On unifying randomized methods for inverse problems

IOP Publishing Inverse Problems
Inverse Problems **39** (2023) 075010 (39pp) https://doi.org/10.1088/1361-6420/acd36e

On unifying randomized methods for inverse problems

Jonathan Wittmer[1,*], **C G Krishnanunni**[2], **Hai V Nguyen**[2] **and Tan Bui-Thanh**[3]

[1] Oden Institute, University of Texas at Austin, Austin, TX 78712, United States of America
[2] Department of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin, Austin, TX 78712, United States of America
[3] Oden Institute, Department of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin, Austin, TX 78712, United States of America

E-mail: jonathan.wittmer@utexas.edu

Received 28 December 2022; revised 26 April 2023
Accepted for publication 9 May 2023
Published 9 June 2023

Abstract

This work unifies the analysis of various randomized methods for solving linear and nonlinear inverse problems with Gaussian priors by framing the problem in a stochastic optimization setting. By doing so, we show that many randomized methods are variants of a sample average approximation (SAA). More importantly, we are able to prove a single theoretical result that guarantees the asymptotic convergence for a variety of randomized methods. Additionally, viewing randomized methods as an SAA enables us to prove, for the first time, a single non-asymptotic error result that holds for randomized methods under consideration. Another important consequence of our unified framework is that it allows us to discover new randomization methods. We present various numerical results for linear, nonlinear, algebraic, and PDE-constrained inverse problems that verify the theoretical convergence results and provide a discussion on the apparently different convergence rates and the behavior for various randomized methods.

Keywords: randomization, Bayesian inversion, ensemble Kalman filter, randomized maximum *a posteriori*

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.

1. Introduction

Solving large-scale ill-posed inverse problems that are governed by partial differential equations (PDEs), though tremendously challenging, is of great practical importance in science and engineering. Classical deterministic inverse methodologies, which provide point estimates of the solution, are not capable of accounting for the uncertainty in the inverse solution in a principled way. The Bayesian formulation provides a systematic quantification of uncertainty by posing the inverse problem as one of statistical inference. The Bayesian framework for inverse problems proceeds as follows: given observational data $\boldsymbol{d} \in \mathbb{R}^k$ and their uncertainty, the governing forward problem and its uncertainty, and a prior probability density function describing uncertainty in the parameters $\boldsymbol{u} \in \mathbb{R}^n$, the solution of the inverse problems is the posterior probability distribution $\pi(\boldsymbol{u}|\boldsymbol{d})$ over the parameters. Bayes' Theorem explicitly gives the posterior density as

$$\pi(\boldsymbol{u}|\boldsymbol{d}) \propto \pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{u}) \times \pi_{\text{prior}}(\boldsymbol{u})$$

which updates the prior knowledge $\pi_{\text{prior}}(\boldsymbol{u})$ using the likelihood $\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{u})$. The prior encodes any knowledge or assumptions about the parameter space that we may wish to impose before any data are observed, while the likelihood explicitly represents the probability that a given set of parameters $\boldsymbol{u}$ might give rise to the observed data $\boldsymbol{d}$.

Even when the prior and noise probability distributions are Gaussian, the posterior need not be Gaussian, due to possible nonlinearity embedded in the likelihood. For large-scale inverse problems, exploring non-Gaussian posteriors in high dimensions to compute statistics is a grand challenge since evaluating the posterior at each point in the parameter space requires a solution of the parameter-to-observable map, including a potentially expensive forward model solve. Using numerical quadrature to compute the mean and covariance matrix, for example, is generally infeasible in high dimensions. Usually the method of choice for computing statistics is Markov chain Monte Carlo (MCMC), which judiciously samples the posterior distribution, so that sample statistics can be used to approximate the exact ones.

The Metropolis-Hastings algorithm, first developed by Metropolis *et al* (1953) and then generalized by Hastings (1970), is perhaps the most popular MCMC method. Its popularity and attractiveness come from the ease of implementation and minimal requirements on the target density and the proposal density (Robert and Casella 2005, Haario *et al* 2006). The problem, however, is that standard MCMC methods often require millions of samples for convergence; since each sample requires an evaluation of the parameter-to-observable map, this could entail millions of expensive forward PDE simulations—a prohibitive proposition. On one hand, with the rapid development of parallel computing, parallel MCMC methods (Wilkinson 2005, Brockwell 2006, Byrd 2010, Strid 2010, Wang 2014) are studied to accelerate the computation. While parallelization allows MCMC algorithms to produce more samples in a shorter time with multiple processors, such accelerations typically do not improve the mixing and convergence of MCMC algorithms. More sophisticated MCMC methods that exploit the gradient and higher derivatives of the log posterior (and hence the parameter-to-observable map) (Duane *et al* 1987, Neal 2010, Beskos *et al* 2011, Girolami and Calderhead 2011, Bui-Thanh and Ghattas 2012, Martin *et al* 2012, Bui-Thanh and Girolami 2014, Cui *et al* 2014, 2016, Petra *et al* 2014) can, on the other hand, improve the mixing, acceptance rate, and convergence of MCMC. Another sample-based family of approaches that provide uncertainty quantification and are well-suited for parallelization on large clusters is the various forms of Stein variational gradient descent (Liu and Wang 2016, Han and Liu 2018, Zhuo *et al* 2018, Chen and Ghattas 2020). Of related interest are particle filter methods such as those found in Carpenter *et al* (1999), Van Der Merwe *et al* (2000), Soto (2005), Yang *et al* (2013)that evolve

particles through a dynamical system over time, updating both an estimate of the state and uncertainty.

One approach to addressing the computational challenge in high-dimensional statistical inverse problems pose is to use randomization, either to reduce the dimension of the optimization problem used in estimating the maximum *a posteriori* (MAP) point (Le *et al* 2017), or to aid in sampling from the posterior distribution (Wang *et al* 2018). Several methods have been proposed which utilize randomization to accelerate the solution of inverse problems (Avron *et al* 2013, Iglesias *et al* 2013, Le *et al* 2017, Wang *et al* 2017, 2018, Chen *et al* 2020a). *As the main contribution of this paper*, we derive unified results of randomized inverse approaches that apply to a broad class of linear and nonlinear inverse problems not only in the asymptotic regime, but also for the non-asymptotic setting. The asymptotic convergence and a non-asymptotic error bound of various existing methods follows immediately as special cases of the general result.

2. A unified analysis of randomized inverse problems through a sample average approximation (SAA) lens

For the remainder of this paper, we will use lower case letters for scalar quantities (α), boldface lower case letters for vectors ($\boldsymbol{u}$) and boldface upper case letters for matrices ($\boldsymbol{A}$). Further, we will use superscript lower case letters to denote sample index and subscript uppercase letters to denote the total number of samples, i.e. $\boldsymbol{\lambda}^i$ is the ith sample and $\boldsymbol{u}_N$ is a quantity depending on N samples. Lastly, descriptions or method names will be in uppercase superscripts, such as $\boldsymbol{u}^{\mathrm{MAP}}$ which is $\boldsymbol{u}$ at the MAP point, for example. This should be clear from the context.

Therefore, let $\boldsymbol{u}, \boldsymbol{u}_0 \in \mathbb{R}^n$. The posterior measure ν in this case has the density $\pi(\boldsymbol{u}|\boldsymbol{d})$ with respect to the Lebesgue measure:

$$\pi(\boldsymbol{u}|\boldsymbol{d}) \propto \pi_{\mathrm{like}}(\boldsymbol{d}|\boldsymbol{u}) \times \pi_{\mathrm{prior}}(\boldsymbol{u}), \tag{1}$$

where the likelihood is given by $\pi_{\mathrm{like}}(\boldsymbol{d}|\boldsymbol{u}) \propto \exp(-\Phi(\boldsymbol{u},\boldsymbol{d})) = \exp(-\frac{1}{2}\|\boldsymbol{d}-\mathcal{F}(\boldsymbol{u})\|^2_{\boldsymbol{\Sigma}^{-1}})$ and the prior by $\pi_{\mathrm{prior}} \propto \exp(-\frac{1}{2}\|\boldsymbol{u}-\boldsymbol{u}_0\|^2_{\boldsymbol{\Gamma}^{-1}})$. Here, $\mathcal{F}(\boldsymbol{u})$ is known as the parameter-to-observable (PtO) map, an evaluation of which typically requires a solution of the forward model (e.g. partial differential equations) governing the underlying physics. Note that this form assumes that both the likelihood and prior are Gaussian distributions. The *maximum a posteriori* (MAP) problem reads

$$\boldsymbol{u}^{\mathrm{MAP}} := \arg\min_{\boldsymbol{u}} \mathcal{J}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d}) := \frac{1}{2}\|\boldsymbol{d}-\mathcal{F}(\boldsymbol{u})\|^2_{\boldsymbol{\Sigma}^{-1}} + \frac{1}{2}\|\boldsymbol{u}-\boldsymbol{u}_0\|^2_{\boldsymbol{\Gamma}^{-1}}, \tag{2}$$

where $\boldsymbol{\Gamma} \in \mathbb{R}^{n\times n}$ is the prior covariance matrix and $\boldsymbol{\Sigma} \in \mathbb{R}^{k\times k}$ is the noise covariance matrix. Though we derive (2) from the Bayesian formulation, this optimization problem also arises in the deterministic setting as a regularized inverse problem, though often the standard l_2-norm is used in the purely deterministic setting. The methodologies discussed here then clearly apply in the deterministic setting with the restriction to l_2-norms, though it is more instructive to consider the Bayesian interpretation.

To the end of the paper, we denote by $\mathbb{E}$ the expectation with subscript as the random variable with respect to which the expectation is taken. When the random variable is clear from the context we simply omit the subscript for brevity. Let $\boldsymbol{\sigma}, \boldsymbol{\varepsilon}, \boldsymbol{\delta}$, and $\boldsymbol{\lambda}$ be finite dimensional independent random variables with bounded second moments such that:

$$\mathbb{E}[\boldsymbol{\sigma}] = 0, \quad \mathbb{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right] = \boldsymbol{\Sigma}^{-1}, \quad \mathbb{E}[\boldsymbol{\delta}] = 0, \quad \mathbb{E}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right] = \boldsymbol{\Gamma}^{-1}. \tag{3}$$

Let us define $\boldsymbol{\xi} = [\boldsymbol{\sigma}, \varepsilon, \boldsymbol{\delta}, \boldsymbol{\lambda}]^{\mathrm{T}} \in \boldsymbol{\Xi}$ with joint probability distribution $\pi = \pi_{\boldsymbol{\sigma}} \times \pi_{\boldsymbol{\varepsilon}} \times \pi_{\boldsymbol{\delta}} \times \pi_{\boldsymbol{\lambda}}$. Consider the following stochastic cost function:

$$\begin{aligned} \mathcal{J}_{\xi}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\xi}) &:= \frac{1}{2}\left\|\varepsilon^{\mathrm{T}}(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))\right\|_2^2 + \frac{1}{2}\left\|\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta})\right\|_2^2 \\ &= \frac{1}{2}(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))^{\mathrm{T}}\varepsilon\varepsilon^{\mathrm{T}}(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u})) \\ &\quad + \frac{1}{2}(\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta})^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta}). \end{aligned} \tag{4}$$

Define

$$\mathscr{J}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d}) := \mathbb{E}_{\pi}\left[\mathcal{J}_{\xi}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\xi})\right],$$

and the SAA of $\mathscr{J}$ to be

$$\mathscr{J}_N := \frac{1}{N}\sum_{j=1}^{N}\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\xi}^j\right) \tag{5}$$

where $\boldsymbol{\xi}^j$ are i.i.d. samples from π. Assume that both $\mathcal{J}$ and $\mathscr{J}_N$ have a minimum and let us define

$$\boldsymbol{u}^{\mathrm{MAP}} := \arg\min_{\boldsymbol{u}} \mathcal{J}, \text{and } \hat{\boldsymbol{u}}_N^{\mathrm{MAP}} := \arg\min_{\boldsymbol{u}} \mathscr{J}_N. \tag{6}$$

Below we study asymptotic and non-asymptotic convergence of $\hat{\boldsymbol{u}}_N^{\mathrm{MAP}}$ to $\boldsymbol{u}^{\mathrm{MAP}}$.

2.1. Asymptotic convergence analysis for inverse problems

Theorem 1 (Asymptotic convergence of randomized nonlinear inverse problems). *Assume that $\mathcal{F}(\boldsymbol{u})$ is such that $\mathcal{J}_{\xi}$ is a convex, twice continuously differentiable function in $\boldsymbol{u}$ for almost every $\boldsymbol{\xi}$, and measurable*[4]*. Then the following hold true:*

(i) Minimizing $\mathscr{J}$ is equivalent to minimizing $\mathcal{J}$ in the sense: $\arg\min_{\boldsymbol{u}} \mathcal{J} = \arg\min_{\boldsymbol{u}} \mathscr{J}$.
(ii) $\hat{\boldsymbol{u}}_N^{MAP} \xrightarrow[N\to\infty]{a.s.} \boldsymbol{u}^{MAP}$.

Proof. For the first assertion, consider only the first term of $\mathscr{J}$ as the second term follows analogously. We have

$$\begin{aligned} &\frac{1}{2}\mathbb{E}_{\pi}\left[(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))^{\mathrm{T}}\varepsilon\varepsilon^{\mathrm{T}}(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))\right] \\ &\quad = \frac{1}{2}\mathbb{E}_{\pi_{\boldsymbol{\sigma}}\times\pi_{\boldsymbol{\varepsilon}}}\left[(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))^{\mathrm{T}}\varepsilon\varepsilon^{\mathrm{T}}(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))\right] \\ &\quad = \frac{1}{2}\mathbb{E}_{\pi_{\boldsymbol{\sigma}}}\left[(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))^{\mathrm{T}}\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\varepsilon\varepsilon^{\mathrm{T}}\right](\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))\right] \\ &\quad = \frac{1}{2}\mathbb{E}_{\pi_{\boldsymbol{\sigma}}}\left[(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u}))\right] \\ &\quad = \frac{1}{2}(\boldsymbol{d}-\mathcal{F}(\boldsymbol{u}))^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{d}-\mathcal{F}(\boldsymbol{u})) + \mathbb{E}_{\pi_{\boldsymbol{\sigma}}}\left[\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\sigma}\right]. \end{aligned} \tag{7}$$

[4] Here, measurable is with respect to the σ-algebra given by the product σ-algebras of the deterministic $(\boldsymbol{u},\boldsymbol{u}_0,\boldsymbol{d})$ and random variables $\boldsymbol{\xi}$.

The final term in (7) is constant with respect to $\boldsymbol{u}$ and can be ignored, leaving only the first term of $\mathcal{J}$. Applying the same procedure to the second term of $\mathscr{J}$ shows that minimizing $\mathscr{J}$ is equivalent to minimizing $\mathcal{J}$.

We invoke Shapiro *et al* (2009), theorem 5.4 to prove the second assertion. It is sufficient to verify the following conditions:

(i) $\mathcal{J}_\xi(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\xi})$ is random lower semicontinuous,
(ii) for almost every $\boldsymbol{\xi}\in\boldsymbol{\Xi}$, $\mathcal{J}_\xi(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\xi})$ is convex in $\boldsymbol{u}$,
(iii) $\mathscr{J}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d})$ is lower semicontinuous in $\boldsymbol{u}$ and there exists a point $\bar{\boldsymbol{u}}\in\mathbb{R}^n$ such that $\mathscr{J}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d})<\infty$ for all $\boldsymbol{u}$ in a neighborhood of $\bar{\boldsymbol{u}}$;
(iv) the set of optimal solutions of the true problem is nonempty and bounded; and
(v) the law of large numbers (LLNs) (Feller 1971, Durrett 2019) holds pointwise for $\mathscr{J}_N$.

Clearly, $\mathcal{J}_\xi$ is a continuous function for every $\boldsymbol{\xi}$, thus random lower semicontinuous as well. By assumption, $\mathcal{J}_\xi$ is also convex for almost every $\boldsymbol{\xi}$. Due to the boundedness assumptions (3) and the fact that $\mathcal{J}$ is a continuous and convex function, $\mathscr{J}$ is also a continuous and convex function. Furthermore, taking, for example, $\bar{\boldsymbol{u}}=\boldsymbol{u}_1$ in (20) it is straightforward to see that $\mathscr{J}(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d})<\infty$ for any ball with finite radius centered at $\bar{\boldsymbol{u}}$. The last two conditions are clear. □

While convergence of the randomized cost function to its expected value is obvious by the LLNs, we prove here the convergence of extremum of the randomized cost function. This does not hold in general, and the rigorous theory for such convergence is the so-called $\Gamma-$convergence. $\Gamma-$convergence theory (Maso 1993, Braides *et al* 2002) is the study of necessary and sufficient conditions for the convergence of the extremum when the cost function converges. Therefore, the novelty of theorem 1 is not in the proof that the cost functions converge, but that the randomized cost functions yield solutions that asymptotically converge to the solution of the non-randomized cost function. While we constructively prove $\Gamma-$convergence for (4), our result follows immediately under the much stronger assumption that the prior term already exhibits $\Gamma-$convergence and the likelihood term converges continuously based on the results presented in Ayanbayev *et al* (2021). In fact, the cost functions themselves do not converge as there is an additional bias term in expected value of the randomized cost function, $\mathscr{J}$, compared to the deterministic cost function, $\mathcal{J}$. However, this bias term is irrelevant as is shown in assertion *i* of theorem 1. This is because the bias term does not depend on $\boldsymbol{u}$ and disappears when computing the optimality condition via taking derivatives with respect to $\boldsymbol{u}$.

Note that very mild constraints are placed on the random variables $\boldsymbol{\xi}$—only those required for the LLN to hold. This allows great freedom in designing a valid randomization scheme. An important special case of this theorem occurs when we consider an inverse problem with a linear parameter-to-observable map. When the forward map $\mathcal{F}(\boldsymbol{u})$ is linear, the convexity and continuous differentiability assumptions are satisfied. While requiring convexity is a strong assumption in general, this is not an insurmountable issue for regularized inverse problems. Note that the Hessian of $\mathcal{J}$ is given by

$$\nabla_{\boldsymbol{u}}^2\mathcal{J}=\nabla_{\boldsymbol{u}}^2\mathcal{F}(\boldsymbol{u})\,\boldsymbol{\Sigma}^{-1}\left(\mathcal{F}(\boldsymbol{u})-\boldsymbol{d}\right)+\nabla_{\boldsymbol{u}}\mathcal{F}(\boldsymbol{u})^{\mathrm{T}}\,\boldsymbol{\Sigma}^{-1}\nabla_{\boldsymbol{u}}\mathcal{F}(\boldsymbol{u})+\boldsymbol{\Gamma}^{-1}.$$

Thus the prior covariance matrix, $\boldsymbol{\Gamma}$, can be chosen such that $\nabla_{\boldsymbol{u}}^2\mathcal{J}$ is semi-positive definite. Indeed, this is the major role that the prior covariance plays in regularizing the ill-posed inverse problem.

One practical shortcoming of theorem 1 is that the cost function $\mathcal{J}_\xi$ is not composed of the sample average of each variable $(\boldsymbol{\sigma}, \boldsymbol{\varepsilon}, \boldsymbol{\lambda}, \boldsymbol{\delta})$ independently, but rather it is the sample average with samples drawn from the joint distribution ξ. This will prove to have theoretical consequences in deriving non-asymptotic error bounds as well as practical impacts when it comes to implementing the randomization schemes. A more practical formulation would compute the sample average of each random variable individually. Fortunately, (Shapiro *et al* 2009, theorem 5.4) can be applied four times, once for each of $\boldsymbol{\sigma}, \boldsymbol{\varepsilon}, \boldsymbol{\lambda}, \boldsymbol{\delta}$, to prove the asymptotic convergence of the following randomized inverse solution:

$$\mathring{\boldsymbol{u}}_N^{\mathrm{MAP}} := \arg\min_{\boldsymbol{u}} \frac{1}{N_1 N_2 N_3 N_4} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} \sum_{l=1}^{N_4} \mathcal{J}_\xi \left(\boldsymbol{u}; \boldsymbol{u}_0, \boldsymbol{d}, \left[\boldsymbol{\sigma}^i; \boldsymbol{\varepsilon}^j; \boldsymbol{\lambda}^k; \boldsymbol{\delta}^l \right] \right). \tag{8}$$

This flexibility in deciding how samples will be drawn will aid in both non-asymptotic convergence analysis and will provide great freedom in designing a variety of randomized methods to solve inverse problems in section 5.

To alleviate some notational burden, let us define a few new quantities.

$$\boldsymbol{S}_N := \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\varepsilon}^i (\boldsymbol{\varepsilon}^i)^{\mathrm{T}}, \qquad \boldsymbol{L}_N := \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\lambda}^i (\boldsymbol{\lambda}^i)^{\mathrm{T}}, \tag{9a}$$

$$\bar{\boldsymbol{\sigma}}_N := \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\sigma}^i, \qquad \bar{\boldsymbol{\delta}}_N := \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\delta}^i. \tag{9b}$$

Written in terms of norms, equation (8) becomes

$$\mathring{\boldsymbol{u}}_N^{\mathrm{MAP}} = \arg\min_{\boldsymbol{u}} \frac{1}{2} \left\| \boldsymbol{d} + \bar{\boldsymbol{\sigma}}_N - \mathcal{F}(\boldsymbol{u}) \right\|_{\boldsymbol{S}_N}^2 + \frac{1}{2} \left\| \boldsymbol{u} - \boldsymbol{u}_0 - \bar{\boldsymbol{\delta}}_N \right\|_{\boldsymbol{L}_N}^2. \tag{10}$$

Note that (10) is equivalent to (5) when at most one of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ are randomized and at most one of $\boldsymbol{\lambda}$ and $\boldsymbol{\delta}$ are randomized. This is because the only difference between the two cost functions is how samples of $\boldsymbol{\varepsilon}$ interact with samples of $\boldsymbol{\sigma}$ and likewise, how samples of $\boldsymbol{\lambda}$ interact with samples of $\boldsymbol{\delta}$. To see this for the $\boldsymbol{\sigma}$-$\boldsymbol{\varepsilon}$ interaction, we compute the gradient of the misfit term for each cost function, assuming $N_1 = N_2 = N$. For (5), this is

$$-\nabla \mathcal{F}(\boldsymbol{u})^{\mathrm{T}} \frac{1}{N} \sum_{i=1}^{N} \left[\boldsymbol{\varepsilon}^i \boldsymbol{\varepsilon}^{i\mathrm{T}} \left(\boldsymbol{d} + \boldsymbol{\sigma}^i - \mathcal{F}(\boldsymbol{u}) \right) \right] \tag{11}$$

and for (10),

$$-\nabla \mathcal{F}(\boldsymbol{u})^{\mathrm{T}} \left(\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\varepsilon}^i \boldsymbol{\varepsilon}^{i\mathrm{T}} \right) \left(\boldsymbol{d} + \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{\sigma}^j - \mathcal{F}(\boldsymbol{u}) \right). \tag{12}$$

Subtracting (11) from (12),

$$(12) - (11) = \nabla \mathcal{F}(\boldsymbol{u})^{\mathrm{T}} \left[\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\varepsilon}^i \boldsymbol{\varepsilon}^{i\mathrm{T}} \left(\boldsymbol{\sigma}^i - \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{\sigma}^j \right) \right]. \tag{13}$$

When $\boldsymbol{\sigma}$ is not randomized, then $\boldsymbol{\sigma}^i = \boldsymbol{\sigma}^j = 0 \; \forall i, j$ and the difference is 0. Likewise, when $\boldsymbol{\varepsilon}$ is not randomized, $\boldsymbol{\varepsilon}^i \boldsymbol{\varepsilon}^{i\mathrm{T}} = \boldsymbol{\Sigma}^{-1}$ and, assuming the samples $\boldsymbol{\sigma}^i = \boldsymbol{\sigma}^j$ when $i = j$, (13) is also 0. However, (13) is in general not 0.

2.2. Non-asymptotic error analysis for nonlinear inverse problems

In addition to proving a general asymptotic convergence of randomized inverse problems, it is also possible to derive a general non-asymptotic error bound with the slightly stronger assumption that the random variables are subgaussian. Subgaussian random variables are random variables whose moment generating function is bounded above by the moment generating function of a Gaussian, i.e. a random variable X with $\mathbb{E}[X] = 0$ is subgaussian with variance proxy σ^2 if

$$\mathbb{E}[tX] \leqslant \exp\left(\frac{\sigma^2 t^2}{2}\right), \quad \forall\, t \in \mathbb{R}.$$

A thorough explanation of the mathematical formalism can be found in Vershynin (2018). Intuitively, subgaussian random variables have tails that decay at least as fast as a Gaussian, allowing for the consideration of a much broader class of random variables than simply Gaussian. Non-asymptotic error analysis is concerned with developing a bound that gives the probability of making an error greater than some tolerance. That is, we are interested in the behavior of the tails of a distribution and the subgaussian assumption enables us to derive such bounds. The general form of this non-asymptotic bound is useful in that it is easy to identify the key components that go into forming the bound—giving insight into the performance of various methods by enabling easy simplification in the case that certain quantities are not randomized. The derived bound gives a probabilistic worst-case for finite sample size N when all of the above randomizations are implemented at the same time. By fixing some of the quantities, e.g. letting $\boldsymbol{\sigma} = 0$, the given bound can be simplified in a straightforward manner—yielding a more insightful bound.

Additionally, we follow the standard vector norm convention of $\|\boldsymbol{u}\|_\infty := \max(|\boldsymbol{u}_1|, \ldots, |\boldsymbol{u}_n|)$ and $\|\boldsymbol{u}\|_1 := \sum_{i=1}^n |\boldsymbol{u}_i|$ for a vector $\boldsymbol{u} \in \mathbb{R}^n$. Matrix norms are understood to be induced norms (Trefethen and Bau III 1997). Due to the equivalence of norms in finite dimensional spaces, all the results are also valid for other norms, albeit with different constants that possibly depend on the dimension. We use w.p. as the abbreviation for *with probability*. Before we can prove the general non-asymptotic error bound given in lemma 6, a few intermediate results are needed.

Proposition 2 (Convergence of mean-zero subgaussian random vector). *Let $\boldsymbol{\nu}^i$, for $i = 1, \ldots, N$, be independent subgaussian random vectors in $\mathbb{R}^n$ such that $\mathbb{E}\left[\boldsymbol{\nu}^i\right] = 0$, $\mathbb{E}\left[\boldsymbol{\nu}^i (\boldsymbol{\nu}^i)^{\mathrm{T}}\right] = \boldsymbol{V}$, and $\mathbb{E}\left[(\boldsymbol{\nu}^i)^{\mathrm{T}} \boldsymbol{\nu}^i\right] < \infty$. Denote the empirical mean $\bar{\boldsymbol{\nu}} := \frac{1}{N}\sum_{i=1}^N \boldsymbol{\nu}^i$. Further, let $\zeta(N, \beta) := \exp\left(-c\beta^2 N\right)$ for some $\beta > 0$ and c is a constant possibly depending on the dimension n but not on N. Then*

$$\|\bar{\boldsymbol{\nu}}\|_\infty \leqslant \beta \left\|\boldsymbol{V}^{\frac{1}{2}}\right\|_\infty \quad \text{w.p. at least } 1 - \zeta(N, \beta). \tag{14}$$

Proof. Define $\boldsymbol{\nu}^i = \boldsymbol{V}^{\frac{1}{2}} \boldsymbol{\tau}^i$, where $\boldsymbol{\tau}^i \sim \mathcal{N}(0, \mathcal{I})$. Thus, $\overline{\boldsymbol{\tau}} = \frac{1}{N}\sum_{i=1}^N \boldsymbol{\tau}^i \sim \mathcal{N}\left(0, \frac{\mathcal{I}}{N}\right)$. First from[5] (Gao *et al* 2022, theorem 1) we have

$$\mathbb{P}\left[\|\boldsymbol{\tau}\|_\infty > \beta_1\right] \leqslant \mathbb{P}\left[\|\boldsymbol{\tau}\|_1 > \beta_1\right] \leqslant \exp\left(-\frac{\beta_1^2}{4cn}\right),$$

[5] While (Gao *et al* 2022, theorem 1) is derived for Gaussian random matrices, it also applies to subgaussian random matrices because subgaussian random variables have the same bound for the expected value of their moment generating function (see (Vershynin 2018, proposition 2.5.2) for the details).

where c is an absolute positive constant. Therefore, abusing notation to consolidate the constant terms, we have

$$\mathbb{P}\left[\|\overline{\boldsymbol{\tau}}\|_\infty > \frac{\beta_1}{\sqrt{N}}\right] \leqslant \exp\left(-c\beta_1^2\right)$$

since $\sqrt{N}\cdot\overline{\boldsymbol{\tau}}$ has identity covariance. Next, set $\beta = \frac{\beta_1}{\sqrt{N}}$ and note that

$$\mathbb{P}\left[\|\overline{\boldsymbol{\nu}}\|_\infty \leqslant \beta\left\|\boldsymbol{V}^{\frac{1}{2}}\right\|_\infty\right] \geqslant \mathbb{P}\left[\|\overline{\boldsymbol{\tau}}\|_\infty \leqslant \beta\right] \geqslant 1-\exp\left(-c\beta^2 N\right),$$

and this concludes the proof. □

Note that we weaken the norm from l_1 to l_∞. Due to the assumption of finite dimensional random variables, all l_p norms are equivalent. We however choose to state the results using l_∞ norms as the l_∞ norm is more natural in the following propositions. In this way, we can simplify the presentation of our final result, lemma 6. Additionally, using the l_∞-norm is consistent with infinite dimensional settings in the context of a Gaussian prior measure. Indeed, as shown in Stuart (2010), a well defined Gaussian prior, and thus the Bayesian posterior (as it is absolutely continuous with respect to the prior), has continuous samples almost surely. The MAP problem that we consider in this paper in fact resides in a smaller subspace: the Cameron-Martin space, which is continuously embedded in the space of continuous functions. On the other hand, using an l_1 norm could be consistent with the Besov(1,1) prior (Saksman and Siltanen 2009).

All results could be stated with l_1 norms, but would be multiplied by constants depending on the dimension. This dependence on the dimension is not unique to the l_1 norm, however, since the probability of failure, ζ, has the dimension of the problem in the constant c, as is shown in the proof of proposition 2. Our theory, regardless of which l_p norm is used, predicts higher error rates and therefore slower convergence as the dimension of the problem increases. This dependence of the error on the problem dimension can be seen numerically in figure 3.

Proposition 3 (Convergence of subgaussian random vector outer product). *Let $\boldsymbol{\nu}^i$ be a subgaussian random vector in $\mathbb{R}^n$ such that $\mathbb{E}\left[\boldsymbol{\nu}^i\left(\boldsymbol{\nu}^i\right)^{\mathrm{T}}\right] = \boldsymbol{V}$ for $i = 1,\ldots,N$. Define Ω_N to be the random matrix formed by stacking $\boldsymbol{\nu}^i$ in the columns and scaling by $1/\sqrt{N}$. It follows that*

$$\left\|\Omega_N\Omega_N^{\mathrm{T}} - \boldsymbol{V}\right\|_\infty \leqslant \beta\|\boldsymbol{V}\|_\infty \text{ w.p. at least } 1-2\zeta(N,\beta), \tag{15}$$

where c is a constant depending on n but not on N.

Proof. This result follows from straightforward algebraic manipulation of (Vershynin 2018, theorem 4.6.1). □

An additional fact needed to prove a non-asymptotic bound is that the product of three subgaussian random variables is α-subexponential with $\alpha = 2/3$. The following discussion on $\alpha-$subexponential random variables is based on Sambale (2020). For a more compete treatment of the topic, Sambale (2020) can be consulted.

It has been established that the product of two subgaussian random variables is subexponential (Vershynin 2018, lemma 2.7.7). A centered random variable X is said to be α-subexponential (or sub-Weibull (Vladimirova *et al* 2020, Zhang and Wei 2022)) if it satisfies

$$\mathbb{P}[|X| \geqslant \beta] \leqslant 2\exp\left(-c\beta^\alpha\right),$$

for any $\beta > 0$, $\alpha \in (0,2]$, and some positive constant c. To show that a random variable satisfies this condition, it is sufficient to show that the following Orlicz (quasi-) norm (Vershynin 2018, Sambale 2020) is finite:

$$\|X\|_{\psi_\alpha} := \inf\left[\beta > 0 : \mathbb{E}\exp\left(\left(|X|/\beta\right)^\alpha\right) \leqslant 2\right] < \infty. \tag{16}$$

When $\alpha < 1$, this is a quasi-norm since it does not satisfy the triangle inequality.

Proposition 4 (α-subexponential from product of three subgaussian random variables). *Let X_1, X_2, X_3 be subgaussian random variables. Then $Y = X_1X_2X_3$ is an α-subexponential random variable with $\alpha = 2/3$.*

Proof. It suffices to show that $\mathbb{E}\exp\left(\left(|Y|\right)^{2/3}\right) \leqslant 2$. Without loss of generality, assume $\|X_i\|_{\psi_2} = 1$. Then $\mathbb{E}\exp\left(X_i^2\right) \leqslant 2$ and we have

$$\begin{aligned}
\mathbb{E}\exp\left(\left(|Y|\right)^{2/3}\right) &= \mathbb{E}\exp\left(\left(|X_1|^{2/3}|X_2|^{2/3}|X_3|^{2/3}\right)\right) \\
&\leqslant \mathbb{E}\exp\left(\frac{{X_1}^2}{3} + \frac{{X_2}^2}{3} + \frac{{X_3}^2}{3}\right) \quad \text{(Young's inequality for 3 variables)} \\
&= \mathbb{E}\exp\left(\frac{{X_1}^2}{3}\right)\exp\left(\frac{{X_2}^2}{3}\right)\exp\left(\frac{{X_3}^2}{3}\right) \\
&\leqslant \mathbb{E}\frac{1}{3}\left[\exp\left({X_1}^2\right) + \exp\left({X_2}^2\right) + \exp\left({X_3}^2\right)\right] \quad \text{(Young's inequality again)} \\
&\leqslant 2.
\end{aligned}$$

□

Corollary 5. *Let $\boldsymbol{\nu} \in \mathbb{R}^n$ be a zero-mean random vector such that $\mathbb{E}[\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}] = \boldsymbol{V}$ with α-subexponential entries. Define $\zeta_\alpha(N,\beta) := \exp\left(-c\beta^\alpha N^{\alpha/2}\right)$. Then*

$$\mathbb{P}\left[\left\|\frac{1}{N}\sum_{i=1}^{n}\boldsymbol{\nu}^i\right\|_\infty \leqslant \beta\left\|\boldsymbol{V}^{\frac{1}{2}}\right\|_\infty\right] \geqslant 1 - 2n\zeta_\alpha(N,\beta). \tag{17}$$

The proof of Young's inequality for 3 variables can be found at (j.j 2013) Note that while inequality (17) has the dimension n in front of the exponential, this is fixed and the probability of committing an error greater than a fixed tolerance still decreases exponentially in the number of samples N. Though the decaying is at the slower rate $\propto \exp\left(-N^{\alpha/2}\right)$ compared to the subgaussian error rate $\propto \exp(-N)$, it is not surprising because subgaussian is a special case of α-subexponential when $\alpha = 2$ (Sambale 2020). With the preceding non-asymptotic error analysis tools, we are finally in position to state and prove a general non-asymptotic error bound for the solution of randomized inverse problems.

Lemma 6 (Non-asymptotic error analysis for randomized nonlinear inverse problems). *Let* $\mathrm{vec}\left(\boldsymbol{\Sigma}^{-1}\right)$ *denote a vectorization of a matrix $\boldsymbol{\Sigma}^{-1}$. Define*

$$\boldsymbol{P} := \left[\mathrm{vec}\left(\boldsymbol{\Sigma}^{-1}\right); \mathrm{vec}\left(\boldsymbol{\Gamma}^{-1}\right); \boldsymbol{e}; \boldsymbol{z}\right]$$

as a vector concatenating all four vectors $\mathrm{vec}\left(\boldsymbol{\Sigma}^{-1}\right), \mathrm{vec}\left(\boldsymbol{\Gamma}^{-1}\right), \boldsymbol{e}$ *and* $\boldsymbol{z}$*, where $\boldsymbol{\Sigma} \in \mathbb{R}^{k\times k}, \boldsymbol{\Gamma} \in \mathbb{R}^{n\times n}, \boldsymbol{e} \in \mathbb{R}^k$, and $\boldsymbol{z} \in \mathbb{R}^n$. Define the function*

$$g(\boldsymbol{P};\boldsymbol{u}) := \nabla_{\boldsymbol{u}}\mathcal{F}(\boldsymbol{u})\left[\boldsymbol{\Sigma}^{-1}\left(\mathcal{F}(\boldsymbol{u}) - \boldsymbol{d}\right) - \boldsymbol{e}\right] + \boldsymbol{\Gamma}^{-1}(\boldsymbol{u} - \boldsymbol{u}_0) - \boldsymbol{z}.$$

Assume that the problem $g(\boldsymbol{P};\boldsymbol{u}) = 0$, with $\boldsymbol{P}$ as parameters and $\boldsymbol{u}$ as solution, is Lipschitz well-posed (Latz 2020) with Lipschitz constant $\mathcal{L}$, and we define $\mathcal{G}(\boldsymbol{P})$ as the solution $\boldsymbol{u}$. Let

$$\begin{aligned}
\boldsymbol{P}^{MAP} &:= \left[\mathrm{vec}\left(\boldsymbol{\Sigma}^{-1}\right);\ \mathrm{vec}\left(\boldsymbol{\Gamma}^{-1}\right);\ 0;\ 0\right],\\
\mathring{\boldsymbol{P}}_N &:= \left[\mathrm{vec}\left(\boldsymbol{S}_N\right);\ \mathrm{vec}\left(\boldsymbol{L}_N\right);\ \boldsymbol{S}_N\bar{\boldsymbol{\sigma}}_N;\ \boldsymbol{L}_N\bar{\boldsymbol{\delta}}_N\right],\\
\hat{\boldsymbol{P}}_N &:= \left[\mathrm{vec}\left(\boldsymbol{S}_N\right);\ \mathrm{vec}\left(\boldsymbol{L}_N\right);\ \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\varepsilon}^i\left(\boldsymbol{\varepsilon}^i\right)^{\mathrm{T}}\boldsymbol{\sigma}^i;\ \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}^i\left(\boldsymbol{\lambda}^i\right)^{\mathrm{T}}\boldsymbol{\delta}^i\right],
\end{aligned}$$

where $\mathbb{E}[\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}] = \boldsymbol{\Sigma}$ *and* $\mathbb{E}[\boldsymbol{\delta}\boldsymbol{\delta}^{\mathrm{T}}] = \boldsymbol{\Gamma}$. *Then*

$$\begin{aligned}
\left\|\boldsymbol{u}^{MAP} - \mathring{\boldsymbol{u}}_N^{MAP}\right\|_\infty &\leqslant \beta\mathcal{L}\left(\left\|\boldsymbol{\Sigma}^{-1}\right\|_\infty + \left\|\boldsymbol{\Gamma}^{-1}\right\|_\infty + (1+\beta)\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty + (1+\beta)\left\|\boldsymbol{\Gamma}^{-\frac{1}{2}}\right\|_\infty\right)\\
&\quad \textit{w.p. at least } 1 - 10\zeta\left(N,\beta\right), \qquad (18)
\end{aligned}$$

and

$$\begin{aligned}
\left\|\boldsymbol{u}^{MAP} - \hat{\boldsymbol{u}}_N^{MAP}\right\|_\infty &\leqslant \beta\mathcal{L}\left(\left\|\boldsymbol{\Sigma}^{-1}\right\|_\infty + \left\|\boldsymbol{\Gamma}^{-1}\right\|_\infty + \left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty + \left\|\boldsymbol{\Gamma}^{-\frac{1}{2}}\right\|_\infty\right)\\
&\quad \textit{w.p. at least } 1 - 4\zeta\left(N,\beta\right) - 2k\zeta_{2/3}\left(N,\beta\right) - 2n\zeta_{2/3}\left(N,\beta\right). \qquad (19)
\end{aligned}$$

Proof. Noting that $\boldsymbol{u}^{\mathrm{MAP}} = \mathcal{G}\left(\boldsymbol{P}^{\mathrm{MAP}}\right)$ and $\mathring{\boldsymbol{u}} = \mathcal{G}\left(\mathring{\boldsymbol{P}}_N\right)$, we have by the Lipschitz well-posedness assumption,

$$\begin{aligned}
\left\|\boldsymbol{u}^{\mathrm{MAP}} - \mathring{\boldsymbol{u}}_N^{\mathrm{MAP}}\right\|_\infty &= \left\|\mathcal{G}\left(\boldsymbol{P}^{\mathrm{MAP}}\right) - \mathcal{G}\left(\mathring{\boldsymbol{P}}_N\right)\right\|_\infty \leqslant \mathcal{L}\left\|\boldsymbol{P} - \mathring{\boldsymbol{P}}_N\right\|_\infty\\
&\leqslant \mathcal{L}\left(\left\|\boldsymbol{\Sigma}^{-1} - \boldsymbol{S}_N\right\|_\infty + \left\|\boldsymbol{\Gamma}^{-1} - \boldsymbol{L}_N\right\|_\infty + \left\|\boldsymbol{S}_N\bar{\boldsymbol{\sigma}}_N\right\|_\infty + \left\|\boldsymbol{L}_N\bar{\boldsymbol{\delta}}_N\right\|_\infty\right).
\end{aligned}$$

We can bound $\boldsymbol{S}_N\bar{\boldsymbol{\sigma}}_N$ (and similarly for $\boldsymbol{L}_N\bar{\boldsymbol{\delta}}_N$) as follows

$$\begin{aligned}
\left\|\boldsymbol{S}_N\bar{\boldsymbol{\sigma}}_N\right\|_\infty &= \left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{S}_N\boldsymbol{\Sigma}^{\frac{1}{2}} - \mathcal{I} + \mathcal{I}\right)\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\sigma}}_N\right\|_\infty\\
&\leqslant \left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{S}_N\boldsymbol{\Sigma}^{\frac{1}{2}} - \mathcal{I}\right)\right\|_\infty\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\sigma}}_N\right\|_\infty + \left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\sigma}}_N\right\|_\infty.
\end{aligned}$$

Note that $\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\sigma}}_N$ is the sample average of a mean-zero subgaussian random variable with identity covariance and $\mathbb{E}[\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{S}_N\boldsymbol{\Sigma}^{\frac{1}{2}}] = \mathcal{I}$. Therefore, applying proposition 3 to $\left\|\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{S}_N\boldsymbol{\Sigma}^{\frac{1}{2}} - \mathcal{I}\right\|_\infty$ and proposition 2 to $\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\sigma}}_N\right\|_\infty$ along with the union bound, we obtain

$$\left\|\boldsymbol{S}_N\bar{\boldsymbol{\sigma}}_N\right\|_\infty \leqslant \left(\beta^2+\beta\right)\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty \quad \text{w.p. at least } 1 - 3\zeta\left(N,\beta\right).$$

The proof of the bound $\left\|\boldsymbol{L}_N\bar{\boldsymbol{\delta}}_N\right\|_\infty \leqslant \left(\beta^2+\beta\right)\left\|\boldsymbol{\Gamma}^{-\frac{1}{2}}\right\|_\infty$ w.p. at least $1 - 3\zeta\left(N,\beta\right)$ follows analogously. Applying proposition 3 to the terms $\left\|\boldsymbol{S}_N - \boldsymbol{\Sigma}^{-1}\right\|_\infty$ and $\left\|\boldsymbol{L}_N - \boldsymbol{\Gamma}^{-1}\right\|_\infty$ along with the union bound, inequality (18) follows immediately.

To prove the bound in (19) for $\left\|\boldsymbol{u}^{\mathrm{MAP}} - \hat{\boldsymbol{u}}_N^{\mathrm{MAP}}\right\|_\infty$, it is sufficient to bound $\|\boldsymbol{e}\|_\infty$ and $\|\boldsymbol{z}\|_\infty$ according to the definition of $\hat{\boldsymbol{P}}_N$ since the bounds for $\left\|\boldsymbol{S}_N - \boldsymbol{\Sigma}^{-1}\right\|_\infty$ and $\left\|\boldsymbol{L}_N - \boldsymbol{\Gamma}^{-1}\right\|_\infty$ follow immediately from proposition 3. By proposition 4, $\boldsymbol{\varepsilon}^i\left(\boldsymbol{\varepsilon}^i\right)^{\mathrm{T}}\boldsymbol{\sigma}^i$ is a zero-mean, 2/3-subexponential random variable with covariance $\boldsymbol{\Sigma}^{-1}$. Then by corollary 5,

$$\mathbb{P}\left[\|\boldsymbol{e}\|_\infty \leqslant \beta\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty\right] \geqslant 1 - 2k\zeta_{2/3}\left(N,\beta\right).$$

Similarly, applying corollary 5 to $\|\boldsymbol{z}\|_\infty$ yields

$$\mathbb{P}\left[\|\boldsymbol{z}\|_\infty \leqslant \beta \left\|\boldsymbol{\Gamma}^{-\frac{1}{2}}\right\|_\infty\right] \geqslant 1 - 2n\zeta_{2/3}(N,\beta).$$

Using the union bound to combine the bounds on $\|\boldsymbol{e}\|_\infty$ and $\|\boldsymbol{z}\|_\infty$ with the bounds previously $\left\|\boldsymbol{S}_N - \boldsymbol{\Sigma}^{-1}\right\|_\infty$ and $\left\|\boldsymbol{L}_N - \boldsymbol{\Gamma}^{-1}\right\|_\infty$, inequality (19) follows. □

Lemma 6 provides a general strategy for analyzing the error of various randomized methods. When a variable is not randomized, we simply drop the corresponding terms in equations (18) and (19), and adjust the probability accordingly. For example, when we do not randomize via $\boldsymbol{\lambda}$ and $\boldsymbol{\delta}$, (18) becomes

$$\left\|\boldsymbol{u}^{\mathrm{MAP}} - \mathring{\boldsymbol{u}}_N^{\mathrm{MAP}}\right\|_\infty \leqslant \beta\mathcal{L}\left(\left\|\boldsymbol{\Sigma}^{-1}\right\|_\infty + (1+\beta)\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty\right) \quad \text{w.p. at least } 1 - 5\zeta(N,\beta),$$

and (19) becomes

$$\left\|\boldsymbol{u}^{\mathrm{MAP}} - \hat{\boldsymbol{u}}_N^{\mathrm{MAP}}\right\|_\infty \leqslant \beta\mathcal{L}\left(\left\|\boldsymbol{\Sigma}^{-1}\right\|_\infty + \left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|_\infty\right) \quad \text{w.p. at least } 1 - 2\zeta(N,\beta) - 2k\zeta_{2/3}(N,\beta).$$

Additionally, as will be discussed in section 4.2.1, the Lipschitz constant may be affected by the choice of randomization strategy. By choosing to randomize only some variables, the inverse solution may have lower worst-case sensitivity (Lipschitz constant). Note that since we are mainly concerned with small deviations, it is sufficient to consider well-posed inverse problems where the solution depends continuously on $\boldsymbol{d}$ everywhere and depends on $\boldsymbol{P}$ in a locally Lipschitz manner in a neighborhood of $\left[\mathrm{vec}\left(\boldsymbol{\Sigma}^{-1}\right); \mathrm{vec}\left(\boldsymbol{\Gamma}^{-1}\right); 0; 0\right]$. Note that it is unlikely for the solution to depend continuously on $\boldsymbol{P}$ everywhere since $\boldsymbol{\Gamma}^{-1}$ acts as regularization. Otherwise, the original problem would not be ill-posed. In particular, the regularizing role that $\boldsymbol{\Gamma}^{-1}$ plays may cause the solution of the inverse problem to be especially sensitive to perturbations of $\boldsymbol{\Gamma}^{-1}$, greatly increasing the (local) Lipschitz constant compared to the case where $\boldsymbol{\Gamma}^{-1}$ is not randomized. This is clear from the linear case where the condition number of the problem takes the place of the Lipschitz constant. It is well-known that the choice of $\boldsymbol{\Gamma}^{-1}$ has a significant impact on the condition number (Chu *et al* 2011, Diao *et al* 2016). Lastly, note that in the linear case, standard perturbation theory for linear systems can be used to find an explicit bound for the relative error in terms of the condition number of $\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A} + \boldsymbol{\Gamma}^{-1}$.

3. Rediscovery of randomized inverse methods and discovery of new methods

In this section, we derive several different known randomization schemes as special cases of (4) and discover new randomization schemes. We begin with a brief survey of existing randomization schemes and then show how our framework naturally rediscovers them in the following sections. Table 1 gives a visual summary of the randomizations used by each existing method as well as the enumeration of a few new randomization schemes that we will discuss. The randomized maximum likelihood (RML) (Kitanidis 1995, Oliver *et al* 2008, Bardsley *et al* 2014), randomized MAP (RMAP) (Wang *et al* 2018), and randomize-then-optimize (RTO) (Bardsley *et al* 2014) methods each aim to reduce the cost of generating samples from the posterior by solving a series of inverse problems—ideally, one for each sample from the posterior. In the linear case, these methods coincide (Wang *et al* 2018). However, for nonlinear

Table 1. Tabular representation of randomization schemes. The top section including RML, RMAP, RTO, RMA, and EnKF are rediscovered methods that exist in the literature. The bottom section gives a few new methods that are discovered through our framework.

	$\boldsymbol{\sigma}$	$\varepsilon\varepsilon^{\mathrm{T}}$	$\boldsymbol{\delta}$	$\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}$
RML	✓		✓	
RMAP	✓		✓	
RTO	✓		✓	
RMA/LS		✓		
EnKF	✓		✓	✓
RMA + RMAP	✓	✓	✓	
RS				✓
ALL	✓	✓	✓	✓

problems, RMAP generates samples from a different proposal distribution than RTO, though both papers propose Metropolization methods to generate bona fide samples of the posterior, while the RML method accepts all samples directly.

In a similar manner, each step of the ensemble Kalman filter (EnKF) pushes random samples through an (indirect) optimization process to provide an estimate of uncertainty (Evensen 2003). While the EnKF is often used as part of an iterative process, either for data assimilation (Houtekamer and Mitchell 1998, Evensen *et al* 2009) or for inversion (Elsheikh *et al* 2013), we show that the EnKF update formula emerges from our framework for randomized inverse problems.

While the RML, RMAP, RTO, and EnKF all solve the inverse problem for multiple samples and use the generated solutions to quantify uncertainty, there are other randomized approaches within our framework that do not require solving many inverse problems. Such an existing method is the randomized misfit approach (RMA) (Le *et al* 2017). RMA, or left sketching (LS), uses randomization to perform dimension reduction of the likelihood. It was shown in Le *et al* (2017) that this approach can reduce the number of PDE solves required to solve PDE constrained inverse problems, thus it accelerates the estimation of the MAP point.

Going beyond the randomized methods existing in the literature, three additional randomizations will be discussed in detail: right sketching (RS) which randomizes the inverse of the prior covariance matrix via $\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}$, the combination of RMA and RMAP, and using the full randomization given in (4), which we denote ALL. The relationship between existing methods and new methods will be discussed, along with their relative strengths and weaknesses.

For the rest of this section, we will explicitly write each of the random variables (subset of $\boldsymbol{\xi} = [\boldsymbol{\sigma}, \varepsilon, \boldsymbol{\delta}, \boldsymbol{\lambda}]^{\mathrm{T}}$) that the stochastic cost function depends on. To keep notation clean, we will use $\mathcal{J}_\xi$ to represent the stochastic cost function for all randomization schemes. To avoid confusion, we put the random variables used in superscript and the number of samples in subscript. For example, a method that only randomizes $\boldsymbol{\sigma}$ and ε with N samples will be denoted $\boldsymbol{u}_N^{\boldsymbol{\sigma},\varepsilon}$. We take the expected value of $\mathcal{J}_\xi$ with respect to the unrandomized variables (i.e. we are not making SAAs). For example, if $\boldsymbol{\sigma}$ is not randomized, we replace it with 0 in (4). Likewise, we replace $\varepsilon\varepsilon^{\mathrm{T}}$ with $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{\delta}$ with 0, and $\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}$ with $\boldsymbol{\Gamma}^{-1}$ in (4) as appropriate. Additionally, we will present each method in the general nonlinear setting, but we will also explicitly write the sample average solution in the linear case as closed form solutions are available, yielding

additional insights. To that end, we write two equivalent formulations of the MAP estimate for linear inverse problems.

Proposition 7. *When the PtO map is linear, i.e.* $\mathcal{F}(\boldsymbol{u}) = \mathcal{A}\boldsymbol{u}$*, the solution of the MAP problem (2) can be written in two forms:*

$$\boldsymbol{u}_1 = \left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A} + \boldsymbol{\Gamma}^{-1}\right)^{-1}\left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{d} + \boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0\right) \tag{20a}$$

and

$$\boldsymbol{u}_2 = \boldsymbol{u}_0 + \boldsymbol{\Gamma}\mathcal{A}^{\mathrm{T}}\left(\boldsymbol{\Sigma} + \mathcal{A}\boldsymbol{\Gamma}\mathcal{A}^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{d} - \mathcal{A}\boldsymbol{u}_0\right). \tag{20b}$$

Proof. The first of these identities is derived directly from the optimality condition of (2). Specifically,

$$\nabla\mathcal{J} = \left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A} + \boldsymbol{\Gamma}^{-1}\right)\boldsymbol{u} - \mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{d} - \boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0 = 0. \tag{21}$$

The second formulation can be derived from $\boldsymbol{u}_1$ using the Sherman-Morrison-Woodbury formula (Deng 2011) under the conditions that $\boldsymbol{\Gamma}^{-1}$ and $\boldsymbol{\Sigma}^{-1}$ are invertible. □

Remark 8. This last assumption concerning the invertibility of $\boldsymbol{\Gamma}^{-1}$, while seemingly trivial in light of the fact that $\boldsymbol{\Gamma}^{-1}$ is written as the inverse of a matrix, will be important in the following discussion of randomization.

Additionally, let us define new random variables:

$$\boldsymbol{d}^i := \boldsymbol{d} + \boldsymbol{\sigma}^i \quad \text{and} \quad \boldsymbol{u}_0^i := \boldsymbol{u}_0 + \boldsymbol{\delta}^i, \tag{22}$$

where $\boldsymbol{\sigma}^i$ and $\boldsymbol{\delta}^i$ are the first and the third components of $\boldsymbol{\xi}^i$ defined in theorem 1. These quantities will be useful in the following discussion. By the LLN we have

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{d}^i \xrightarrow[N\to\infty]{a.s.} \boldsymbol{d} \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{u}_0^i \xrightarrow[N\to\infty]{a.s.} \boldsymbol{u}_0.$$

3.1. Randomizing via $\boldsymbol{\sigma}$ and $\boldsymbol{\delta}$

Assuming that the order of minimization and expectation can be interchanged[6], we can write

$$\arg\min_{\boldsymbol{u}} \mathbb{E}_{\pi_{\boldsymbol{\sigma}}\times\pi_{\boldsymbol{\delta}}}\left[\mathcal{J}_\xi\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\sigma},\boldsymbol{\delta}\right)\right] = \mathbb{E}_{\pi_{\boldsymbol{\sigma}}\times\pi_{\boldsymbol{\delta}}}\left[\arg\min_{\boldsymbol{u}} \mathcal{J}_\xi\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\sigma},\boldsymbol{\delta}\right)\right]. \tag{23}$$

The SAA of the RHS can then be written

$$\boldsymbol{u}_N^{\mathrm{RMAP}} := \boldsymbol{u}_N^{\boldsymbol{\sigma},\boldsymbol{\delta}} = \frac{1}{N}\sum_{i=1}^{N}\arg\min_{\boldsymbol{u}} \mathcal{J}_\xi\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\sigma}^i,\boldsymbol{\delta}^i\right). \tag{24}$$

This randomization approach coincides with the RMAP approach (Wang *et al* 2018) when $\mathbb{E}_{\pi_{\boldsymbol{\sigma}}}\left[\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}\right] = \boldsymbol{\Sigma}$ and $\mathbb{E}_{\pi_{\boldsymbol{\delta}}}\left[\boldsymbol{\delta}\boldsymbol{\delta}^{\mathrm{T}}\right] = \boldsymbol{\Gamma}$ (also known as the RML (Kitanidis 1995, Oliver *et al* 2008, Bardsley *et al* 2014)). In the linear case, we can write

$$\boldsymbol{u}_N^{\mathrm{RMAP}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{u}^{\boldsymbol{\sigma}^i,\boldsymbol{\delta}^i},$$

[6] The conditions under which the interchange is valid can be consulted in (Rockafellar and Wetts 1998, theorem 14.60).

where, thanks to proposition 7,

$$\begin{aligned}\left(\boldsymbol{u}^{\mathrm{RMAP}}\right)^{i} := \boldsymbol{u}^{\boldsymbol{\sigma}^{i},\boldsymbol{\delta}^{i}} &= \left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A}+\boldsymbol{\Gamma}^{-1}\right)^{-1}\left[\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{d}+\boldsymbol{\sigma}^{i}\right)+\boldsymbol{\Gamma}^{-1}\left(\boldsymbol{u}_0+\boldsymbol{\delta}^{i}\right)\right]\\ &= \left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A}+\boldsymbol{\Gamma}^{-1}\right)^{-1}\left[\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{d}^{i}+\boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0^{i}\right].\end{aligned}$$

Since the SAA (24) of the right hand side of (23) converges to its expectation, the analysis from section 2 applies.

Note that if *only* the MAP point $\boldsymbol{u}^{\mathrm{MAP}}$ is needed, then the RMAP approach is not useful: in fact it is very expensive while only giving an approximate solution for $\boldsymbol{u}^{\mathrm{MAP}}$. However, the approach could be appealing for Bayesian settings. By choosing $\mathbb{E}_{\pi_{\boldsymbol{\sigma}}}\left[\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}\right]=\boldsymbol{\Sigma}$ and $\mathbb{E}_{\pi_{\boldsymbol{\delta}}}\left[\boldsymbol{\delta}\boldsymbol{\delta}^{\mathrm{T}}\right]=\boldsymbol{\Gamma}$, each solution $\left(\boldsymbol{u}^{\mathrm{RMAP}}\right)^{i}$ is a bona fide sample of the posterior distribution in the linear case. For nonlinear cases, $\left(\boldsymbol{u}^{\mathrm{RMAP}}\right)^{i}$ are biased samples of the posterior (Wang *et al* 2018), but can be corrected via Metropolization (Bui-Thanh and Nguyen 2016, Chen *et al* 2020b). Note that for linear inverse problems, the RMAP approach is the same as the RTO approach in Bardsley *et al* (2014) (see an explanation from Wang *et al* 2018). For nonlinear problems, the theory in section 2 proves that the sample average solution converges to the MAP point. Typical Metropolization methods feature an accept/reject step where proposals are generated and then accepted or rejected as samples from the posterior based on some criteria. With this in mind, our theory and the specific methods discussed can be considered as simply ignoring the accept/reject step. There has been some recent work on this front showing that the error incurred due to accepting all samples may not be significant and substantial computational advantage can be achieved (Blatter *et al* 2022a, b). The RMAP method is embarrassingly parallel and is well-suited for implementation on distributed computing systems. While we could randomize the data and prior mean without exchanging expectation and optimization and convergence would be maintained, such a method would be of little use because we would obtain only an inaccurate approximation of $\boldsymbol{u}^{\mathrm{MAP}}$ while not reducing the cost of solving the inverse problem.

3.2. Randomizing the likelihood via $\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}$

In this section we show that the RMA (Le *et al* 2017) is a special case of our randomization in (4). Indeed, if we let $\boldsymbol{\varepsilon}\sim\pi_{\boldsymbol{\varepsilon}}$ where $\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}[\boldsymbol{\varepsilon}]=0$ and $\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right]=\boldsymbol{\Sigma}^{-1}$, then

$$\begin{aligned}\boldsymbol{u}^{\mathrm{RMA}} := \boldsymbol{u}^{\boldsymbol{\varepsilon}} &= \arg\min_{\boldsymbol{u}}\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\varepsilon}\right)\right]\\ &= \arg\min_{\boldsymbol{u}}\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\frac{1}{2}\left\|\boldsymbol{\varepsilon}^{\mathrm{T}}\left(\boldsymbol{d}-\mathcal{F}(\boldsymbol{u})\right)\right\|_2^2+\frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}_0\right\|_{\boldsymbol{\Gamma}^{-1}}^2\right].\end{aligned}$$

The SAA of $\boldsymbol{u}^{\mathrm{RMA}}$ can be written as

$$\boldsymbol{u}_N^{\mathrm{RMA}} := \boldsymbol{u}_N^{\boldsymbol{\varepsilon}} = \arg\min_{\boldsymbol{u}}\frac{1}{N}\sum_{i=1}^{N}\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\varepsilon}^{i}\right) \tag{25a}$$

$$= \arg\min_{\boldsymbol{u}}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left\|\tilde{\boldsymbol{d}}^{i}-\tilde{\mathcal{F}}^{i}(\boldsymbol{u})\right\|_2^2+\frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}_0\right\|_{\boldsymbol{\Gamma}^{-1}}^2, \tag{25b}$$

where

$$\tilde{\mathcal{F}}^{i} := \boldsymbol{\varepsilon}^{i\mathrm{T}}\mathcal{F} \quad \text{and} \quad \tilde{\boldsymbol{d}}^{i} := \boldsymbol{\varepsilon}^{i\mathrm{T}}\boldsymbol{d}.$$

That is, the random samples, which can be combined into a random matrix, sketch the PtO map and the data from the left. Rather than working with a potentially high dimensional misfit term, the dimension of the problem is reduced by first multiplying through by a so-called *sketching matrix*. Random sketching has been used extensively to reduce the cost of solving inverse problems Clarkson and Woodruff (2017), Liu *et al* (2018), Chen *et al* (2020a). It was shown in Le *et al* (2017) that sketching can reduce the number of required PDE solves for nonlinear PDE constrained inverse problems. A review of randomized sketching from a statistical perspective can be found in Raskutti and Mahoney (2016). Calculating the optimality condition in the linear case results in

$$\boldsymbol{u}^{\mathrm{RMA}} = \left(\mathcal{A}^{\mathrm{T}}\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right]\mathcal{A} + \boldsymbol{\Gamma}^{-1}\right)^{-1}\left(\mathcal{A}^{\mathrm{T}}\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right]\boldsymbol{d} + \boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0\right). \tag{26}$$

By letting

$$\tilde{\mathcal{A}} := \varepsilon^{\mathrm{T}}\mathcal{A} \quad \text{and} \quad \tilde{\boldsymbol{d}} := \varepsilon^{\mathrm{T}}\boldsymbol{d},$$

we can rewrite (26) as

$$\boldsymbol{u}^{\mathrm{RMA}} = \left(\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\tilde{\mathcal{A}}^{\mathrm{T}}\tilde{\mathcal{A}}\right] + \boldsymbol{\Gamma}^{-1}\right)^{-1}\left(\mathbb{E}_{\pi_{\boldsymbol{\varepsilon}}}\left[\tilde{\mathcal{A}}^{\mathrm{T}}\tilde{\boldsymbol{d}}\right] + \boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0\right)$$

3.3. Randomizing via $\boldsymbol{\sigma}, \boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}$

If we combine the RMA and RMAP approaches into a single stochastic optimization problem, we discover a new method which we will denote RMA+RMAP. Specifically, consider the problem directly arising from randomization of (4) and define the solution using the RMA+RMAP method to be

$$\begin{aligned}\boldsymbol{u}^{\mathrm{RMA+RMAP}} := \boldsymbol{u}^{\boldsymbol{\sigma},\boldsymbol{\varepsilon},\boldsymbol{\delta}} &= \arg\min_{\boldsymbol{u}} \mathbb{E}_{\pi_{\boldsymbol{\sigma}}\times\pi_{\boldsymbol{\varepsilon}}\times\pi_{\boldsymbol{\delta}}}\left[\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\sigma},\boldsymbol{\varepsilon},\boldsymbol{\delta}\right)\right] \\ &= \arg\min_{\boldsymbol{u}} \mathbb{E}_{\pi_{\boldsymbol{\sigma}}\times\pi_{\boldsymbol{\varepsilon}}\times\pi_{\boldsymbol{\delta}}}\left[\frac{1}{2}\left\|\boldsymbol{\varepsilon}^{\mathrm{T}}\left(\boldsymbol{d}+\boldsymbol{\sigma}-\mathcal{F}(\boldsymbol{u})\right)\right\|_2^2 + \frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta}\right\|_{\boldsymbol{\Gamma}^{-1}}^2\right],\end{aligned}$$

and the corresponding SAA solution

$$\begin{aligned}\boldsymbol{u}_N^{\mathrm{RMA+RMAP}} := \boldsymbol{u}_N^{\boldsymbol{\sigma},\boldsymbol{\varepsilon},\boldsymbol{\delta}} &= \arg\min_{\boldsymbol{u}} \frac{1}{N}\sum_{i=1}^{N}\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\sigma}^i,\boldsymbol{\varepsilon}^i,\boldsymbol{\delta}^i\right) \\ &= \arg\min_{\boldsymbol{u}} \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{2}\left\|\boldsymbol{\varepsilon}^{i\mathrm{T}}\left(\boldsymbol{d}+\boldsymbol{\sigma}^i-\mathcal{F}(\boldsymbol{u})\right)\right\|_2^2 + \frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta}^i\right\|_{\boldsymbol{\Gamma}^{-1}}^2\right].\end{aligned}$$

Allowing for the interchange of optimization and expectation as in the RMAP approach along with independent SAAs of each random variable, the following variant sequences also converge.

$$\boldsymbol{u}_N^{\mathrm{RMA+RMAP_1}} = \frac{1}{N}\sum_{i=1}^{N}\arg\min_{\boldsymbol{u}}\left[\frac{1}{2}\left\|\boldsymbol{\varepsilon}^{i\mathrm{T}}\left(\boldsymbol{d}+\boldsymbol{\sigma}^i-\mathcal{F}(\boldsymbol{u})\right)\right\|_2^2 + \frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta}^i\right\|_{\boldsymbol{\Gamma}^{-1}}^2\right]. \tag{27}$$

$$\boldsymbol{u}_{N,M}^{\mathrm{RMA+RMAP_2}} = \frac{1}{M}\sum_{i=1}^{M}\arg\min_{\boldsymbol{u}}\frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{2}\left\|\boldsymbol{\varepsilon}^{j\mathrm{T}}\left(\boldsymbol{d}+\boldsymbol{\sigma}^i-\mathcal{F}(\boldsymbol{u})\right)\right\|_2^2 + \frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}_0-\boldsymbol{\delta}^i\right\|_{\boldsymbol{\Gamma}^{-1}}^2\right]. \tag{28}$$

We would like to point out (28) is perhaps the most intuitive way to combine RMA and RMAP. Randomization of the noise covariance matrix acts as a random projection (LS) while the randomized prior mean and data aid in sampling from the posterior. Note that (28) arises as a variant of the loss function defined in equation (10) by exchanging the optimization and expectation of only $\boldsymbol{\sigma}$ and $\boldsymbol{\delta}$. On the other hand, (27) would likely yield inaccurate results as it is the sum of solutions where the PtO map and data have been projected onto a one dimensional subspace: thus the prior dominates each solution. In the linear case, $\boldsymbol{u}_{N,M}^{\text{RMA+RMAP}_2}$ can be written as

$$\boldsymbol{u}_{N,M}^{\text{RMA+RMAP}_2} = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{1}{N}\sum_{j=1}^{N}\left(\tilde{\mathcal{A}}^j\right)^{\mathrm{T}}\tilde{\mathcal{A}}^j + \boldsymbol{\Gamma}^{-1}\right)^{-1}\left(\frac{1}{N}\sum_{j=1}^{N}\left(\tilde{\mathcal{A}}^j\right)^{\mathrm{T}}\left(\boldsymbol{\varepsilon}^j\right)^{\mathrm{T}}\boldsymbol{d}^i + \boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0^i\right).$$

Clearly we also have convergence to the MAP point of other combinations, such as randomizing only one of the data or prior mean, but their enumeration here is omitted in the interest of space.

3.4. Randomizing the prior via $\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}$

Here we propose a randomization scheme based on randomizing $\boldsymbol{\Gamma}^{-1}$ though $\boldsymbol{\lambda} \sim \pi_{\boldsymbol{\lambda}}$ where $\mathbb{E}_{\pi_{\boldsymbol{\lambda}}}[\boldsymbol{\lambda}] = 0$ and $\mathbb{E}_{\pi_{\boldsymbol{\lambda}}}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right] = \boldsymbol{\Gamma}^{-1}$. Let

$$\begin{aligned}\boldsymbol{u}^{\text{RS\_U1}} := \boldsymbol{u}^{\boldsymbol{\lambda}} &= \arg\min_{\boldsymbol{u}} \mathbb{E}_{\pi_{\boldsymbol{\lambda}}}\left[\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\lambda}\right)\right] \\ &= \arg\min_{\boldsymbol{u}} \mathbb{E}_{\pi_{\boldsymbol{\lambda}}}\left[\frac{1}{2}\left\|\boldsymbol{d}-\mathcal{F}(\boldsymbol{u})\right\|_{\boldsymbol{\Sigma}^{-1}}^2 + \frac{1}{2}\left\|\boldsymbol{\lambda}^{\mathrm{T}}\left(\boldsymbol{u}-\boldsymbol{u}_0\right)\right\|_2^2\right]. \end{aligned} \tag{29}$$

The reason for designating this method 'RS_U1' is due to its relationship with the RS approach (see section 4.1). Then the SAA reads

$$\begin{aligned}\boldsymbol{u}_N^{\text{RS\_U1}} := \boldsymbol{u}_N^{\boldsymbol{\lambda}} &= \arg\min_{\boldsymbol{u}} \frac{1}{N}\sum_{i=1}^{N}\mathcal{J}_{\xi}\left(\boldsymbol{u};\boldsymbol{u}_0,\boldsymbol{d},\boldsymbol{\varepsilon}^i\right) \\ &= \arg\min_{\boldsymbol{u}} \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left\|\boldsymbol{d}-\mathcal{F}(\boldsymbol{u})\right\|_{\boldsymbol{\Sigma}^{-1}}^2 + \frac{1}{2}\left\|\left(\boldsymbol{\lambda}^i\right)^{\mathrm{T}}\left(\boldsymbol{u}-\boldsymbol{u}_0\right)\right\|_2^2. \end{aligned} \tag{30}$$

In the linear case, the optimality condition yields the following solution

$$\boldsymbol{u}^{\text{RS\_U1}} = \left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A} + \mathbb{E}_{\pi_{\boldsymbol{\lambda}}}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\right)^{-1}\left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{d} + \mathbb{E}_{\pi_{\boldsymbol{\lambda}}}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\boldsymbol{u}_0\right).$$

This approach can be thought of as sketching the prior from the left.

4. Optimize, transform, then randomize

An important observation about the methods discussed so far is that the linear settings are solved using $\boldsymbol{u}_1$ given in (20*a*). That is, to show the equivalence of the solution of the randomized cost function to the solution of the corresponding method in the literature, one only needs to consider the optimal solution (20*a*). Additionally, the SAA of the cost function is exactly the same as replacing the expectations in form $\boldsymbol{u}_1$ with their respective SAAs. The next methods require form $\boldsymbol{u}_2$ in (20*b*) to see the equivalence of the randomized solution and the corresponding method given in the literature—where the Sherman-Morrison-Woodbury

formula is applied to the optimality condition before making SAAs. *As $\boldsymbol{u}_2$ is only equivalent to $\boldsymbol{u}_1$ in the linear case, we will restrict the following discussion to linear inverse problems.*

4.1. Randomizing via $\omega\omega^{\mathrm{T}}$—revisiting $\lambda\lambda^{\mathrm{T}}$

Since we are now considering schemes derived from randomizing $\boldsymbol{u}_2$, we introduce a new random variable, $\boldsymbol{\omega}$, defined such that $\mathbb{E}_{\pi_\omega}[\boldsymbol{\omega}] = 0$ and $\mathbb{E}_{\pi_\omega}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\right] = \boldsymbol{\Gamma}$. By taking advantage of the asymptotic convergence of the SAA of $\mathbb{E}_{\pi_\omega}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\right]$, we have

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\omega}^i\left(\mathcal{A}\boldsymbol{\omega}^i\right)^{\mathrm{T}} \xrightarrow[N\to\infty]{a.s.} \mathbb{E}_{\boldsymbol{\omega}}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\mathcal{A}^{\mathrm{T}}\right] = \boldsymbol{\Gamma}\mathcal{A}^{\mathrm{T}}, \tag{31a}$$

$$\frac{1}{N}\sum_{i=1}^{N}\left(\mathcal{A}\boldsymbol{\omega}^i\right)\left(\mathcal{A}\boldsymbol{\omega}^i\right)^{\mathrm{T}} \xrightarrow[N\to\infty]{a.s.} \mathbb{E}_{\boldsymbol{\omega}}\left[\mathcal{A}\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\mathcal{A}^{\mathrm{T}}\right] = \mathcal{A}\boldsymbol{\Gamma}\mathcal{A}^{\mathrm{T}}. \tag{31b}$$

Combining (20*b*) and (31) gives

$$\boldsymbol{u}_N^{\mathrm{RS}} := \boldsymbol{u}_N^{\boldsymbol{\omega}} = \boldsymbol{u}_0 + \left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\omega}^i\left(\mathcal{A}\boldsymbol{\omega}^i\right)^{\mathrm{T}}\right)\left(\boldsymbol{\Sigma} + \frac{1}{N}\sum_{i=1}^{N}\left(\mathcal{A}\boldsymbol{\omega}^i\right)\left(\mathcal{A}\boldsymbol{\omega}^i\right)^{\mathrm{T}}\right)^{-1}(\boldsymbol{d} - \mathcal{A}\boldsymbol{u}_0), \tag{32}$$

which is the same as sketching the PtO map $\mathcal{A}$ from the right or sketching the transpose of the PtO map from the left.

Lemma 9 (Asymptotic convergence of RS). *Let $\boldsymbol{u}_N^{RS}$ be defined in (32) and assume that $\mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]$ (where $\boldsymbol{\lambda}$ is defined in section 2) is invertible. Then*

$$\boldsymbol{u}_N^{RS} \xrightarrow{a.s.} \boldsymbol{u}^{MAP} \quad \text{as } N \to \infty.$$

Proof. Beginning with equation (2) and randomizing only $\boldsymbol{\Gamma}^{-1}$ through $\boldsymbol{\lambda}$, the optimality condition is

$$\boldsymbol{u}^* = \left(\mathcal{A}\boldsymbol{\Sigma}^{-1}\mathcal{A} + \mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\right)^{-1}\left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{d} + \mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\boldsymbol{u}_0\right).$$

Since $\mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]$ is assumed to be invertible, this can be rewritten using the Sherman-Morrison-Woodbury formula in the form $\boldsymbol{u}_2$ as

$$\boldsymbol{u}^* = \boldsymbol{u}_0 + \left(\mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\right)^{-1}\mathcal{A}^{\mathrm{T}}\left(\boldsymbol{\Sigma} + \mathcal{A}\left(\mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\right)^{-1}\mathcal{A}^{\mathrm{T}}\right)^{-1}(\boldsymbol{d} - \mathcal{A}\boldsymbol{u}_0).$$

Before making an SAA, note that

$$\left(\mathbb{E}_{\pi_\lambda}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right]\right)^{-1} = \left(\boldsymbol{\Gamma}^{-1}\right)^{-1} = \boldsymbol{\Gamma} = \mathbb{E}_{\pi_\omega}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\right].$$

Then,

$$\boldsymbol{u}^* = \boldsymbol{u}^{\mathrm{RS}} = \boldsymbol{u}_0 + \mathbb{E}_{\pi_\omega}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\right]\mathcal{A}^{\mathrm{T}}\left(\boldsymbol{\Sigma} + \mathcal{A}\mathbb{E}_{\pi_\omega}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\right]\mathcal{A}^{\mathrm{T}}\right)^{-1}(\boldsymbol{d} - \mathcal{A}\boldsymbol{u}_0).$$

Since matrix multiplication and matrix inversion are continuous functions,

$$\boldsymbol{u}_N^{\mathrm{RS}} \xrightarrow{a.s.} \boldsymbol{u}^{\mathrm{RS}} = \boldsymbol{u}^*,$$

by the continuous mapping theorem (Van der Vaart 2000, theorem 2.3). □

The key step here is recognizing that, asymptotically, sampling from $\pi_{\boldsymbol{\lambda}}$ and solving using form $\boldsymbol{u}_1$ (20*a*) gives the same results as sampling from $\pi_{\boldsymbol{\omega}}$ and solving using form $\boldsymbol{u}_2$ (20*b*). However, lemma 9 does not imply that $\boldsymbol{u}_N^{\boldsymbol{\omega}}$ using form $\boldsymbol{u}_2$ is equivalent to $\boldsymbol{u}_N^{\boldsymbol{\lambda}}$ using form $\boldsymbol{u}_1$ for a finite N. Indeed, when $N < \dim(\boldsymbol{u})$, $\boldsymbol{u}_N^{\boldsymbol{\omega}}$ cannot be rewritten in the form $\boldsymbol{u}_1$ since

$$\operatorname{rank}\left(\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\omega}^j\left(\boldsymbol{\omega}^j\right)^{\mathrm{T}}\right) \leqslant N < \dim(\boldsymbol{u}).$$

This implies that the sample average of $\mathbb{E}_{\pi_{\boldsymbol{\omega}}}\left[\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}\right]$ is not invertible, breaking an assumption of lemma 9 and showing that $\boldsymbol{u}_N^{\boldsymbol{\omega}}$ does not satisfy the optimality condition of (2). Here, it is important that $\boldsymbol{\Gamma}^{-1}$ is invertible, otherwise $\boldsymbol{u}_1 \neq \boldsymbol{u}_2$ and the randomized schemes discussed here have no hope of converging to $\boldsymbol{u}^*$.

4.2. Randomizing via $\boldsymbol{\sigma}, \boldsymbol{\delta}$, and $\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}$

If in addition to RS, we use (22) to randomize $\boldsymbol{d}$ and $\boldsymbol{u}_0$ in (32), and define

$$\left(\boldsymbol{u}_N^{\mathrm{ENKF}}\right)^i := \boldsymbol{u}_N^{\boldsymbol{\sigma}^i,\boldsymbol{\delta}^i,\boldsymbol{\omega}} = \boldsymbol{u}_0^i + \left(\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\omega}^j\left(\mathcal{A}\boldsymbol{\omega}^j\right)^{\mathrm{T}}\right)\left(\boldsymbol{\Sigma} + \frac{1}{N}\sum_{j=1}^{N}\left(\mathcal{A}\boldsymbol{\omega}^j\right)\left(\mathcal{A}\boldsymbol{\omega}^j\right)^{\mathrm{T}}\right)^{-1} \times \left(\boldsymbol{d}^i - \mathcal{A}\boldsymbol{u}_0^i\right), \tag{33}$$

we rediscover the well-known EnKF update formula for a single member of the ensemble (Evensen 2003). Notice here that the sketching of $\mathcal{A}$ from the right is fixed for each random sample $\boldsymbol{d}^i$ and $\boldsymbol{u}_0^i$. In the language of the EnKF, the sample prior covariance matrix is fixed for all members of the ensemble. Our framework recovers the update formula for a single step of the EnKF for each member of the ensemble. The EnKF is often used in the data assimilation community to propagate the state of a dynamical system while incorporating any measurements. At each timestep, (33) is applied to each member of the ensemble. In the inverse problems community, the iterative technique using the EnKF is referred to as ensemble Kalman inversion (EKI) (Iglesias *et al* 2013, Chada *et al* 2020). The EKI approach yields samples from the posterior in the linear case after just a single application of (33) (Iglesias *et al* 2013).

As with RS, convergence to the MAP point only holds asymptotically, with special sensitivity to N, since the validity of $\left(\boldsymbol{u}_N^{\mathrm{ENKF}}\right)^i$ as an optimal solution of (2) requires N to be large enough to ensure invertibility of all matrices involved. The reason for this can be understood by investigating what RS (and thus the EnKF) is doing to the prior covariance matrix and viewing this through the lens of regularization.

4.2.1. RS from the left as randomized regularization. Consider again the form $\boldsymbol{u}_1$ given in (20*a*):

$$\boldsymbol{u}_1 = \left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A} + \boldsymbol{\Gamma}^{-1}\right)^{-1}\left(\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{d} + \boldsymbol{\Gamma}^{-1}\boldsymbol{u}_0\right).$$

While the randomized prior (30), RS (32) and EnKF (33) methods still fall under the asymptotic analysis given in section 2.1 for linear inverse problems, a practical and theoretical issue arises due to the regularizing role that $\boldsymbol{\Gamma}^{-1}$ plays. The inverse of the prior covariance, $\boldsymbol{\Gamma}^{-1}$, can be considered to be a regularization operator when viewed through the lens of deterministic inverse problems and is indeed equivalent to a Tikhonov regularization strategy (Engl *et al* 1989). In the deterministic setting, the role of regularization is often to 'damp out' highly

oscillatory modes caused by the rapidly decaying spectrum of $\mathcal{A}$—modes that are highly polluted by noise. While asymptotic analysis (see theorem 1) establishes the convergence of these aforementioned methods, it is incapable of explaining why these methods could fail for finite sample size N. This is where non-asymptotic analysis shines. Indeed, lemma 6 shows that the successful (small error) probability requires quite a large number of samples. According to remark 4.7.2 of Vershynin (2018), the number of samples required to accurately estimate the covariance matrix is proportional to n/β^2 where n is the dimension of the matrix and β is the tolerance. This is not surprising from a regularization point of view as the sample covariance needs to closely approximate the true covariance in order to adequately perform its role as a regularizer.

As a concrete example, consider the simple case where $\boldsymbol{\Gamma}^{-1} = \alpha\mathcal{I}$ with $\alpha > 0$. Letting $\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathcal{A} = \boldsymbol{USV}^{\mathrm{T}}$ be the SVD of the whitened PtO map, the first term of $\boldsymbol{u}_1$ can be written

$$\left(\boldsymbol{VS}^2\boldsymbol{V}^{\mathrm{T}} + \boldsymbol{\Gamma}^{-1}\right)^{-1} = \left(\boldsymbol{V}\left(\boldsymbol{S}^2 + \alpha\mathcal{I}\right)\boldsymbol{V}^{\mathrm{T}}\right)^{-1} = \boldsymbol{VDV}^{\mathrm{T}},$$

where $\mathbf{D}$ is the diagonal matrix with the ith diagonal element given by $\mathbf{D}_{ii} = \frac{1}{\boldsymbol{S}_{ii}^2+\alpha}$. Comparing to the case of no regularization, we can see that the inverse of the prior covariance shifts the spectrum of $\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A}$ upward by the constant α. Furthermore, upon inverting, $\alpha > 0$ ensures that the denominator of $\frac{1}{\boldsymbol{S}_{ii}^2+\alpha}$ is not too close to 0, keeping the inverse solution from blowing up as $\boldsymbol{S}_{ii}^2 \to 0$. Now, consider the RS_U1 randomization of $\boldsymbol{\Gamma}^{-1}$ proposed in (30)—the same randomization as RS when viewed in the $\boldsymbol{u}_1$ form (sketching the prior from the left):

$$\boldsymbol{\Gamma}^{-1} = \mathbb{E}_{\pi_{\boldsymbol{\lambda}}}\left[\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right] \approx \frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\lambda}^i\right)\left(\boldsymbol{\lambda}^i\right)^{\mathrm{T}}.$$

As we saw before, this randomization converges as $N \to \infty$, but the convergence rate $\mathcal{O}(1/\sqrt{N})$ of a SAA is notoriously slow. So how does this slow convergence affect the regularization strategy? Clearly when $N < \dim(\boldsymbol{u})$, the regularization is not full rank and there may be $\dim(\boldsymbol{u}) - N$ modes of $\mathcal{A}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathcal{A}$ left unregularized, assuming the random matrix has linearly independent columns. Even in the case when $N \geqslant \dim(\boldsymbol{u})$, slow convergence of the SAA leaves modes underregularized leading to oscillatory solutions as seen in figures 12(d) and (e) for the 1D deconvolution problem. This can also be seen explicitly in figure 1(a) where the spectrum of the sample average inverse covariance is plotted against the spectrum of the true prior inverse covariance for various N.

In figure 1(a), we consider the case where $\boldsymbol{\Gamma} = \mathcal{I}$ and $\dim(\boldsymbol{u}) = 1000$. Because randomizing the inverse of the prior covariance results in a poor performing regularizer, solutions using RS or a single step of the EnKF exhibit highly oscillatory behavior when choosing N to be of reasonable size, at least for the identity prior. In problems where a decaying prior spectrum is desirable, randomization of the prior has a less pronounced effect on the quality of the inverse solution. For example, the advection-diffusion PDE constrained inverse problem detailed in section 5.3 with the BiLaplacian prior shows good results with RS. The similarity of the sample average spectrum to the spectrum of the true prior inverse covariance can be seen in figure 1(b) for the BiLaplacian prior. Additionally, problems where the PtO map has a slowly decaying spectrum as in the x-ray tomography problem (section 5.2) may also be less sensitive to inaccurate approximations to $\boldsymbol{\Gamma}^{-1}$.

Lest we give the impression that all hope of obtaining high quality inverse solutions is lost when randomizing the prior covariance with few samples, there is extensive research

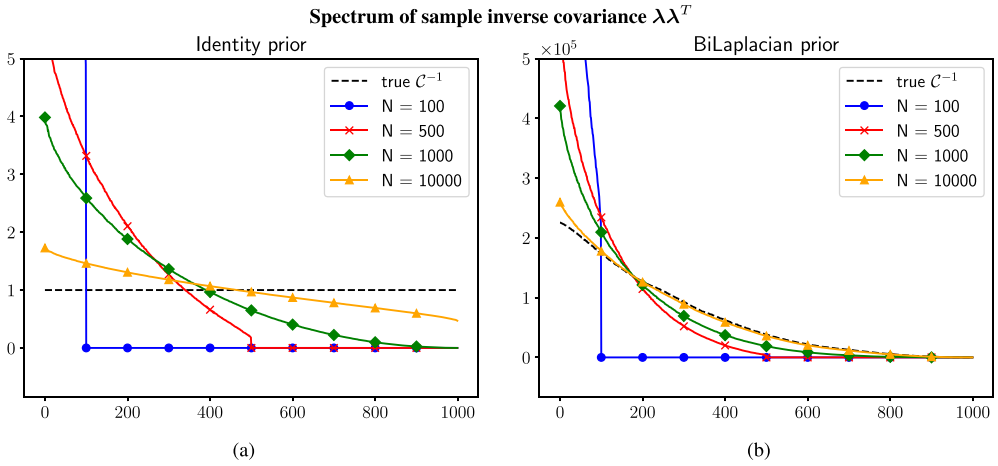


Figure 1. Convergence of spectrum of the sample average approximation of the inverse prior covariance for the case where $\mathbf{\Gamma} = \mathcal{I} \in \mathbb{R}^{1000\times 1000}$ (a) and $\mathbf{\Gamma}$ is the BiLaplacian (b). When $N = 100$, there are fewer samples than the dimension of the parameter and some modes are left completely unregularized. Even when there are more samples than the dimension of the parameter, this does not guarantee acceptable convergence for SAA. This shows that the sample average of the inverse prior covariance converges slowly to the true inverse prior covariance. However, when the spectrum of the inverse prior covariance decays, the sample average approximation more closely matches the true inverse prior covariance with fewer samples.

into methods of modifying the sample covariance matrix of the EnKF to address the rank-deficiency problem. While such methods do not fall under our framework for asymptotic convergence, it is nonetheless useful to mention them here, though we defer to the references for detailed treatment. The two main approaches are localization and covariance inflation. Localization addresses the issue of spurious long-range correlations induced by having few random samples and is performed by element-wise multiplication of the sample prior covariance matrix by a sparse covariance matrix ensuring locality of the correlations (Petrie 2008, Pourahmadi 2011, Farchi and Bocquet 2019). The second approach of covariance inflation adds a full-rank covariance matrix to the sample covariance matrix at each iteration of the EnKF to prevent ensemble collapse and solve the rank-deficiency problem (Anderson 2007, Whitaker *et al* 2008, Elsheikh *et al* 2013, Schillings and Stuart 2017). In the limit of infinite samples, however, these approaches are not guaranteed to recover the MAP point.

5. Numerical results

In this section we show numerical results for a variety of inverse problems demonstrating the asymptotic convergence of various methods. As the possible number of randomized variants would be unnecessarily burdensome to enumerate, we will focus on a few key methods: RMA (25), RMAP (24), the combination of RMA and RMAP (28), RS (32), the EnKF (33), and randomizing everything (10) (listed as ALL). It is important to keep in mind that we are not advocating for or against the use of any particular method. The purpose of this section is two-fold: to serve as numerical validation of the asymptotic convergence of each method and to provide practical insight into the behavior of each method on a variety of problem types as

we show no one scheme is best for all problems. In particular, we find empirically that methods randomizing $\mathbf{\Gamma}^{-1}$ such as RS, the EnKF, and ALL generally have very poor performance and require many more samples than the dimension of the problem in order to provide suitable results for several problems. The reason for this has been discussed at length in section 4.2.1. These methods do however exhibit asymptotic convergence to the MAP solution as predicted by our theoretical results.

To explore the performance and convergence of the various methods, we consider a variety of prototype problems with different characteristics. The 1D deconvolution problem with scaled identity prior covariance is a relatively simple inverse problem that provides easily digestible visualizations of the convergence for each method. X-ray tomography is a mildly ill-posed two dimensional imaging problem with fewer observations than parameters. The fact that it is only mildly ill-posed exposes interesting effects in the context of randomization. We also show the convergence of each method for a linear time dependent PDE-constrained inverse problem with PDE-based prior covariance on a domain with a hole. Finally, we conclude with an example demonstrating convergence on a non-linear elliptic PDE-constrained inverse problem.

In problems with more than one randomization, such as EnKF and RMA+RMAP, each expectation can be approximated by a separate sample average. However, exploring the effect of choosing a different number of samples for each random variable is outside the scope of this paper and serves only to obscure the asymptotic convergence property that we aim to show in this section. Therefore, all methods assume that the number of random samples is the same for all random variables, i.e. $N = N_1 = N_2 = N_3 = N_4$. In addition, the relative errors presented are with respect to $\boldsymbol{u}^{\text{MAP}}$, not the true solution, emphasizing the errors induced by randomization rather than errors due to other effects. This is due to the fact that the theory presented shows convergence to $\boldsymbol{u}^{\text{MAP}}$.

5.1. 1D Deconvolution problem

Deconvolution, the inverse problem associated with the convolution process, finds enormous application in the signal and image processing domains (Kundur and Hatzinakos 1996, Ryan and Debbah 2007, Swedlow 2013). For demonstration, we consider the 1D deconvolution problem with a 1-periodic function given by:

$$f(x) = \sin(2\pi x) + \cos(2\pi x) \quad x \in [0,1].$$

The domain is divided into $n = 1000$ sub-intervals. The kernel is constructed as (Mueller and Siltanen 2012):

$$\Psi(x) = C_a \left(x + a\right)^2 \left(x - a\right)^2,$$

where $a = 0.235$ and the constant C_a is chosen to enforce the normalization condition (Mueller and Siltanen 2012). Synthetic observations are generated with 5% additive Gaussian noise. We choose to randomize using the Achlioptas distribution (Achlioptas 2003, Le *et al* 2017), an example of an $l-$percent sparse random variable with $l = 2/3$ and entries in $\{-1, 0, 1\}$ with equal probability. The reconstructed functions obtained by different randomization approaches are shown in figure 12 and the relative errors are given in table 2. It can be seen that the RS and EnKF methods give the least accurate results as evidenced in table 2. This is because randomizing the inverse of the prior covariance results in poor performance as a regularizer, providing numerical confirmation of the discussion in section 4.2.1. Other methods perform reasonably well. While not all methods perform equally well, all methods converge as more

Table 2. Relative error for various randomized methods compared to the $\boldsymbol{u}^{\text{MAP}}$ solution for 1D deconvolution.

	Relative error (%)				
Method	$N = 10$	$N = 100$	$N = 1000$	$N = 10000$	$N = 100000$
RMAP	7.74	2.39	0.80	0.25	0.07
RMA	35.5	4.2	1.8	0.42	0.18
RMA+RMAP	41.7	9.0	1.7	0.80	0.15
RS	917	295	106	31.4	9.8
ENKF	896	352	100	31.9	9.7
ALL	158	33 875	844	33.3	9.3

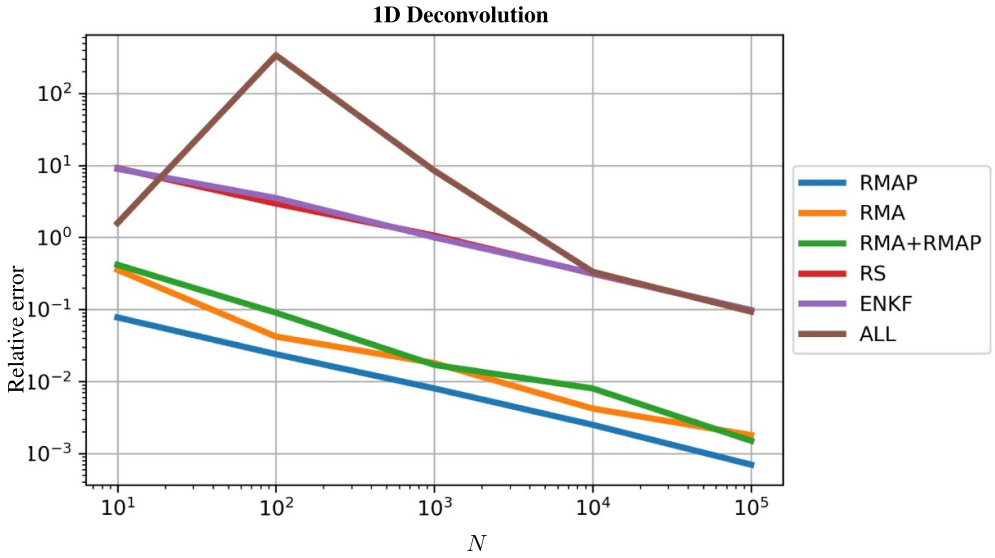


Figure 2. Relative error plot for 1D Deconvolution problem with Achlioptas random variable.

samples are taken and this is consistent with our asymptotic convergence results as seen in figure 2.

According to lemma 6, the probability of committing an error greater than some tolerance, say β, depends on the number of degrees of freedom, or mesh size, denoted n. This arises from the fact that a factor like $\frac{1}{n}$ is buried inside the constant c. Additionally, recall that we can move constants between the probability of failure, $\zeta(N, \beta)$, and the tolerance by relabeling the tolerance as was done in proposition 2 to move $\frac{1}{N}$ into ζ. That is to say, by letting $\beta = n\beta_1$, we can move the mesh dependence onto the tolerance rather than the probability of failure. Therefore, as the mesh is refined and there are greater degrees of freedom, our theory predicts that greater errors will be made with a fixed probability of failure. This is indeed the case as is shown in figure 3 the relative error between the EnKF solution as a function of number of samples is shown for several mesh sizes. As the mesh is refined, we still see asymptotic convergence toward the MAP point, though the relative error value is higher for more refined meshes. Each of the other randomized methods exhibit similar behavior, validating the non-asymptotic error analysis given in section 2.2.

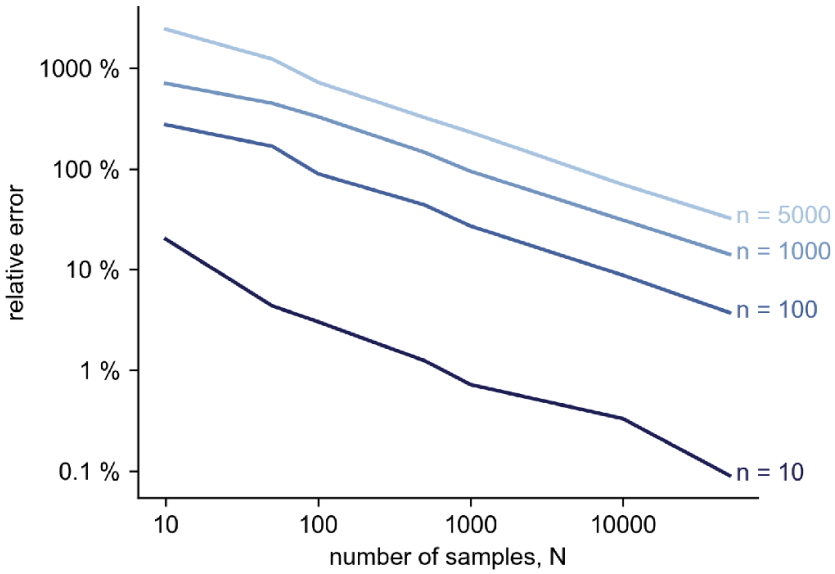


Figure 3. Shown here is the relative error between the EnKF solution as a function of number of samples, N, and the MAP solution for the 1D deconvolution problem. As the mesh degrees of freedom, n, increases, so does the error. This behavior is predicted by lemma 6.

Table 3. Relative error for various randomized methods compared to the $\boldsymbol{u}^{\text{MAP}}$ solution for the x-ray tomography problem.

	Relative error (%)				
Method	$N = 10$	$N = 100$	$N = 1000$	$N = 10000$	$N = 50000$
RMAP	6.44	6.44	6.44	6.44	0.04
RMA	94.17	74.95	39.42	31.07	4.02
RMA+RMAP	96.64	77.94	40.35	30.89	4.02
RS	191.87	373.07	176.02	51.87	22.43
ENKF	324.27	352.58	178.58	52.25	21.97
ALL	95.12	71.30	80.50	60.36	23.60

5.2. *X-ray tomography*

In x-ray tomographic imaging, x-ray projections of an object are captured at multiple angles and the inverse problem is to recover the internal structure of the object from the projection data (Mueller and Siltanen 2012). We consider the canonical *phantom* image of size 64×64 pixels with 45 measurement angles uniformly divided over the range $[0, \pi]$. With this number of measurement angles, the PtO map has shape 64×45 by 64^2 resulting in fewer observations than parameters (pixels). A scaled identity prior covariance is once again considered. Measurements are corrupted with 1% additive Gaussian noise.

The results are shown in figure 13 while table 3 and figure 4 show the relative error for different methods. Two observations are in order. First, results show asymptotic convergence of all methods, though convergence is noticeably slower for RMA and RMA+RMAP than in previous problems. This occurs because x-ray tomography is only a mildly ill-posed inverse problem

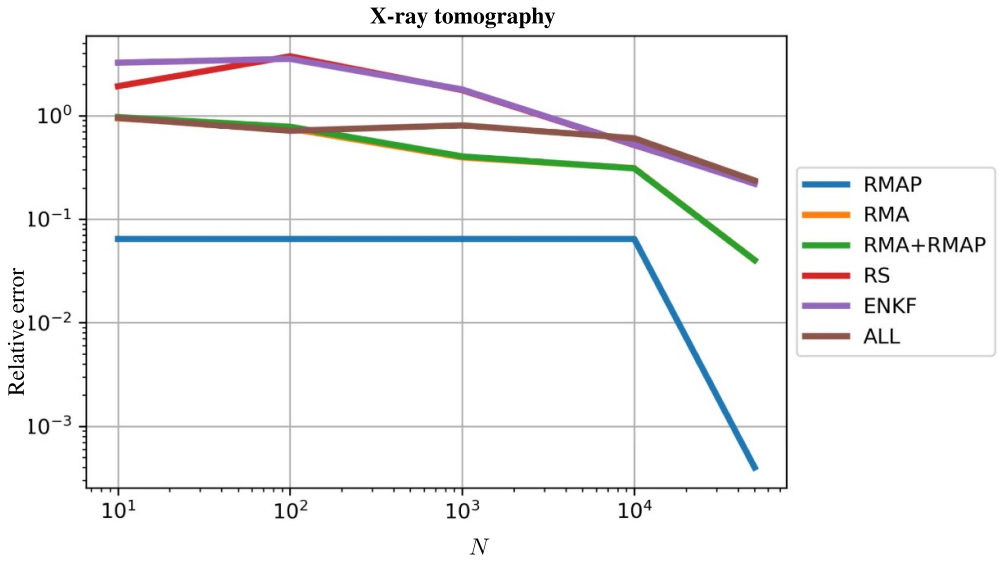


Figure 4. Relative error plot for x-ray tomography problem.

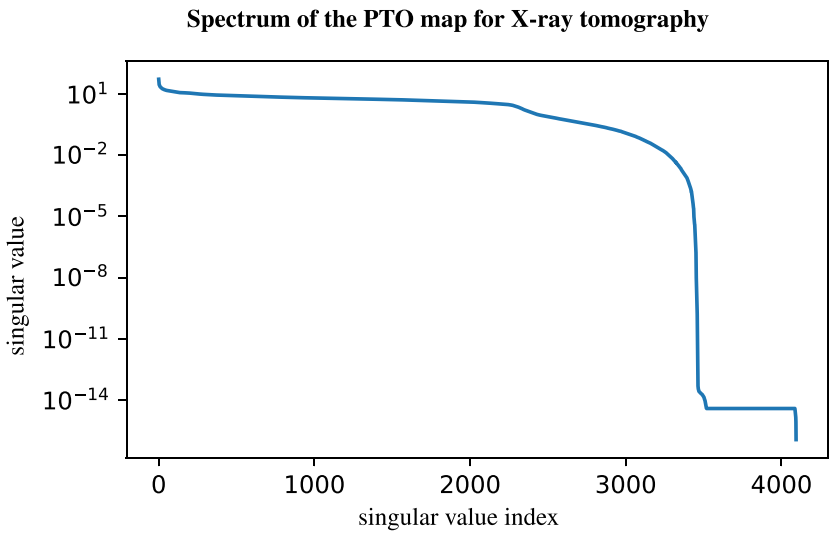


Figure 5. The singular values of the parameter-to-observable map for an x-ray tomography problem decay rapidly at first and then slowly until the last few singular vectors. This shows that the effective rank of the PtO map is close to the minimum dimension.

with the spectrum of the PtO map decaying slowly after an initial fast decay (figure 5). This means that the effective rank of the PtO map is close to the dimension of the data in the case presented. While mildly ill-posed problems are usually easier to work with, this can present a challenge for randomized methods, particularly methods such as RMA that randomize the misfit term. Recall that for any two matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\mathrm{rank}(\boldsymbol{XY}) \leqslant \mathrm{rank}(\boldsymbol{X})$. By projecting the misfit term onto a lower dimensional subspace, important information is lost in the case where

$\mathcal{A}$ is has effective rank close to the dimension of the data. This indicates that such a method is better suited for problems that are severely ill-posed.

The second observation is that the error for randomized prior methods initially increases then decreases as the number of samples increases. In previous problems, we have used a direct linear solver to find the solution to the stochastic optimization problem. In this problem, we use an iterative conjugate gradient (CG) solver to showcase how solver choice interacts with the randomized approaches. The main difference that can be seen here between a direct solver and an iterative solver such as CG is that a direct solver will invert possibly tiny eigenvalues of an ill-conditioned matrix while an iterative solver will often stop early, depending on the convergence parameters set, acting as an iterative regularizer (Hanke and Nagy 1996, Piccolomini and Zama 1999, Landi *et al* 2016). This effect is particularly pronounced on the randomized prior methods such as RS and EnKF where the low rank randomized prior causes the iterative solver to stop earlier with fewer samples. This causes the error of RS and EnKF to increase initially until the regularization has sufficient rank, then the methods converge asymptotically to $\boldsymbol{u}^{\text{MAP}}$. In the case of x-ray tomography, full-rank regularization is not needed due to the mildly ill-posed nature of the problem.

5.3. *Initial condition inversion in an advection-diffusion problem*

We now consider a linear inverse problem governed by a parabolic PDE based on the method used in UT Austin and UC Merced (2017). The parameter to observable map (advection-diffusion equation) maps an initial condition $\boldsymbol{u} \in L^2(\Omega)$ to pointwise spatio-temporal observations of the concentration field $\mathbf{y}(x,t)$. The advection-diffusion equation is given by:

$$\begin{aligned} \mathbf{y}_t - \kappa\Delta\mathbf{y} + \vec{v}\cdot\nabla\mathbf{y} = 0 \quad & \text{in } \Omega\times(0,T), \\ \mathbf{y}(.,0) = \boldsymbol{u} \quad & \text{in } \Omega, \\ \kappa\nabla\mathbf{y}\cdot\mathbf{n} = 0 \quad & \text{on } \partial\Omega\times(0,T), \end{aligned} \tag{34}$$

where, $\Omega \subset \mathbb{R}^2$ is a bounded domain, $\kappa > 0$ is the diffusion coefficient, $T > 0$ is the final time. The velocity field $\vec{v}$ is computed by solving the following steady-state Navier–Stokes equation with the side walls driving the flow:

$$\begin{aligned} -\frac{1}{\text{Re}}\Delta\vec{v} + \nabla q + \vec{v}\cdot\nabla\vec{v} = 0 \quad & \text{in } \Omega, \\ \nabla\cdot\vec{v} = 0 \quad & \text{in } \Omega, \\ \vec{v} = \vec{g} \quad & \text{on } \partial\Omega. \end{aligned} \tag{35}$$

where q is the pressure, and Re is the Reynolds number. The Dirichlet boundary condition $\vec{g} \in \mathbb{R}^2$ is prescribed as $\vec{g} = [0,1]$ on the left side of the domain, and $\vec{g} = [0,0]$ elsewhere. Velocity boundary conditions are not prescribed on the right side of the boundary. The values of the forward solution $\mathbf{y}$ on a set of locations $\{x_1, x_2, \ldots, x_m\}$ at the final time T are extracted and used as the observation vector $\mathbf{d} \in \mathbb{R}^k$ for solving the initial condition inverse problem. Synthetic observations are generated by corrupting this observation vector with 1% additive Gaussian noise. The observation data and the velocity profile used in the study are shown in figure 6. Upon discretization, the operator $\mathcal{A}$ maps the initial condition $\boldsymbol{u} \in \mathbb{R}^n$ to the observation $\mathbf{d} \in \mathbb{R}^k$.

In addition, we define the prior covariance matrix to be the PDE-based BiLaplacian prior defined as:

$$\boldsymbol{\Gamma} = (\delta I + \gamma\nabla\cdot(\theta\nabla))^{-2}, \tag{36}$$

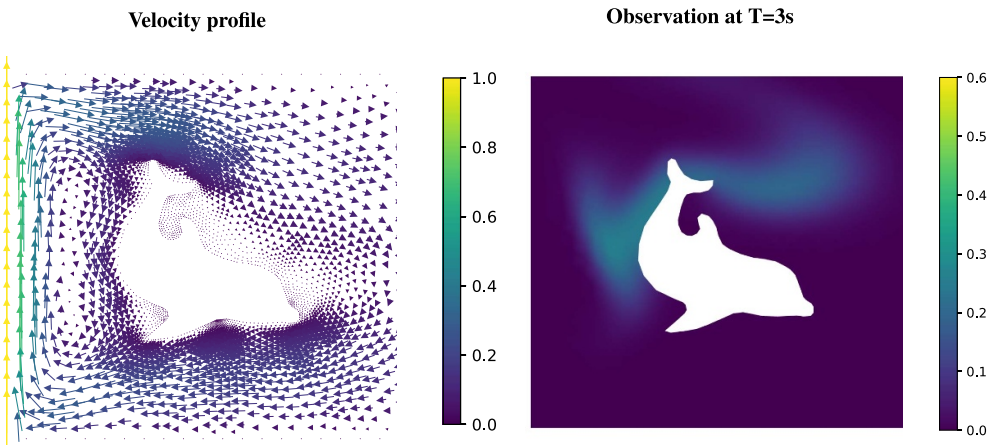


Figure 6. The velocity profile and observation data used for inversion.

Table 4. Relative error for various randomized methods compared to the $\boldsymbol{u}^{\mathrm{MAP}}$ solution for 2D linear advection-diffusion initial condition inverse problem.

	Relative error (%)			
Method	$N = 10$	$N = 100$	$N = 1000$	$N = 10000$
RMAP	5.02	1.96	0.49	0.15
RMA	53.38	15.16	5.30	1.15
RMA+RMAP	80.14	14.48	7.05	1.30
RS	59.58	26.78	7.33	5.06
ENKF	74.09	18.76	7.10	5.43
ALL	91.58	197.66	193.25	9.66

where, δ governs the variance of the samples, while the ratio $\frac{\gamma}{\delta}$ governs the correlation length. θ is a symmetric positive definite tensor to introduce anisotropy in the correlation length.

Following UT Austin and UC Merced (2017), a mixed formulation employing $P2$ Lagrange elements for approximating the velocity field and $P1$ elements for pressure is adopted for solving (35) to obtain the velocity field. The computed velocity field is then used to solve the advection-diffusion equation, (34). $P1$ Lagrange elements are used for the variational formulation of the advection-diffusion equation.

The observation vector $\boldsymbol{d}$ is computed at time $t = 3$ s with $m = 200$ observation points. For this problem, there are $n = 2868$ degrees of freedom. The diffusion coefficient is $\kappa = 0.001$ and the parameters of the BiLaplacian prior (36) are $\delta = 8$, $\gamma = 1$, and $\Theta = \mathcal{I}$.

The MAP solution $\boldsymbol{u}^{\mathrm{MAP}}$ is shown in figure 8. The condition number of $\boldsymbol{\Gamma}^{-1}$ is of the order of 10^6. The results of different randomization schemes are shown in figure 14 in the appendix. Table 4 and figure 7 give the relative error with respect to $\boldsymbol{u}^{\mathrm{MAP}}$. As expected, RMAP gives the most accurate results followed by LS and the RMA. In contrast to the previous examples considered, the RS and EnKF approaches gives reasonably good results as evident from table 4 and figure 14 in the appendix.

This is due to the faster convergence of the randomized prior covariance to the true prior covariance for the BiLaplacian prior. This also points to the fact that care should be exercised when choosing a randomized method for a particular problem. For inverse problems

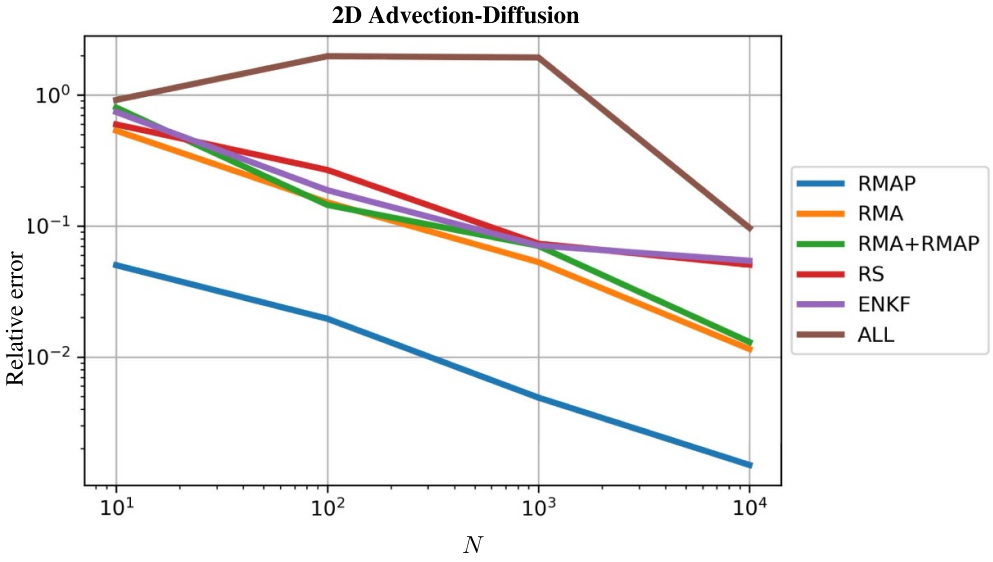


Figure 7. Relative error plot for 2D linear advection-diffusion initial condition inversion.

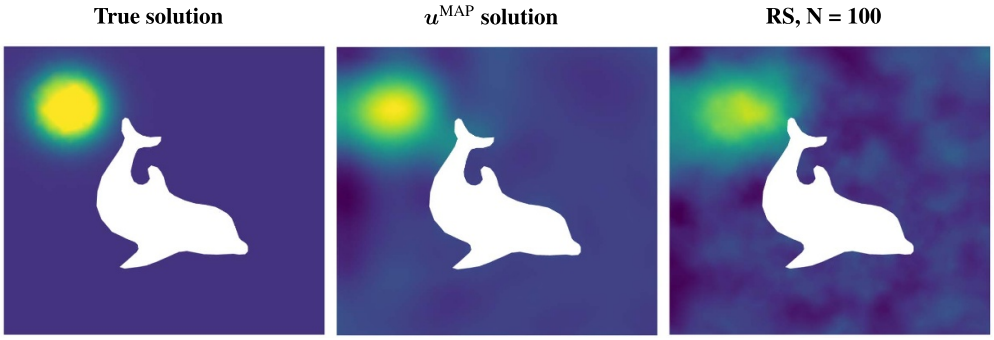


Figure 8. From left to right are true solution, $\boldsymbol{u}^{\mathrm{MAP}}$ solution and right sketching solution for linear advection-diffusion initial condition inverse problem. With only 100 random samples, right sketching can obtain a reasonably good initial condition reconstruction (26% relative error). This behavior indicates the fast convergence of the randomly sampled prior inverse covariance to the true prior inverse covariance.

where a prior with a decaying spectrum is desirable, the EnKF or RS approach may perform well.

5.4. Nonlinear parameter inversion in a steady-state heat equation

To show the convergence of various methods for nonlinear inverse problems, we consider a nonlinear PDE constrained parameter inversion problem. In the previous section, we considered an initial condition problem where the data depended linearly on the parameter, even though the statement of the problem itself was rather involved. Now we consider a simple to state but extremely ill-posed nonlinear inverse problem. Given a steady-state temperature

Table 5. Relative error of MAP solution for various randomization schemes compared to the $\boldsymbol{u}^{\mathrm{MAP}}$ solution.

	Relative error (%)			
Method	$N = 10$	$N = 100$	$N = 1000$	$N = 5000$
RMAP	19.50	14.66	16.39	16.17
RMA	24.97	8.12	4.19	3.08
RMA+RMAP	17.51	17.82	16.27	16.49
RS	47.06	110.88	107.46	14.44
ENKF	20.82	18.26	17.54	19.55
ALL	26.13	15.63	19.45	19.62

distribution $\boldsymbol{T}(x,y)$ and boundary conditions, invert for the conductivity everywhere in the domain. The governing equations are given by

$$\begin{aligned} \nabla \cdot (e^{\kappa} \nabla \boldsymbol{T}) &= f \quad \text{in } \Omega, \\ \boldsymbol{T}(x,0) &= 2(1-x), \\ \boldsymbol{T}(x,1) &= 2x, \\ \nabla \boldsymbol{T} \cdot \boldsymbol{n} &= 0 \quad \text{on } \partial\Omega \setminus \{y = 0, y = 1\}. \end{aligned}$$

While this equation is linear in the temperature distribution $\boldsymbol{T}$, the (log-) thermal conductivity that we are inverting for, κ, appears non-linearly. That is, the parameter-to-observable map is nonlinear. The heat equation makes for an excellent test problem for inverse solvers since the dependence of the steady-state temperature distribution on the conductivity is rather weak.

We again follow a mixed formulation where the temperature distribution is modeled using $P2$ Lagrange elements and the parameter is modeled with $P1$ Lagrange elements. With a mesh size of 64×64 elements, this results in discrete variables $\boldsymbol{y} \in \mathbb{R}^{16,641}$ and $\kappa \in \mathbb{R}^{4,225}$. The BiLaplacian prior defined in (36) is also used here with $\delta = 0.5$, $\gamma = 0.1$, and the anisotropic diffusion tensor

$$\theta = \begin{bmatrix} \theta_1 \sin^2 \alpha & (\theta_1 - \theta_2) \sin \alpha \cos \alpha \\ (\theta_1 - \theta_2) \sin \alpha \cos \alpha & \theta_2 \cos^2 \alpha \end{bmatrix},$$

where $\theta_1 = 2.0$, $\theta_2 = 0.5$ and $\alpha = \pi/4$. Lastly, we consider an inhomogeneous case where

$$f = 50 \sin^2(\pi x) \cos^2(\pi y).$$

A few remarks are in order to understand the rather unimpressive results in table 5 and figure 10. First, recall that the results shown here are for an extremely difficult problem. The inverse of the diffusion equation is notoriously ill-posed, as it amounts to the inverse of a compact operator. That is, small perturbations in the data can lead to drastically different inversion results. Indeed, we show an even more difficult problem where the task is to infer 4225

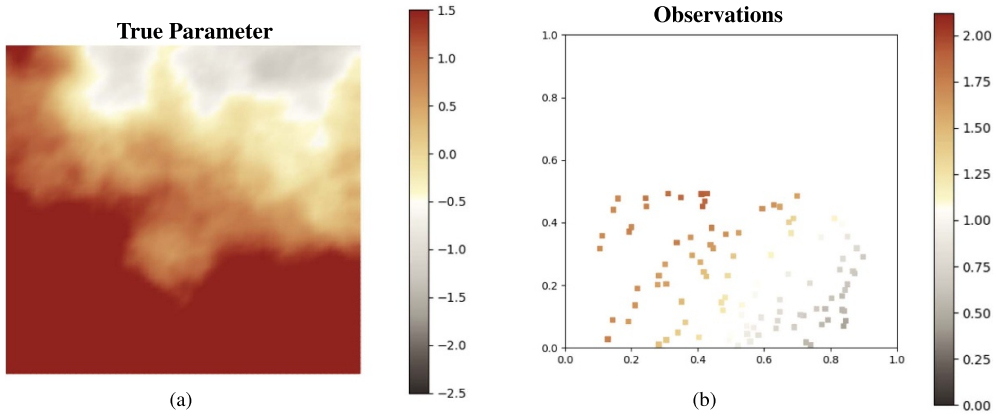


Figure 9. The true log-conductivity (κ) and 100 sparse observations. Observing the temperature distribution only in the lower half of the domain makes inverting for the log-conductivity in the entire domain a more difficult task.

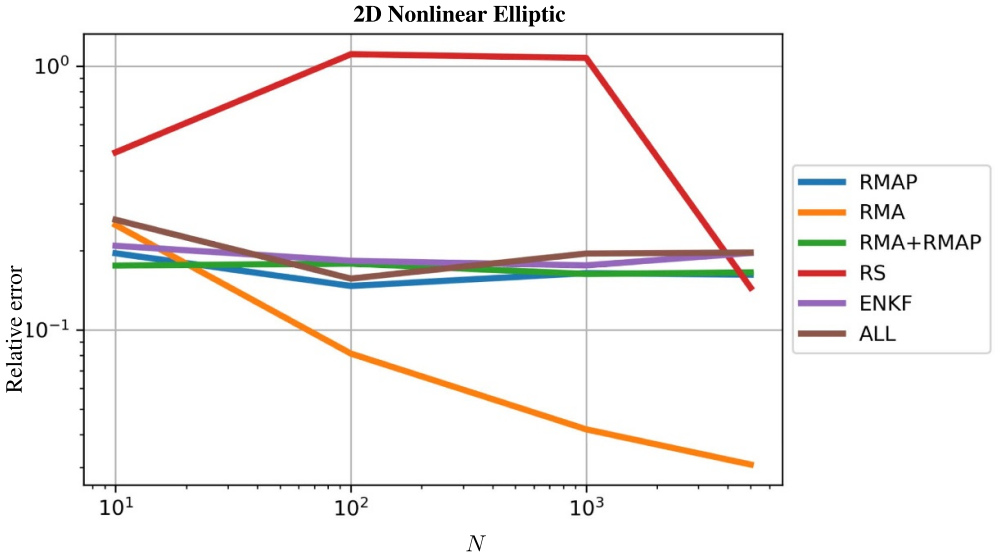


Figure 10. Relative error plot for nonlinear elliptic parameter inverse problem.

parameters from only 100 measurements which are recorded in only half of the domain as shown in figure 9. With so few measurements, adding noise to the data as in RMAP, RMA+RMAP, ENKF, and ALL may not lead to desirable results.

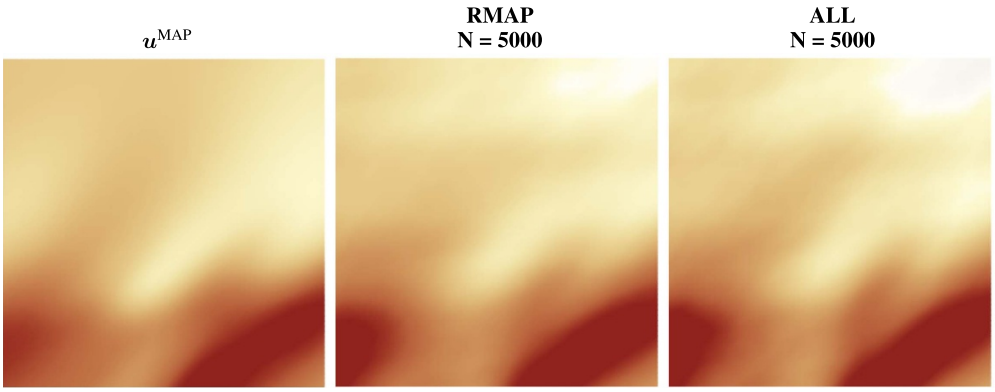


Figure 11. Solutions to the nonlinear diffusion inverse problem for two different methods. Visually, these methods give nearly identical results compared to the $\boldsymbol{u}^{\mathrm{MAP}}$ solution even though numerically, they have relative error $\sim$20%.

Secondly, while there is still $\sim$20% error for these methods, the MAP estimate for each of these is still reasonably good visibly as seen in figure 11. The high relative error is in part due to the small norm of the $\boldsymbol{u}^{\mathrm{MAP}}$ solution. Simply shifting the parameter up by a constant changes the norm of the denominator in the relative error formula

$$\text{relative error} := \left\| \boldsymbol{u}_N^{\langle \text{method} \rangle} - \boldsymbol{u}^{\mathrm{MAP}} \right\| / \left\| \boldsymbol{u}^{\mathrm{MAP}} \right\|.$$

In addition to the numerical relative error, it is important to consider the 'eyeball norm'. Figure 11 shows that the estimated solutions are still quite close to the $\boldsymbol{u}^{\mathrm{MAP}}$ solution, especially given the extreme ill-posedness. Despite this shortcoming of data-randomization methods, recall that the main advantage of additively randomizing the data and prior mean is to *aid in sampling* from the posterior. In other words, we are most interested in accelerating uncertainty quantification, not getting the SAA MAP estimate error down to machine precision. These randomization schemes may still find use in such applications. Comprehensive visual results can be found in figures 15 and 16 of the appendix.

6. Conclusions

By viewing the randomized solution of inverse problems through the lens of stochastic programming and the SAA, we developed a unified framework though which we can analyze the asymptotic convergence of randomized solutions of linear inverse problems to the solution obtained with its deterministic counterpart. This framework allowed us to prove the asymptotic and non-asymptotic convergence of the minimizer of a general stochastic cost function to the minimizer of the expected value of the stochastic cost function. Several well-known methods for introducing randomness into linear and nonlinear inverse problems were recovered as special cases of this general framework. Viewing the solution to randomized inverse problems through the lens of the SAA also allowed us to prove a novel non-asymptotic error analysis that applies to all randomized methods discussed. We also show that while all of the methods presented converge asymptotically, the results can be quite poor if an insufficient number of samples are drawn. While this observation is easily understood through our non-asymptotic

error analysis, it is not possible from an asymptotic view point. In particular, we showed that randomizing the prior covariance matrix may not be a good idea for certain priors due to the regularizing role that the prior plays in the solution of inverse problems. This is due to the potentially slow convergence of random matrices to their expected value, depending on the spectrum of the expected value matrix.

The convergence of all schemes was shown numerically for a variety of linear and nonlinear inverse problems, including 1D and 2D problems governed by algebraic and PDE constraints. Through these sample problems, we explored the performance of 6 methods, 3 rediscoveries and 3 new methods, numerically studying the convergence of each method to the MAP point. The results indicate that the RMAP approach has the fastest convergence. While this is evident from the numerical studies, each sample of the RMAP method requires the solution of an inverse problem, making the overall cost of the RMAP method potentially very high. While the main contribution of this work is in unifying the theory of convergence to the MAP point, we note that the utility of RMAP, RMA+RMAP and EnKF, is in generating (approximate) samples from the posterior as shown in other cited works. The variety of sample problems considered shows the varying performance of each algorithm and explores the sensitivity of each method to different problem features, such as the poor performance of RMA on the x-ray tomography problem. These numerical observations along with the asymptotic and non-asymptotic theory combine to give practitioners the tools to design new randomized methods for solving inverse problems or to choose the most appropriate existing method for their particular problem.

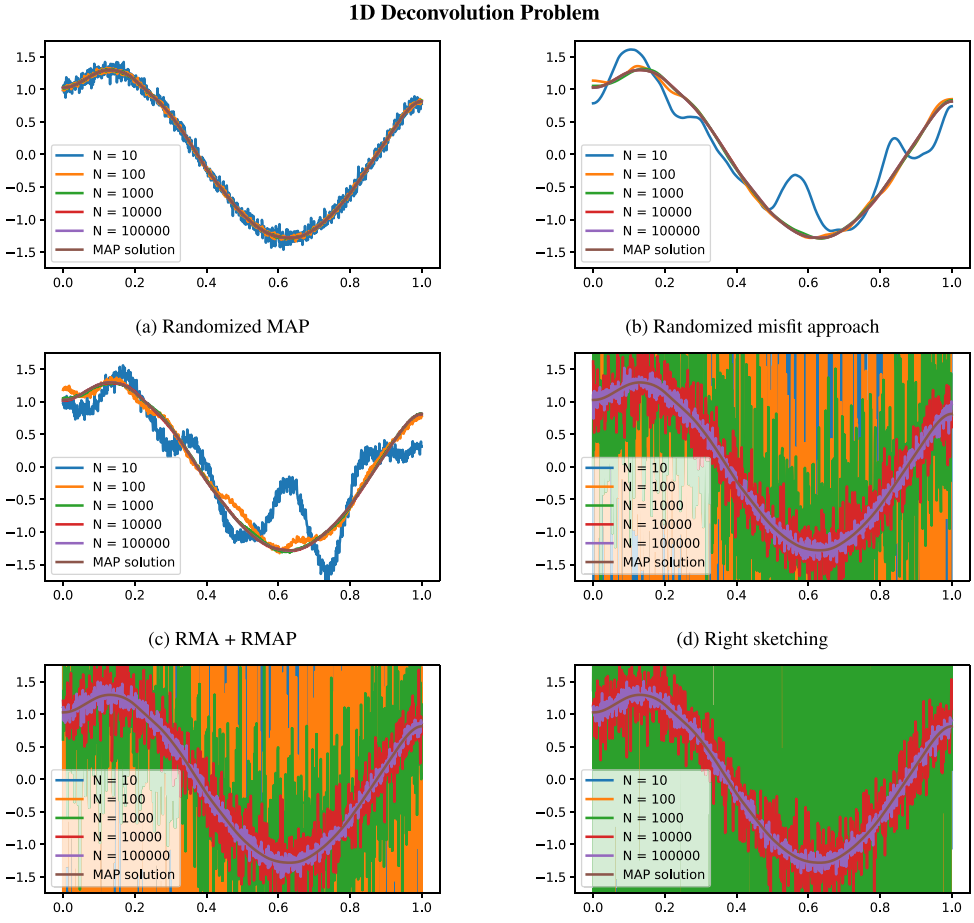
Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/jonwittmer/randomized-inversion-toolkit.

Acknowledgments

We would like to thank the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing HPC resources that contributed to the results presented in this work. URL: www.tacc.utexas.edu. This research is partially funded by the National Science Foundation Awards NSF-OAC-2212442, NSF-2108320, NSF-1808576 and NSF-CAREER-1845799; and by the Department of Energy Awards DE-SC0018147 and DE-SC0022211.

Appendix. Figures



(a) Randomized MAP

(b) Randomized misfit approach

(c) RMA + RMAP

(d) Right sketching

(e) EnKF

(f) All

Figure 12. Solutions to 1D deconvolution problem with mesh size $n = 1000$ using various randomization schemes with scaled identity prior. This prior works sufficiently well for those randomization schemes that do not randomize the prior covariance (RMAP, RMA, RMA+RMAP), but performs poorly for RS, EnKF, and ALL which randomize the prior covariance. Random sampling is performed via an Achlioptas ($2/3-$sparse) random variable.

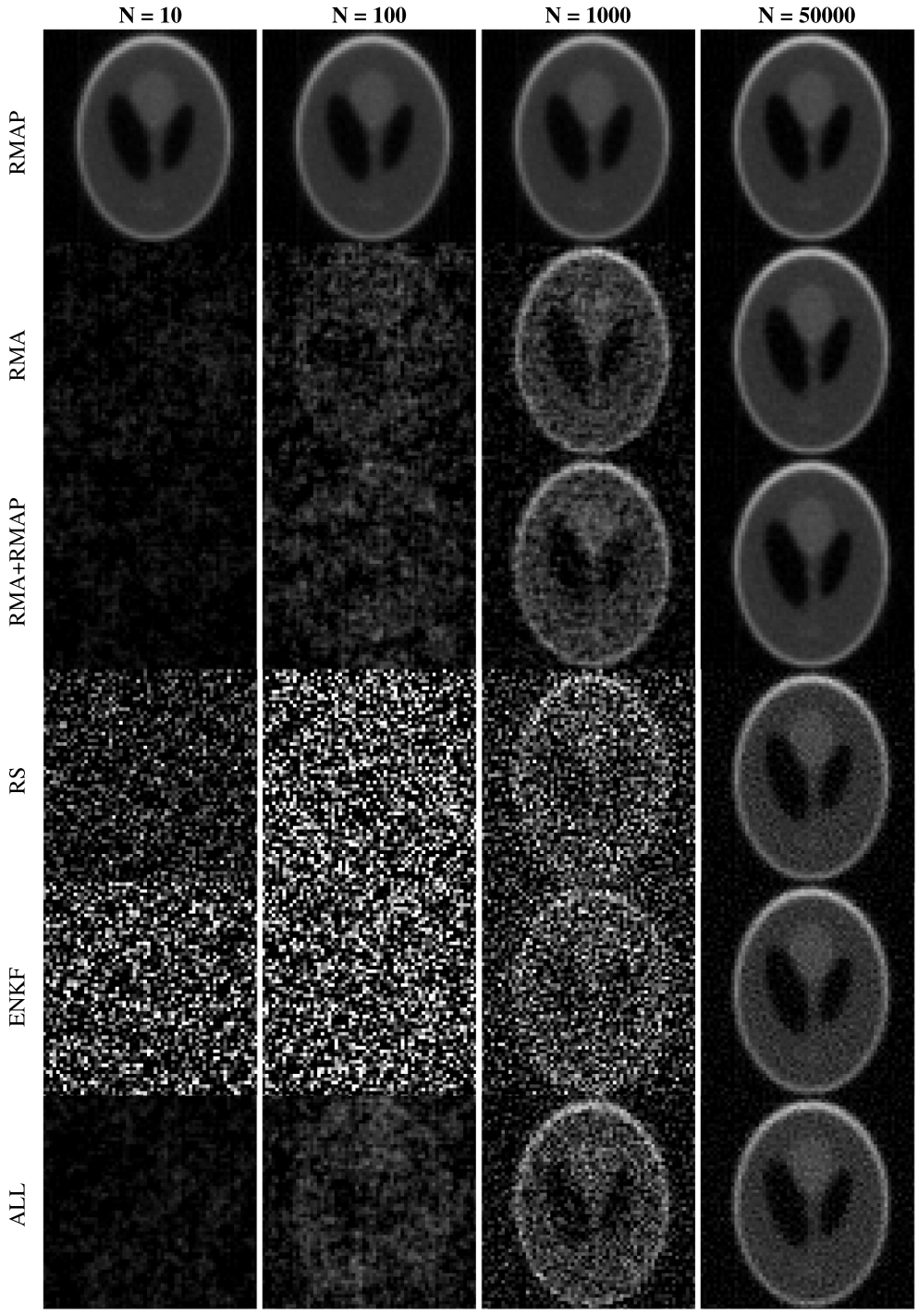


Figure 13. Solutions for various randomization approaches for an x-ray tomography problem with Gaussian random variables.

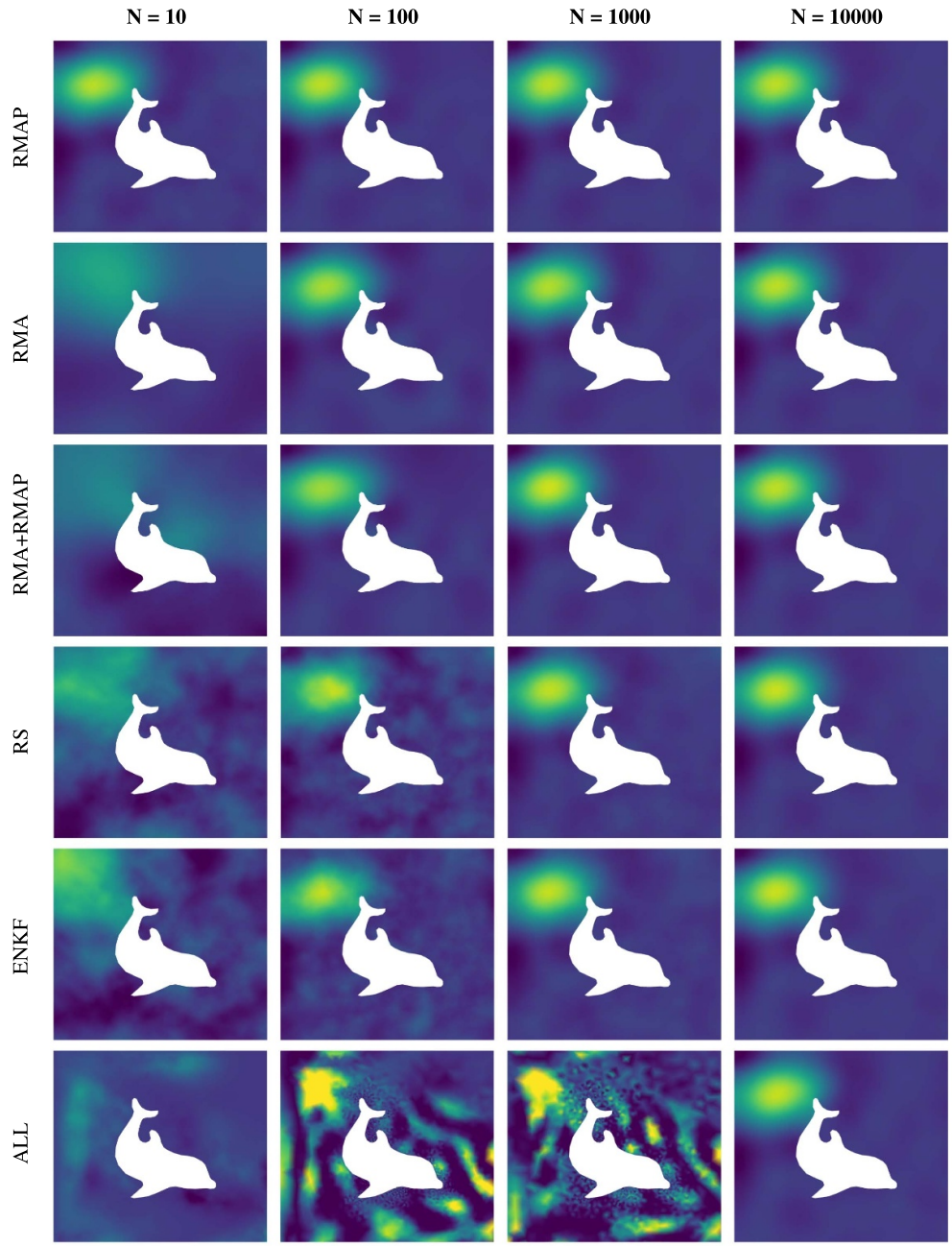


Figure 14. Solutions for various randomization approaches for a linear advection-diffusion initial condition inverse problem with Gaussian random variables.

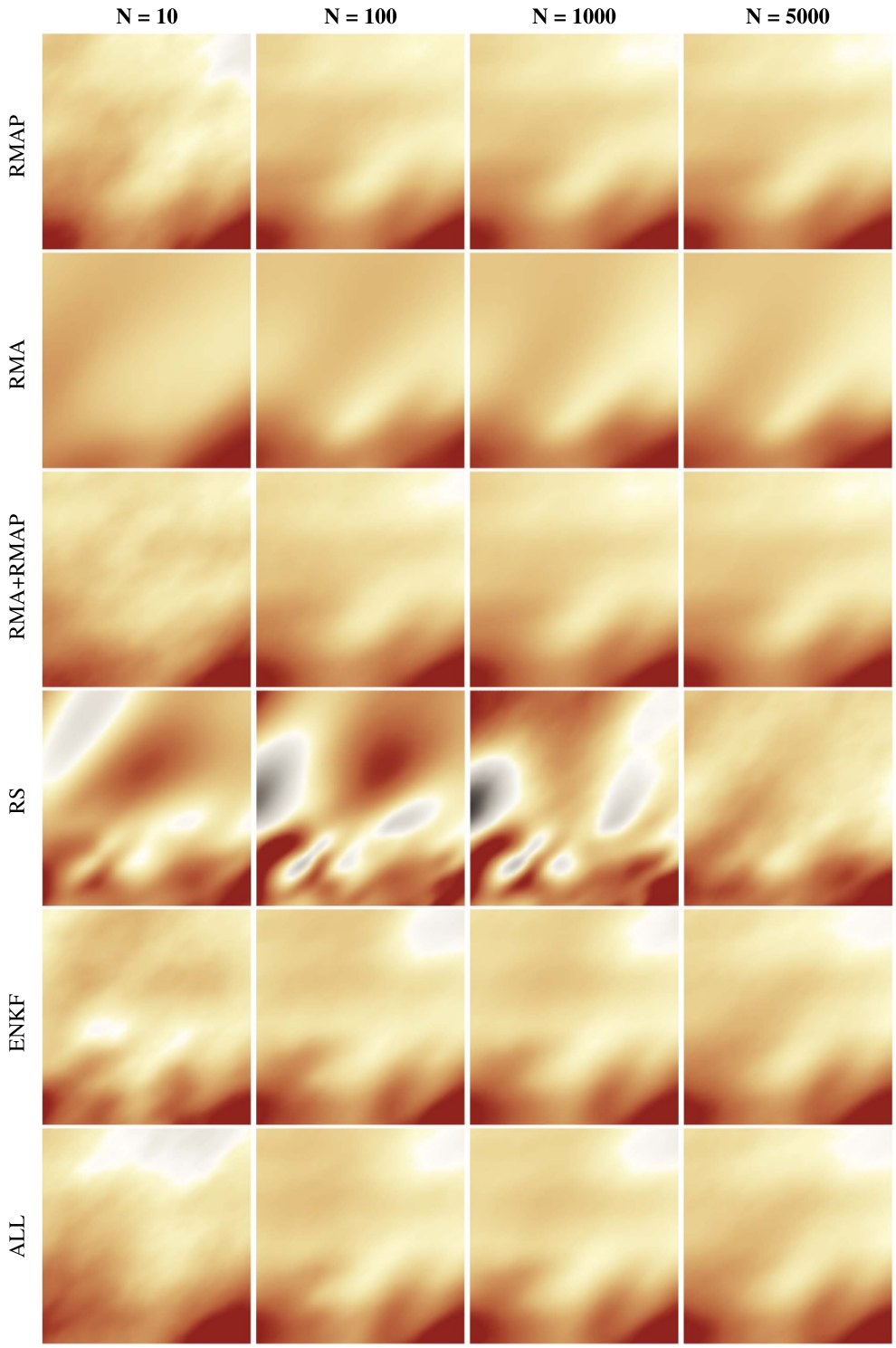


Figure 15. Solutions for various randomization approaches for a nonlinear diffusion parameter inversion problem with Gaussian random variables.

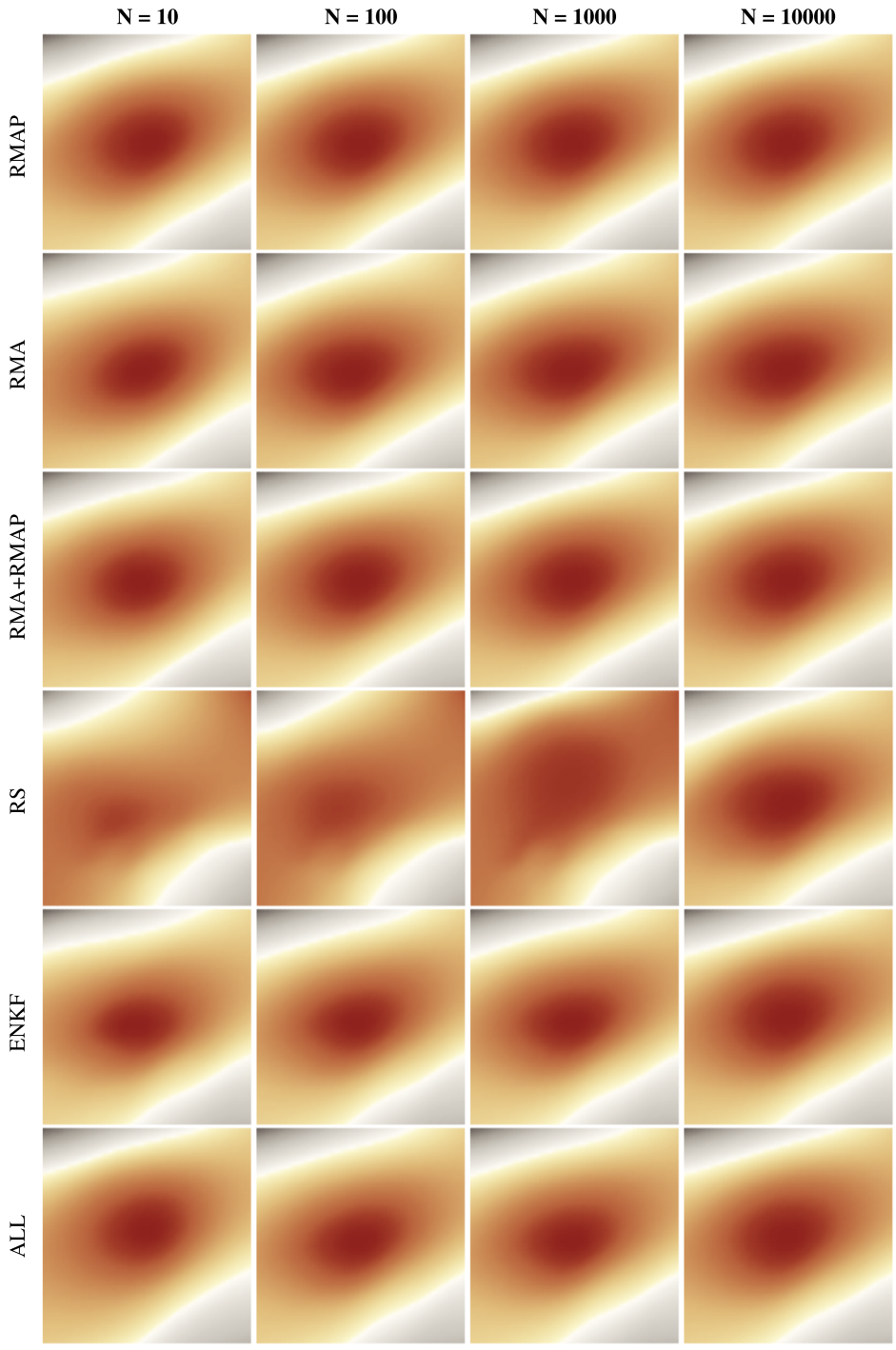


Figure 16. Reconstructed state for various randomization approaches for a nonlinear diffusion problem. Even for the right sketching method which did not give good parameter reconstructions until $N = 10\,000$ samples, the state in the lower half of the domain, where the 100 measurements are taken, look similar to all the other methods.

ORCID iD

Jonathan Wittmer https://orcid.org/0000-0002-7538-5932

References

Achlioptas D 2003 Database-friendly random projections: Johnson-Lindenstrauss with binary coins *J. Comput. Syst. Sci.* **66** 671–87
Anderson J L 2007 An adaptive covariance inflation error correction algorithm for ensemble filters *Tellus* A **59** 210–24
Avron H, Sindhwani V and Woodruff D 2013 Sketching structured matrices for faster nonlinear regression *Advances in Neural Information Processing Systems* vol 26
Ayanbayev B, Klebanov I, Lie H C and Sullivan T J 2021 γ-convergence of Onsager–Machlup functionals: I. With applications to maximum *a posteriori* estimation in Bayesian inverse problems *Inverse Problems* **38** 025005
Bardsley J, Solonen A, Haario H and Laine M 2014 Randomize-then-optimize: a method for sampling from posterior distributions in nonlinear inverse problems *SIAM J. Sci. Comput.* **36** A1895–910 submitted
Beskos A, Pinski F J, Sanz-Serna J M and Stuart A M 2011 Hybrid Monte Carlo on Hilbert spaces *Stoch. Process. Appl.* **121** 2201–30
Blatter D, Morzfeld M, Key K and Constable S 2022a Uncertainty quantification for regularized inversion of electromagnetic geophysical data–part II: application in 1-D and 2-D problems *Geophys. J. Int.* **231** 1075–95
Blatter D, Morzfeld M, Key K and Constable S 2022b Uncertainty quantification for regularized inversion of electromagnetic geophysical data—part I: motivation and theory *Geophys. J. Int.* **231** 1057–74
Braides A *et al* 2002 *Gamma-Convergence for Beginners* vol 22 (Oxford: Clarendon)
Brockwell A 2006 Parallel Markov chain Monte Carlo simulation by pre-fetching *J. Comput. Graph. Stat.* **15** 246–61
Bui-Thanh T and Ghattas O 2012 A scaled stochastic Newton algorithm for Markov chain Monte Carlo simulations *SIAM J. Uncertain. Quantification* 1–25 submitted
Bui-Thanh T and Girolami M A 2014 Solving large-scale PDE-constrained Bayesian inverse problems with Riemann manifold Hamiltonian Monte Carlo *Inverse Problems* **30** 114014
Bui-Thanh T and Nguyen Q P 2016 FEM-based discretization-invariant MCMC methods for PDE-constrained bayesian inverse problems *Inverse Problems Imaging* **10** 943
Byrd J 2010 Parallel Markov chain Monte Carlo *PhD Thesis* University of Warwick
Carpenter J, Clifford P and Fearnhead P 1999 Improved particle filter for nonlinear problems *IEE Proc., Radar Sonar Navig.* **146** 2–7
Chada N K, Chen Y and Sanz-Alonso D 2020 Iterative ensemble Kalman methods: a unified perspective with some new variants (arXiv:2010.13299)
Chen K, Li Q, Newton K and Wright S J 2020a Structured random sketching for PDE inverse problems *SIAM J. Matrix Anal. Appl.* **41** 1742–70
Chen P and Ghattas O 2020 Projected Stein variational gradient descent *Advances in Neural Information Processing Systems* vol 33 pp 1947–58
Chen Y, Dwivedi R, Wainwright M J and Yu B 2020b Fast mixing of metropolized Hamiltonian Monte Carlo: benefits of multi-step gradients *J. Mach. Learn. Res.* **21** 92–91
Chu D, Lin L, Tan R C E and Wei Y 2011 Condition numbers and perturbation analysis for the Tikhonov regularization of discrete ill-posed problems *Numer. Linear Algebra Appl.* **18** 87–103
Clarkson K L and Woodruff D P 2017 Low-rank approximation and regression in input sparsity time *J. ACM* **63** 1–45
Cui T, Law K J H and Marzouk Y M 2016 Dimension-independent likelihood-informed MCMC *J. Comput. Phys.* **304** 109–37
Cui T, Martin J, Marzouk Y M, Solonen A and Spantini A 2014 Likelihood-informed dimension reduction for nonlinear inverse problems (arXiv:1403.4680)
Deng C Y 2011 A generalization of the Sherman–Morrison–Woodbury formula *Appl. Math. Lett.* **24** 1561–4
Diao H-A, Wei Y and Qiao S 2016 Structured condition numbers of structured Tikhonov regularization problem and their estimations *J. Comput. Appl. Math.* **308** 276–300

Duane S, Kennedy A D, Pendleton B and Roweth D 1987 Hybrid Monte Carlo *Phys. Lett.* B **195** 216–22
Durrett R 2019 *Probability: Theory and Examples* (*Cambridge Series in Statistical and Probabilistic Mathematics*) (Cambridge: Cambridge University Press)
Elsheikh A H, Pain C, Fang F, Gomes J L and Navon I M 2013 Parameter estimation of subsurface flow models using iterative regularized ensemble Kalman filter *Stoch. Environ. Res. Risk Assess.* **27** 877–97
Engl H W, Kunisch K and Neubauer A 1989 Convergence rates for Tikhonov regularisation of non-linear ill-posed problems *Inverse Problems* **5** 523
Evensen G 2003 The ensemble Kalman filter: theoretical formulation and practical implementation *Ocean Dyn.* **53** 343–67
Evensen G *et al* 2009 *Data Assimilation: The Ensemble Kalman Filter* vol 2 (Berlin: Springer)
Farchi A and Bocquet M 2019 On the efficiency of covariance localisation of the ensemble Kalman filter using augmented ensembles *Front. Appl. Math. Stat.* **5** 3
Feller W 1971 *An Introduction to Probability Theory and Its Applications* vol 2, 2nd edn (New York: Wiley)
Gao X, Zhang M, Luo J and Fan J 2022 Tail bounds for norm of Gaussian random matrices with applications *J. Math.* **2022** 1–5
Girolami M and Calderhead B 2011 Riemann manifold Langevin and Hamiltonian Monte Carlo methods *J. R. Stat. Soc.* B **73** 123–214
Haario H, Laine M, Miravete A and Saksman E 2006 DRAM: efficient adaptive MCMC *Stat. Comput.* **16** 339–54
Han J and Liu Q 2018 Stein variational gradient descent without gradient *Int. Conf. on Machine Learning* (PMLR) pp 1900–8
Hanke M and Nagy J G 1996 Restoration of atmospherically blurred images by symmetric indefinite conjugate gradient techniques *Inverse Problems* **12** 157–73
Hastings W K 1970 Monte Carlo sampling methods using Markov chains and their applications *Biometrika* **57** 97–109
Houtekamer P L and Mitchell H L 1998 Data assimilation using an ensemble Kalman filter technique *Mon. Weather Rev.* **126** 796–811
Iglesias M A, Law K J H and Stuart A M 2013 Ensemble Kalman methods for inverse problems *Inverse Problems* **29** 045001
J. J. 2013 Young's inequality for three variables (version: 2013-05-30) (Mathematics Stack Exchange) (available at: https://math.stackexchange.com/q/406922)
Kitanidis P K 1995 Quasi-linear geostatistical theory for inversing *Water Resour. Res.* **31** 2411–9
Kundur D and Hatzinakos D 1996 Blind image deconvolution *IEEE Signal Process. Mag.* **13** 43–64
Landi G, Loli Piccolomini E and Tomba I 2016 A stopping criterion for iterative regularization methods *Appl. Numer. Math.* **106** 53–68
Latz J 2020 On the well-posedness of Bayesian inverse problems *SIAM/ASA J. Uncertain. Quantification* **8** 451–82
Le E B, Myers A, Bui-Thanh T and Nguyen Q P 2017 A data-scalable randomized misfit approach for solving large-scale PDE-constrained inverse problems *Inverse Problems* **33** 065003
Liu M, Kumar R, Haber E and Aravkin A Y 2018 Simultaneous shot inversion for nonuniform geometries using fast data interpolation (arXiv:1804.08697 [math.OC])
Liu Q and Wang D 2016 Stein variational gradient descent: a general purpose Bayesian inference algorithm *Advances in Neural Information Processing Systems* vol 29
Martin J, Wilcox L C, Burstedde C and Ghattas O 2012 A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion *SIAM J. Sci. Comput.* **34** A1460–87
Maso G 1993 *An Introduction to Γ-Convergence* vol 113 (Berlin: Springer)
Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 Equation of state calculations by fast computing machines *J. Chem. Phys.* **21** 1087–92
Mueller J L and Siltanen S 2012 *Linear and Nonlinear Inverse Problems With Practical Applications* (Philadelphia, PA: SIAM)
Neal R M 2010 MCMC using Hamiltonian dynamics *Handbook of Markov Chain Monte Carlo* (London/Boca Raton, FL: Chapman & Hall/CRC Press)
Oliver D S, Reynolds A C and Liu N 2008 *Inverse Theory for Petroleum Reservoir Characterization and History Matching* (Cambridge: Cambridge University Press)

Petra N, Martin J, Stadler G and Ghattas O 2014 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** A1525–55
Petrie R 2008 Localization in the ensemble Kalman filter *MSc Atmosphere, Ocean and Climate University of Reading* vol 460
Piccolomini E L and Zama F 1999 The conjugate gradient regularization method in computed tomography problems *Appl. Math. Comput.* **102** 87–99
Pourahmadi M 2011 Covariance estimation: The GLM and regularization perspectives *Stat. Sci* . **26** 369–87
Raskutti G and Mahoney M W 2016 A statistical perspective on randomized sketching for ordinary least-squares *J. Mach. Learn. Res.* **17** 7508–38
Robert C P and Casella G 2005 *Monte Carlo Statistical Methods* (*Springer Texts in Statistics*) (Secaucus, NJ: Springer New York, Inc.)
Rockafellar R T and Wetts R J-B 1998 *Variational Analysis* (Berlin: Springer)
Ryan O and Debbah M 2007 Free deconvolution for signal processing applications *2007 IEEE Int. Symp. on Information Theory* (IEEE) pp 1846–50
Saksman M L and Siltanen S 2009 Discretization-invariant Bayesian inversion and Besov space priors (arXiv:0901.4220)
Sambale H 2020 Some notes on concentration for α-subexponential random variables (arXiv:2002.10761)
Schillings C and Stuart A M 2017 Analysis of the ensemble Kalman filter for inverse problems *SIAM J. Numer. Anal.* **55** 1264–90
Shapiro A, Dentcheva D and Ruszczynski A 2009 *Lectures on Stochastic Programming: Modeling and Theory* (Philadelphia, PA: Society for Industrial and Applied Mathematics)
Soto A 2005 Self adaptive particle filter *IJCAI* (Citeseer) pp 1398–406
Strid I 2010 Efficient parallelisation of Metropolis–Hastings algorithms using a prefetching approach *Comput. Stat. Data Anal.* **54** 2814–35
Stuart A M 2010 Inverse problems: a Bayesian perspective *Acta Numer.* **19** 451–559
Swedlow J R 2013 Quantitative fluorescence microscopy and image deconvolution *Methods in Cell Biology* vol 114 (Amsterdam: Elsevier) pp 407–26
Trefethen L N and Bau D III 1997 *Numerical Linear Algebra* vol 50 (Philadelphia, PA: SIAM)
UT Austin and UC Merced 2017 hIPPYlib: inverse problem Python library (available at: https://hippylib.github.io/)
Van Der Merwe R, Doucet A, De Freitas N and Wan E 2000 The unscented particle filter *Advances in Neural Information Processing Systems* vol 13
Van der Vaart A W 2000 *Asymptotic Statistics* vol 3 (Cambridge: Cambridge University Press)
Vershynin R 2018 *High-Dimensional Probability: An Introduction With Applications in Data Science* (*Cambridge Series in Statistical and Probabilistic Mathematics*) (Cambridge: Cambridge University Press)
Vladimirova M, Girard S, Nguyen H and Arbel J 2020 Sub-Weibull distributions: generalizing sub-Gaussian and sub-Exponential properties to heavier tailed distributions *Stat* **9** e318
Wang J, Lee J, Mahdavi M, Kolar M and Srebro N 2017 Sketching meets random projection in the dual: a provable recovery algorithm for big and high-dimensional data *Artificial Intelligence and Statistics* (PMLR) pp 1150–8
Wang K 2014 Parallel Markov chain Monte Carlo methods for large scale statistical inverse problems *PhD Thesis* Texas A&M University
Wang K, Bui-Thanh T and Ghattas O 2018 A randomized maximum *a posteriori* method for Posterior sampling of high dimensional nonlinear Bayesian inverse problems *SIAM J. Sci. Comput.* **40** A142–71
Whitaker J S, Hamill T M, Wei X, Song Y and Toth Z 2008 Ensemble data assimilation with the NCEP global forecast system *Mon. Weather Rev.* **136** 463–82
Wilkinson D J 2005 Parallel Bayesian computation *Handbook of Parallel Computing and Statistics* (London/Boca Raton, FL: Marcel Dekker/CRC Press) pp 481–512
Yang T, Mehta P G and Meyn S P 2013 Feedback particle filter *IEEE Trans. Autom. Control* **58** 2465–80
Zhang H and Wei H 2022 Sharper sub-Weibull concentrations *Mathematics* **10** 2252
Zhuo J, Liu C, Shi J, Zhu J, Chen N and Zhang B 2018 Message passing Stein variational gradient descent *Int. Conf. on Machine Learning* (PMLR) pp 6018–27