Journal of Computational Physics 380 (2019) 1–28

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Localization for MCMC: sampling high-dimensional posterior distributions with local structure

M. Morzfeld *, X.T. Tong, Y.M. Marzouk

ARTICLE INFO

Article history:
Received 29 May 2018
Received in revised form 10 November 2018
Accepted 7 December 2018
Available online 3 January 2019

Keywords:
Markov chain Monte Carlo
Bayesian inverse problems
High dimensions
Localization
Dimension-independent convergence

ABSTRACT

We investigate how ideas from covariance localization in numerical weather prediction can be used in Markov chain Monte Carlo (MCMC) sampling of high-dimensional posterior distributions arising in Bayesian inverse problems. To localize an inverse problem is to enforce an anticipated "local" structure by (*i*) neglecting small off-diagonal elements of the prior precision and covariance matrices; and (*ii*) restricting the influence of observations to their neighborhood. For linear problems we can specify the conditions under which posterior moments of the localized problem are close to those of the original problem. We explain physical interpretations of our assumptions about local structure and discuss the notion of high dimensionality in local problems, which is different from the usual notion of high dimensionality in function space MCMC. The Gibbs sampler is a natural choice of MCMC algorithm for localized inverse problems and we demonstrate that its convergence rate is independent of dimension for localized linear problems. Nonlinear problems can also be tackled efficiently by localization and, as a simple illustration of these ideas, we present a localized Metropolis-within-Gibbs sampler. Several linear and nonlinear numerical examples illustrate localization in the context of MCMC samplers for inverse problems.

1. Introduction

We consider inverse problems in the Bayesian setting. Let $\mathbf{x}$ be an n-dimensional real-valued random vector. The observations are defined by

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) + \mathbf{v},$$

where $\mathbf{y}$ is a k-dimensional vector, $\mathbf{h}$ is a given function, and $\mathbf{v}$ is a random variable with known distribution. The observations $\mathbf{y}$, along with the distribution of $\mathbf{v}$, define a likelihood $p_l(\mathbf{y}|\mathbf{x})$. A prior probability density p_0 describes prior knowledge about $\mathbf{x}$. For example, one may know that the variables are likely to be within a certain interval. The prior distribution often also describes the smoothness of a random field whose discretization is the vector $\mathbf{x}$. The prior and likelihood together define the posterior density:

$$p(\mathbf{x}|\mathbf{y}) \propto p_0(\mathbf{x}) p_l(\mathbf{y}|\mathbf{x}).$$

* Corresponding author.
E-mail address: mmo@math.arizona.edu (M. Morzfeld).

https://doi.org/10.1016/j.jcp.2018.12.008
0021-9991/

Throughout this paper, we assume Gaussian errors, $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ and Gaussian priors $p_0(\mathbf{x}) = \mathcal{N}(\mathbf{m}, \mathbf{C})$, as is common; see, e.g., [58]. In addition, we assume that

(i) the state dimension n is large and the number of observations k is also large, i.e., $k = O(n)$;
(ii) the prior covariance and precision matrices are nearly banded;
(iii) each predicted observation $[\mathbf{h}(\mathbf{x})]_j$ has significant dependence on only $\ell \ll n$ components of $\mathbf{x}$ (i.e., the contribution from the remaining components of $\mathbf{x}$ is small), and $\mathbf{R}$ is diagonal.

In simple terms, nearly banded means that the elements away from the diagonal are small, but we make the meaning of nearly banded and significant dependence more precise below. We call problems that satisfy the above assumptions "local."

In numerical weather prediction (NWP) and ensemble Kalman filtering (EnKF), local problems arise frequently. During a typical EnKF step, the covariance of 10^8 variables needs to be estimated accurately, but the number of samples used is usually 100 or less. This seemingly impossible task is made possible by "localization" [28,30,31]. During localization, a nearly banded forecast covariance matrix is transformed into an (exactly) banded matrix by setting small off-diagonal elements to zero, e.g., by multiplying each entry of the covariance with a suitable "localization function" [22]. In addition, the influence of each observation is restricted to its neighborhood. Practitioners agree that localization is a key requirement for making EnKF applicable to large-scale NWP problems. Note that localization trades numerical efficiency against errors which can be controlled: problems with banded forecast covariance structure and local observations are more easily solved, and the errors introduced by localization are controllable, e.g., they vanish when localization thresholds are sufficiently small. From a more theoretical perspective, it was shown in [6] that one can estimate a covariance matrix of bandwidth l with $O(l + \log n)$ samples. Reference [61] explains how this result is used in the context of EnKF.

This paper examines localization in the context of Bayesian inverse problems and Markov chain Monte Carlo (MCMC) samplers for the associated posterior distributions. It is known that sampling *generic* high dimensional (posterior) distributions is challenging; see, e.g., [1,51]. We suggest, however, that one can design relatively simple MCMC algorithms to sample high dimensional posterior distributions of *localized* inverse problems efficiently. Specifically, we discuss the following three questions:

(a) Is the localized problem near the local problem (Section 2)?
(b) Can one solve a localized problem efficiently by MCMC (Section 3)?
(c) Are local inverse problems of practical importance (Section 4)?

In Section 4 we also explain that the notion of high dimensionality of a local inverse problem is different from what is usually considered in the MCMC literature [8,12–14,17,45,57]. Numerical illustrations are provided in Section 5. We summarize our conclusions in Section 6.

2. Localization of inverse problems

To localize an inverse problem means to enforce that prior interactions (correlations and/or conditional dependencies) and the effects of observations are confined to a neighborhood. For local problems, in the sense of Assumptions (*i*)–(*iii*) in Section 1, one may thus expect that errors introduced by localization are small. We prove this intuitive result for linear and Gaussian problems, and then discuss how localization can be used in nonlinear problems. Note that localization as described here can also be interpreted in the context of a more general robust Bayesian analysis [32], which examines how perturbations to the prior and likelihood affect the posterior distribution.

2.1. Localization of prior covariance and precision matrices

Let $[\mathbf{C}]_{i,j}$ be the i, j entry of a covariance matrix $\mathbf{C}$. Suppose that $|[\mathbf{C}]_{i,j}| \ll |[\mathbf{C}]_{i,i}|$ for $|i - j| \geq l$. During localization, these "small" off-diagonal elements are set to zero. The resulting localized covariance matrix $\mathbf{C}_{\text{loc}}$ has entries

$$[\mathbf{C}_{\text{loc}}]_{i,j} = [\mathbf{C}]_{i,j} \mathbb{1}_{|i-j|\leq l}, \tag{1}$$

where $\mathbb{1}_{|i-j|\leq l}$ is an indicator function. We define the bandwidth l of a $m \times m$ matrix $\mathbf{A}$ by

$$l = \min\{r : [\mathbf{A}]_{i,j} = 0 \text{ if } |i - j| > r\}.$$

With these definitions and notation, it becomes clear that localization turns the prior covariance matrix $\mathbf{C}$ with small off-diagonal elements into a banded matrix $\mathbf{C}_{\text{loc}}$, whose bandwidth is less or equal to the threshold l used during localization. Moreover, the localized prior covariance matrix $\mathbf{C}_{\text{loc}}$ is positive definite if the minimum eigenvalue of $\mathbf{C}$ is above δ_C, where

$$\delta_C := \max_i \sum_{j:|i-j|>l} |[\mathbf{C}]_{i,j}|. \tag{2}$$

We show in Appendix A (Proposition A.2) that

$$\|\mathbf{C} - \mathbf{C}_{\text{loc}}\| \leq \delta_C, \tag{3}$$

i.e., the localized prior covariance matrix is a small perturbation of the prior covariance matrix. Note that, throughout this paper, we use $\|\cdot\|$ to denote the l_2 norm for a vector $\mathbf{x}$ with elements $x_1, \ldots, x_n$, $\|\mathbf{x}\| = \sqrt{\sum_{j=1}^n x_j^2}$, as well as the l_2 operator norm for a matrix $\mathbf{A}$, $\|\mathbf{A}\| = \sup_{v\in\mathbb{R}^n, \|v\|=1} \|\mathbf{A}\mathbf{v}\|$.

Working with prior precision matrices, rather than prior covariance matrices, is sometimes more natural. In this case, one can write the prior distribution as $p_0(\mathbf{x}) = \mathcal{N}(\mathbf{m}, \mathbf{\Omega}^{-1})$, where $\mathbf{\Omega}$ is the prior precision matrix. Precision matrices can be localized in the same way as covariance matrices. We assume, as above, that $|[\mathbf{\Omega}]_{i,j}| \ll |[\mathbf{\Omega}]_{i,i}|$ for $|i-j| \geq l$ and localize the prior precision matrix by setting small off-diagonal elements equal to zero. The result is a localized prior precision matrix with entries

$$[\mathbf{\Omega}_{\text{loc}}]_{i,j} = [\mathbf{\Omega}]_{i,j} \mathbb{1}_{|i-j|\leq l}.$$

Under our assumptions, the localized precision matrix is a small perturbation of the precision matrix:

$$\|\mathbf{\Omega} - \mathbf{\Omega}_{\text{loc}}\| \leq \delta_\Omega, \quad \delta_\Omega := \max_i \sum_{j:|i-j|>l} |[\mathbf{\Omega}]_{i,j}|. \tag{4}$$

Localization of the covariance or precision matrices results in localized prior distributions $p_{0,\text{loc}}(\mathbf{x}) = \mathcal{N}(\mathbf{m}, \mathbf{C}_{\text{loc}})$ or $p_{0,\text{loc}} = \mathcal{N}(\mathbf{m}, \mathbf{\Omega}_{\text{loc}}^{-1})$.

2.2. Localization of the observation matrix

We first consider the case $\mathbf{h}(\mathbf{x}) = \mathbf{H}\mathbf{x}$, where $\mathbf{H}$ is a given $k \times n$ matrix. Assumption (ii) in Section 1 implies that each observation $[\mathbf{H}\mathbf{x}]_j$ may depend on all components of $\mathbf{x}$, but the contributions of many components are negligible. In this case, the observation matrix $\mathbf{H}$ can be localized similarly to how we localized the prior covariance.

For a given observation matrix $\mathbf{H}$ and a given threshold l_H, define a localized observation matrix by

$$[\mathbf{H}_{\text{loc}}]_{j,i} = [\mathbf{H}]_{j,i} \mathbb{1}_{|o_j - i|\leq l_H},$$

where o_j represents the "center" of the j-th observation. Following (2), we can quantify the difference between $\mathbf{H}$ and $\mathbf{H}_{\text{loc}}$ by

$$\delta_H = \max_i \left\{ \sum_{j:|i-o_j|>l_H} |[\mathbf{H}]_{j,i}|, \sum_{i:|i-o_j|>l_H} |[\mathbf{H}]_{j,i}| \right\}. \tag{5}$$

As before, if δ_H is small, then errors due to the localization are expected to be small.

2.3. Localized posterior distributions of linear-Gaussian inverse problems

We continue to assume that $\mathbf{h}(\mathbf{x}) = \mathbf{H}\mathbf{x}$, so that the true (original) posterior distribution is a Gaussian with mean and covariance given by

$$\begin{aligned}
\hat{\mathbf{m}} &= \mathbf{m} + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}), \\
\widehat{\mathbf{C}} &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{C}, \\
\mathbf{K} &= \mathbf{C}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{C}\mathbf{H}^T)^{-1},
\end{aligned}$$

where $\mathbf{K}$ is the Kalman gain.

Localization of the observation matrix leads to the localized likelihood $p_{\text{loc},l}(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{H}_{\text{loc}}\mathbf{x}, \mathbf{R})$. If we also localize the prior covariance matrix, we obtain the localized prior $p_{0,\text{loc}}(\mathbf{x}) = \mathcal{N}(\mathbf{m}, \mathbf{C}_{\text{loc}})$. The localized prior and likelihood then define the *localized posterior* $p_{\text{loc}}(\mathbf{x}|\mathbf{y}) \propto p_{\text{loc},l}(\mathbf{y}|\mathbf{x}) p_{\text{loc},0}(\mathbf{x})$, whose mean and covariance are given by

$$\begin{aligned}
\hat{\mathbf{m}}_{\text{loc}} &= \mathbf{m} + \mathbf{K}_{\text{loc}}(\mathbf{y} - \mathbf{H}_{\text{loc}}\mathbf{x}), \\
\widehat{\mathbf{C}}_{\text{loc}} &= (\mathbf{I} - \mathbf{K}_{\text{loc}}\mathbf{H}_{\text{loc}})\mathbf{C}_{\text{loc}}, \\
\mathbf{K}_{\text{loc}} &= \mathbf{C}_{\text{loc}}\mathbf{H}_{\text{loc}}^T(\mathbf{R} + \mathbf{H}_{\text{loc}}\mathbf{C}_{\text{loc}}\mathbf{H}_{\text{loc}}^T)^{-1}.
\end{aligned}$$

We prove in Appendix A (Proposition A.2) that the means and covariance matrices of the localized and original posterior distributions satisfy

$$\|\hat{\mathbf{m}} - \hat{\mathbf{m}}_{\text{loc}}\| \leq (\delta_C + \delta_H) \cdot D_1(\|\mathbf{R}^{-1}\|, \|\widehat{\mathbf{C}}\|, \|\mathbf{C}^{-1}\|) \cdot (\|\mathbf{m}\| + \|\mathbf{y}\|),$$
$$\|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\text{loc}}\| \leq (\delta_C + \delta_H) \cdot D_2(\|\mathbf{R}^{-1}\|, \|\widehat{\mathbf{C}}\|, \|\mathbf{C}^{-1}\|),$$

where the functions D_1 and D_2 are defined in Proposition A.2. Similarly, if we work with the prior precision matrix, we obtain, after localization, the prior $p_{0,\text{loc}}(\mathbf{x}) = \mathcal{N}(\mathbf{m}, \mathbf{\Omega}_{\text{loc}}^{-1})$, which leads to a localized posterior whose mean and covariance are given by

$$\hat{\mathbf{m}}_{\text{p,loc}} = \mathbf{m} + \mathbf{K}_{\text{p,loc}} (\mathbf{y} - \mathbf{H}_{\text{loc}}\mathbf{x}), \tag{6}$$
$$\widehat{\mathbf{C}}_{\text{p,loc}} = (\mathbf{I} - \mathbf{K}_{\text{p,loc}}\mathbf{H}_{\text{loc}})\mathbf{\Omega}_{\text{loc}}^{-1}, \tag{7}$$
$$\mathbf{K}_{\text{p,loc}} = \mathbf{\Omega}_{\text{loc}}^{-1}\mathbf{H}_{\text{loc}}^T(\mathbf{R} + \mathbf{H}_{\text{loc}}\mathbf{\Omega}_{\text{loc}}^{-1}\mathbf{H}_{\text{loc}}^T)^{-1}. \tag{8}$$

We find in Proposition A.3 that:

$$\|\hat{\mathbf{m}} - \hat{\mathbf{m}}_{\text{p,loc}}\| \leq (\delta_\Omega + \delta_H) \cdot D_3(\|\mathbf{R}^{-1}\|, \|\widehat{\mathbf{C}}\|, \|\mathbf{\Omega}\|) \cdot (\|\mathbf{m}\| + \|\mathbf{y}\|),$$
$$\|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\text{p,loc}}\| \leq (\delta_\Omega + \delta_H) \cdot D_4(\|\mathbf{R}^{-1}\|, \|\widehat{\mathbf{C}}\|, \|\mathbf{\Omega}\|),$$

where D_3 and D_4 are defined in Proposition A.3.

In summary, the difference between the localized and unlocalized posterior distributions depends on the localization thresholds we chose, on the structure of the prior covariance or precision matrix, as well as on the observation matrix $\mathbf{H}$ and its localization. The bandwidth of localizations of $\mathbf{C}$ or $\mathbf{\Omega}$ and $\mathbf{H}$ should therefore be tuned to obtain small differences between the localized and unlocalized problems.

We use a threshold for localization because it makes our proofs simpler. In practice, one may localize more effectively using suitable localization functions [22], which set small off-diagonal elements to zero smoothly and which can preserve positive-definiteness during localization. Moreover, we only consider (nearly) banded covariance matrices, which arise, for instance, in linear-Gaussian Bayesian inverse problems on one-dimensional spatial domains. The conceptual ideas of localization, however, can be used in problems with 2D spatial domains (see Section 5 for a numerical example). In fact, we anticipate that many of our ideas can be adapted to matrices with more general sparsity patterns but defer this investigation to future work.

2.4. Localization of nonlinear problems

In a nonlinear inverse problem, the function $\mathbf{h}$ is not linear, and thus the posterior distribution is non-Gaussian even when the prior distribution is Gaussian. However, the function $\mathbf{h}$ is usually well understood because it is the result of a careful modeling effort. Thus, it is not unreasonable to assume that it is known whether $\mathbf{h}$ is local in the sense of Assumption (*iii*) in Section 1. If $\mathbf{h}$ is local, then we can localize it by neglecting the some of the components of $\mathbf{x}$ when computing the components, $[\mathbf{h}]_i$, of $\mathbf{h}$ (see Section 3.2.3 for more detail on how to do this). The localized $\mathbf{h}$, along with the localized Gaussian prior, defines the localized posterior distribution of a nonlinear inverse problem. The localized posterior distribution may be close to the posterior distribution of the unlocalized problem, but we do not prove this statement in the general, nonlinear setting. The lack of theoretical results, however, does not prevent us from *using* localization in nonlinear problems, and we present such an example in Section 5. Moreover, more than a decade of experience with using localization in EnKF and NWP can also be viewed as numerical and empirical evidence that localization is indeed applicable in nonlinear problems (see also [40]).

3. MCMC for local inverse problems

MCMC is often used for the numerical solution of Bayesian inverse problems. To illustrate the behavior of some MCMC algorithms on localized problems, we first consider the extreme case of a linear problem with *diagonal* covariance and precision matrix and with a diagonal observation function $\mathbf{h}(\mathbf{x}) = \mathbf{x}$. Specifically, suppose the target distribution is the n-dimensional Gaussian distribution $p(x) = \mathcal{N}(0, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix of dimension n. Suppose that the current state of the Markov chain is $\mathbf{x}^k$. The Metropolis–Hastings (MH) algorithm proposes a move to $\mathbf{x}'$ by drawing from a proposal distribution $q(\mathbf{x}'|\mathbf{x}^k)$, and accepts or rejects the move with probability

$$a^{k+1} = \min\left\{1, \frac{p(\mathbf{x}')q(\mathbf{x}^k|\mathbf{x}')}{p(\mathbf{x}^k)q(\mathbf{x}'|\mathbf{x}^k)}\right\};$$

see, e.g., [33,38,43]. Averages over the samples generated in this way converge to expected values with respect to the target distribution p as $k \to \infty$. A question of practical importance is: how many samples are needed to accurately estimate expectations with respect to the target distribution? The answer depends on the proposal distribution. For a Gaussian proposal distribution $q(\mathbf{x}^{k+1}|\mathbf{x}^k) = \mathcal{N}(\mathbf{x}^k, \sigma^2\mathbf{I})$, which produces a so-called random walk Metropolis (RWM) chain, one must

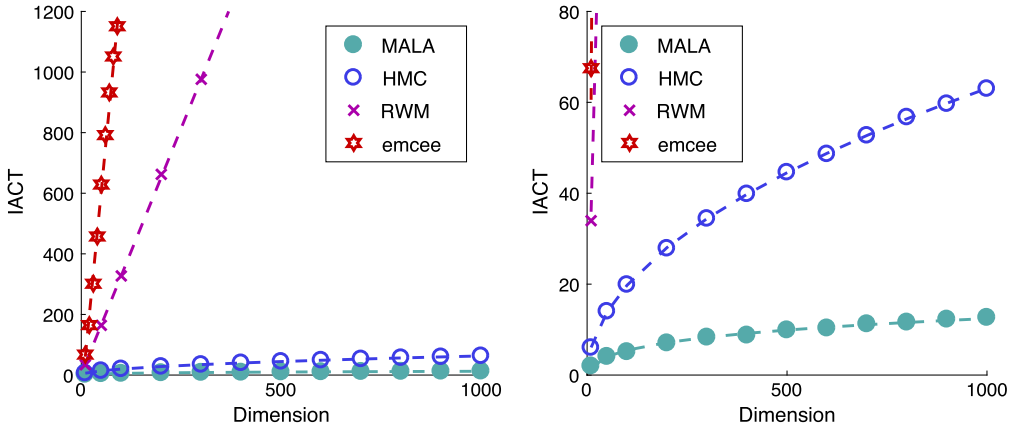

Fig. 1. IACT as a function of dimension for various MCMC samplers applied to an isotropic Gaussian. Dots represent IACT, averaged over all n variables, for emcee (red), RWM (purple), Hamiltonian MCMC (blue), and MALA (teal). The dashed lines represent linear fits (emcee and RWM), fits to a square root (Hamiltonian MCMC), and a fit to a 1/3-degree polynomial (MALA). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

choose a proposal variance σ^2 such that the acceptance probability is reasonably large while, at the same time, the accepted MCMC moves are large enough to explore the space appropriately. An optimal choice that achieves this trade-off for RWM is $\sigma^2 = O(n^{-1})$, where n is the dimension of the problem; see, e.g., [5,51,52]. Optimal scalings of the proposal variance with problem dimension are also known for other MCMC algorithms. For example, the proposal distribution of the Metropolis-adjusted Langevin algorithm (MALA) is defined by

$$\mathbf{x}' = \mathbf{x}^k + \frac{\sigma^2}{2}\nabla \log p(\mathbf{x}^k) + \sigma\,\xi$$

where ∇ denotes a gradient and ξ is a vector of n standard normal variates. An optimal choice is $\sigma = O(n^{-1/3})$. For Hamiltonian Monte Carlo (see, e.g., [16,41]) an optimal step size is $\sigma = O(n^{-1/4})$ [4] (here and below we refer to σ as a step size and to σ^2 as the proposal variance).

Setting aside issues of transient behavior [11], the efficiency of an MCMC algorithm can be assessed by computing the integrated auto-correlation time (IACT). Throughout this paper we use the definitions and numerical approximations of IACT discussed in [63]. Heuristically, the number of effective samples is the number of samples divided by IACT; see, e.g., [33,38]. In Fig. 1 we compute IACT for various MCMC algorithms applied to the n-dimensional isotropic Gaussian, as a function of n. We observe that IACT grows with dimension for all algorithms we consider, but at different rates. For both RWM and an affine invariant sampler [18,26] called emcee (or the MCMC Hammer), IACT grows linearly with the dimension n; for Hamiltonian MCMC, we observe that IACT grows with the square root of n, while for MALA, IACT grows as $n^{1/4}$. Similar tests were done for the "t-walk" [10], a general purpose ensemble sampler. The numerical results of [10] suggest a linear scaling of t-walk's IACT with dimension.

However, sampling an isotropic Gaussian is trivial, and MCMC should be independent of dimension for this problem because it can be decomposed into n independent sub-problems (see also [40,49]). An MCMC sampler that naturally makes use of local problem structure is the Gibbs sampler (sometimes Gibbs samplers are also called "heat bath" or "partial resampling," and these can be viewed as examples of "single-component Metropolis algorithms," see, e.g., [23]). The basic Gibbs sampler is as follows. Let the target density be $p(\mathbf{x})$, where $\mathbf{x}$ is shorthand for the vector with n elements $x_1, x_2, \ldots, x_n$. Set $j = 1$, and let $\mathbf{x}^k$ be the current state of a Markov chain. Then a Gibbs sampler proceeds as follows:

(i) Set $k \to k+1$.
(ii) Sample x_j^{k+1} from the conditional distribution $p(x_j | x_1^{k+1}, \ldots x_{j-1}^{k+1}, x_{j+1}^k, \ldots, x_n^k)$.
(iii) Repeat (ii) for all n elements x_j of $\mathbf{x}$.

Repeating this process N_e times, one obtains samples such that averages over these samples converge to expected values with respect to the target distribution p as $N_e \to \infty$.

The Gibbs sampler generates independent samples, independently of dimension, for the isotropic Gaussian (an extreme example of a local problem). Similarly, a block Gibbs sampler generates independent samples, independently of dimension, if the covariance or precision matrices are block-diagonal. This suggests that MCMC based on Gibbs samplers may be more effective *for local problems* than the MCMC algorithms we considered above. In this section, we investigate this idea in more detail and study convergence rates of Gibbs samplers for Gaussian distributions with banded covariance and precision matrices.

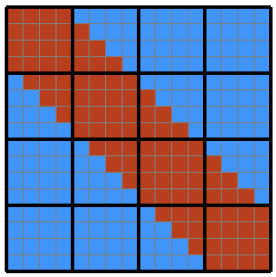

Fig. 2. Sparsity pattern of a 4-block-tridiagonal matrix with bandwidth $l = 4$. The red color indicates nonzero entries and the blue color indicates zero entries. The black squares define the $q = 4$ blocks $\mathbf{C}_{i,j}$.

Many of the results we present below may be known, but we decided to summarize what is important about Gibbs sampling for our purposes because (*i*) the Gibbs sampler and its effectiveness in local problems is essential to the understanding of how MCMC can function in high-dimensional problems with local structure; (*ii*) we could not find references on the connection between MCMC convergence rates and local problem structure, perhaps because relevant results are spread over several papers and books in different disciplines (applied mathematics, physics, and statistics) and over several decades.

3.1. Dimension independent convergence rates of Gibbs samplers

For simplicity, we assume that there are m blocks of the same size q so that $n = mq$. We divide the n elements of $\mathbf{x}$ according to the m blocks and write $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_m)$, with the understanding that each $\mathbf{x}_j$ consists of q consecutive elements in $\mathbf{x}$. A blocked Gibbs sampler uses the conditionals defined for each $\mathbf{x}_j$, i.e., at the kth step, we sample the block $\mathbf{x}_j$ using the q-dimensional conditional $p(\mathbf{x}_j|\mathbf{x}_1^{k+1}, \ldots \mathbf{x}_{j-1}^{k+1}, \mathbf{x}_{j+1}^{k}, \ldots, \mathbf{x}_m^{k})$. We call the resulting algorithm a "Gibbs sampler with block-size q" because the "standard" Gibbs sampler above is a Gibbs sampler of block-size one in this terminology.

We first consider a Gibbs sampler with block-size q for Gaussian distributions with q-block-tridiagonal covariance matrices. We say that a matrix $\mathbf{A}$ is q-block-tridiagonal, if

$$\mathbf{A}_{i,j} = \mathbf{0} \quad \text{for} \quad (i,j) \notin \{(i,i), (i,i+1), (i,i-1), i = 1, \cdots, m\},$$

where $\mathbf{A}_{i,j}$ is the (i,j)-th $q \times q$ block of $\mathbf{A}$. It is straightforward to see that a matrix with bandwidth l is l-block-tridiagonal, and a q-block-tridiagonal matrix has bandwidth less than $2q$. Fig. 2 illustrates a 4-block-tridiagonal matrix with bandwidth $l = 4$. The convergence rate of the Gibbs sampler is dimension-independent if the covariance matrix is q-block-tridiagonal, as detailed in the following theorem.

Theorem 3.1. *Suppose the Gibbs sampler with block-size q is applied to a Gaussian target distribution $p = \mathcal{N}(\mathbf{m}, \mathbf{C})$ with m blocks of size q. Suppose $\mathbf{C}$ is q-block-tridiagonal. Then the distribution of $\mathbf{x}^k$ converges to p geometrically fast in all coordinates, and we can couple $\mathbf{x}^k$ and a sample $\mathbf{z} \sim \mathcal{N}(\mathbf{m}, \mathbf{C})$ such that*

$$\mathbb{E}\|\mathbf{C}^{-1/2}(\mathbf{x}^k - \mathbf{z})\|^2 \le \beta^k n(1 + \|\mathbf{C}^{-1/2}(\mathbf{x}^0 - \mathbf{m})\|^2), \tag{9}$$

where

$$\beta \le \frac{2(1-\mathcal{C}^{-1})^2\mathcal{C}^4}{1 + 2(1-\mathcal{C}^{-1})^2\mathcal{C}^4}, \tag{10}$$

with $\mathcal{C}$ being the condition number of $\mathbf{C}$.

Similarly, the convergence rate of the Gibbs sampler is dimension-independent if the precision matrix, rather than the covariance matrix, is block-tridiagonal. One can modify Theorem 3.1 to address the case of banded precision matrices.

Theorem 3.2. *Suppose the Gibbs sampler with block-size q is applied to a Gaussian target distribution $p = \mathcal{N}(\mathbf{m}, \mathbf{\Omega}^{-1})$ with m blocks of size q. Suppose $\mathbf{\Omega}$ is q-block-tridiagonal. Then the distribution of $\mathbf{x}^k$ converges to p geometrically fast in all coordinates, and we can couple $\mathbf{x}^k$ and a sample $\mathbf{z} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Omega}^{-1})$ such that equation (9) holds with*

$$\beta \le \frac{\mathcal{C}(1-\mathcal{C}^{-1})^2}{1 + \mathcal{C}(1-\mathcal{C}^{-1})^2}, \tag{11}$$

where $\mathcal{C}$ is the condition number of $\mathbf{\Omega}$.

The proofs of Theorems 3.1 and 3.2 can be found in Appendix A.

While upper bounds for the rate of convergence β are independent of the dimension, the upper bounds themselves are linear in n and involve a norm of the initial condition, which may also scale linearly in n. This does not cause practical difficulties because the dimension independent scaling of the convergence *rate* implies that the number of iterations required to reach a given error level scales logarithmically in n, which is essentially a constant in practice (even when n is large). For example, suppose one wants that the l^2 error $\mathbb{E}\|\mathbf{C}^{-1/2}(\mathbf{x}^k - \mathbf{z})\|^2$ be bounded by a threshold ε. To reach this goal, the Gibbs sampler must perform k iterations, where

$$\beta^k n(1 + \|\mathbf{C}^{-1/2}(\mathbf{x}^0 - \mathbf{m})\|^2) \leq \epsilon \quad \Rightarrow \quad k \geq \frac{\log n - \log \epsilon + \log(1 + \|\mathbf{C}^{-1/2}(\mathbf{x}^0 - \mathbf{m})\|^2)}{-\log \beta}.$$

Finally, we emphasize that dimension-independent convergence for linear problems is not a new observation, and that there are multiple ways to accelerate this convergence [20,25]. Yet it is often assumed that the spectral gap of an associated linear operator, e.g., the Gauss–Seidel operator, is dimension independent. Our theoretical contribution here is to show that this dimension-independent gap indeed exists when the inverse problem is local.

3.1.1. Banded covariance matrices vs. banded precision matrices

The upper bounds on the convergence rates depend on the condition number of the covariance matrix, or, equivalently, on the condition number of the precision matrix. This means that convergence is fast, in any dimension, only if this condition number is moderate. However, if the condition number of the covariance matrix is moderate, a banded covariance matrix implies that the precision matrix can be approximated by a banded matrix, and vice versa. This equivalence is made precise in Lemma 2.1 of [7]. Thus, our assumptions and Theorems 3.1 and 3.2 describe and apply to one unified class of problems: convergence of the Gibbs sampler is fast, in any dimension, if the covariance *and* precision matrices are banded, i.e., if statistical interactions (correlations and conditional dependencies) are local.

3.1.2. Computational costs

While the upper bound of the convergence rate is independent of dimension, the actual computational cost of sampling may not be independent of dimension. The Gibbs sampler for Gaussians with banded covariance matrix requires matrix square roots and linear solves of matrices of sizes $m \times m$ and $d \times d$, where $d = n - m$. These square roots need only be computed once (not once per sample), but the cost of this computation increases with n at a rate that depends on the bandwidth of the covariance matrix. Generating a sample requires, at each block, solution of a banded linear problem, and some matrix-vector multiplications and vector-vector operations. The cost-per-sample is dominated by the linear solve, and this cost also increases with n and at a rate that depends on the bandwidth of the covariance matrix. The computational cost for one sample is thus, roughly, m times the cost of the linear solve.

If the precision matrix is banded, the conditional distributions used during Gibbs sampling simplify and it is sufficient to condition on neighboring blocks (see, e.g., equation (21) in the Appendix). The required matrix operations (square roots, linear solves) depend on the size of the blocks q, but not on the number of blocks. Assuming that $q \ll m \ll n$, the computational cost of a Gibbs sampler for a Gaussian with banded precision matrix is, roughly, m times the cost of computations with blocks of size q.

Our discussion of the Gibbs sampler so far has applied to generic (localized) Gaussian targets. Thus the Gibbs sampler can, in principle, be used to draw samples from posterior distributions of linear inverse problems. Assuming that the prior covariance or precision matrices are localized, the localized *posterior* covariance and precision matrices are also banded, because, as shown in [7], the basic arithmetic operations in (7) preserve bandedness. The convergence rate of the Gibbs sampler is independent of dimension in this case. However, linear inverse problems can be solved by a variety of other specialized MCMC techniques; see, e.g., [8,20]. We do not claim that the Gibbs sampler is necessarily a competitive computational strategy for large scale (million or more variables in $\mathbf{x}$) linear inverse problems with local structure, but we do anticipate that effective samplers can be built from a combination of localization, Gibbs sampling, acceleration, and preconditioning methods.

3.2. Metropolis-within-Gibbs for localized nonlinear inverse problems

In Section 3.1 we described the Gibbs sampler for generic Gaussian target distributions with banded covariance and precision matrices. When $\mathbf{h}(\mathbf{x})$ is nonlinear, however, the posterior distribution is not Gaussian and in general may not have tractable full conditionals, which makes the direct use of Gibbs sampling infeasible. In this section we continue to assume that the prior is Gaussian and use the Gibbs sampler to draw from this prior and then Metropolize. This simple Metropolis-within-Gibbs (MwG) sampler can handle nonlinear $\mathbf{h}$ and samples non-Gaussian posterior distributions. The sampler can be localized (l-MwG) by localizing the Gaussian prior and the likelihood.

3.2.1. Localized Metropolis-within-Gibbs sampling: banded covariance

Suppose the prior covariance matrix is block-tridiagonal (after localization), with m blocks of size q, and further suppose that $\mathbf{x}^k$ is the current state of the Markov chain. One iteration of the l-MwG sampler is as follows. Start with the first of m

blocks. Use the Gibbs proposal with block-size q (see Section 3.1) to propose a local move $\mathbf{x}'$ by drawing a sample from the localized Gaussian prior, conditioned on the current state $\mathbf{x}^k$. Accept or reject the (local) move by taking the observations into account, i.e., accept with probability

$$a = \min\left\{1, \frac{\exp\left(-0.5\left(\mathbf{y}-\mathbf{h}(\mathbf{x}')\right)^T \mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{h}(\mathbf{x}')\right)\right)}{\exp\left(-0.5\left(\mathbf{y}-\mathbf{h}(\mathbf{x})\right)^T \mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{h}(\mathbf{x})\right)\right)}\right\}. \tag{12}$$

Iterating these steps over all m blocks completes one move of l-MwG. The l-MwG sampler converges to the localized posterior distribution. This follows from the usual theory of Metropolis-within-Gibbs sampling.

3.2.2. Localized Metropolis-within-Gibbs sampling: banded precision

The l-MwG sampler can also be applied if the precision matrix, rather than the covariance matrix, is given. The sampler is as described above, but the implementation using precision matrices can be numerically more efficient. If the precision matrix has bandwidth l, then the conditional distribution of a block depends only on a few neighboring blocks, so that the computations required for drawing a sample from the conditional distributions require only matrices of size much less than n (see also Section 3.1.1, and equation (21) in the appendix). This conditional independence also implies that one can sample the block independently of other far away blocks. This provides opportunities for leveraging parallel computing to reduce overall wall-clock time.

3.2.3. Localization of the likelihood

The l-MwG sampler requires that $\mathbf{h}(\mathbf{x})$ be evaluated for each block even though only a small number of the components of $\mathbf{x}$ are changed during one of the local moves. Since we assume that $\mathbf{h}(\mathbf{x})$ is a local function, i.e., each element of $\mathbf{h}(\mathbf{x})$ depends only on a few components of $\mathbf{x}$, one may want to use the local structure of $\mathbf{h}(\mathbf{x})$ during sampling. This localization can accelerate the computations of $\mathbf{h}(\mathbf{x})$, and can also increase acceptance rates and shorten the burn-in period. Yet changing $\mathbf{h}(\mathbf{x})$ alters the likelihood and, therefore, the posterior distribution of the inverse problem. For linear $\mathbf{h}$, we showed in Section 2 that this change in the posterior distribution can be small. We now show how to localize the likelihood under the assumption that the localized posterior distribution can be written as

$$p_{\text{loc}}(\mathbf{x}|\mathbf{y}) \propto p_{0,\text{loc}}(\mathbf{x}) \exp\left(-\sum_j H_j(\mathbf{x}_{I_j}, \mathbf{y}_j)\right), \tag{13}$$

where the observations $\mathbf{y}_j$ are the observations "assigned" to the set of variables $\mathbf{x}_{I_j}$, $p_{0,\text{loc}}$ is the localized Gaussian prior distribution, and the functions H_j can be nonlinear.

Recall that the Gibbs sampler for the Gaussian prior produces a local update $\mathbf{x}'$ from $\mathbf{x}$ such that $\mathbf{x}_j = \mathbf{x}'_j$ for $j \neq i$. Localization of the likelihood means to accept the proposed local adjustment of the sample, $\mathbf{x}'$, with probability

$$\min\left\{1, \frac{\exp(-\sum_{j:i\in I_j} H_j(\mathbf{x}'_{I_j}, \mathbf{y}_j))}{\exp(-\sum_{j:i\in I_j} H_j(\mathbf{x}_{I_j}, \mathbf{y}_j))}\right\}, \tag{14}$$

i.e., we only use the observations assigned to the block we are sampling when we consider acceptance of the move.

The stationary distribution of the l-MwG sampler is the localized posterior distribution (13). To prove this statement, it suffices to check the detailed balanced relation

$$p_{\text{loc}}(\mathbf{x}|\mathbf{y})\, Q_i(\mathbf{x}, \mathbf{x}') = p_{\text{loc}}(\mathbf{x}'|\mathbf{y})\, Q_i(\mathbf{x}', \mathbf{x})$$

for $\mathbf{x} \neq \mathbf{x}'$, where Q_i is the transition density resulting from the above two steps (propose a local sample using the Gibbs proposal for the prior, then accept or reject it using the local criterion (14)). For $\mathbf{x} \neq \mathbf{x}'$, the transition density has the explicit form:

$$Q_i(\mathbf{x}, \mathbf{x}') = K_i(\mathbf{x}, \mathbf{x}') \min\left\{1, \frac{\exp(-\sum_{j:i\in I_j} H_j(\mathbf{x}'_{I_j}, \mathbf{y}_j))}{\exp(-\sum_{j:i\in I_j} H_j(\mathbf{x}_{I_j}, \mathbf{y}_j))}\right\},$$

where the transition density $K_i(\mathbf{x}, \mathbf{x}')$ is defined by the Gibbs move. The Gaussian prior $p_{0,\text{loc}}(\mathbf{x})$ is the invariant distribution of this transition, in the sense that

$$p_{0,\text{loc}}(\mathbf{x}) K_i(\mathbf{x}, \mathbf{x}') = p_{0,\text{loc}}(\mathbf{x}') K_i(\mathbf{x}', \mathbf{x}).$$

Without loss of generality, we assume $\sum_{j:i\in I_j} H_j(\mathbf{x}'_{I_j}, \mathbf{y}_j) \le \sum_{j:i\in I_j} H_j(\mathbf{x}_{I_j}, \mathbf{y}_j)$ which leads to

$$p_{\rm loc}(\mathbf{x}|\mathbf{y})\, Q_i(\mathbf{x},\mathbf{x}') = p_{0,\rm loc}(\mathbf{x}) K_i(\mathbf{x},\mathbf{x}') \exp\left(-\sum_{j:i\in I_j} H_j(\mathbf{x}_{I_j},\mathbf{y}_j)\right),$$

$$p_{\rm loc}(\mathbf{x}'|\mathbf{y})\, Q_i(\mathbf{x}',\mathbf{x}) = p_{0,\rm loc}(\mathbf{x}') K_i(\mathbf{x}',\mathbf{x}) \exp\left(-\sum_{j:i\in I_j} H_j(\mathbf{x}_{I_j},\mathbf{y}_j)\right) \frac{\exp(-\sum_{j:i\in I_j} H_j(\mathbf{x}_{I_j},\mathbf{y}_j))}{\exp(-\sum_{j:i\in I_j} H_j(\mathbf{x}'_{I_j},\mathbf{y}_j))}.$$

Because we also have that

$$p_{0,\rm loc}(\mathbf{x}) K_i(\mathbf{x},\mathbf{x}') = p_{0,\rm loc}(\mathbf{x}') K_i(\mathbf{x}',\mathbf{x}), \quad \mathbf{x}'_j = \mathbf{x}_j \quad \forall j \neq i,$$

detailed balance, i.e., $p_{\rm loc}(\mathbf{x}|\mathbf{y})\, Q_i(\mathbf{x},\mathbf{x}') = p_{\rm loc}(\mathbf{x}'|\mathbf{y})\, Q_i(\mathbf{x}',\mathbf{x})$, is now verified.

3.2.4. Computational requirements of l-MwG

Recall that the convergence rate of the Gibbs sampler for the localized prior is independent of dimension (Theorems 3.1 and 3.2). This means that the size of the proposed moves does not decrease as the dimension of $\mathbf{x}$ increases, since the size of the move depends only on the local properties of the prior in one of the blocks, rather than the overall number of blocks. In contrast, the step sizes of many other MCMC algorithms decrease with the dimension of $\mathbf{x}$ (see Fig. 1). Moreover, the proposed l-MwG moves are local, i.e., $\mathbf{x}'$ is different from $\mathbf{x}$ only in a few (much less than n) of its components, independent of the number of observations k and of the dimension n of $\mathbf{x}$. Because we assume that $k = O(n)$, the number of observations grows with the dimension, but the number of observations per block remains fixed. This suggests that the acceptance ratio in equation (12) may be independent of dimension.

If the size of the proposed moves and the acceptance ratio do not decrease with dimension, then the convergence rate of the Gibbs sampler could be independent of dimension. Indicators of sampling efficiency, e.g., IACT, would depend only on the *properties* of the blocks rather than on the *number* of blocks, with properties of each block defined by the local structure of the prior covariance and precision matrices and local properties of the observation function. The overall computational cost per sample of l-MwG is then linear in the number of blocks. We do not provide a proof for the dimension independence convergence of l-MwG, but we provide numerical examples (linear and nonlinear) in which IACT is indeed independent of dimension, while IACT increases with dimension for other MCMC algorithms. We also present an example in which we deliberately violate some of our assumptions to demonstrate the limitations of these ideas.

3.2.5. Limitations of l-MwG

We note that l-MwG has the format of "sample from the prior and correct (Metropolize) to account for the likelihood." This approach is generally not efficient and, in particular, degenerates as the observational noise diminishes and the posterior concentrates with respect to the prior. Localization can somewhat mitigate the effect of posterior concentration (at least, relative to non-localized samplers) as Metropolization is applied only to low-dimensional blocks. Nonetheless, practical MCMC samplers will require more sophisticated proposal distributions *and* localization. The l-MwG presented here should be viewed as a first and simple example of an MCMC sampler that can achieve dimension independent performance by exploiting underlying local structure. We remark that l-MwG using "likelihood-informed proposals" is feasible by mixing localization with other MCMC strategies, such as more sophisticated proposal distributions [24], acceleration by matrix splittings and analogies of Gibbs samplers to linear solvers (see [20]), multigrid methods [25], or preconditioning. This is beyond the scope of this paper, and will be investigated in the future.

4. Discussion of assumptions and effective dimension

4.1. Physical interpretation of local problems

We have shown, for linear problems, that localization causes small errors if a problem is local (see Assumptions (*i–iii*) in Section 1). We argued in Section 3 that localized problems can be solved efficiently by MCMC. All this is relevant only if there are indeed inverse problems which are local and which can be localized, i.e., if Assumptions (*i–iii*) in Section 1 are valid for some interesting problems.

The assumptions in Section 1 correspond to prior distributions with short correlation lengths (e.g., in a Gaussian process) and small neighborhood sizes (e.g., in a Gaussian Markov random field [54]) relative to the dimensions of the physical domain. A common choice is Gaussian prior distributions with precision matrices defined via Laplace-like operators; see, e.g., [8,15,36,39,45,58] and our examples in Section 5. These priors typically have banded precision matrices and, in many cases, also have (nearly) banded covariance matrices. The priors are updated to posterior distributions by likelihoods involving observations that depend largely on local properties. This means in particular that the set of observations, $\mathbf{y}$, may inform *all* components of $\mathbf{x}$, but each individual component of the observation, $[\mathbf{y}]_j$, $j = 1,\ldots,k$, may only inform a subset of the

components of $\mathbf{x}$. An example of a local observation function is when some or all components of the "quantity of interest" $\mathbf{x}$ are directly observed, which is often the case in state estimation problems.

We expect that the assumptions are valid in many geophysical and engineering applications, where the target distribution is the posterior distribution of a physical quantity defined over a spatial domain. For example, observation matrices $\mathbf{H}$ in image deblurring are often constructed through the discretization of kernels that have (nearly) compact support, and, for that reason, are typically local. Observations of diffusion processes are local on sufficiently short time scales. PDEs where information is carried mostly along characteristics (e.g., transport equations) give rise to local observations when the quantity of interest is an initial condition. We already brought up NWP and the EnKF as an example of a local problem in which localization enables efficient computations in high dimensional problems. Similarly, exploiting localization in importance sampling (particle filtering) is also a current topic in NWP, see, e.g., [34,35,44,46–50,60,62]. NWP, however, is usually not considered an inverse problem due to its sequential-in-time nature.

It is important to realize that many important problems are not local, and that localization is not useful for such problems. An example is computed tomography, where each observation might depend on material properties along an entire tomographic ray, leading to observation matrices which are not local in the sense we define here.

4.2. Connections with infinite dimensional inverse problems and effective dimensions

One can think of $\mathbf{x}$ as the discretization of a physical quantity in some domain, for instance on a grid with n degrees of freedom or in Fourier basis of n modes. For a given domain, the dimension n of $\mathbf{x}$ grows as the discretization is refined. If the number of observations k is held constant and we let $n \to \infty$, then we describe what happens as the discretization is refined while the domain and observation network remain fixed. This leads to the concept of an effective dimension, which may be small (finite) even when the apparent dimension is large (infinite); see, e.g., [1,9].

Related to a small effective dimension are low-rank updates from prior to posterior distributions [14,17,57]. A low-rank update means that the posterior distribution differs from the prior distribution in $n_{\text{eff}} \ll n$ directions. More precisely, for Gaussian problems, a low-rank update means that the difference between prior and posterior covariance is low rank. In practice this occurs, for example, when the number of observations is much less than the dimension of $\mathbf{x}$, or when the observations constrain only a few linear combinations of the components of $\mathbf{x}$. Indeed, some definitions of effective dimension are directly related to the difference of prior and posterior covariance. In [1], an effective dimension is defined by:

$$n_{\text{eff},1} = \text{tr}((\mathbf{C} - \hat{\mathbf{C}})\mathbf{C}^{-1}), \tag{15}$$

where $\mathbf{C}$ and $\hat{\mathbf{C}}$ are prior and posterior covariance, respectively, and where $\text{tr}(\cdot)$ is the trace of a matrix. Alternatively, one can consider precision matrices and define an effective dimension by

$$n_{\text{eff},2} = \text{tr}((\hat{\mathbf{\Omega}} - \mathbf{\Omega})\mathbf{\Omega}^{-1}), \tag{16}$$

where $\mathbf{\Omega}$ and $\hat{\mathbf{\Omega}}$ are the prior and posterior precision matrices, respectively. Other effective dimensions have been defined for specific importance sampling algorithms; see, e.g., [1,9]. Certain MCMC and importance sampling algorithms can exploit a small effective dimension or low rank updates, and can be made discretization invariant—such that indicators of computational efficiency become independent of the chosen grid; see, e.g., [8,12,13,17,45].

Our assumptions and problem setting describe a different mechanism for reaching large dimensions, because we assume that the number of observations is on the order of the dimension, $k = O(n)$. The assumptions translate to a problem for which the discretization and the number of observation per unit "length" remain fixed, while the size of the domain grows. As an example, consider a Gaussian process defined on an interval of length L. Further suppose that this process is observed every r units of distance, i.e., we have L/r observations. We investigate the situation where the domain size L gets bigger but the number of observations per unit length remains constant. The limit $L \to \infty$ may not be meaningful because it would correspond to an infinitely large domain. Considering large L, however, is meaningful because it describes a large domain and a large number of observations. Moreover, as L increases, the dimension n, the number of observations k, and the effective dimension all increase. The cartoon example of sampling a high-dimensional isotropic Gaussian illustrates the performance one can expect from MCMC when the size of the domain is large, and the number of observations is on the order of the dimension (i.e., large, well-observed domains).

The assumptions defining a local problem (see Section 1) do not in general imply that updates from prior to posterior are low-rank. Rather, we replace assumptions about low effective dimension and/or low-rank updates by assumptions about locality of the precision/covariance matrices and local observations. The Gibbs samplers discussed above are efficient when precision and covariance matrices are banded, and if observations are local. If these assumptions are only approximately satisfied, we propose to localize the problem, i.e., to replace covariance/precision matrices by banded matrices. The resulting localized problem can be solved efficiently, but at the cost of additional errors due to the localization procedure.

In summary, we suggest that while MCMC for generic high-dimensional problems might remain difficult, one can effectively solve two classes of problems. If an effective dimension is small and/or an update from prior to posterior distribution is low rank, then one can exploit this structure and create effective MCMC algorithms. This case has been discussed extensively over the past years. Our contribution is to suggest that MCMC can also be made efficient if effective dimensions

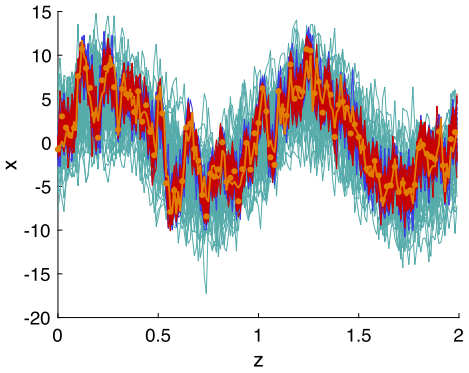

Fig. 3. Illustration of the Gaussian inverse problem of example 1. Teal – 50 samples of the prior distribution. Red – 50 samples of l-MwG. Blue (often hidden) – 50 samples of posterior distribution. Orange line – "true state" that gives rise to the data. Orange dots – data.

are large or if updates from prior to posterior distributions are high rank, as long as the prior covariance and precision matrices are banded and the observations are local. In fact, the situation we describe is analogous to linear algebra. High-dimensional matrices are easy to deal with if they are either low-rank or banded (sparse). The same seems true in MCMC for Bayesian inverse problems—these problems are manageable if either an effective dimension is small, or if the covariance and precision matrices have banded structure and if the observations are local.

5. Numerical illustrations

We consider several linear and nonlinear examples to illustrate local inverse problems, their localization, as well as the dimension independence of Gibbs and l-MwG samplers.

5.1. Example 1: Gaussian prior with exponential covariance function

We consider a Gaussian prior on the interval $z \in [0, L]$ with mean and covariance function

$$\mu(z) = 5\,\sin(2\pi z), \quad k(z,z') = C\,\exp\left(-\frac{|z-z'|}{2\rho}\right).$$

where $\rho = 0.02$, $C = 10$, and fix a discretization of the domain with $\Delta z = 0.01$. This results in a slowly varying prior mean and, for a given discretization, the covariance matrix has off-diagonal elements that decay quickly away from the diagonal.

For a given domain size L, the dimension of the discretized Gaussian is $n = L/\Delta z$, i.e., we have 100 state variables per unit length. For numerical stability we add $10^{-6}\mathbf{I}$ to the discretized covariance matrix. Samples from the prior are shown (in teal) for a problem with $L = 2$ in Fig. 3. The data are measurements of a prior sample, collected every $2\Delta z$ length units. The measurements are perturbed by Gaussian noise with mean zero and identity covariance matrix, i.e., $\mathbf{R} = \mathbf{I}$, the identity matrix of size $n/2$. Note that the observation network is such that no localization of the observation matrix $\mathbf{H}$ is required, because the observations are already local (point-wise measurements and diagonal $\mathbf{R}$). For a given L we have n variables to estimate and $n/2$ data points, or, equivalently, we have 100 variables and 50 measurements per unit length. The true state and measurements are illustrated (in orange) for a problem with $L = 2$ in Fig. 3. The data and prior distribution define a Bayesian posterior distribution, and we show 50 samples of this posterior distribution (for $L = 2$) in Fig. 3.

To illustrate the banded structure and relative size of the elements of the prior and posterior covariance/precision matrices, we plot the absolute value of the elements of these matrices, scaled by their largest element, in the left panels of Fig. 4. These scaled absolute values are between 0 and 1, and indicate where off-diagonal elements are small. The prior and posterior covariance matrix have nearly banded structure in the sense that the elements near the diagonal are larger than the off-diagonal elements. The condition number of the prior precision and covariance matrices is about 64 (independently of the domain length). We localize the prior covariance matrix by setting all elements smaller than 0.1 equal to zero. Since the prior precision matrix is tridiagonal, we do not need to localize it. The l-MwG sampler in "precision matrix implementation" already makes use of the banded structure of the precision matrix because we condition only on a few neighbors, rather than the full state. Since the observation matrix $\mathbf{H}$ is already local (direct observations of the state variables), $\mathbf{H}$ does not need to be localized. The likelihood has the form (13) and we make use of this structure by considering blocks of size 2 with one observation in each block.

For a fixed domain length L, we apply RWM, pCN [12], MALA, Hamiltonian MCMC, and l-MwG (with covariance localization or in precision matrix implementation). We tune the proposal variance of RWM, pCN, MALA, Hamiltonian MCMC by considering a scaling of the step size with n^{-k}, for several different values of k, e.g., $k = \{1, 2, 3, 4, 5, 6\}$ (recall that the dimension of our discretization is $n = L/\Delta z$). We call the step size that lead to the smallest IACT "optimal." All chains are

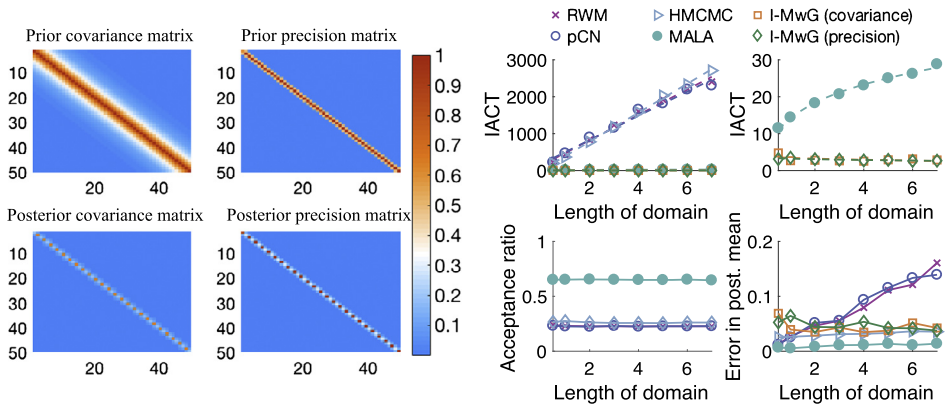

Fig. 4. Covariance and precision matrices, IACT, acceptance rate and errors for Example 1. *Left panels*. Prior and posterior covariance and precision matrices of example 1 with $L = 0.5$. Top row: prior covariance (left) and precision matrix (right). Bottom row: posterior covariance (left) and precision matrix (right). *Right panels*. Top row: IACT as a function of domain length for RWM, pCN, Hamiltonian MCMC (left) and MALA and l-MwG (right). Bottom row: acceptance rate as a function of domain length (left) and mean squared error in posterior mean (right).

Table 1
Configurations and IACT scalings for examples 1, 2 and 3. In each cell, left to right, are the configurations and results for examples 1, 2 and 3.

Algorithm	Tuned step size	Sample size	IACT scaling
RWM	$n^{-1/2}, n^{-1/2}, n^{-1/2}$	$10^6, 10^6, 10^6$	n, n, n
MALA	$n^{-1/6}, n^{-1/4}, n^{-1/4}$	$10^4, 10^4, 10^4$	$n^{1/4}, n^{1/4}, n^{1/4}$
Hamiltonian MCMC	$n^{-1/2}, n^{-1/2}, n^{-1/3}$	$10^5, 10^5, 10^5$	n, n, n
pCN	$n^{-1/2}, n^{-1/2}, n^{-1/3}$	$10^6, 10^6, 10^6$	n, n, n
l-MwG (precision)	n/a	500, 500, 5000	const., const., $n^{1/2}$
l-MwG (covariance)	n/a	500, 500, 5000	const., const., $n^{1/2}$

initialized by a random sample from the prior and we produce 10^6 samples with pCN and RWM, 10^5 with Hamiltonian MCMC, 10^3 with MALA, and 500 with the l-MwG samplers. The set-up for this problem is also summarized in Table 1.

We perform numerical experiments for domain lengths ranging from $L = 0.5$ to $L = 7$, which leads to problems of dimension $n = 50$ to $n = 700$, and with 25 to 350 observations. For each domain length and algorithm we compute the corresponding IACT. This leads to the scalings, obtained by least-squares fitting of polynomials, of IACT with dimension, as shown in Table 1, and as illustrated in Fig. 4. We observe that RWM, pCN, Hamiltonian MCMC, and MALA exhibit an increasing IACT (at different rates) as the domain size and the number of observations increase, while both implementations of l-MwG are characterized by an IACT that remains constant.

We further compute the acceptance rate (i.e., the acceptance ratio, averaged over all moves) for RWM, pCN, Hamiltonian MCMC, and MALA. Results are shown in Fig. 4, and we note that our tuning of the step-size for each algorithm keeps the acceptance rate constant as dimension increases. Finally, we compute an error to check that each algorithm indeed samples the correct distribution. We define an "error in posterior mean" by

$$e = \frac{1}{n}\sum_{j=1}^{n}([\hat{\mathbf{m}}]_j - [\bar{\mathbf{x}}]_j)^2, \tag{17}$$

where $\hat{\mathbf{m}}$ is the posterior mean and where $\bar{\mathbf{x}}$ is the average over the MCMC samples. We compute this error, which describes how far our estimated posterior mean is from the actual posterior mean, for each algorithm and show the results in Fig. 4. We note that all algorithms lead to a small error. Since this is also true for l-MwG with prior covariance localization, we conclude that the localized problem is indeed nearby the unlocalized problem we set out to solve.

In this example, it is not the overall dimension, the overall number of observations, or the rank of the update of prior to posterior covariance matrix that defines performance bounds for l-MwG. Because the local problem structure is used during problem formulation and during its MCMC solution, the characteristics of each loosely coupled block define the behavior of the sampler. The overall number of blocks is irrelevant.

One may wonder why the IACT of pCN increases with dimension, even though pCN is "by design" dimension independent. The reason is *how* dimension increases in this example (see the discussion in Section 4.2). The pCN algorithm is dimension independent if the dimension increases because a discretization is refined, while the size of the domain, the number of observations, and an effective dimension remain constant. In the current example, pCN is dimension independent for fixed domain size L and a fixed observation network (fixed k), while we decrease the discretization parameter Δz. In such a scenario, an effective dimension remains constant. However, as we increase the domain size L *and* the number of ob-

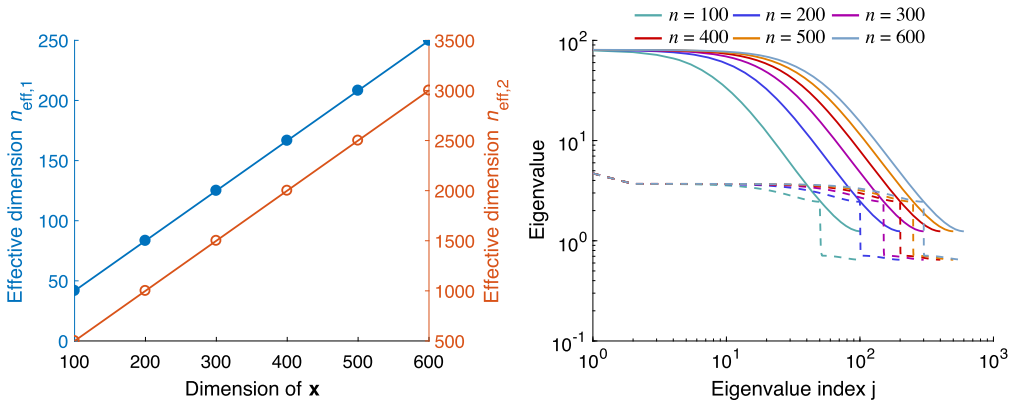

Fig. 5. *Left panel.* Effective dimensions (see Section 4.2) as a function of dimension (equivalently domain size and number of observations). *Right panel.* Eigenvalues of prior (solid) and posterior (dashed) covariance matrices for domains of sizes between $L = 1$ and $L = 6$.

servations k, while keeping the discretization parameter Δz fixed, the apparent and effective dimensions both increase, and the performance of pCN deteriorates (see left panel of Fig. 5 and also Section 4.2). Similarly, the number of non-negligible eigenvalues of prior or posterior covariance matrices increases with the domain size and the number of observations (see right panel of Fig. 5). The usual scenario considered for pCN (and other function space MCMC algorithms) is that the number of "relevant" eigendirections remains constant as dimension increases. This, too, is not the case for this example and may not be true for other local inverse problems.

The original convergence analysis of pCN can be found in [27], and applies to a setting where dimension increases because the discretization of the unknown function is refined. The analysis requires that the limiting ($n \to \infty$) observation operator be well defined. This is not the case in the above example, with increasing domain size and an increasing number of observations. The limit of an increasing domain size would be a domain of infinite size, which is difficult to describe with mathematical rigor.

5.2. Example 2: Gaussian prior with squared Laplacian as precision matrix

We now consider a Gaussian prior on $z \in [0, L]$ with mean zero and with a precision matrix derived from the squared Laplacian, which leads to a pentadiagonal precision matrix after discretization. As before, we fix the discretization and chose $\Delta z = 0.01$. The Laplacian is approximated by the $n = L/\Delta z$ dimensional matrix $\mathbf{L} = 1/(\Delta z)^2\mathbf{A}$, where $\mathbf{A}$ is a matrix with 2 on the main diagonal and -1 on the first upper and lower diagonal. We define the prior precision matrix by $\mathbf{\Omega} = (1/\rho^2\mathbf{I} + \mathbf{L})^2$, where $\rho = 0.06$. As in example 1, we collect data by collecting noisy measurements of a prior sample every $L/(2\Delta z)$ units. The data are perturbed by a Gaussian random variable with mean zero and covariance $\mathbf{R} = \mathbf{I}$. As in Example 1, the discretization and observation network yields 100 discrete state variables and 50 data points per unit length. The condition number of the prior precision or covariance matrix is about $21 \cdot 10^3$.

In Fig. 6, we illustrate the banded structure of the prior and posterior covariance and precision matrices for a problem with domain length $L = 0.5$. As in example 1, the prior covariance matrix is nearly banded (off-diagonal elements are small compared to diagonal elements), while the prior precision matrix is banded (pentadiagonal). Thus, as before, we localize the prior covariance matrix by setting all elements below a threshold of 0.01 equal to zero. The l-MwG sampler in prior precision matrix implementation does not require localization since the prior precision is banded. We also apply RMW, pCN, MALA, Hamiltonian MCMC, and tune their step-size as described in Example 1, with our findings summarized in Table 1. The number of samples we consider for each algorithm is also given in Table 1.

We vary the domain length L from $L = 0.5$ to $L = 7$, apply the various samplers and compute the corresponding IACTs. This leads to the scalings of IACT with domain length (or, equivalently, dimension) as shown in Table 1, and as illustrated in Fig. 6. While the scalings of IACT with domain size for RWM, Hamiltonian MCMC and pCN are equal (all scale linearly), we find that IACT of pCN and Hamiltonian MCMC is reduced compared to RWM. As in Example 1, we find that the l-MwG samplers and MALA exhibit smaller IACT than the other algorithms we tested, and that IACT of the l-MwG samplers remains constant as we increase the domain length. We further find that the acceptance rate of MALA first increases with L, but then levels off (due to our tuning of the step-size). Similarly, the acceptance rate of RWM, pCN, and Hamiltonian MCMC first decreases, but then levels off for large L. All algorithms yield a small error (see equation (17)), which, in the case of l-MwG in covariance matrix implementation, supports our claim that the localized problem is indeed a small perturbation of the unlocalized problem. The reasons for increase in IACT with dimension for pCN are the same as in example 1.

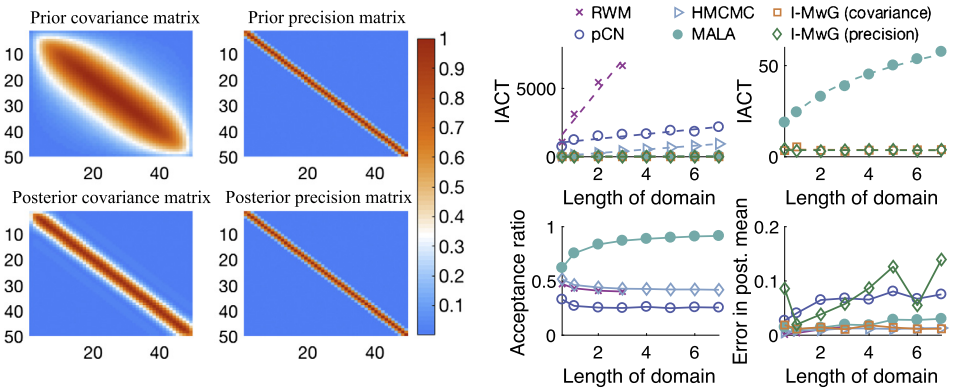

Fig. 6. Covariance and precision matrices, IACT, acceptance rate and errors for Example 2. *Left panels.* Prior and posterior covariance and precision matrices of Example 2 with $L = 0.5$. Top row: prior covariance (left) and precision matrix (right). Bottom row: posterior covariance (left) and precision matrix (right). *Right panels.* Top row: IACT as a function of domain length for RWM, pCN, Hamiltonian MCMC (left) and MALA and l-MwG (right). Bottom row: acceptance rate as a function of domain length (left) and mean squared error in posterior mean (right).

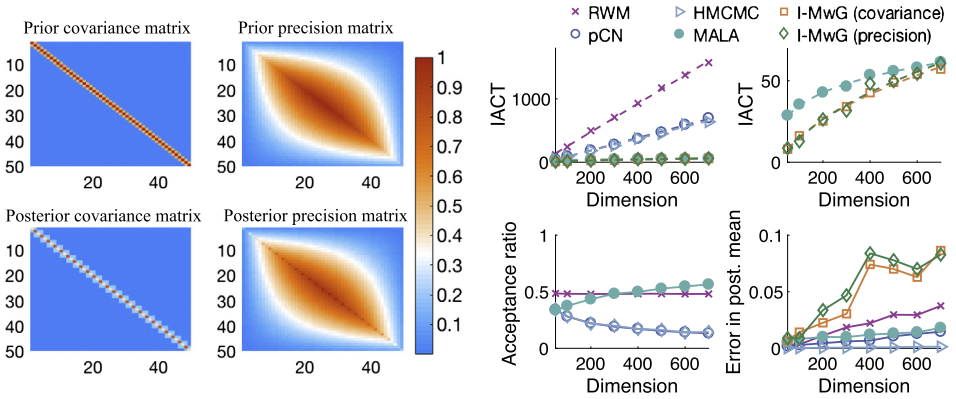

Fig. 7. Covariance and precision matrices, IACT, acceptance rate and errors for Example 3. *Left panels.* Prior and posterior covariance and precision matrices of Example 3 with $L = 0.5$. Top row: prior covariance (left) and precision matrix (right). Bottom row: posterior covariance (left) and precision matrix (right). *Right panels.* Top row: IACT as a function of domain length for RWM, pCN, Hamiltonian MCMC (left) and MALA and l-MwG (right). Bottom row: acceptance rate as a function of domain length (left) and mean squared error in posterior mean (right).

5.3. Example 3: banded covariance, but full precision matrix

We now consider a discrete Gaussian prior and Gaussian inverse problem that does not necessarily have a limit as the discretization is refined. We pick this perhaps unphysical problem to illustrate limitations of l-MwG when the condition number is large and when it increases with dimension.

We pick a dimension, n, and chose a zero mean and a covariance matrix given by an $n \times n$ matrix $\mathbf{C}$ which has 2 on the main diagonal and -1 on the first upper and lower diagonal. We vary the dimension between $n = 50$ and $n = 700$ which leads to condition numbers between $1 \cdot 10^3$ and $2 \cdot 10^5$ (the condition number strictly increases with dimension). Because of the large condition number, the banded covariance does not imply that the precision matrix is also banded. In fact, for this prior, the covariance matrix is tridiagonal, but the precision matrix is full, as illustrated in Fig. 7. As in examples 1 and 2, we observe every other variable and perturb these measurements by a Gaussian with mean zero and covariance $\mathbf{R} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix of dimension $n/2$. Because the prior precision matrix is not banded, the posterior precision is also not banded as shown in Fig. 7.

We solve the Gaussian inverse problem by RWM, MALA, pCN, Hamiltonian MCMC and l-MwG. The set-up and tuning of the algorithms are as described in Examples 1 and 2, and summarized in Table 1. The l-MwG sampler in covariance matrix implementation does not require localization since the prior covariance is banded (tridiagonal). The precision matrix on the other hand cannot be localized efficiently because the bandwidth of the prior precision matrix is large (see Fig. 7), and also increases with n. We thus do *not* localize the l-MwG sampler in precision matrix implementation and condition on all variables during sampling, not only on nearby variables.

For each algorithm and dimension, we compute IACT (see Table 1 for the rates with which IACT increases with dimension for the various algorithms). We observe that IACT increases for all algorithms we consider, as illustrated by Fig. 7. We observe that IACT for the l-MwG samplers now also increases. The rate is $n^{1/2}$ which is larger than the rate of MALA which, as in examples 1 and 2, is equal to $n^{1/4}$. The reasons for increase in IACT with dimension for pCN are the same as in

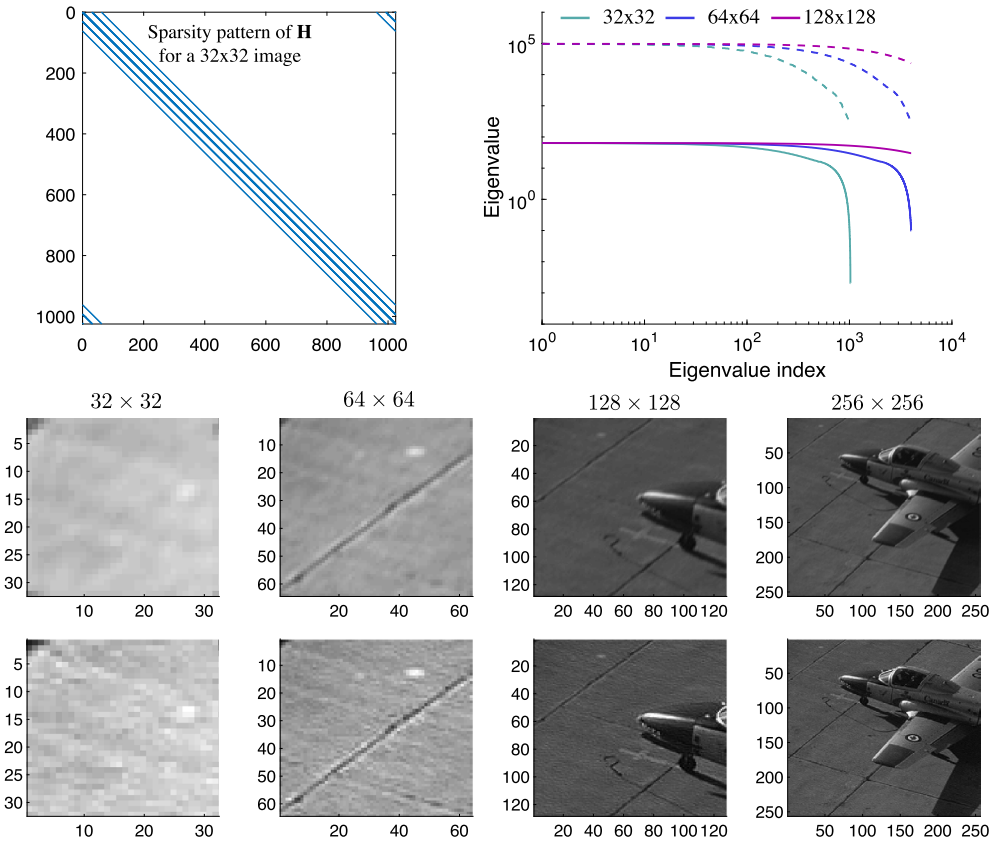

Fig. 8. Top left: sparsity pattern of the matrix **H** for a 32×32 image. Top right: the largest eigenvalues of the prior precision (solid) and posterior precision (dashed) matrices of 32×32 (turquoise), 64×64 (blue) and 128×128 (purple) images. Center row: blurred images. Bottom row: mean of 10^4 samples of the Gibbs sampler.

examples 1 and 2. The increase in IACT of l-MwG with dimension can be understood within our theory by the increase in condition number with dimension. This numerical experiment thus suggests that our assumption of moderate condition number is indeed necessary (rather than being a tool for proving the theorems), which corroborates our claim that l-MwG is effective (nearly dimension independent) for a unified class of problems with banded precision *and* covariance matrix. The bandedness of one implies bandedness of the other by the small condition number.

5.4. Example 4: image deblurring

We consider a linear inverse problem similar to an inverse problem in image deblurring and assume that **x** is a column-stack of the pixels of a 2D image. For example, for an image with 64×64 pixels, the dimension of **x** is $n = 64^2 = 4{,}096$. Thus, for high-resolution image deblurring (images with a large number of pixels), the dimension of **x** is huge (10^7 for a $4{,}000 \times 4{,}000$ image). We consider the images of sizes 32×32, 64×64, 128×128, 256×256 and shown in Fig. 8. These images are obtained by cutting square images out of a given, slightly larger image. We use this example to illustrate that IACT of the Gibbs sampler can remain constant as the dimension of **x** and the number of observations increase with image size. Many practical difficulties, e.g., estimation of noise and regularization parameters (see, e.g., [3,19,42]), are neglected in this example.

We assume that the image **x** is blurred by multiplication from the left with a given $n \times n$ matrix **H** and that a noisy version of the blurred image, **y**, is available. Thus,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \eta, \quad \eta \sim \mathcal{N}(\mathbf{0}, \lambda^{-1}\mathbf{I}),$$

where $\lambda = 10^5$ describes a (small) Gaussian measurement noise. The matrix **H** represents blurring the image by a Gaussian kernel with standard deviation 0.7, which is "realistic" because it effectively averages only a few pixels in each direction. Throughout this example we assume periodic boundary conditions. The prior is a Gaussian with mean zero and the prior

Table 2
Dimension, effective dimension and IACT of a Gibbs sampler as a function of the image size.

Image size:	32×32	64×64	128×128	256×256
Dimension:	1,024	4,096	16,384	65,536
$n_{\text{eff},2}$:	$4.8 \cdot 10^8$	$7.4 \cdot 10^9$	$1.2 \cdot 10^{11}$	-
IACT (l-MwG):	2.92	2.97	1.74	1.11

precision is the 2D Laplacian, i.e., $p_0(\mathbf{x}) = \mathcal{N}(0, \delta^{-1}\mathbf{L}^{-1})$, where $\delta = 10$. For this set-up, the posterior distribution is the Gaussian

$$p(\mathbf{x}|\mathbf{b}) \propto \exp\left(-\frac{\lambda}{2}||\mathbf{H}\mathbf{x} - \mathbf{b}||^2 - \frac{\delta}{2}||\mathbf{L}^{1/2}\mathbf{x}||^2\right),$$

and the posterior precision matrix is given by

$$\mathbf{\Omega} = \lambda \mathbf{H}^T\mathbf{H} + \delta\mathbf{L}.$$

The prior precision matrix $\mathbf{L}$ is banded and, thus, no localization of the prior is required. In this example, localization of the observation matrix is straightforward: since blurring occurs locally, only a few bands near the diagonal of $\mathbf{H}$ are "significant" and all other elements of $\mathbf{H}$ are "small." We can thus localize $\mathbf{H}$ by setting all elements below a threshold equal to zero. The threshold is 1% of the maximum value of all elements of $\mathbf{H}$. This gives rise to a sparse matrix $\mathbf{H}_{\text{loc}}$ whose sparsity pattern is illustrated for a 32×32 image in Fig. 8. Since $\mathbf{H}_{\text{loc}}$ and $\mathbf{L}$ are sparse, the posterior precision matrix $\mathbf{\Omega}_{\text{loc}}$ is also sparse (see also [7]). Errors due to localization of $\mathbf{H}$ are small. We compute that $||\mathbf{H} - \mathbf{H}_{\text{loc}}|| \approx 0.02$ and that the differences in the posterior covariance are $||\hat{\mathbf{C}} - \hat{\mathbf{C}}_{\text{loc}}|| \approx 3.3 \cdot 10^{-3}$, the error in the posterior mean is $||\hat{\mathbf{m}} - \hat{\mathbf{m}}_{\text{loc}}|| \approx 1$ (independently of image size).

We can compute the effective dimension $n_{\text{eff},2}$ in (16) and the largest (at most 4000) eigenvalues of the prior and posterior precision matrices for some of the localized problems. Our results are shown in Fig. 8 and displayed in Table 2. We note that the effective dimension and the number of non-negligible eigenvalues increase with image size and dimension. Thus, as in the previous examples, the high-dimension of this local inverse problem is not caused by a large apparent, but small effective dimension, or by an increasing number of negligible eigenvalues, but rather *all* dimensions of this problem are large. We did not compute the effective dimension or eigenvalues corresponding to the 256×256 image because the required matrix operations were difficult to do (on our laptop) even when exploiting the sparsity of the matrices.

We implement a Gibbs sampler for the posterior distribution using the posterior precision matrix $\mathbf{\Omega}_{\text{loc}}$. We define blocks using the 2D structure of the image rather than consecutive elements of the column stacks. In specific, the component index is two dimensional, i.e. $[\mathbf{x}]_{i,j}$ denotes the (i, j)-th pixel of an $n' \times n'$ image. And the (a, b)-th block of size $q = q' \times q'$ consists of entries

$$\mathbf{x}_{a,b} = ([\mathbf{x}]_{a+i,b+j}, i, j = 1, \cdots, q').$$

Two blocks $\mathbf{x}_{a_1,b_1}$ and $\mathbf{x}_{a_2,b_2}$ are neighbors if $|a_1 - a_2| \leq 1$ and $|b_1 - b_2| \leq 1$. During one iteration of the Gibbs sampler, we sample one block conditioned on the neighboring blocks. The blocks are of size 16×16 for the 32×32 and 64×64 images, for the 128×128 image we use blocks of size 32×32 and for the 256×256 image we use blocks of size 64×64. For each image, we draw 10,000 samples from the posterior distribution using the Gibbs sampler. We then compute IACT for every $8th$ pixel. The averages of these IACTs are displayed in Table 2. IACT is near one for all four images we consider and we emphasize that, as in previous examples, IACT does not increase with dimension. The slight decrease in IACT as dimension increases can be attributed to the larger block size we use. For example, if the block size is the size of the image, we draw independent samples and IACT equals one.

We checked the results we obtained by the Gibbs sampler as follows. The sample average is, to three digits, the same as the posterior mean of the localized problem (as computed by linear algebra, rather than sampling). We also compute the trace of the sample covariance matrix and compared it to the trace of the sample covariance of 10,000 samples obtained by a "direct" sampler that makes use of the sparse Cholesky factorization of the posterior precision matrix. The average of these variances agrees to the first three digits. This is perhaps not surprising in view of the small IACT.

We also applied pCN and MALA for this problem, but could not obtain useable results even for the smallest image (dimension $n = 1{,}024$). We tuned the step sizes of these algorithms, but for all choices we considered, we either accepted often, in which case the moves were too small (large IACT), or we observed that larger moves were almost never accepted. The reason for the difficulties with MALA and pCN is the large effective dimension of this problem.

5.5. Example 5: a nonlinear inverse problem

We consider a nonlinear inverse problem whose likelihood involves numerical solution of the Lorenz'96 (L96) model [37]

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + 8,$$

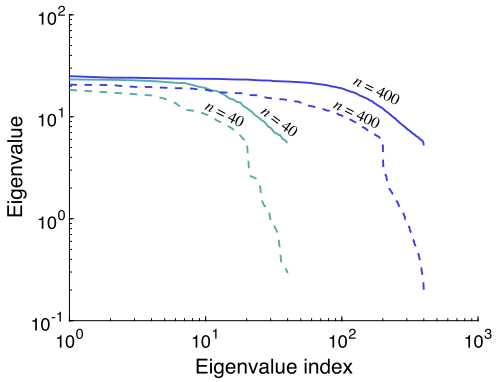

Fig. 9. Eigenvalues of the prior covariance matrices for the L96 model. Turquoise: $n = 40$. Blue: $n = 400$. Solid: prior covariance matrix. Dashed: approximate posterior covariance.

where $i = 1, \ldots, n$ and $x_{-1} = x_{n-1}$, $x_0 = x_n$, $x_{n+1} = x_1$. We consider L96 models with $n = 40$ and $n = 400$. Specifically, we try to estimate the initial condition $\mathbf{x}_0 = (x_1(0), \ldots, x_n(0))^T$ given noisy observations of every other state variable at time $T = 0.2$. We write this as

$$\mathbf{y} = \mathbf{H}\mathcal{M}_{0\to T}(\mathbf{x}_0) + \eta, \quad \eta \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{18}$$

where $\mathbf{H}$ is a $n/2 \times n$ matrix which selects every other component of $\mathbf{x}_T = \mathcal{M}_{0\to T}(\mathbf{x}_0)$; $\mathcal{M}_{0\to T}$ is the numerical solver of the L96 model from $t = 0$ to $t = T$, i.e., $\mathcal{M}_{0\to T}(\mathbf{x}_0)$ maps the initial condition $\mathbf{x}_0$ to $\mathbf{x}_T$. We use a fourth order Runge–Kutta scheme with time step $\Delta t = 0.01$.

We assume a Gaussian prior whose mean and covariance we obtain as follows. We initialize the L96 model with an arbitrary initial condition near the attractor and perform a simulation for 100 time units (10^4 time steps). The mean of this L96 trajectory is used as the prior mean, and we use a localized version of the covariance matrix computed from the trajectory as the prior covariance matrix. The localization is done as follows. We first compute the Hadarmard (element by element) product of the sample covariance and a localization matrix whose elements are $[\mathbf{A}]_{i,j} = \exp\left(-\left(d(i,j)/(3\sqrt{2})\right)^2\right)$, where d is a periodic distance: $d(i,j) = \min\{|i-j|, |i-j+n|, |i-j-n|\}$. Subsequently, off-diagonal elements below a threshold (1% of the largest element) are set to zero. This results in a localized, sparse prior covariance matrix. The eigenvalues of the prior covariances for problems of dimensions $n = 40$ and $n = 400$ are shown in Fig. 9. As in the above examples, the eigenvalues do not decay quickly and the number of non-negligible eigenvalues increases with dimension.

The localized prior and a likelihood defined by equation (18) define the (non-Gaussian) posterior distribution

$$p(\mathbf{x}_0|\mathbf{y}) \propto p_0(\mathbf{x}_0) \exp\left(-\frac{1}{2}||\mathbf{H}\mathcal{M}_{0\to T}(\mathbf{x}_0) - \mathbf{y}||^2\right). \tag{19}$$

We minimize $-\log(p(\mathbf{x}|\mathbf{y}))$ by Gauss–Newton, using adjoints for gradient calculations, and use the optimization result as a reference solution for the MCMC results. We compute the inverse of the Gauss–Newton approximation of the Hessian of $-\log(p(\mathbf{x}|\mathbf{y}))$, evaluated at its minimizer, and use this matrix, $\mathbf{P}$, as an approximation of posterior covariance (see, e.g., [2,29,55,59] for discussion of these approximations in NWP). The eigenvalues of the approximate posterior covariance matrices for problems of dimension $n = 40$ and $n = 400$ are shown in Fig. 9. Note that the number of non-negligible eigenvalues increases with dimension. With the approximate posterior covariance $\mathbf{P}$ we can also compute an effective dimension. Here we use $n_{\text{eff},1}$ as in (15), because the problem is naturally formulated in terms of covariance matrices, rather than precision matrices. We compute $n_{\text{eff},1} = 17.94$ when $n = 40$ and $n_{\text{eff},1} = 181.36$ when $n = 400$. As in the above examples, this effective dimension increases with dimension.

We use MALA, pCN and l-MwG to draw samples from the posterior distribution (19), which is not Gaussian because the L96 model is nonlinear, i.e., $\mathcal{M}_{0\to T}(\mathbf{x}_0)$ is not a linear function of $\mathbf{x}_0$. The MALA proposal we use is

$$\mathbf{x}'_{k+1} = \mathbf{x}_k + \frac{\sigma^2}{2}\nabla \log p(\mathbf{x}_k) + \sigma\,\xi$$

where $\xi \sim \mathcal{N}(0, \mathbf{C}_{\text{MALA}})$. We chose the covariance matrix $\mathbf{C}_{\text{MALA}}$ to be a diagonal $n \times n$ matrix whose diagonal elements are the diagonal elements of $\mathbf{P}$. For this example, this choice for the covariance of ξ normalizes each dimension, and it gives better results, in terms of acceptance ratio and IACT, than using the identity matrix. We consider several choices for the step size σ, and for each choice generate 10^5 samples. For fixed σ, we compute IACT for all n variables, and then average. We also compute the average (along the chain) of the acceptance ratio. The results for L96 models of dimensions $n = 40$ and $n = 400$ are shown in Fig. 10. We note that a small step-size is required in order to retain a reasonable acceptance ratio. As we increase the dimension from $n = 40$ to $n = 400$, we note that the acceptance ratio for a given step size σ drops. For

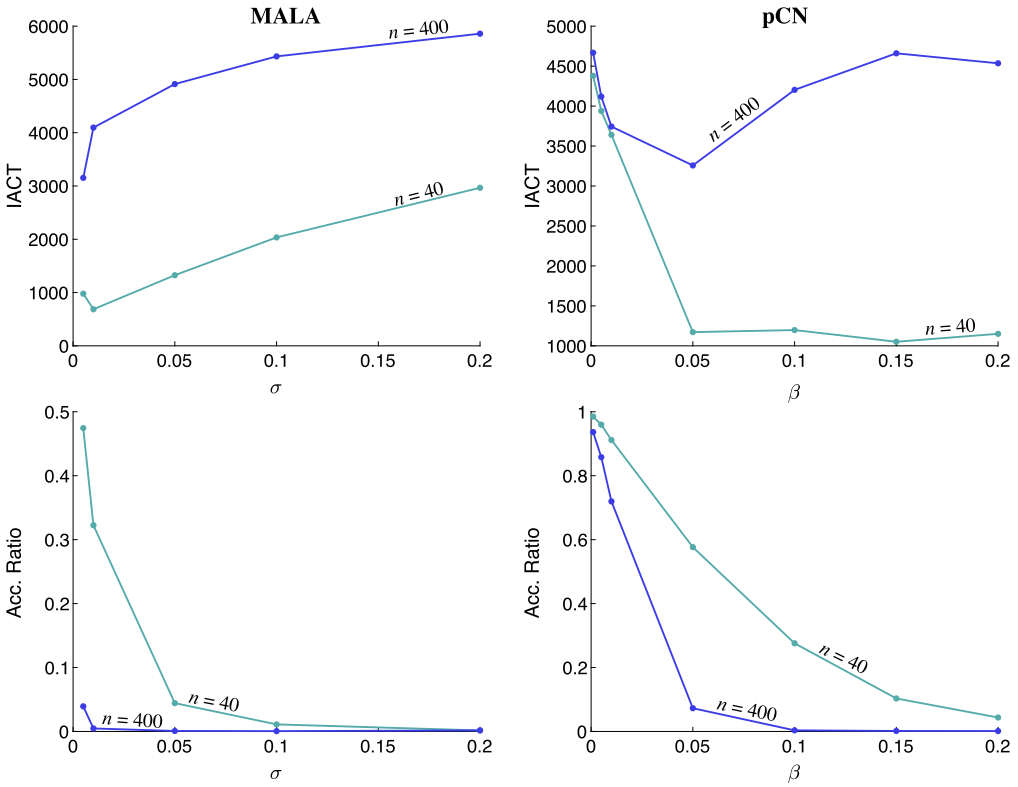

Fig. 10. Average IACT (top) and average acceptance ratio (bottom) as a function of the step-size. Left: MALA. Right: pCN.

Table 3
IACT of MCMC methods: The IACT of MALA and pCN are minimized with σ and β. The q in l-MwG-q is the block size.

	MALA	pCN	l-MwG-2	l-MwG-4	l-MwG-8
$n = 40$	686	1,051	55	60	266
$n = 400$	3,153	3,257	43	81	257

this reason, IACT increases as the dimension increases. The smallest (average) IACTs we could achieve for the $n = 40$ and $n = 400$ dimensional problems are shown in Table 3. It is evident that IACT increases with dimension in this problem.

The pCN proposal we use is

$$\Delta \mathbf{x}'_{k+1} = \sqrt{1-\beta^2}\Delta \mathbf{x}_k + \beta\, \xi_k$$

where $\xi_k \sim \mathcal{N}(0, \mathbf{C})$ and where $\Delta\mathbf{x}$ are perturbations from the prior mean. We consider several choices for β and, for each choice, we generate 10^5 samples and compute the average IACT and average acceptance ratio as above. Our results are shown in the right panels of Fig. 10. We find that the acceptance ratio of pCN is generally higher than that of MALA, but IACT of pCN and MALA are comparable. We also observe that IACT of pCN increases with dimension. The smallest IACT we found are shown in comparison with those of MALA in Table 3.

We implement the l-MwG sampler using the prior covariance matrix and observation localization. The blocks are consecutive components of $\mathbf{x}_0$ (but taking into account the periodicity of the L96 model) and the block sizes we tested are 2, 4 and 8. Localization of the observation function is done by using, for each block, all observations within the block as well as the neighboring two observations on each side (accounting for the periodicity of L96). For each block size, we generate 10^4 samples and compute an average IACT as above. Results are shown in Table 3. Note that IACT of l-MwG is significantly smaller than the IACT we computed for pCN or MALA. More importantly we observe that IACT does not increase significantly when we increase the dimension.

We illustrate the localized prior and posterior distributions in Fig. 11. Samples of the prior distribution serve as a reference and 500 samples of the prior are shown in turquoise; 500 samples of the posterior distribution, obtained by l-MwG with block size eight, are shown in blue. The true state is shown in red. Localization of the prior and using l-MwG to draw samples from the localized problem produces a meaningful posterior distribution: the uncertainty is reduced and the posterior samples are centered around the true state. The lower panel of Fig. 11 illustrates how the uncertainty in the localized

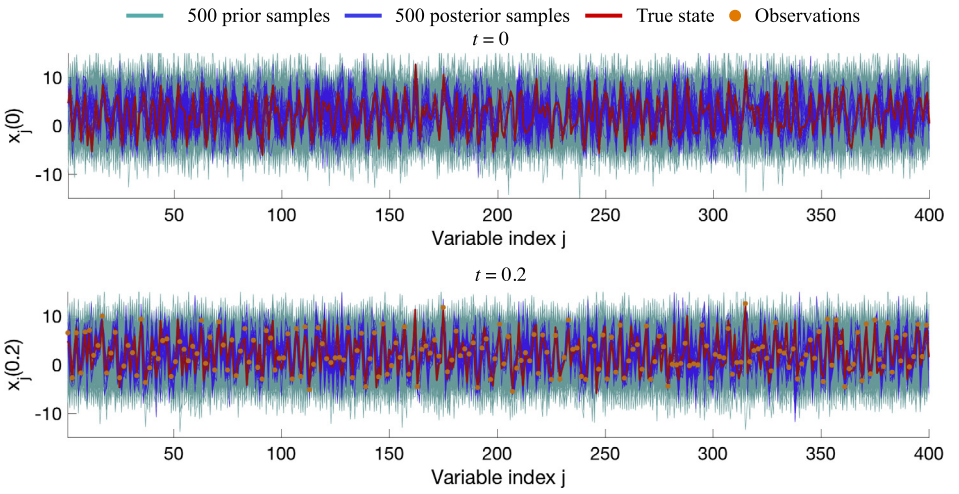

Fig. 11. Top: 500 prior samples (turquoise), the true state (red) and 500 posterior samples obtained by l-MwG (blue). Bottom: 500 model states at time $t = 0.2$ obtained by using the L96 model on the prior (turquoise) and posterior (blue) samples. Shown in red is the true state and the orange dots are the observations.

prior and posterior distributions propagates through time. We show the 500 samples of the prior (turquoise) and posterior (blue) mapped to time $T = 0.2$ by the L96 dynamics $\mathcal{M}_{0\to T}(\mathbf{x}_0)$. Again, the true state (at time T) is shown in red and lies well within the cloud of posterior samples. Plotting states at time T also allows us to compare to the observations, shown as orange dots, and we see good agreement between the model states at time T, generated from posterior samples, and the observations.

We emphasize that l-MwG, or in fact any MCMC method, may not be appropriate for solving this inverse problem. The minimizer of $-\log(p(\mathbf{x}_0|\mathbf{y}))$ and the associated Hessian-based covariance are good approximations of the posterior mean and covariances we computed by MCMC, but Gauss–Newton optimization is more efficient than MCMC for this problem. Our main messages for this example are that (*i*) l-MwG can be used on nonlinear problems; (*ii*) l-MwG, or other MCMC samplers that make use of the local structure of the problem, may be more effective than MCMC methods that do not make use of local problem structure (pCN and MALA); and (*iii*) increasing the dimension of this problem has almost no effect on the performance of l-MwG, but MCMC samplers that do not make use of local problem structure (pCN and MALA) are ineffective if the dimension and effective dimension are large.

6. Summary and conclusions

The main goal of this paper is to demonstrate that ideas of localization in numerical weather prediction (NWP) and the ensemble Kalman filter (EnKF) have relevance in inverse problems and MCMC. During localization, one restricts prior statistical interactions (correlations and/or conditional dependencies) and the effects of observations to neighborhoods. We have discussed conditions under which such ideas can be used in the context of inverse problems and their solution by MCMC. For example, we proved that localization introduces small errors for a class of linear "local" problems, but expect that this also holds for nonlinear local problems.

We reviewed the Gibbs sampler as a natural MCMC sampler that exploits local problem structure. We observed that its performance is dimension independent when sampling Gaussian distributions with banded precision and covariance matrices. We presented a Metropolis-within-Gibbs sampler that can be applied to linear or nonlinear problems. We demonstrated our ideas in several numerical examples in which Gibbs samplers outperformed MALA, Hamiltonian MCMC, RWM and pCN.

Localization is useful for inverse problems only if there are interesting applications in which localization can be applied. We speculate that local structure is common in physics and engineering, with NWP being the most spectacular example. Finally, we have discussed that the notion of high dimensionality in local problems is different from what is usually assumed in function space MCMC. The literature on function space MCMC focuses on problems with a large apparent dimension, but few observations and low-rank updates from prior to posterior. Localization is useful in a different scenario, in which the apparent and effective dimensions, and the number of observations, are large, and updates from prior to posterior are not low-rank.

Our study neglects many practical challenges. For example, localization may require some tuning, which in itself can be computationally expensive. More importantly, we have not rigorously defined localization for non-Gaussian problems, where enforcing bandedness of the covariance matrices may not be sufficient because higher moments become important. Taking inspiration instead from the sparsity of the precision matrix, a useful route may be to consider the conditional independence structure of more general non-Gaussian Markov random fields [56]. We hope to address these issues in future work.

Acknowledgements

M. Morzfeld gratefully acknowledges support by the Office of Naval Research (grant number N00173-17-2-C003), by the National Science Foundation (grant DMS-1619630), and by the Alfred P. Sloan Foundation (Sloan Research Fellowship). X.T. Tong gratefully acknowledges support by the National University of Singapore (grant R-146-000-226-133). Y.M. Marzouk gratefully acknowledges support from the DOE Office of Advanced Scientific Computing Research (grant DE-SC0009297).

Appendix A. Bias caused by localization

In Section 2, we propose to localize the prior covariance or precision matrix, and apply the l-MwG. This leads to samples from the localized posterior distribution, not from the exact posterior distribution. In this section we show that the difference between these two distributions can be small.

First we need the following estimate.

Lemma A.1. *Suppose* $\mathbf{A}$ *is a symmetric positive definite matrix, and* $\mathbf{B}$ *is a symmetric matrix such that* $\|\mathbf{B}\| \le \|\mathbf{A}^{-1}\|^{-1}$*, then*

$$\|(\mathbf{A}+\mathbf{B})^{-1} - \mathbf{A}^{-1}\| \le \frac{\|\mathbf{A}^{-1}\|^2\|\mathbf{B}\|}{1-\|\mathbf{B}\|\|\mathbf{A}^{-1}\|}.$$

Proof. Let $\mathbf{B} = \Psi L \Psi^T$ be the eigenvalue decomposition of matrix $\mathbf{B}$. Remove columns of Ψ that are eigenvectors of value 0, so L only contains nonsingular terms. Then by the Woodbury matrix identity

$$\mathbf{A}^{-1} - (\mathbf{A}+\Psi L\Psi^T)^{-1} = \mathbf{A}^{-1}\Psi(L^{-1}+\Psi^T\mathbf{A}^{-1}\Psi)^{-1}\Psi^T\mathbf{A}^{-1}.$$

Let v be the eigenvector corresponds to the largest absolute eigenvalue of $(L^{-1}+\Psi^T\mathbf{A}^{-1}\Psi)^{-1}$. Then

$$|v^T(L^{-1}+\Psi^T\mathbf{A}^{-1}\Psi)v| \ge |v^TL^{-1}v| - |v^T\Psi^T\mathbf{A}^{-1}\Psi v| \ge \|L\|^{-1} - \|\mathbf{A}^{-1}\|.$$

Moreover, $\|L\| = \|\mathbf{B}\|$, so $\|(L^{-1}+\Psi^T\mathbf{A}^{-1}\Psi)^{-1}\| \le (\|\mathbf{B}\|^{-1} - \|\mathbf{A}^{-1}\|)^{-1}$, and

$$\|(\mathbf{A}+\mathbf{B})^{-1} - \mathbf{A}^{-1}\| \le \frac{\|\mathbf{A}^{-1}\|^2}{\|\mathbf{B}\|^{-1} - \|\mathbf{A}^{-1}\|}. \quad \square$$

Proposition A.2. *Let* $\mathbf{C}$ *be a prior covariance matrix and* $\mathbf{H}$ *be an observation matrix. Let* $\mathbf{C}_{\rm loc}$ *and* $\mathbf{H}_{\rm loc}$ *be localized covariance and observation matrices, and let* δ_C *and* δ_H *be as defined in* (2) *and* (5)*.*

a) The localized prior covariance and observation matrices are small perturbations of the unlocalized covariance and observation matrices in the sense that

$$\|\mathbf{C} - \mathbf{C}_{\rm loc}\| \le \delta_C, \quad \|\mathbf{H}_{\rm loc} - \mathbf{H}\| \le \delta_H.$$

Moreover, if $\delta_C \le \|\mathbf{C}^{-1}\|^{-1}$*, then* $\mathbf{C}_{\rm loc}$ *is positive semidefinite.*

b) If $\delta_C \le \|\mathbf{C}^{-1}\|^{-1}$, $\Delta_1 \le \|\widehat{\mathbf{C}}\|^{-1}$*, the localization creates a small perturbation of the posterior covariance matrix in the sense that*

$$\|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\rm loc}\| \le \frac{\|\widehat{\mathbf{C}}\|^2\Delta_1}{1-\Delta_1\|\widehat{\mathbf{C}}\|}, \quad \Delta_1 := \frac{\|\mathbf{C}^{-1}\|^2\delta_C}{1-\delta_C\|\mathbf{C}^{-1}\|} + (2\delta_H + \delta_H^2)\|\mathbf{R}^{-1}\|.$$

c) Under the same conditions as in b), the localization creates a small perturbation of the posterior mean in the sense that

$$\|\hat{\mathbf{m}}_{\rm loc} - \hat{\mathbf{m}}\| \le \left(\frac{\|\mathbf{C}^{-1}\|^2\|\widehat{\mathbf{C}}\|\delta_C}{(1-\delta_C\|\mathbf{C}^{-1}\|)(1-\Delta_1\|\widehat{\mathbf{C}}\|)} + \frac{\|\mathbf{C}^{-1}\|\|\widehat{\mathbf{C}}\|^2\Delta_1}{1-\Delta_1\|\widehat{\mathbf{C}}\|}\right)\|\mathbf{m}\| + \frac{\|\widehat{\mathbf{C}}\|\|\mathbf{R}^{-1}\|}{1-\Delta_1\|\widehat{\mathbf{C}}\|}(\|\widehat{\mathbf{C}}\|\Delta_1 + \delta_H)\|\mathbf{y}\|.$$

Proof. For claim a), Define $\Delta = \mathbf{C} - \mathbf{C}_{\rm loc}$, apply Lemma B.2 with block size 1,

$$\|\mathbf{C} - \mathbf{C}_{\rm loc}\| \le \left(\max_i \sum_j |[\Delta]_{i,j}|\right) = \delta_C.$$

The bound $\|\mathbf{H} - \mathbf{H}_{\rm loc}\| \le \delta_H$ follows the same argument.

For claim b), we will exploit the identity $\widehat{\mathbf{C}}^{-1} = \mathbf{C}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$. By Lemma A.1,

$$\|\mathbf{C}^{-1} - \mathbf{C}_{\rm loc}^{-1}\| \le \frac{\|\mathbf{C}^{-1}\|^2\delta_C}{1-\delta_C\|\mathbf{C}^{-1}\|}.$$

Under our normalization assumption that $\|\mathbf{H}\| = 1$,

$$\begin{aligned}\|\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} - \mathbf{H}_{\rm loc}^T\mathbf{R}^{-1}\mathbf{H}_{\rm loc}\| &\le \|(\mathbf{H}^T - \mathbf{H}_{\rm loc}^T)\mathbf{R}^{-1}\mathbf{H}\| + \|\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H} - \mathbf{H}_{\rm loc})\| + \|(\mathbf{H}^T - \mathbf{H}_{\rm loc}^T)\mathbf{R}^{-1}(\mathbf{H} - \mathbf{H}_{\rm loc})\| \\ &\le 2\delta_H\|\mathbf{R}^{-1}\mathbf{H}\| + \delta_H^2\|\mathbf{R}^{-1}\| \le (2\delta_H + \delta_H^2)\|\mathbf{R}^{-1}\|. \end{aligned} \tag{20}$$

Therefore

$$\|(\mathbf{C}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) - (\mathbf{C}_{\rm loc}^{-1} + \mathbf{H}_{\rm loc}^T\mathbf{R}^{-1}\mathbf{H}_{\rm loc})\| \le \frac{\|\mathbf{C}^{-1}\|^2\delta_C}{1 - \delta_C\|\mathbf{C}^{-1}\|} + (2\delta_H + \delta_H^2)\|\mathbf{R}^{-1}\| =: \Delta_1.$$

We apply Lemma A.1 to $\widehat{\mathbf{C}} = (\mathbf{C}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}$, and obtain

$$\|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\rm loc}\| \le \frac{\|\widehat{\mathbf{C}}\|^2\Delta_1}{1 - \Delta_1\|\widehat{\mathbf{C}}\|}.$$

As a consequence, we also have that $\|\widehat{\mathbf{C}}_{\rm loc}\| \le \frac{\|\widehat{\mathbf{C}}\|^2\Delta_1}{1-\Delta_1\|\widehat{\mathbf{C}}\|} + \|\widehat{\mathbf{C}}\| = \frac{\|\widehat{\mathbf{C}}\|}{1-\Delta_1\|\widehat{\mathbf{C}}\|}$.

As for claim c), the difference is given by

$$\hat{\mathbf{m}}_{\rm loc} - \hat{\mathbf{m}} = (\widehat{\mathbf{C}}_{\rm loc}\mathbf{C}_{\rm loc}^{-1} - \widehat{\mathbf{C}}\mathbf{C}^{-1})\mathbf{m} + (\widehat{\mathbf{C}}_{\rm loc}\mathbf{H}_{\rm loc}^T - \widehat{\mathbf{C}}\mathbf{H}^T)\mathbf{R}^{-1}\mathbf{y}.$$

The norm of the matrix $(\widehat{\mathbf{C}}_{\rm loc}\mathbf{H}_{\rm loc}^T - \widehat{\mathbf{C}}\mathbf{H}^T)\mathbf{R}^{-1}$ can be bounded using claim b)

$$\begin{aligned}\|(\widehat{\mathbf{C}}_{\rm loc}\mathbf{H}_{\rm loc}^T - \widehat{\mathbf{C}}\mathbf{H}^T)\mathbf{R}^{-1}\| &\le \|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\rm loc}\|\|\mathbf{H}^T\mathbf{R}^{-1}\| + \|\widehat{\mathbf{C}}_{\rm loc}\|\|\mathbf{H}_{\rm loc} - \mathbf{H}\|\|\mathbf{R}^{-1}\| \\ &\le \frac{\|\widehat{\mathbf{C}}\|\|\mathbf{R}^{-1}\|}{1 - \Delta_1\|\widehat{\mathbf{C}}\|}(\|\widehat{\mathbf{C}}\|\Delta_1 + \delta_H).\end{aligned}$$

As for the norm of the matrix $\widehat{\mathbf{C}}_{\rm loc}\mathbf{C}_{\rm loc}^{-1} - \widehat{\mathbf{C}}\mathbf{C}^{-1}$, we use

$$\|\widehat{\mathbf{C}}_{\rm loc}\mathbf{C}_{\rm loc}^{-1} - \widehat{\mathbf{C}}\mathbf{C}^{-1}\| \le \|(\mathbf{C}_{\rm loc})^{-1} - \mathbf{C}^{-1}\|\|\widehat{\mathbf{C}}_{\rm loc}\| + \|\widehat{\mathbf{C}}_{\rm loc} - \widehat{\mathbf{C}}\|\|\mathbf{C}^{-1}\|.$$

From part (b), we have

$$\|\mathbf{C}_{\rm loc}^{-1} - \mathbf{C}^{-1}\|\|\widehat{\mathbf{C}}_{\rm loc}\| \le \frac{\|\mathbf{C}^{-1}\|^2\|\widehat{\mathbf{C}}\|\delta_C}{(1 - \delta_C\|\mathbf{C}^{-1}\|)(1 - \Delta_1\|\widehat{\mathbf{C}}\|)}, \quad \|\widehat{\mathbf{C}}_{\rm loc} - \widehat{\mathbf{C}}\|\|\mathbf{C}^{-1}\| \le \frac{\|\mathbf{C}^{-1}\|\|\widehat{\mathbf{C}}\|^2\Delta_1}{1 - \Delta_1\|\widehat{\mathbf{C}}\|}.$$

In summary:

$$\|\hat{\mathbf{m}}_{\rm loc} - \hat{\mathbf{m}}\| \le \left(\frac{\|\mathbf{C}^{-1}\|^2\|\widehat{\mathbf{C}}\|\delta_C}{(1 - \delta_C\|\mathbf{C}^{-1}\|)(1 - \Delta_1\|\widehat{\mathbf{C}}\|)} + \frac{\|\mathbf{C}^{-1}\|\|\widehat{\mathbf{C}}\|^2\Delta_1}{1 - \Delta_1\|\widehat{\mathbf{C}}\|}\right)\|\mathbf{m}\| + \frac{\|\widehat{\mathbf{C}}\|\|\mathbf{R}^{-1}\|}{1 - \Delta_1\|\widehat{\mathbf{C}}\|}(\|\widehat{\mathbf{C}}\|\Delta_1 + \delta_H)\|\mathbf{y}\|. \quad \square$$

Proposition A.3. *Let $\boldsymbol{\Omega}$ be a prior precision matrix and $\mathbf{H}$ be an observation matrix. Let δ_Ω and δ_H be as defined in (4) and (5). Let $\boldsymbol{\Omega}_{\rm loc}$ and $\mathbf{H}_{\rm loc}$ be the localized precision and observation matrices, then*

a) If $\delta_\Omega \le \|\boldsymbol{\Omega}\|$, $\Delta_2 \le \|\widehat{\mathbf{C}}\|^{-1}$, the localization creates a small perturbation of the posterior precision and covariance matrix in the sense that $\|\widehat{\mathbf{C}}^{-1} - \widehat{\mathbf{C}}_{\rm p,loc}^{-1}\| \le \Delta_2$ and

$$\|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\rm p,loc}\| \le \frac{\|\widehat{\mathbf{C}}\|^2\Delta_2}{1 - \Delta_2\|\widehat{\mathbf{C}}\|}, \quad \textit{where} \quad \Delta_2 := \delta_\Omega + (2\delta_H + \delta_H^2)\|\mathbf{R}^{-1}\|.$$

b) Under the same conditions as in a), the localization creates a small perturbation of the posterior mean in the sense that

$$\|\hat{\mathbf{m}}_{\rm p,loc} - \hat{\mathbf{m}}\| \le \frac{\|\widehat{\mathbf{C}}\|\delta_\Omega + \|\boldsymbol{\Omega}\|\|\widehat{\mathbf{C}}\|^2\Delta_2}{1 - \Delta_2\|\widehat{\mathbf{C}}\|}\|\mathbf{m}\| + \frac{\|\widehat{\mathbf{C}}\|\|\mathbf{R}^{-1}\|}{1 - \Delta_2\|\widehat{\mathbf{C}}\|}(\|\widehat{\mathbf{C}}\|\Delta_2 + \delta_H)\|\mathbf{y}\|.$$

Proof. The proofs are similar to the proofs of Proposition A.2. To prove a), we follow the proof of Proposition A.2 a) and find that $\|\boldsymbol{\Omega} - \boldsymbol{\Omega}_{\rm loc}\| \le \delta_\Omega$, and (20)

$$\|\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} - \mathbf{H}_{\rm loc}^T\mathbf{R}^{-1}\mathbf{H}_{\rm loc}\| \le (2\delta_H + \delta_H^2)\|\mathbf{R}^{-1}\|.$$

Therefore

$$\|(\boldsymbol{\Omega} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) - (\boldsymbol{\Omega}_{\rm loc} + \mathbf{H}_{\rm loc}^T\mathbf{R}^{-1}\mathbf{H}_{\rm loc})\| \le \delta_\Omega + (2\delta_H + \delta_H^2)\|\mathbf{R}^{-1}\| =: \Delta_2.$$

We apply Lemma A.1 to $\widehat{\mathbf{C}} = (\mathbf{\Omega} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}$,

$$\|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\mathrm{p,loc}}\| \leq \frac{\|\widehat{\mathbf{C}}\|^2\Delta_2}{1-\Delta_2\|\widehat{\mathbf{C}}\|}.$$

As a consequence, we have that $\|\widehat{\mathbf{C}}_{\mathrm{p,loc}}\| \leq \frac{\|\widehat{\mathbf{C}}\|}{1-\Delta_2\|\widehat{\mathbf{C}}\|}$.

To prove claim b), we write

$$\hat{\mathbf{m}}_{\mathrm{p,loc}} - \hat{\mathbf{m}} = (\widehat{\mathbf{C}}_{\mathrm{p,loc}}\mathbf{\Omega}_{\mathrm{loc}} - \widehat{\mathbf{C}}\mathbf{\Omega})\mathbf{m} + (\widehat{\mathbf{C}}_{\mathrm{p,loc}}\mathbf{H}_{\mathrm{loc}}^T - \widehat{\mathbf{C}}\mathbf{H}^T)\mathbf{R}^{-1}\mathbf{y}.$$

The norm of the matrix $(\widehat{\mathbf{C}}_{\mathrm{p,loc}}\mathbf{H}_{\mathrm{loc}}^T - \widehat{\mathbf{C}}\mathbf{H}^T)\mathbf{R}^{-1}$ can be bounded directly using claim a)

$$\begin{aligned}\|(\widehat{\mathbf{C}}_{\mathrm{p,loc}}\mathbf{H}_{\mathrm{loc}}^T - \widehat{\mathbf{C}}\mathbf{H}^T)\mathbf{R}^{-1}\| &\leq \|\widehat{\mathbf{C}} - \widehat{\mathbf{C}}_{\mathrm{p,loc}}\|\|\mathbf{H}^T\mathbf{R}^{-1}\| + \|\widehat{\mathbf{C}}_{\mathrm{p,loc}}\|\|\mathbf{H}_{\mathrm{loc}} - \mathbf{H}\|\|\mathbf{R}^{-1}\| \\ &\leq \frac{\|\widehat{\mathbf{C}}\|\|\mathbf{R}^{-1}\|}{1-\Delta_2\|\widehat{\mathbf{C}}\|}(\|\widehat{\mathbf{C}}\|\Delta_2 + \delta_H).\end{aligned}$$

The norm of the matrix $\widehat{\mathbf{C}}_{\mathrm{p,loc}}\mathbf{\Omega}_{\mathrm{loc}} - \widehat{\mathbf{C}}\mathbf{\Omega}$ can be bounded by

$$\|\widehat{\mathbf{C}}_{\mathrm{p,loc}}\mathbf{\Omega}_{\mathrm{loc}} - \widehat{\mathbf{C}}\mathbf{\Omega}\| \leq \|\mathbf{\Omega}_{\mathrm{loc}} - \mathbf{\Omega}\|\|\widehat{\mathbf{C}}_{\mathrm{p,loc}}\| + \|\widehat{\mathbf{C}}_{\mathrm{p,loc}} - \widehat{\mathbf{C}}\|\|\mathbf{\Omega}\|.$$

From part (a), we have

$$\|\mathbf{\Omega}_{\mathrm{loc}} - \mathbf{\Omega}\|\|\widehat{\mathbf{C}}_{\mathrm{p,loc}}\| \leq \frac{\|\widehat{\mathbf{C}}\|\delta_\Omega}{1-\Delta_2\|\widehat{\mathbf{C}}\|}, \quad \|\widehat{\mathbf{C}}_{\mathrm{p,loc}} - \widehat{\mathbf{C}}\|\|\mathbf{\Omega}\| \leq \frac{\|\mathbf{\Omega}\|\|\widehat{\mathbf{C}}\|^2\Delta_2}{1-\Delta_2\|\widehat{\mathbf{C}}\|}.$$

In summary:

$$\|\hat{\mathbf{m}}_{\mathrm{p,loc}} - \hat{\mathbf{m}}\| \leq \frac{\|\widehat{\mathbf{C}}\|\delta_\Omega + \|\mathbf{\Omega}\|\|\widehat{\mathbf{C}}\|^2\Delta_2}{1-\Delta_2\|\widehat{\mathbf{C}}\|}\|\mathbf{m}\| + \frac{\|\widehat{\mathbf{C}}\|\|\mathbf{R}^{-1}\|}{1-\Delta_2\|\widehat{\mathbf{C}}\|}(\|\widehat{\mathbf{C}}\|\Delta_2 + \delta_H)\|\mathbf{y}\|. \quad \square$$

Appendix B. Convergence rates for Gibbs samplers

In this section, we prove Theorems 3.1 and 3.2.

B.1. Review of Gibbs and Gauss–Seidel

Our strategy for proving these theorems relies on the Gauss–Seidel operator associated with the Gibbs sampler. The connection between the Gibbs sampler and the Gauss–Seidel operator is well documented [21,25,53], but we briefly review it here so that our paper can be read and understood independently.

We consider sampling $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Omega}^{-1})$ using the Gibbs sampler of block size q, assuming $\mathbf{\Omega}$ is of dimension $mq \times mq$. We will use $\mathbf{\Omega}_{i,j}$ to denote its (i,j)-$q \times q$ sub-block, which should not be confused with the matrix entry $[\mathbf{\Omega}]_{i,j}$. We say an $mq \times mq$ matrix $\mathbf{\Omega}$ is block-strictly-lower-triangular (BSLT), if $\mathbf{\Omega}_{i,j}$ is nonzero only if $i > j$. We define block-diagonal and block-strictly-upper-triangular (BSUT) in an analogous way. Let $\mathbf{\Omega} = \mathbf{\Omega}_L + \mathbf{\Omega}_D + \mathbf{\Omega}_U$ be the BSLT + block-diagonal + BSUT decomposition of the precision matrix $\mathbf{\Omega}$. In other words, $\mathbf{\Omega}_L$ consists of the blocks $\{\mathbf{\Omega}_{i,j}\}_{i>j}$ of $\mathbf{\Omega}$, $\mathbf{\Omega}_D$ consists of the diagonal blocks of $\mathbf{\Omega}$, and $\mathbf{\Omega}_U = \mathbf{\Omega}_L^T$.

We investigate how the j-th coordinate is updated at the $k+1$-th iteration. For simplicity, let us write the state before this coordinate update as

$$(\mathbf{z}_1, \cdots, \mathbf{z}_m) = (\mathbf{x}_1^{k+1}, \cdots, \mathbf{x}_{j-1}^{k+1}, \mathbf{x}_j^k, \cdots, \mathbf{x}_m^k).$$

Then $\log(p(\mathbf{z}))$, ignoring a constant term, can be written as

$$\begin{aligned}-\frac{1}{2}(\mathbf{z}-\mathbf{m})^T\mathbf{\Omega}(\mathbf{z}-\mathbf{m}) &= -\frac{1}{2}\sum_{i,i'}(\mathbf{z}_i - \mathbf{m}_i)^T\mathbf{\Omega}_{i,i'}(\mathbf{z}_{i'} - \mathbf{m}_{i'}) \\ &= -\frac{1}{2}(\mathbf{z}_j - \mathbf{m}_j)^T\mathbf{\Omega}_{j,j}(\mathbf{z}_j - \mathbf{m}_j) - \frac{1}{2}(\mathbf{z}_j - \mathbf{m}_j)^T\sum_{i\neq j}\mathbf{\Omega}_{j,i}(\mathbf{z}_i - \mathbf{m}_i) \\ &\quad - \frac{1}{2}\sum_{i\neq j}(\mathbf{z}_i - \mathbf{m}_i)^T\mathbf{\Omega}_{i,j}(\mathbf{z}_j - \mathbf{m}_j) + F_j'(\mathbf{z}) \\ &= -\frac{1}{2}(\mathbf{z}_j - \mathbf{m}_j')^T\mathbf{\Omega}_{j,j}(\mathbf{z}_j - \mathbf{m}_j')^T + F_j(\mathbf{z}).\end{aligned}$$

Here F_j and F_j' are functions independent of $\mathbf{z}_j$, and

$$\mathbf{m}_j' := \mathbf{m}_j - \sum_{i\neq j} \mathbf{\Omega}_{j,j}^{-1}\mathbf{\Omega}_{j,i}(\mathbf{z}_i - \mathbf{m}_i) = \mathbf{m}_j - \sum_{i<j} \mathbf{\Omega}_{j,j}^{-1}\mathbf{\Omega}_{j,i}(\mathbf{x}_i^{k+1} - \mathbf{m}_i) - \sum_{i>j} \mathbf{\Omega}_{j,j}^{-1}\mathbf{\Omega}_{j,i}(\mathbf{x}_i^{k} - \mathbf{m}_i).$$

Using Bayes' formula, we find that the probability of $\mathbf{x}_j^{k+1}$ conditioned on $\mathbf{x}_i^{k+1}$, $i<j$ and $\mathbf{x}_i^k$, $i>j$ is proportional to $\exp(-\frac{1}{2}(\mathbf{z}_j - \mathbf{m}_j')^T\mathbf{\Omega}_{j,j}(\mathbf{z}_j - \mathbf{m}_j'))$. In other words, the updated coordinate has the distribution

$$\mathbf{x}_j^{k+1} \sim \mathcal{N}\left(\mathbf{m}_j - \sum_{i<j} \mathbf{\Omega}_{j,j}^{-1}\mathbf{\Omega}_{j,i}(\mathbf{x}_i^{k+1} - \mathbf{m}_i) - \sum_{i>j} \mathbf{\Omega}_{j,j}^{-1}\mathbf{\Omega}_{j,i}(\mathbf{x}_i^{k} - \mathbf{m}_i), \mathbf{\Omega}_{j,j}^{-1}\right). \tag{21}$$

One way to obtain a sample of this distribution is to find the solution, $\mathbf{x}_j^{k+1}$, of the linear equation

$$\mathbf{\Omega}_{j,j}(\mathbf{x}_j^{k+1} - \mathbf{m}_j) + \sum_{i<j} \mathbf{\Omega}_{j,i}(\mathbf{x}_i^{k+1} - \mathbf{m}_i) + \sum_{i>j} \mathbf{\Omega}_{j,i}(\mathbf{x}_i^{k} - \mathbf{m}_i) = \xi_j^{k+1} \tag{22}$$

where ξ_j^{k+1} is an independent sample from $\mathcal{N}(0, \mathbf{\Omega}_{j,j})$.

Let $\mathbf{\Omega} = \mathbf{\Omega}_L + \mathbf{\Omega}_D + \mathbf{\Omega}_U$ be the BSLT + block-diagonal + BSUT decomposition of the precision matrix $\mathbf{\Omega}$. If we concatenate (22) for all coordinates j, the equation becomes

$$(\mathbf{\Omega}_L + \mathbf{\Omega}_D)(\mathbf{x}^{k+1} - \mathbf{m}) + \mathbf{\Omega}_U(\mathbf{x}^k - \mathbf{m}) = \xi^{k+1},$$

where ξ^{k+1} is the concatenation of ξ_j^{k+1}, and distributed as $\xi^{k+1} \sim \mathcal{N}(0, \mathbf{\Omega}_D)$. An equivalent representation is

$$(\mathbf{x}^{k+1} - \mathbf{m}) = -(\mathbf{\Omega}_L + \mathbf{\Omega}_D)^{-1}\mathbf{\Omega}_U(\mathbf{x}^k - \mathbf{m}) + (\mathbf{\Omega}_L + \mathbf{\Omega}_D)^{-1}\xi^{k+1}. \tag{23}$$

The correlation between two consecutive iterations then is determined by

$$\mathbf{G} := -(\mathbf{\Omega}_L + \mathbf{\Omega}_D)^{-1}\mathbf{\Omega}_U.$$

The spectral radius of $\mathbf{G}$ decides how fast the Gibbs sampler converges in l_2 norm.

The matrix $\mathbf{G}$ is known as the "Gauss–Seidel operator," which is also used in the iterative solution of linear equations. Recall that $\mathcal{N}(0, \mathbf{C})$ is the invariant distribution of $\mathbf{x}^k - \mathbf{m}$. By comparing covariances on both sides of (23), we find that

$$\mathbf{C} = \mathbf{G}\mathbf{C}\mathbf{G}^T + (\mathbf{\Omega}_L + \mathbf{\Omega}_D)^{-1}\mathbf{\Omega}_D(\mathbf{\Omega}_L + \mathbf{\Omega}_D)^{-T}. \tag{24}$$

It follows that $\mathbf{C} \preceq \mathbf{G}\mathbf{C}\mathbf{G}^T$, which in turn implies that the spectral radius of $\mathbf{G}$ is less than 1, which implies convergence. Here and below, for two symmetric matrices $\mathbf{A}$ and $\mathbf{B}$, we use $\mathbf{A} \preceq \mathbf{B}$ to indicate that $\mathbf{B} - \mathbf{A}$ is positive semidefinite. However, in order to show the convergence rate is dimension-independent, we need to exploit the banded structure of $\mathbf{C}$ or $\mathbf{\Omega}$, which will be the purpose of the next section.

B.2. Gauss–Seidel with localized structures

First we need two estimates.

Lemma B.1. *For any positive definite $qm \times qm$ matrix $\mathbf{C}$, denote its maximum eigenvalue as λ_{max}, its minimum eigenvalue as λ_{min}, its condition number as $\mathcal{C} = \lambda_{max}/\lambda_{min}$, its inverse $\mathbf{\Omega} = \mathbf{C}^{-1}$. Then the $q \times q$ blocks satisfy:*

$$\lambda_{min}\mathbf{I} \preceq \mathbf{C}_{i,i} \preceq \lambda_{max}\mathbf{I}, \quad \lambda_{max}^{-1}\mathbf{I} \preceq \mathbf{\Omega}_{i,i} \preceq \lambda_{min}^{-1}\mathbf{I}, \tag{25}$$

$$\|\mathbf{\Omega}_{i,i}^{-1/2}\mathbf{\Omega}_{i,j}\mathbf{\Omega}_{j,j}^{-1/2}\| \leq 1 - \mathcal{C}^{-1}, \quad \|\mathbf{C}_{i,i}^{-1/2}\mathbf{C}_{i,j}\mathbf{C}_{j,j}^{-1/2}\| \leq 1 - \mathcal{C}^{-1}. \tag{26}$$

Proof. Let λ_i be an eigenvalue of $\mathbf{C}_{i,i}$ and $v \in \mathbb{R}^q$ be one of its eigenvectors with norm 1. Let $\mathbf{v}$ be the $\mathbb{R}^{qm}$ vector with its i-th block being v. Then

$$\lambda_i = v^T\mathbf{C}_{i,i}v = \mathbf{v}^T\mathbf{C}\mathbf{v} \in [\lambda_{\min}, \lambda_{\max}].$$

The left inequality in (25) follows. The right inequality of equation (25) can be derived in a similar fashion.

Let $\mathbf{x}$ and $\mathbf{y}$ be the left and right singular vectors corresponding to the largest singular value of $\Gamma_{i,j} = \mathbf{\Omega}_{i,i}^{-1/2}\mathbf{\Omega}_{i,j}\mathbf{\Omega}_{j,j}^{-1/2}$. The vectors $\mathbf{x}$ and $\mathbf{y}$ are of dimension q, have norm one, $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$, and $\mathbf{x}^T\Gamma_{i,j}\mathbf{y} = \|\Gamma_{i,j}\|$. Now consider an qm dimensional vector $\mathbf{v}$, where its i-th block is $\mathbf{\Omega}_{i,i}^{-1/2}\mathbf{x}$, its j-th block is $-\mathbf{\Omega}_{j,j}^{-1/2}\mathbf{y}$, and all other blocks are zero. Then

$$\mathbf{v}^T\boldsymbol{\Omega}\mathbf{v} = \mathbf{x}^T\boldsymbol{\Omega}_{i,i}^{-1/2}\boldsymbol{\Omega}_{i,i}\boldsymbol{\Omega}_{i,i}^{-1/2}\mathbf{x} - 2\mathbf{x}^T\boldsymbol{\Omega}_{i,i}^{-1/2}\boldsymbol{\Omega}_{i,j}\boldsymbol{\Omega}_{j,j}^{-1/2}\mathbf{y} + \mathbf{y}^T\boldsymbol{\Omega}_{j,j}^{-1/2}\boldsymbol{\Omega}_{j,j}\boldsymbol{\Omega}_{j,j}^{-1/2}\mathbf{y}$$
$$= 2 - 2\mathbf{x}^T\Gamma_{i,j}\mathbf{y} = 2(1 - \|\Gamma_{i,j}\|).$$

On the other hand,

$$\|\mathbf{v}\|^2 = \|\boldsymbol{\Omega}_{i,i}^{-1/2}\mathbf{x}\|^2 + \|\boldsymbol{\Omega}_{j,j}^{-1/2}\mathbf{y}\|^2 \geq 2\lambda_{\min}.$$

Thus

$$2(1 - \|\Gamma_{i,j}\|) = 2\mathbf{v}^T\boldsymbol{\Omega}\mathbf{v} \geq 2\lambda_{\min}\lambda_{\max}^{-1} = 2\mathcal{C}^{-1}.$$

The left inequality in (26) follows. Since the derivation above uses nothing of $\boldsymbol{\Omega}$ other than its eigenvalues, so the right inequality in (26) also holds. □

For our proofs below, we need the following bound for an operator norm. For $q = 1$, this bound appeared in [6] as inequality (A2). Note that this bound is well-suited for block-sparse $\mathbf{A}$ since then the right hand side consists of only a few terms.

Lemma B.2. *For any $qm \times qm$ matrix $\mathbf{A}$, the following holds*

$$\|\mathbf{A}\| \leq \left(\max_{i=1,\cdots,m}\sum_{j=1}^{m}\|\mathbf{A}_{i,j}\|\right)^{1/2}\left(\max_{i=1,\cdots,m}\sum_{j=1}^{m}\|\mathbf{A}_{j,i}\|\right)^{1/2}.$$

Proof. First we show the claim for symmetric $\mathbf{A} = \mathbf{A}^T$. In this case, the bound for the norm of $\mathbf{A}$ is

$$\|\mathbf{A}\| \leq \max_{i=1,\cdots,m}\sum_{j=1}^{m}\|\mathbf{A}_{i,j}\|. \tag{27}$$

Let $\mathbf{v} = [\mathbf{v}_1, \ldots, \mathbf{v}_m] \in \mathbb{R}^{qm}$ be an eigenvector so that $\lambda\mathbf{v} = \mathbf{A}\mathbf{v}$ while $|\lambda| = \|\mathbf{A}\|$. Suppose $\|\mathbf{v}_{i_*}\| = \max_i\{\|\mathbf{v}_i\|\}$, then

$$|\lambda|\|\mathbf{v}_{i_*}\| = \|\lambda\mathbf{v}_{i_*}\| = \left\|\sum_{j=1}^{m}\mathbf{A}_{i_*,j}\mathbf{v}_j\right\| \leq \sum_{j=1}^{m}\|\mathbf{A}_{i_*,j}\|\|\mathbf{v}_j\| \leq \|\mathbf{v}_{i_*}\|\sum_{j=1}^{m}\|\mathbf{A}_{i_*,j}\|.$$

This leads to $\|\mathbf{A}\| \leq \sum_{j=1}^{m}\|\mathbf{A}_{i_*,j}\|$ and hence (27).

For a general $\mathbf{A}$, note that $\|\mathbf{A}\| = \|\mathbf{A}^T\mathbf{A}\|^{1/2}$. The matrix $\mathbf{P} = \mathbf{A}^T\mathbf{A}$ is symmetric, and its blocks are

$$\mathbf{P}_{i,j} = \sum_k \mathbf{A}_{k,i}^T\mathbf{A}_{k,j}.$$

Applying (27) to $\mathbf{P}$, we obtain

$$\begin{aligned}\|\mathbf{P}\| \leq \max_i\sum_{j=1}^{m}\|\mathbf{P}_{i,j}\| &\leq \max_i\sum_j\sum_k\|\mathbf{A}_{k,i}^T\mathbf{A}_{k,j}\| \\ &\leq \max_i\sum_k\sum_j\|\mathbf{A}_{k,i}\|\|\mathbf{A}_{k,j}\| \\ &\leq \max_i\sum_k\|\mathbf{A}_{k,i}\|\sum_j\|\mathbf{A}_{k,j}\| \\ &\leq \left(\max_i\sum_k\|\mathbf{A}_{k,i}\|\right)\left(\max_i\sum_j\|\mathbf{A}_{i,j}\|\right).\end{aligned}$$

This leads to our general claim. □

Now we are at the position to establish bounds for the Gauss Seidel operator.

Lemma B.3. *If $\boldsymbol{\Omega}$ is block-tridiagonal with $\mathcal{C}$ being its condition number, then*

$$\mathbf{G}\mathbf{C}\mathbf{G}^T \preceq \frac{\mathcal{C}(1-\mathcal{C}^{-1})^2}{1+\mathcal{C}(1-\mathcal{C}^{-1})^2}\mathbf{C}.$$

Proof. Since $\mathbf{\Omega}$ is block-tridiagonal, $\mathbf{\Omega}_U$ has at most one nonzero block in each row and each column, and, likewise, $\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2}$ has at most one nonzero block in each row and each column. Therefore, by Lemma B.2

$$\left\|\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2}\right\| \le \max_{i=1,\dots,m-1}\left\|\mathbf{\Omega}_{i,i}^{-1/2}\mathbf{\Omega}_{i,i+1}\mathbf{\Omega}_{i+1,i+1}^{-1/2}\right\|,$$

which by Lemma B.1 is bounded by $1-\mathcal{C}^{-1}$.

To continue, we look at the right hand side of (24). We want to show that for some $\gamma > 0$,

$$\mathbf{G}\mathbf{C}\mathbf{G}^T \preceq \gamma(\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-1}\mathbf{\Omega}_D(\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-T} \tag{28}$$

For this purpose, note that

$$\mathbf{G}\mathbf{C}\mathbf{G}^T = (\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-1}\mathbf{\Omega}_U\mathbf{C}\mathbf{\Omega}_U^T(\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-T},$$

and that

$$\mathbf{\Omega}_U\mathbf{C}\mathbf{\Omega}_U^T = \mathbf{\Omega}_D^{1/2}(\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2})(\mathbf{\Omega}_D^{1/2}\mathbf{C}\mathbf{\Omega}_D^{1/2})(\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2})^T\mathbf{\Omega}_D^{1/2}.$$

Thus, in order to prove (28), it suffices to find a γ such that

$$\left\|(\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2})(\mathbf{\Omega}_D^{1/2}\mathbf{C}\mathbf{\Omega}_D^{1/2})(\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2})^T\right\| \le \gamma.$$

By Lemma B.1, this is straight forward, since we have that

$$\left\|(\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2})(\mathbf{\Omega}_D^{1/2}\mathbf{C}\mathbf{\Omega}_D^{1/2})(\mathbf{\Omega}_D^{-1/2}\mathbf{\Omega}_U\mathbf{\Omega}_D^{-1/2})^T\right\| \le (1-\mathcal{C}^{-1})^2\|\mathbf{\Omega}_D\|\|\mathbf{C}\| \le \mathcal{C}(1-\mathcal{C}^{-1})^2 =: \gamma,$$

since $\mathbf{\Omega}_D$ is diagonal with blocks bounded in operator norm by $\lambda_{\min}^{-1}$.

Finally, we can plug (28) into (24), and find that

$$\mathbf{C} = \mathbf{G}\mathbf{C}\mathbf{G}^T + (\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-1}\mathbf{\Omega}_D(\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-T} \succeq (\gamma^{-1}+1)\mathbf{G}\mathbf{C}\mathbf{G}^T. \quad \square$$

Lemma B.4. *If* $\mathbf{C}$ *is block-tridiagonal with* $\mathcal{C}$ *being its condition number, then*

$$\mathbf{G}\mathbf{C}\mathbf{G}^T \preceq \frac{2(1-\mathcal{C}^{-1})^2\mathcal{C}^4}{1+2(1-\mathcal{C}^{-1})^2\mathcal{C}^4}\mathbf{C}.$$

Proof. Since $\mathbf{\Omega}^{-1} = \mathbf{C}$, we apply the Woodbury's formula to $(\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-1} = (\mathbf{\Omega}-\mathbf{\Omega}_U)^{-1}$,

$$(\mathbf{\Omega}_L+\mathbf{\Omega}_D)^{-1} = (\mathbf{\Omega}-\mathbf{\Omega}_U)^{-1} = \mathbf{C}+\mathbf{C}\mathbf{\Omega}_U(\mathbf{I}-\mathbf{C}\mathbf{\Omega}_U)^{-1}\mathbf{C} = (\mathbf{I}-\mathbf{C}\mathbf{\Omega}_U)^{-1}\mathbf{C}.$$

Consequentially, $\mathbf{G} = (\mathbf{I}-\mathbf{C}\mathbf{\Omega}_U)^{-1}\mathbf{C}\mathbf{\Omega}_U$.

Next, we claim that $\mathbf{C}\mathbf{\Omega}_U$ is BUT, and that only the blocks $(\mathbf{C}\mathbf{\Omega}_U)_{i,i}$, $(\mathbf{C}\mathbf{\Omega}_U)_{i,i+1}$ are nonzero. To see it is BUT, recall that $\mathbf{C}$ is block-tridiagonal, $(\mathbf{\Omega}_U)_{i,j} = \mathbf{\Omega}_{i,j}\mathbb{1}_{i\le j-1}$,

$$(\mathbf{C}\mathbf{\Omega}_U)_{i,j} = \mathbf{C}_{i,i-1}\mathbf{\Omega}_{i-1,j}\mathbb{1}_{i\le j} + \mathbf{C}_{i,i}\mathbf{\Omega}_{i,j}\mathbb{1}_{i\le j-1} + \mathbf{C}_{i,i+1}\mathbf{\Omega}_{i+1,j}\mathbb{1}_{i\le j-2}.$$

Thus $(\mathbf{C}\mathbf{\Omega}_U)_{i,j}$ is nonzero only if $i \le j$, i.e., $\mathbf{C}\mathbf{\Omega}_U$ is BUT. Moreover, if $j \ge i+2$, by the identity $\mathbf{C}\mathbf{\Omega} = \mathbf{I}$, we have

$$\mathbf{0} = (\mathbf{C}\mathbf{\Omega})_{i,j} = \mathbf{C}_{i,i-1}\mathbf{\Omega}_{i-1,j} + \mathbf{C}_{i,i}\mathbf{\Omega}_{i,j} + \mathbf{C}_{i,i+1}\mathbf{\Omega}_{i+1,j} = (\mathbf{C}\mathbf{\Omega}_U)_{i,j},$$

proving our claim.

Next, not that for $j = i$, we have

$$(\mathbf{C}\mathbf{\Omega}_U)_{i,i} = \mathbf{C}_{i,i-1}\mathbf{\Omega}_{i-1,i}\mathbb{1}_{i\le i} + \mathbf{C}_{i,i}\mathbf{\Omega}_{i,i}\mathbb{1}_{i\le i-1} + \mathbf{C}_{i,i+1}\mathbf{\Omega}_{i+1,i}\mathbb{1}_{i\le i-2} = \mathbf{C}_{i,i-1}\mathbf{\Omega}_{i-1,i}.$$

Likewise, for $j = i+1$, we have

$$\mathbf{0} = (\mathbf{C}\mathbf{\Omega})_{i,i+1} = \mathbf{C}_{i,i-1}\mathbf{\Omega}_{i-1,i+1} + \mathbf{C}_{i,i}\mathbf{\Omega}_{i,i+1} + \mathbf{C}_{i,i+1}\mathbf{\Omega}_{i+1,i+1} = (\mathbf{C}\mathbf{\Omega}_U)_{i,i+1} + \mathbf{C}_{i,i+1}\mathbf{\Omega}_{i+1,i+1}.$$

In other words, $(\mathbf{C}\mathbf{\Omega}_U)_{i,i+1} = -\mathbf{C}_{i,i+1}\mathbf{\Omega}_{i+1,i+1}$.

Applying Lemma B.2 leads to

$$\|\mathbf{C}\mathbf{\Omega}_U\| \le \left(\max_{i=1,\cdots,m}\{\|(\mathbf{C}\mathbf{\Omega}_U)_{i,i}\| + \|(\mathbf{C}\mathbf{\Omega}_U)_{i,i+1}\|\}\right)^{1/2}\left(\max_{i=1,\cdots,m}\|(\mathbf{C}\mathbf{\Omega}_U)_{i,i}\|\right)^{1/2}.$$

Applying Lemma B.1 to $(\mathbf{C}\boldsymbol{\Omega}_U)_{i,i} = \mathbf{C}_{i,i-1}\boldsymbol{\Omega}_{i-1,i}$ gives

$$\begin{aligned}\|(\mathbf{C}\boldsymbol{\Omega}_U)_{i,i}\| &\le \|\mathbf{C}_{i,i}^{1/2}\|\|\mathbf{C}_{i,i}^{-1/2}\mathbf{C}_{i,i-1}\mathbf{C}_{i-1,i-1}^{-1/2}\|\|\mathbf{C}_{i-1,i-1}^{1/2}\|\|\boldsymbol{\Omega}_{i-1,i-1}^{1/2}\|\|\boldsymbol{\Omega}_{i-1,i-1}^{-1/2}\boldsymbol{\Omega}_{i-1,i}\boldsymbol{\Omega}_{i,i}^{-1/2}\|\|\boldsymbol{\Omega}_{i,i}^{1/2}\| \\ &\le (1-\mathcal{C}^{-1})^2\mathcal{C} \le (1-\mathcal{C}^{-1})\mathcal{C}.\end{aligned}$$

Similarly, $(\mathbf{C}\boldsymbol{\Omega}_U)_{i,i+1} = -\mathbf{C}_{i,i+1}\boldsymbol{\Omega}_{i+1,i+1}$, which by Lemma B.1 implies that

$$\|(\mathbf{C}\boldsymbol{\Omega}_U)_{i,i+1}\| \le \|\mathbf{C}_{i,i}^{1/2}\|\|\mathbf{C}_{i,i}^{-1/2}\mathbf{C}_{i,i+1}\mathbf{C}_{i+1,i+1}^{-1/2}\|\|\mathbf{C}_{i+1,i+1}^{1/2}\|\|\boldsymbol{\Omega}_{i+1,i+1}\| \le (1-\mathcal{C}^{-1})\mathcal{C}.$$

Consequentially, another application of Lemma B.2 implies that

$$\|\mathbf{C}\boldsymbol{\Omega}_U\| \le \sqrt{2}(1-\mathcal{C}^{-1})\mathcal{C}.$$

To continue, we again want to use (24) by exploiting relations like (28). We first note that

$$\begin{aligned}\mathbf{G}\mathbf{C}\mathbf{G}^T &= (\mathbf{I}-\mathbf{C}\boldsymbol{\Omega}_U)^{-1}\mathbf{C}\boldsymbol{\Omega}_U\mathbf{C}\boldsymbol{\Omega}_U^T\mathbf{C}(\mathbf{I}-\mathbf{C}\boldsymbol{\Omega}_U)^{-T}, \\ (\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-1}\boldsymbol{\Omega}_D(\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-T} &= (\mathbf{I}-\mathbf{C}\boldsymbol{\Omega}_U)^{-1}\mathbf{C}\boldsymbol{\Omega}_D\mathbf{C}(\mathbf{I}-\mathbf{C}\boldsymbol{\Omega}_U)^{-T}.\end{aligned}$$

Using $\|\mathbf{C}\boldsymbol{\Omega}_U\| \le \sqrt{2}(1-\mathcal{C}^{-1})\mathcal{C}$, we have

$$\mathbf{C}\boldsymbol{\Omega}_U\mathbf{C}\boldsymbol{\Omega}_U^T\mathbf{C} \preceq 2(1-\mathcal{C}^{-1})^2\mathcal{C}^2\lambda_{\max}\mathbf{I}.$$

Moreover,

$$\mathbf{C}\boldsymbol{\Omega}_D\mathbf{C} \succeq \lambda_{\max}^{-1}\mathbf{C}\mathbf{C}\mathbf{I} \succeq \lambda_{\max}^{-1}\lambda_{\min}^2\mathbf{I}.$$

Consequentially, $\mathbf{C}\boldsymbol{\Omega}_U\mathbf{C}\boldsymbol{\Omega}_U^T\mathbf{C} \preceq 2(1-\mathcal{C}^{-1})^2\mathcal{C}^4\mathbf{C}\boldsymbol{\Omega}_D\mathbf{C}$ and

$$\mathbf{G}\mathbf{C}\mathbf{G}^T \preceq 2(1-\mathcal{C}^{-1})^2\mathcal{C}^4(\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-1}\boldsymbol{\Omega}_D(\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-T}.$$

Combining the above inequality with (24) leads to

$$\mathbf{C} = \mathbf{G}\mathbf{C}\mathbf{G}^T + (\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-1}\boldsymbol{\Omega}_D(\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-T} \succeq ((2(1-\mathcal{C}^{-1})^2\mathcal{C}^4)^{-1}+1)\mathbf{G}\mathbf{C}\mathbf{G}^T. \quad \square$$

B.3. Proofs of the main theorems

Armed with these results, the proofs for the main theorems follow from an elementary coupling argument.

Proof of Theorems 3.1 and 3.2. As discussed above, and illustrated in Section B.1, we can generate iterates from the Gibbs sampler with block-size q by solving the linear equations

$$(\mathbf{x}^{k+1}-\mathbf{m}) = \mathbf{G}(\mathbf{x}^k-\mathbf{m}) + (\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-1}\xi^{k+1} \quad k=0,1,\cdots,$$

where ξ^{k+1} are i.i.d. samples from $\mathcal{N}(0,\boldsymbol{\Omega}_D)$.

Next we consider a random sample $\mathbf{z}^0$ from $\mathcal{N}(\mathbf{m},\mathbf{C})$. We can apply the block-Gibbs sampler with $\mathbf{z}^0$ as the initial condition, while using the same sequence ξ^k. In other words, we generate $\mathbf{z}^{k+1}$ by letting

$$(\mathbf{z}^{k+1}-\mathbf{m}) = \mathbf{G}(\mathbf{z}^k-\mathbf{m}) + (\boldsymbol{\Omega}_L+\boldsymbol{\Omega}_D)^{-1}\xi^{k+1} \quad k=0,1,\cdots.$$

Since $\mathcal{N}(\mathbf{m},\mathbf{C})$ is the invariant measure for the Gibbs sampler, marginally $\mathbf{z}^k \sim \mathcal{N}(\mathbf{m},\mathbf{C})$.

Next we look at the difference between the two Gibbs samplers, $\Delta^k = \mathbf{x}^k - \mathbf{z}^k$. Note that

$$\Delta^0 \sim \mathcal{N}(\mathbf{x}^0-\mathbf{m},\mathbf{C}), \quad \Delta^{k+1} = \mathbf{G}\Delta^k, \quad k=0,1\cdots.$$

Consequentially, $\Delta^k \sim \mathcal{N}(\mathbf{G}^k\Delta_0, \mathbf{G}^k\mathbf{C}(\mathbf{G}^T)^k)$, where $\Delta_0 := \mathbf{x}^0-\mathbf{m}$. Since

$$\Delta^0\Delta^{0T} = \mathbf{C}^{1/2}(\mathbf{C}^{-1/2}\Delta_0)(\mathbf{C}^{-1/2}\Delta_0)^T\mathbf{C}^{1/2} \preceq \|\mathbf{C}^{-1/2}\Delta_0\|^2\mathbf{C},$$

we find that

$$\begin{aligned}\mathbb{E}\|\mathbf{C}^{-1/2}\Delta^k\|^2 &= \|\mathbf{C}^{-1/2}\mathbf{G}^k\Delta_0\|^2 + \mathrm{tr}(\mathbf{C}^{-1/2}\mathbf{G}^k\mathbf{C}(\mathbf{G}^T)^k\mathbf{C}^{-1/2}) \\ &= \mathrm{tr}(\mathbf{C}^{-1/2}\mathbf{G}^k\Delta_0\Delta_0^T(\mathbf{G}^T)^k\mathbf{C}^{-1/2} + \mathbf{C}^{-1/2}\mathbf{G}^k\mathbf{C}(\mathbf{G}^T)^k\mathbf{C}^{-1/2}) \\ &\le (1+\|\mathbf{C}^{-1/2}\Delta_0\|^2)\cdot\mathrm{tr}(\mathbf{C}^{-1/2}\mathbf{G}^k\mathbf{C}(\mathbf{G}^T)^k\mathbf{C}^{-1/2}). \qquad (29)\end{aligned}$$

If $\mathbf{\Omega}$ is block-tridiagonal, then, by Lemma B.3,

$$\mathbf{C}^{-1/2}\mathbf{G}^k\mathbf{C}(\mathbf{G}^T)^k\mathbf{C}^{-1/2} \leq \beta^k\mathbf{I}, \quad \beta = \frac{\mathcal{C}(1-\mathcal{C}^{-1})^2}{1+\mathcal{C}(1-\mathcal{C}^{-1})^2}.$$

Combining the above inequality with (29) proves Theorem 3.2.

If $\mathbf{C}$ is block-tridiagonal, then, by Lemma B.4,

$$\mathbf{C}^{-1/2}\mathbf{G}^k\mathbf{C}(\mathbf{G}^T)^k\mathbf{C}^{-1/2} \leq \beta^k\mathbf{I}, \quad \beta = \frac{2(1-\mathcal{C}^{-1})^2\mathcal{C}^4}{1+2(1-\mathcal{C}^{-1})^2\mathcal{C}^4}.$$

Combining the above inequality with (29) proves Theorem 3.1. □

References

[1] S. Agapiou, O. Papaspiliopoulos, D. Sanz-Alonso, A. Stuart, Importance sampling: computational complexity and intrinsic dimension, Stat. Sci. 32 (2017) 405–431.
[2] T. Auligné, B. Ménétrier, A. Lorenc, M. Buehner, Ensemble-variational integrated localized data assimilation, Mon. Weather Rev. 144 (2016) 3677–3696.
[3] J. Bardsley, MCMC-based image reconstruction with uncertainty quantification, SIAM J. Sci. Comput. 34 (2012) A1316–A1332.
[4] A. Beskos, N. Pillai, G. Roberts, J.-M. Sanz-Serna, A. Stuart, Optimal tuning of the hybrid Monte Carlo algorithm, Bernoulli 19 (2013) 1501–1534.
[5] A. Beskos, G. Roberts, A. Stuart, Optimal scalings for local Metropolis–Hastings chains on nonproduct targets in high dimensions, Ann. Appl. Probab. 19 (2009) 863–898.
[6] P. Bickel, E. Levina, Regularized estimation of large covariance matrices, Ann. Stat. 36 (2008) 199–227.
[7] P. Bickel, M. Lindner, Approximating the inverse of banded matrices by banded matrices with applications to probability and statistics, Theory Probab. Appl. 56 (2012) 1–20.
[8] T. Bui-Thanh, O. Ghattas, J. Martin, G. Stadler, A computational framework for infinite-dimensional Bayesian inverse problems. Part I: the linearized case, with application to global seismic inversion, SIAM J. Sci. Comput. 36 (2013) A2494–A2523.
[9] A. Chorin, M. Morzfeld, Conditions for successful data assimilation, J. Geophys. Res. 118 (2013) 11,522–11,533.
[10] J. Christen, C. Fox, A general purpose sampling algorithm for continuous distributions (the t-walk), Bayesian Anal. 5 (2010) 263–282.
[11] O.F. Christensen, G.O. Roberts, J.S. Rosenthal, Scaling limits for the transient phase of local Metropolis–Hastings algorithms, J. R. Stat. Soc., Ser. B, Stat. Methodol. 67 (2005) 253–268.
[12] S. Cotter, G. Roberts, A. Stuart, D. White, MCMC methods for functions: modifying old algorithms to make them faster, Stat. Sci. 28 (2013) 424–446.
[13] T. Cui, K.J.H. Law, Y.M. Marzouk, Dimension-independent likelihood-informed MCMC, J. Comput. Phys. 304 (2016) 109–137.
[14] T. Cui, J. Martin, Y.M. Marzouk, A. Solonen, A. Spantini, Likelihood-informed dimension reduction for nonlinear inverse problems, Inverse Probl. 29 (2014) 114015.
[15] T. Cui, Y.M. Marzouk, K. Willcox, Scalable posterior approximations for large-scale Bayesian inverse problems via likelihood-informed parameter and state reduction, J. Comput. Phys. 315 (2016) 363–387.
[16] S. Duane, A. Kennedy, B. Pendleton, D. Roweth, Hybrid Monte Carlo, Phys. Lett. B 195 (1987) 216–222.
[17] H. Flath, L. Wilcox, V. Akçelik, J. Hill, B. van Bloemen Waander, O. Ghattas, Fast algorithms for Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations, SIAM J. Sci. Comput. 33 (2011) 407–432.
[18] D. Foreman-Mackey, A. Conley, W. Meierjurgen Farr, D.W. Hogg, D. Long, P. Marshall, A. Price-Whelan, J. Sanders, J. Zuntz, emcee: the MCMC hammer, Astrophysics Source Code Library, https://arxiv.org/abs/1303.002, Mar. 2013.
[19] M. Fowler, M. Howard, A. Luttman, S. Mitchell, T. Webb, A stochastic approach to quantifying the blur with uncertainty estimation for high-energy X-ray imaging systems, Inverse Probl. Sci. Eng. 34 (2016) 353–371.
[20] C. Fox, A. Parker, Accelerated Gibbs sampling of normal distributions using matrix splittings and polynomials, Bernoulli 23 (2017) 3711–3743.
[21] A. Galli, H. Gao, Rate of convergence of the Gibbs sampler in the Gaussian case, Math. Geol. 33 (2001) 653–677.
[22] G. Gaspari, S. Cohn, Construction of correlation functions in two and three dimensions, Q. J. R. Meteorol. Soc. 125 (1999) 723–757.
[23] W. Gilks, S. Richardson, D. Spiegelhalter, Introducing Markov chain Monte Carlo, in: W. Gilks, S. Richardson, D. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, Srpinger-Science+Business Media, 1996, pp. 1–20, ch. 1.
[24] M. Girolami, B. Calderhead, Riemann manifold Langevin and Hamiltonian Monte Carlo methods, J. R. Stat. Soc. B 73 (2011) 123–214.
[25] J. Goodman, A.D. Sokal, Multigrid Monte Carlo method. Conceptual foundation, Phys. Rev. D 40 (1989) 2035–2071.
[26] J. Goodman, J. Weare, Ensemble samplers with affine invariance, Commun. Appl. Math. Comput. Sci. 5 (2010) 65–80.
[27] M. Hairer, A.M. Stuart, S.J. Vollmer, Spectral gaps for a Metropolis–Hastings algorithm in infinite dimensions, Ann. Appl. Probab. 24 (2014) 2455–2490.
[28] T.M. Hamill, J. Whitaker, C. Snyder, Distance-dependent filtering of background covariance estimates in an ensemble Kalman filter, Mon. Weather Rev. 129 (2001) 2776–2790.
[29] D. Hodyss, C. Bishop, M. Morzfeld, To what extent is your data assimilation scheme designed to find the posterior mean, the posterior mode or something else? Tellus A: dynamic meteorology and oceanography 68 (2016) 30625.
[30] P. Houtekamer, H. Mitchell, A sequential ensemble Kalman filter for atmospheric data assimilation, Mon. Weather Rev. 129 (2001) 123–136.
[31] P.L. Houtekamer, H.L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, B. Hansen, Atmospheric data assimilation with an ensemble Kalman filter: results with real observations, Mon. Weather Rev. 133 (2005) 604–620.
[32] D.R. Insua, F. Ruggeri, Robust Bayesian Analysis, vol. 152, Springer Science & Business Media, 2012.
[33] M. Kalos, P. Whitlock, Monte Carlo Methods, 1st ed., vol. 1, John Wiley & Sons, 1986.
[34] Y. Lee, A. Majda, State estimation and prediction using clustered particle filters, Proc. Natl. Acad. Sci. USA 113 (2016) 14609–14614.
[35] J. Lei, P. Bickel, A moment matching ensemble filter for nonlinear non-Gaussian data assimilation, Mon. Weather Rev. 139 (2011) 3964–3973.
[36] F. Lindgren, H. Rue, J. Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach, J. R. Stat. Soc., Ser. B, Stat. Methodol. 73 (2011) 423–498.
[37] E.N. Lorenz, Predictability: a problem partly solved, in: Proc. ECMWF Seminar on Predictability, vol. 1, 1996, pp. 1–18.
[38] D. MacKay, Introduction to Monte Carlo methods, in: M.I. Jordan (Ed.), Learning in Graphical Models, in: NATO Science Series, Kluwer Academic Press, 1998, pp. 175–204.
[39] J. Martin, L.C. Wilcox, C. Burstedde, O. Ghattas, A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion, SIAM J. Sci. Comput. 34 (2012) A1460–A1487.
[40] M. Morzfeld, D. Hodyss, C. Snyder, What the collapse of the ensemble Kalman filter tells us about particle filters, Tellus A 69 (2017) 1283809.

[41] R. Neal, MCMC using Hamiltonian dynamics, in: S. Brooks, A. Gelman, G. Jones, X.-L. Meng (Eds.), Handbook of Markov Chain Monte Carlo, Chapman & Hall, Oxford, 2011, ch. 5.
[42] R. Norton, C. Fox, Fast sampling in a linear-Gaussian inverse problem, SIAM/ASA J. Uncertain. Quantificat. 4 (2016) 1191–1218.
[43] A. Owen, Monte Carlo Theory, Methods and Examples, 2013.
[44] S. Penny, T. Miyoshi, A local particle filter for high dimensional geophysical systems, Nonlinear Process. Geophys. 2 (2015) 1631–1658.
[45] N. Petra, J. Martin, G. Stadler, O. Ghattas, A computational framework for infinite-dimensional Bayesian inverse problems. Part II: stochastic Newton MCMC with application to ice sheet flow inverse problems, SIAM J. Sci. Comput. 36 (2013) A1525–A1555.
[46] J. Poterjoy, A localized particle filter for high-dimensional nonlinear systems, Mon. Weather Rev. 144 (2015) 59–76.
[47] J. Poterjoy, J. Anderson, Efficient assimilation of simulated observations in a high-dimensional geophysical system using a localized particle filter, Mon. Weather Rev. 144 (2016) 2007–2020.
[48] J. Poterjoy, R. Sobash, J. Anderson, Convective-scale data assimilation for the weather research and forecasting model using the local particle filter, Mon. Weather Rev. 145 (2017) 1897–1918.
[49] P. Rebeschini, R. van Handel, Can local particle filters beat the curse of dimensionality?, Ann. Appl. Probab. 25 (2015) 2809–2866.
[50] S. Reich, A nonparametric ensemble transform method for Bayesian inference, Mon. Weather Rev. 35 (2013) 1337–1367.
[51] G. Roberts, A. Gelman, W. Gilks, Weak convergence and optimal scaling of random walk Metropolis algorithms, Ann. Appl. Probab. 7 (1997) 110–120.
[52] G. Roberts, J. Rosenthal, Optimal scaling of discrete approximations to Langevin diffusions, J. R. Stat. Soc., Ser. B, Stat. Methodol. 60 (1998) 255–268.
[53] G.O. Roberts, S.K. Sahu, Updating schemes, correlation structure, blocking and parameterization for the Gibbs sampler, J. R. Stat. Soc., Ser. B, Stat. Methodol. 59 (1997) 291–317.
[54] H. Rue, L. Held, Gaussian Markov Random Fields: Theory and Applications, CRC Press, 2005.
[55] P. Sakov, D.S. Oliver, L. Bertino, An iterative EnKF for strongly nonlinear systems, Mon. Weather Rev. 140 (2012) 1988–2004.
[56] A. Spantini, D. Bigoni, Y.M. Marzouk, Inference via low-dimensional couplings, arXiv:1703.06131, 2017.
[57] A. Spantini, A. Solonen, T. Cui, J. Martin, L. Tenorio, Y.M. Marzouk, Optimal low-rank approximations of Bayesian linear inverse problems, SIAM J. Sci. Comput. 37 (2015) A2451–A2487.
[58] A. Stuart, Inverse problems: a Bayesian perspective, Acta Numer. 19 (2010) 451–559.
[59] O. Talagrand, P. Courtier, Variational assimilation of meteorological observations with the adjoint vorticity equation. I: theory, Q. J. R. Meteorol. Soc. 113 (1987) 1311–1328.
[60] J. Tödter, B. Ahrens, A second-order exact ensemble square root filter for nonlinear data assimilation, Mon. Weather Rev. 143 (2015) 1337–1367.
[61] X.T. Tong, Performance analysis of local ensemble Kalman filter, arXiv:1705.10598, 2019, Accepted by J. Nonlinear Science.
[62] P. van Leeuwen, Y. Cheng, S. Reich, Nonlinear Data Assimilation, Springer, 2015.
[63] U. Wolff, Monte Carlo errors with less errors, Comput. Phys. Commun. 156 (2004) 143–153.