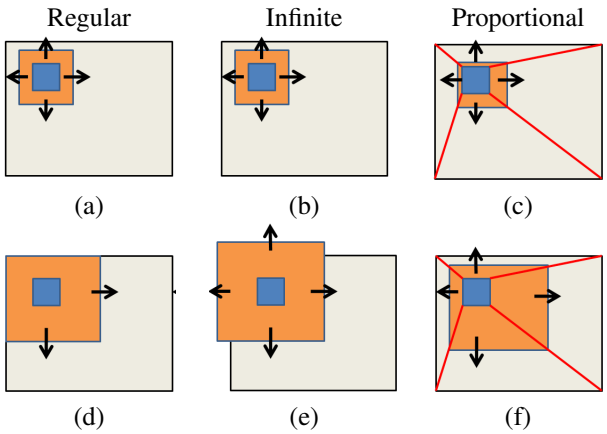


Figure 2: Algorithms of three approaches to approximate the marginal precision: Regular increase approach (a, d), infinite increase approach (b, e) and proportional increase approach (c, f). White square and blue square represent a large $N \times N$ image and a blocked $n \times n$ image. An orange square represent a chopped image to obtain an approximated precision where its volume is $V \approx L^2$. (a, b, c) and (d, e, f) are for smaller orange squares and for larger orange squares.

to validate the performance comparison, we generated 100 random samples with N_i, n_i, V where n is the size of a chopped area, N is the size of a whole area and V is the volume of the areas of the orange area obtained by the heuristic approximation. Afterwards, we calculated the averages of Area Under Curves (AUC) of Fig. 3 in terms of n as shown in Fig. 4 for the 100 random samples. In these figures, we find that proportional increase model is superior to approximate the marginalized precision matrix.

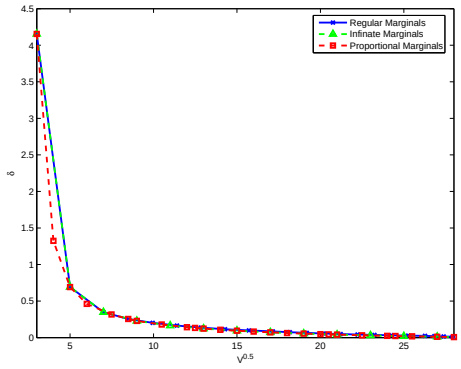


Figure 3: Comparison of standard deviations of differences of elements between a truly marginalized precision and an approximately marginalized precision by three different approaches

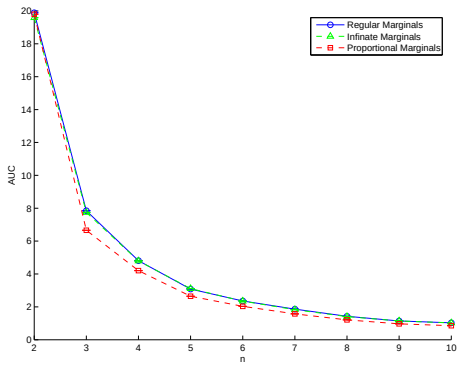


Figure 4: Comparison of AUCs of Fig. 3 in terms of randomly selected parameters including N, n and V.

3 Algorithms

Our main interest is to obtain a reconstructed NMR spectra $\mathbf{f}$ and to estimate other hidden variables, including scaling factors φ and noise precision κ and τ. We obtain the underlying surface by inferring the marginal posterior distribution $p(\mathbf{x}|\mathbf{y}) = \frac{1}{p(\mathbf{y})} \int_\theta p(\mathbf{y}|\mathbf{x},\theta)p(\mathbf{x}|\theta)p(\theta)d\theta$. However, it is rather difficult to jointly estimate $\mathbf{x}$ from $p(\mathbf{x}|\mathbf{y})$ because (1) the size of $\mathbf{x}$ is too large to calculate inverse matrix and (2) marginalizing the system parameters θ is not so easy since it often does not have any closed form. To address the two problems, we applied two strategies: we sequentially estimate partial regions using conditional Gibbs sampling and then we throw away and ignore samples of the θ after convergence.

Algorithm 1 Reconstruction by Blocked Gibbs sampler

for $k = 1$ to N_{conv} **do**
  Sample $\mathbf{x}^{(k)}$ from $p(\mathbf{x}|\mathbf{y},\theta)$.
  Partition $\mathbf{x}$ into N_p regions $\{A_i\}_{i=1}^{N_p}$.
  for $i = 1$ to $i = N_p$ **do**
    $\mathbf{x}_{A_i} \sim q(\mathbf{x}_{A_i}|\mathbf{x}_{-A_i},\mathbf{y},\theta)$ where $\cup_{i=1}^{N_p} A_i = A$ and $A_i \cap A_j = \{\}$ for $i \neq j$.
  end for
  $\mathbf{x}^{(k)} \leftarrow \cup_{i=1}^{N_p} \mathbf{x}_{A_i}$
  $\varphi^{(k)} \sim p(\varphi|\mathbf{x},\mathbf{y},\theta_{-\varphi})$.
  $\rho^{(k)} \sim p(\rho|\mathbf{x},\mathbf{y},\theta_{-\rho})$.
  $\kappa^{(k)} \sim p(\kappa|\mathbf{x},\mathbf{y},\theta_{-\kappa})$.
  $\tau^{(k)} \sim p(\tau|\mathbf{x},\mathbf{y},\theta_{-\tau})$.
  $\beta^{(k)} \sim p(\beta|\mathbf{x},\mathbf{y},\theta_{-\beta})$.
end for
$\hat{\mathbf{x}} = \frac{1}{\frac{N_{conv}}{2}-1} \sum_{k=\frac{N_{conv}}{2}+1}^{N_{conv}} \mathbf{x}^{(k)}$.

4 Experimental Results

Fig. 5 shows the experimental projected spectral dataset for reconstruction of NMR spectra using the Gibbs sampler reconstruction. They are projections of $^{13}C^{15}N$ correlation peaks in the 700MHz HNCO spectrum of the protein HasA. Fig. 6 (a) shows the contour images of a desired target map. The F_1F_2 plane extracted from the full three-dimensional HNCO experiment on HasA, performed by the conventional method where both evolution times are incremented independently.

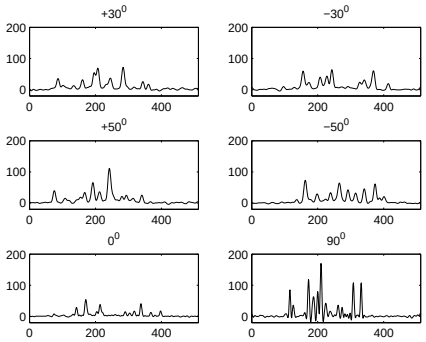
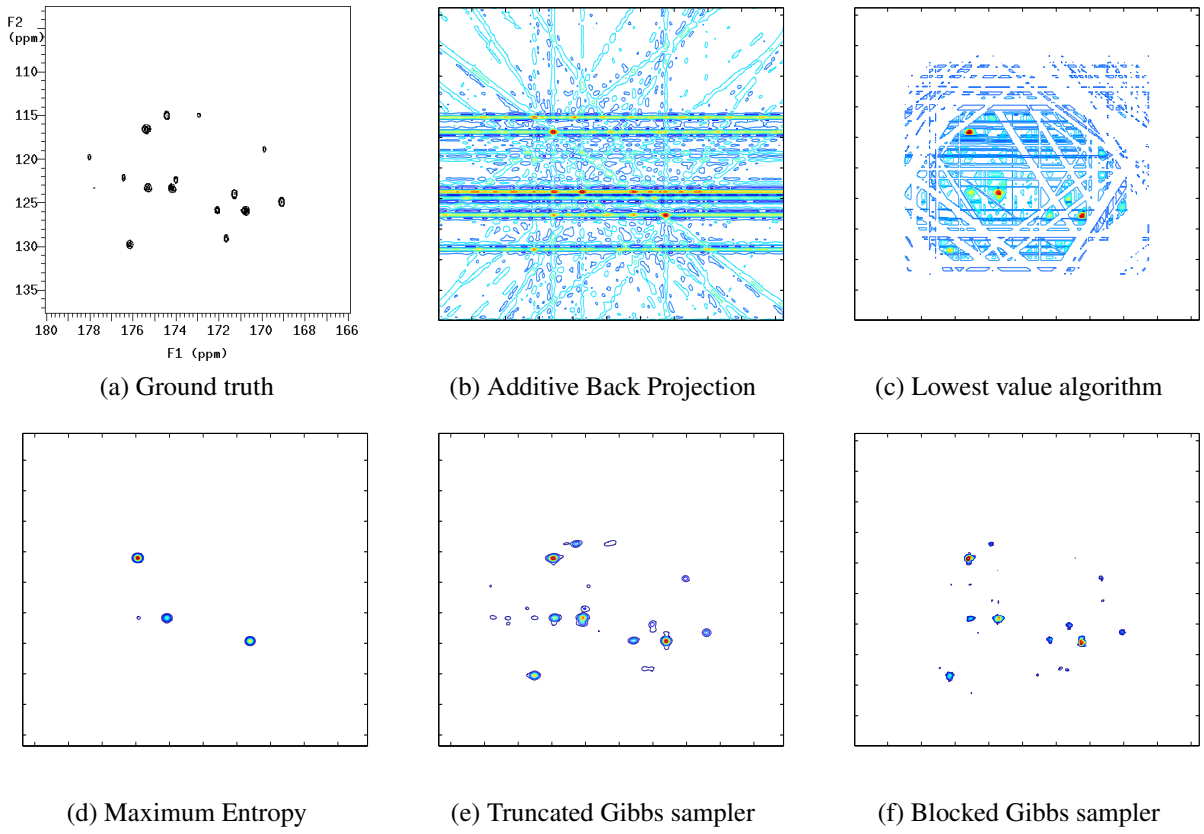


Figure 5: Experimental NMR spectra with its 6 Projections

We compared the results of the blocked Gibbs sampler reconstruction with those of back projection reconstruction (Kupce & Freeman, 2003b), lowest value reconstruction, Maximum Entropy reconstruction (Reis & Roberty, 1992) which are well-known methods in NMR and truncated Gibbs sampler reconstruction. A contour of an experimental target map is shown in Fig. 6 (a). The additive back projection and lowest value reconstructions of Fig. 6 (b) and (c) show that there are a lot of artifacts and unwanted ridges and we cannot distinguish the false alarms from real signals. Fig. 6 (d) shows that the reconstructed spectra from Maximum Entropy has serious problems with poor detection of peaks, since weak signals are regarded as noise (Yoon & Godsill, 2006). Truncated Gibbs sampler and blocked Gibbs sampler detect almost all peaks including very small peaks as in Fig. 6 (e) and (f). We compared the detected positions with varying thresholds into the reconstructed spectra. Fig. 7 shows the comparsion of ROC and accuracy the reconstructed spectra by the several approaches. Blocked Gibbs sampler has higher specificity given sensitivity and high accuracy with varying thresholds.

5 Discussion

We focused on how to apply Gibbs sampler to the latent Gaussian model with Gaussian Markov Random Fields (GMRF) prior in a large image for projection reconstruction in this paper. It was our initial intention to adapt Integrated Nested Laplace Approximation (INLA) rather than Gibbs sampler to our model because INLA had seemed promising (Rue et al., 2009). However, the difficulty of



(a) Ground truth (b) Additive Back Projection (c) Lowest value algorithm

(d) Maximum Entropy (e) Truncated Gibbs sampler (f) Blocked Gibbs sampler

Figure 6: Contour maps of reconstructed spectra for experimental dataset

processing data by INLA comes from two factors: the very large size and high dependency. The original INLA cannot work in an extremely large size dataset such as ours. A blocking technique is introduced to make INLA feasible; however in doing so the blocking technique prevented us from managing stationary systematic parameters and from keeping the dependency between neighbours since the parameters were differently estimated per each block. Therefore we eventually decided on using Gibbs sampling. To reduce the information loss by the blocking technique, a heuristic algorithm is used to calculate the marginalized precision matrix. We also reduce the information loss by enlarging the size of the block since the precision matrix $\mathbf{Q}$ is sparse matrix. In addition, we can also employ the well-known library *GMRFLib* to update blocks in the Gibbs scheme (Rue & Held, 2005).

6 Conclusion

With a small number of projections, Projection Reconstruction NMR (PR-NMR) can reconstruct multi-dimensional NMR spectra efficiently. In this paper, Blocked Gibbs sampler is applied to reconstruct discrete NMR spectra from a small number of projections. The blocked Gibbs sampler reconstruction algorithm with reasonable heuristic approximation is efficient in reconstructing multi-dimensional NMR spectra compared to other pixel-by-pixel based approaches.

Acknowledgement

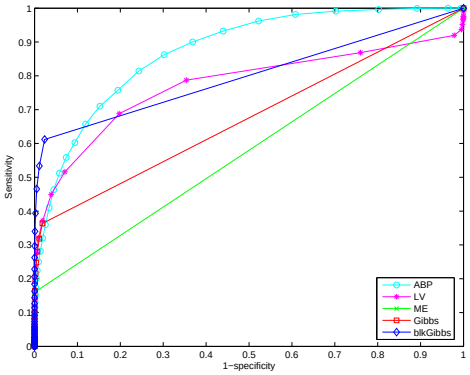
This work is supported by STATICA project which is funded by the Principal Investigator programme of Science Foundation Ireland, Grant number $08/IN.1/I1879$. In addition, the first author would like to thank Prof. R. Freeman who introduced PR-NMR.

References

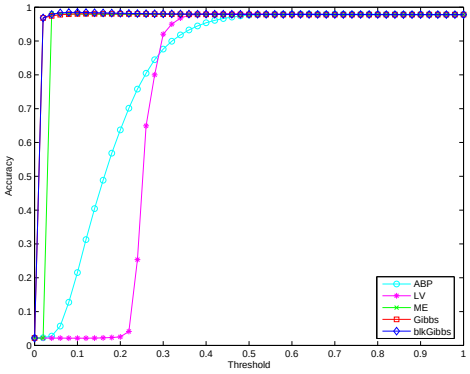
Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of Royal Statistical Society B*, 36, 192–236.

Besag, J., York, J., & Mollie, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1), 1–20.

Bodenhausen, G. & Ernst, R. R. (1982). Direct Determination of Rate Constants of Slow Dynamic Processes by Two-Dimensional Accordion SpectroscopyinNucle-



(a) ROC



(b) Accuracy

Figure 7: ROC and Accuracy of reconstructed spectra for experimental dataset

arMagneticResonance. *Journal of American Chemical Society*, 104, 1304–1309.

Felli, I. & Brutscher, B. (2009). Recent advances in solution nmr: Fast methods and heteronuclear direct detection. *ChemPhysChem*, 10, 1356–1368.

Freeman, R. & Kupce, E. (2004). Distant Echoes of the Accordion - Reduced Dimensionality, GFT-NMR, and Projection-Reconstruction of Multidimensional Spectra. *Concepts in Magnetic Resonance Part A*, 23A.

Hanson, K. M. & Wecksung, G. W. (1983). Bayesian approach to limited-angle reconstruction in computed tomography. *Journal of the Optical Society of America*, 73, 1501–1509.

Hunt, B. R. (1977). Bayesian methods in nonlinear digital image restoration. *IEEE Transactions on Computers*, 26(3), 219–229.

Kim, S. & Szyperski, T. (2003). GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR Spectral information. *Journal of American Chemistry Society*, 125, 1385–1393.

Kupce, E. & Freeman, R. (2003a). Projection-Reconstruction of Three-Dimensional NMR Spectra. *Journal of Americal Chemistry Society*, 125(46), 13958–13959.

Kupce, E. & Freeman, R. (2003b). The Radon Transform: A new Scheme for Fast Multidimensional NMR. *Concepts in Magnetic Resonance Part A*, 22A, 4–11.

Kupce, E. & Freeman, R. (2004a). Fast Multidimensional NMR Spectroscopy by the Projection-Reconstruction Technique. *Spectroscopy*.

Kupce, E. & Freeman, R. (2004b). Projection-Reconstruction Technique for Speeding up Multidimensional NMR Spectroscopy. *Journal of American Chemical Society*, 126, 6429–6440.

Mardia, K. V. (1988). Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *J. Multivar. Anal.*, 24(2), 265–284.

Mishkovsky, M. & Frydman, L. (2009). Principles and progress in ultrafast multidimensional nuclear magnetic resonance. *Ann Rev Phys Chem*, 60, 429–448.

Mobli, M., Stern, A. S., & Hoch, J. C. (2006). Spectral reconstruction methods in fast NMR: Reduced dimensionality, random sampling and maximum entropy. *Journal of Magnetic Resonance*, 182(1), 96–105.

Mok, K., Kuhn, L. T., Goez, M., Day, I. J., Lin, J., Andersen, N. H., & Hore, P. J. (2007). A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature*, 447, 106–109.

Reis, M. L. & Roberty, N. C. (1992). Maximum entropy algorithms for image reconstruction from projections. *IOP publishing Ltd*, Inverse Problems 8(4), 623–644.

Rue, H. & Held, L. (2005). *Gaussian Markov Random Field: Theory and Applications*. Chapman & Hall/CRC.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of The Royal Statistical Society Series B*, 71(2), 319–392.

Simon, B. & Sattler, M. (2004). Speeding Up Biomolecular NMR Spectroscopy. *Angew. Chem. Int. Ed.*, 43, 782–786.

Therrien, C. (1993). An approximate method of evaluating the joint likelihood for first-order gmrfs. *IEEE Transactions on Image Processing*, 2(4), 520–523.

Weir, I. S. (1997). Fully bayesian reconstruction from single-photon emission computed tomography data. *Journal of the Ameerican Statistical Association*, 92(437), 49–60.

West, M. & Harrison, J. (1997). *Bayesian forecasting and dynamic models (2nd ed.)*. New York, NY, USA: Springer-Verlag New York, Inc.

Yoon, J. & Fitzgerald, W. J. (2009). Multi-dimensional NMR spectra reconstruction using Gibbs sampler. *Computers in Biology and Medicine*. (submitted).

Yoon, J. & Godsill, S. J. (2006). Bayesian Inference for Multidimensional NMR image reconstruction. In *European Signal Processing Conference (EUSIPCO 2006)* Florence, Italy.

Yoon, J., Godsill, S. J., Kupce, E., & Freeman, R. (2006). Deterministic and statistical methods for reconstructing multidimensional NMR spectra. *Magnetic Resonance in Chemistry*, 44(3), 197–209.

Appendix

A Conditional Prior for a blocked area, $\mathbf{x}_A$

The conditional prior for a blocked (chopped) area is defined by

$$p(\mathbf{x}_A|\mathbf{x}_{-A},\theta) \propto$$
$$\exp\left\{-\tfrac{\tau}{2}\mathbf{f}_A^T\mathbf{Q}_{f_A}\mathbf{f}_A - \tfrac{\rho}{2}\left(\eta_A - \mathbf{z}_A^T\beta - \mathbf{f}_A\right)^T\left(\eta_A - \mathbf{z}_A^T\beta - \mathbf{f}_A\right)\right\}$$
$$= \exp\left\{-\tfrac{1}{2}(\mathbf{x}_A-\mu_A)^T\mathbf{Q}_A(\mathbf{x}_A-\mu_A)\right\} \text{ where } \mu_A = \begin{bmatrix}\mathbf{0}\\ -\frac{\mathbf{z}_A^T\beta}{\rho}\end{bmatrix} \text{ and } \mathbf{Q}_A = \begin{bmatrix}\rho\mathbf{I}_{K\times K} & -\rho\mathbf{I}_{K\times K}\\ -\rho\mathbf{I}_{K\times K} & \rho\mathbf{I}_{K\times K}+\tau\mathbf{Q}_{f_A}\end{bmatrix}.$$

B Gibbs update for θ

B.1 Sampling β from the conditionals

The posterior of the regression parameter β has the following analytical form and we can draw sample from the distribution.

$$\begin{aligned}\pi(\beta) &= p(\beta|\mathbf{x},\mathbf{y},\theta_{-\beta}) \propto p(\mathbf{x}|\theta)p(\beta)\\ &\propto \exp\left\{-\frac{\rho}{2}(\eta-\mathbf{z}^T\beta-\mathbf{f})^T(\eta-\mathbf{z}^T\beta-\mathbf{f})\right.\\ &\quad \left.-\frac{1}{2}\beta^T\mathbf{Q}_\beta\beta\right\} = \mathcal{N}(\cdot;\mu_\beta,Q_\beta^{*-1})\end{aligned}$$

where $Q_\beta^* = \rho\mathbf{z}\mathbf{z}^T + \mathbf{Q}_\beta$ and $\mu_\beta = Q_\beta^{*-1}\rho(\eta-\mathbf{f})^T\mathbf{z}^T$.

B.2 Sampling $\varphi_r^{(i)}$ from the conditionals

The posterior of the scaling factor $\varphi_r^{(i)}$ also has the following analytical equation and we can generate samples from the distribution:

$$\begin{aligned}\log\pi(\varphi) &= \log p(\varphi|\mathbf{x},\mathbf{y},\theta_{-\varphi}) \propto \log p(\mathbf{y}|\mathbf{x},\theta) + \log p(\varphi)\\ &\propto -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta\right)^T\kappa_r\mathbf{I}\left(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta\right)\\ &= -\frac{1}{2}\sum_{r=1}^{|\Lambda|}\left(\varphi_r-(\mathbf{G}_r\eta)^{-1}\mathbf{y}^{(r)}\right)^T\\ &\quad\times\kappa_r\eta^T\mathbf{G}_r^T\mathbf{G}_r\eta\left(\varphi_r-(\mathbf{G}_r\eta)^{-1}\mathbf{y}^{(r)}\right)\\ &= -\frac{1}{2}(\varphi-\mu_\varphi)^T\mathbf{Q}_\varphi(\varphi-\mu_\varphi)\end{aligned}$$

where $\mathbf{Q}_\varphi = \mathrm{diag}(\kappa_1\eta^T\mathbf{G}_1^T\mathbf{G}_1\eta,\cdots,\kappa_{|\Lambda|}\eta^T\mathbf{G}_{|\Lambda|}^T\mathbf{G}_{|\Lambda|}\eta)$ and $\mu_\varphi = \left[(\mathbf{G}_1\eta)^{-1}\mathbf{y}^{(1)},\cdots,(\mathbf{G}_{|\Lambda|}\eta)^{-1}\mathbf{y}^{(|\Lambda|)}\right]^T$ where we assume the prior distribution is uniform distribution.

B.3 Sampling κ_r from the conditionals

The posterior of the noise precision for each projection, κ_r also has the following analytical equation and we can generate samples from the distribution:

$$\begin{aligned}\pi(\kappa) &= p(\kappa|\mathbf{x},\mathbf{y},\theta_{-\kappa}) = p(\mathbf{y}|\mathbf{x},\theta)p(\kappa)\\ &\propto \prod_{r=1}^{|\Lambda|}\kappa_r^{m/2}\\ &\quad\times\exp\left\{-\frac{\kappa_r}{2}\left(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta\right)^T\left(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta\right)\right\}\\ &\quad\times\kappa_r^{\alpha_\kappa-1}\exp(-\kappa_r/\beta_\kappa)\\ &= \prod_{r=1}^{|\Lambda|}\kappa_r^{\alpha_\kappa+m/2-1}\exp\left\{-\kappa_r\right.\\ &\quad\left.\times\left[\frac{(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta)^T(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta)}{2}+\frac{1}{\beta_\kappa}\right]\right\}\\ &= \mathcal{G}(\cdot;[\alpha_1^*,\cdots,\alpha_{|\Lambda|}^*]^T,[\beta_1^*,\cdots,\beta_{|\Lambda|}^*]^T)\end{aligned}$$

where

$$\begin{aligned}\alpha_r^* &= \alpha_\kappa + m/2\\ \beta_r^* &= \left[\frac{\left(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta\right)^T\left(\mathbf{y}^{(r)}-\varphi_r\mathbf{G}_r\eta\right)}{2}+\frac{1}{\beta_\kappa}\right]^{-1}\end{aligned}$$

B.4 Sampling ρ from the conditionals

The posterior of the precision for Gaussian Markov Random Field, ρ also has the following analytical equation:

$$\pi(\rho) = p(\rho|\mathbf{x},\mathbf{y},\theta_{-\rho}) = \mathcal{G}(\cdot;\alpha_\rho^*,\beta_\rho^*)$$

where $\alpha_\rho^* = \alpha_\rho + N/2$ and

$$\beta_\rho^* = \left[\frac{(\eta-\mathbf{f}-\mathbf{z}^T\beta)^T(\eta-\mathbf{f}-\mathbf{z}^T\beta)}{2}+\frac{1}{\beta_\rho}\right]^{-1}.$$

B.5 Sampling τ from the conditionals

The posterior of the noise precision for reconstructed spectra, τ also has the following analytical equation:

$$\pi(\tau) = p(\tau|\mathbf{x},\mathbf{y},\theta_{-\tau}) = \mathcal{G}(\cdot;\alpha_\tau^*,\beta_\tau^*)$$

where $\alpha_\tau^* = \alpha_\tau + N/2$ and $\beta_\tau^* = \left[\frac{\mathbf{f}^T\mathbf{Q}_f\mathbf{f}}{2}+\frac{1}{\beta_\tau}\right]^{-1}$. Here $\mathbf{f}^T\mathbf{Q}_f\mathbf{f}$ is calculated easily without building $\mathbf{Q}$ by using the characteristics of the intrinsic GMRF: $\mathbf{f}^T\mathbf{Q}_f\mathbf{f} = \sum_{i=1}^{n-1}\sum_{i=1}^{n-1}\Delta_{ij}^2$ according to Eq. (2).