SIAM J. SCI. COMPUT.
Vol. 42, No. 3, pp. A1714–A1740
© 2020 Society for Industrial and Applied Mathematics

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

RANDOMIZATION AND REWEIGHTED ℓ_1-MINIMIZATION FOR A-OPTIMAL DESIGN OF LINEAR INVERSE PROBLEMS*

ELIZABETH HERMAN†, ALEN ALEXANDERIAN†, AND ARVIND K. SAIBABA†

Abstract. We consider optimal design of PDE-based Bayesian linear inverse problems with infinite-dimensional parameters. We focus on the A-optimal design criterion, defined as the average posterior variance and quantified by the trace of the posterior covariance operator. We propose using structure exploiting randomized methods to compute the A-optimal objective function and its gradient, and we provide a detailed analysis of the error for the proposed estimators. To ensure sparse and binary design vectors, we develop a novel reweighted ℓ_1-minimization algorithm. We also introduce a modified A-optimal criterion and present randomized estimators for its efficient computation. We present numerical results illustrating the proposed methods on a model contaminant source identification problem, where the inverse problem seeks to recover the initial state of a contaminant plume using discrete measurements of the contaminant in space and time.

Key words. Bayesian inversion, A-optimal experimental design, large-scale ill-posed inverse problems, randomized matrix methods, reweighted ℓ_1 minimization, uncertainty quantification

AMS subject classifications. 35R30, 62K05, 68W20, 35Q62, 65C60, 62F15

DOI. 10.1137/19M1267362

1. Introduction.

A central problem in scientific computing involves estimating parameters that describe mathematical models, such as initial conditions, boundary conditions, or material parameters. This is often addressed by using experimental measurements and a mathematical model to compute estimates of unknown model parameters. In other words, one can estimate the parameters by solving an inverse problem. Experimental design involves specifying the experimental setup for collecting measurement data, with the goal of accurately recovering the parameters of interest. As such, optimal experimental design (OED) is an important aspect of effective and efficient parameter estimation. Namely, in applications where collecting experimental measurements is expensive (e.g., because of budget, labor, or physical constraints), deploying experimental resources has to be done efficiently and in a parsimonious manner. Even when collecting large amounts of data is feasible, OED is still important; the computational cost of processing all the data may be prohibitive or a poorly designed experiment with many measurements may miss important information about the parameters of interest.

To make matters concrete, we explain the inverse problem and the experimental design in the context of an application. Consider the transport of a contaminant in an urban environment or the subsurface. The forward problem involves forecasting the spread of the contaminant, whereas the inverse problem involves using the measurements of the contaminant concentration at discrete points in space and time to recover the source of the contaminant (i.e., the initial state). In this application, OED involves optimal placement of sensors, at which measurement data is collected, to reconstruct the initial state.

*Submitted to the journal's Methods and Algorithms for Scientific Computing section June 10, 2019; accepted for publication (in revised form) March 30, 2020; published electronically June 9, 2020.

https://doi.org/10.1137/19M1267362

Funding: This work was partially supported by the NSF through grant DMS-1745654.

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205 (eaherman@ncsu.edu, alexanderian@ncsu.edu, asaibab@ncsu.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We focus on OED for Bayesian linear inverse problems governed by PDEs. In our formulation of the OED problem, the goal is to find an optimal subset of sensors from a fixed array of n_s candidate sensor sites. The experimental design is parameterized by assigning nonnegative weights to each candidate sensor location. Ideally, we seek a binary weight vector $\boldsymbol{w}$; if $w_i = 1$, a sensor will be placed at the ith candidate location, and if $w_i = 0$, no sensor will be placed at the location. However, formulating an optimization problem over binary weight vectors leads to a problem with combinatorial complexity that is computationally prohibitive. A common approach to address this issue is to relax the binary requirement on design weights by letting the weights take values in the interval $[0, 1]$. The sparsity of the design will then be controlled using a penalty method; see, e.g., [2, 13, 23]. This results in an optimization problem of the following form:

$$\min_{\boldsymbol{w}\in[0,1]^{n_s}} \Phi(\boldsymbol{w}) + \gamma P(\boldsymbol{w}), \tag{1}$$

where Φ denotes the design criterion, $\gamma > 0$ is a penalty parameter, and P is a penalty function.

Adopting a Bayesian approach to the inverse problem, the design criterion will be a measure of the uncertainty in the estimated parameters. In this article, we focus on a popular choice known as the A-optimal criterion [8, 22]. That is, we seek a sensor configuration that results in a minimized average posterior variance. The design criterion, in this case, is given by the trace of the posterior covariance operator.

One major challenge in solving (1), specifically for PDE-based inverse problems, is the computational cost of objective function and gradient evaluations. Namely, the posterior covariance operator $\mathbf{\Gamma}_{\text{post}}$ is dense, high-dimensional, and computationally challenging to explicitly form—computing applications of $\mathbf{\Gamma}_{\text{post}}$ to vectors requires solving multiple PDEs. Furthermore, these computations must be performed at each iteration of an optimization algorithm used to solve (1). To address this computational challenge, efficient and accurate matrix-free approaches for computing the OED objective and its gradient are needed. Another challenge in solving the OED problem (1) is the need for a suitable penalty method that is computationally tractable and results in sparse and binary optimal weight vectors. This article is about methods for overcoming these computational challenges.

Related work. For an extensive review of the OED literature, we refer the reader to [3]. We focus here on works that are closely related to the present article. Algorithms for A-optimal designs for ill-posed linear inverse problems were proposed in [10, 11, 13] and more specifically for infinite-dimensional Bayesian linear inverse problems in [2]. In the majority of these articles, Monte Carlo trace estimators are used to approximate the A-optimal design criterion and its gradient. Also, [2, 13] advocate use of low-rank approximations using the Lanczos algorithm or the randomized SVD [14]. We refer to our previous work [20] for comparison of Monte Carlo trace estimators and those based on randomized subspace iteration; it was shown that the latter are significantly more accurate than Monte Carlo trace estimators. Regarding sparsity control, various techniques have been used to approximate the ℓ_0-"norm" to enforce sparse and binary designs. For example, [11, 12, 13] use the ℓ_1-penalty function with an appropriate threshold to postprocess the solution. In [2], a continuation approach is proposed that involves solving a sequence of optimization problems with nonconvex penalty functions that approximate the ℓ_0-"norm." More recently, in [23], a sum-up rounding approach is proposed to obtain binary optimal designs.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Our approach and contributions. In this article, we make the following advances in methods for A-optimal sensor placements in infinite-dimensional Bayesian linear inverse problems:

1. We present efficient and accurate randomized estimators of the A-optimal criterion and its gradient, based on randomized subspace iteration. This is accompanied by a detailed algorithm that guides efficient implementations, discussion of computational cost, as well as theoretical error analysis; see section 3. Our estimators are structure exploiting, in that they use the low-rank structure embedded in the posterior covariance operator. To quantify the accuracy of the estimators we present rigorous error analysis, significantly advancing the methods in [20]. A desirable feature of our analysis is that the bounds are independent of the dimension of the discretized inversion parameter. Furthermore, the computational cost (measured in the number of PDE solves) of the A-optimal objective and gradient using our proposed estimators is independent of the discretized parameter dimension.
2. We propose a new algorithm for optimal sensor placement that is based on solving a sequence of reweighted ℓ_1-optimization problems; see section 4. An important benefit of this approach is that one works with convex penalty functions, and since the A-optimal criterion itself is a convex function of $\boldsymbol{w}$, in each step of the reweighted ℓ_1 algorithm a convex optimization problem is solved. We derive this algorithm by using the majorization-minimization (MM) principle applied to a novel penalty function that promotes binary designs. The solution of the reweighted ℓ_1-optimization problems is accelerated by the efficient randomized estimators for the optimality criterion and its gradient. To our knowledge, the presented framework, based on reweighted ℓ_1-minimization, is the first of its kind in the context of OED.
3. Motivated by reducing computational cost, we propose a new criterion known as modified A-optimal criterion; see section 5. This criterion is derived by considering a suitably weighted A-optimal criterion. We present randomized estimators with complete error analysis for computing the modified A-optimal criterion and its gradient.

We illustrate the benefits of the proposed algorithms on a model problem from contaminant source identification. A comprehensive set of numerical experiments is provided to test various aspects of the presented approach; see section 6.

Finally, we remark that the randomized estimators and the reweighted ℓ_1 approach for promoting sparse and binary weights are of independent interest beyond the application to OED.

2. Preliminaries.

In this section, we recall the background material needed in the remainder of the article.

2.1. Bayesian linear inverse problems.

We consider a linear inverse problem of estimating m, using the model

$$Fm + \boldsymbol{\eta} = \boldsymbol{y}.$$

Here F is a linear parameter-to-observable map (also called the forward operator), $\boldsymbol{\eta}$ represents the measurement noise, and $\boldsymbol{y}$ is a vector of measurement data. The inversion parameter m is an element of $\mathcal{V} = L^2(\mathcal{D})$, where $\mathcal{D}$ is a bounded domain.

The setup of the inverse problem. To fully specify the inverse problem, we need to describe the prior law of m and our choice of data likelihood. For the prior,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

we choose a Gaussian measure $\mu_{\text{pr}} = \mathcal{N}(m_{\text{pr}}, \Gamma_{\text{pr}})$. We assume the prior mean m_{pr} is a sufficiently regular element of $\mathcal{V}$ and that the covariance operator $\Gamma_{\text{pr}} : \mathcal{V} \to \mathcal{V}$ is a strictly positive self-adjoint trace-class operator. Following the developments in [6, 9, 21], we use $\Gamma_{\text{pr}} = \mathcal{A}^{-2}$ with $\mathcal{A}$ taken to be a Laplacian-like operator [21]. This ensures that Γ_{pr} is trace-class in two and three space dimensions.

We consider the case where F represents a time-dependent PDE and we assume observations are taken at n_s sensor locations at n_t points in time. Thus, the vector of experimental data $\boldsymbol{y}$ is an element of $\mathbb{R}^{n_s n_t}$. An application of the parameter-to-observable map, $F : \mathcal{V} \to \mathbb{R}^{n_s n_t}$, involves a PDE solve followed by an application of a spatiotemporal observation operator. We assume a Gaussian distribution on the experimental noise, $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\text{noise}})$. Given this choice of the noise model—additive and Gaussian—the likelihood probability density function is

$$\pi_{\text{like}}(\boldsymbol{y} \mid m) \propto \exp\left\{-\frac{1}{2}\big(Fm - \boldsymbol{y}\big)^{\top} \boldsymbol{\Gamma}_{\text{noise}}^{-1}\big(Fm - \boldsymbol{y}\big)\right\}.$$

Furthermore, the solution of the Bayesian inverse problem—the posterior distribution law $\mu_{\text{post}}^{\boldsymbol{y}}$—is given by the Gaussian measure $\mu_{\text{post}}^{\boldsymbol{y}} = \mathcal{N}(m_{\text{post}}^{\boldsymbol{y}}, \Gamma_{\text{post}})$ with

$$m_{\text{post}}^{\boldsymbol{y}} = \Gamma_{\text{post}}\left(F^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{y} + \Gamma_{\text{pr}}^{-1} m_{\text{pr}}\right), \qquad \Gamma_{\text{post}} = \left(F^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} F + \Gamma_{\text{pr}}^{-1}\right)^{-1}. \tag{2}$$

Note that here the posterior mean $m_{\text{post}}^{\boldsymbol{y}}$ coincides with the maximum a posteriori probability (MAP) estimator. We refer to [21] for further details.

Discretization. We use a continuous Galerkin finite element discretization approach for the governing PDEs, as well as the inverse problem. Specifically, our discretization of the Bayesian inverse problem follows the developments in [6]. The discretized parameter space in the present case is $\mathcal{V}_n = \mathbb{R}^n$ equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathbf{M}}$ and norm $\|\cdot\|_{\mathbf{M}} = \langle \cdot, \cdot \rangle_{\mathbf{M}}^{1/2}$, where $\mathbf{M}$ is the finite element mass matrix. Note that $\langle \cdot, \cdot \rangle_{\mathbf{M}}$ is the discretized $L^2(\mathcal{D})$ inner product. The discretized parameter-to-observable map is a linear transformation $\mathbf{F} : \mathcal{V}_n \to \mathbb{R}^{n_s n_t}$ with adjoint $\mathbf{F}^*$ discussed below. The discretized prior measure $\mathcal{N}(\boldsymbol{m}_{\text{pr}}, \boldsymbol{\Gamma}_{\text{pr}})$ is obtained by discretizing the prior mean and covariance operator, and the discretized posterior measure is given by $\mathcal{N}(\boldsymbol{m}_{\text{post}}, \boldsymbol{\Gamma}_{\text{post}})$, with

$$\boldsymbol{\Gamma}_{\text{post}} = \left(\mathbf{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{F} + \boldsymbol{\Gamma}_{\text{pr}}^{-1}\right)^{-1}, \quad \boldsymbol{m}_{\text{post}}^{\boldsymbol{y}} = \boldsymbol{\Gamma}_{\text{post}}\left(\mathbf{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{y} + \boldsymbol{\Gamma}_{\text{pr}}^{-1} \boldsymbol{m}_{\text{pr}}\right).$$

We point out that the operator $\mathbf{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{F}$ is the Hessian of the data-misfit cost functional whose minimizer is the MAP point and is thus referred to as the data-misfit Hessian; see, e.g., [2].

It is important to keep track of the inner products in the domains and ranges of the linear mappings, appearing in the above expressions, when computing the respective adjoint operators. For the readers' convenience, in Figure 1, we summarize the different spaces that are important, the respective inner products, and the adjoints of the linear transformations defined between these spaces.

Using the fact that $\boldsymbol{\Gamma}_{\text{pr}}$ is a self-adjoint operator on $\mathcal{V}_n$ and the form of the adjoint operator $\mathbf{F}^*$ (see Figure 1), we can rewrite the expression for $\boldsymbol{\Gamma}_{\text{post}}$ as follows:

$$\boldsymbol{\Gamma}_{\text{post}} = \boldsymbol{\Gamma}_{\text{pr}}^{1/2} \mathbf{M}^{-1/2} \left(\mathbf{I} + \boldsymbol{\mathcal{F}}^{\top} \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{\mathcal{F}}\right)^{-1} \mathbf{M}^{1/2} \boldsymbol{\Gamma}_{\text{pr}}^{1/2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

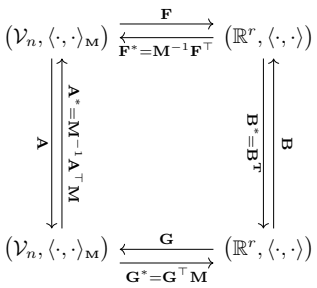

FIG. 1. *Different spaces, their inner products, and the adjoints of linear transformations between them. Here $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product and $\langle\cdot,\cdot\rangle_{\mathbf{M}}$ is the mass-weighted inner product.*

with

$$\boldsymbol{\mathcal{F}} = \mathbf{F}\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\mathbf{M}^{-1/2}. \tag{3}$$

Note that the operator $\boldsymbol{\mathcal{H}}_{\text{m}} = \boldsymbol{\mathcal{F}}^\top\boldsymbol{\Gamma}_{\text{noise}}^{-1}\boldsymbol{\mathcal{F}}$ is a symmetric positive semidefinite matrix and is a similarity transform of the prior-preconditioned data-misfit Hessian $\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\mathbf{F}^*\boldsymbol{\Gamma}_{\text{noise}}^{-1}\mathbf{F}\boldsymbol{\Gamma}_{\text{pr}}^{1/2}$. In many applications (including the application considered in section 6), $\boldsymbol{\mathcal{H}}_{\text{m}}$ has rapidly decaying eigenvalues and therefore it can be approximated by a low-rank matrix. This is a key insight that will be exploited in our estimators for the OED criterion and its gradient.

2.2. Randomized subspace iteration algorithm. In this article, we develop and use randomized estimators to efficiently compute the design criteria and their derivatives. We first explain how to use randomized algorithms for computing low-rank approximations of a symmetric positive semidefinite matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$. To draw connection with the previous subsection, in our application $\mathbf{A}$ will stand for $\boldsymbol{\mathcal{H}}_{\text{m}}$. We first draw a random Gaussian matrix $\boldsymbol{\Omega} \in \mathbb{R}^{n\times\ell}$ (i.e., the entries are independent and identically distributed standard normal random variables). We then perform q steps of subspace iteration on $\mathbf{A}$ with the starting guess $\boldsymbol{\Omega}$ to obtain the matrix $\mathbf{Y}$. If, for example, the matrix has rank $k \le \ell$, or the eigenvalues decay sufficiently, then the range of $\mathbf{Y}$ is a good approximation to the range of $\mathbf{A}$ under these suitable conditions. This is the main insight behind randomized algorithms. We now show how to obtain a low-rank approximation of $\mathbf{A}$. A thin-QR factorization of $\mathbf{Y}$ is performed to obtain the matrix $\mathbf{Q}$, which has orthonormal columns. We then form the "projected" matrix $\mathbf{T} = \mathbf{Q}^\top\mathbf{A}\mathbf{Q}$ and obtain the low-rank approximation

$$\mathbf{A} \approx \mathbf{Q}\mathbf{T}\mathbf{Q}^\top. \tag{4}$$

This low-rank approximation can be manipulated in many ways depending on the desired application. An alternative low-rank approximation can be computed using the Nyström approximation; see, e.g., [14]. In addition, once the matrix $\mathbf{T}$ is computed, it can be used in various ways. In [20], $\mathsf{tr}\,(\mathbf{T})$ was used as an estimator for $\mathsf{tr}\,(\mathbf{A})$, whereas $\log\det(\mathbf{I}+\mathbf{T})$ was used as an estimator for $\log\det(\mathbf{I}+\mathbf{A})$. The main idea behind these estimators is that the eigenvalues of $\mathbf{T}$ are good approximations to the eigenvalues of $\mathbf{A}$, when $\mathbf{A}$ is sufficiently low-rank or has rapidly decaying eigenvalues. Our estimators for the A-optimal criterion and its gradient utilize the same idea but in a slightly different form.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

2.3. A-optimal design of experiments. As mentioned in the introduction, an experimental design refers to a placement of sensors used to collect measurement data for the purposes of parameter inversion. Here we describe the basic setup of the optimization problem for finding an A-optimal design.

Experimental design and A-optimal criterion. We seek to find an optimal subset of a network of n_s candidate sensor locations, which collect measurements at n_t points in time. The experimental design is parameterized by a vector of design weights $\boldsymbol{w} \in [0,1]^{n_s}$. In the present work, we use the A-optimal criterion to find the optimal design. That is, we seek designs that minimize the average posterior variance, as quantified by $\mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w})\right)$. (The precise nature of the dependence of $\boldsymbol{\Gamma}_{\text{post}}$ on $\boldsymbol{w}$ will be explained below.) Note that $\mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w})\right) = \mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w}) - \boldsymbol{\Gamma}_{\text{pr}}\right) + \mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{pr}}\right)$ and thus, minimizing the trace of the posterior covariance operator is equivalent to minimizing

$$\Phi_{\text{aopt}}(\boldsymbol{w}) \equiv \mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w}) - \boldsymbol{\Gamma}_{\text{pr}}\right). \tag{5}$$

This is the objective function we seek to minimize for finding A-optimal designs. As seen below, this formulation of the A-optimal criterion is well suited for approximations via randomized matrix methods in subsection 2.2. Note also that minimizing $\Phi_{\text{aopt}}(\boldsymbol{w})$ amounts to maximizing $\mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{pr}}\right) - \mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w})\right)$, which can be thought of as a measure of uncertainty reduction.

We can also understand (5) from a decision theoretic point of view. It is well known [1, 4, 8] that for Bayesian linear inverse problems with Gaussian prior and additive Gaussian noise models, $\mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}\right)$ coincides with Bayes risk (with respect to the L^2 loss function):

$$\begin{aligned}\mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}}\right) &= \mathbb{E}_{\mu_{\text{pr}}}\Big(\mathbb{E}_{\pi_{\text{like}}(\boldsymbol{y}|\boldsymbol{m})}\big(\|\boldsymbol{m}_{\text{post}}^{\boldsymbol{y}} - \boldsymbol{m}\|_{\text{M}}^2\big)\Big) \\ &= \int_{\mathcal{V}_n}\int_{\mathbb{R}^{n_s n_t}} \|\boldsymbol{m}_{\text{post}}^{\boldsymbol{y}} - \boldsymbol{m}\|_{\text{M}}^2\, \pi_{\text{like}}(\boldsymbol{y} \mid \boldsymbol{m}) d\boldsymbol{y}\, \mu_{\text{pr}}(d\boldsymbol{m}).\end{aligned}$$

Here μ_{pr} denotes the discretized prior measure, $\mu_{\text{pr}} = \mathcal{N}(\boldsymbol{m}_{\text{pr}}, \boldsymbol{\Gamma}_{\text{pr}})$. Using

$$\int_{\mathcal{V}_n} \|\boldsymbol{m}_{\text{pr}} - \boldsymbol{m}\|_{\text{M}}^2\, \mu_{\text{pr}}(d\boldsymbol{m}) = \mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{pr}}\right),$$

we see that

$$\mathsf{tr}\left(\boldsymbol{\Gamma}_{\text{post}} - \boldsymbol{\Gamma}_{\text{pr}}\right) = \mathbb{E}_{\mu_{\text{pr}}}\Big(\mathbb{E}_{\pi_{\text{like}}(\boldsymbol{y}|\boldsymbol{m})}\big(\|\boldsymbol{m}_{\text{post}}^{\boldsymbol{y}} - \boldsymbol{m}\|_{\text{M}}^2 - \|\boldsymbol{m}_{\text{pr}} - \boldsymbol{m}\|_{\text{M}}^2\big)\Big);$$

this provides an alternate interpretation of $\Phi_{\text{aopt}}(\cdot)$ as a Bayes risk with respect to a modified loss function.

Dependence of the A-optimal criterion to the design weights. We follow the same setup as [2]. Namely, the design weights enter the Bayesian inverse problem through the data likelihood, resulting in the $\boldsymbol{\mathcal{H}}_{\text{m}}$ operator being dependent on $\boldsymbol{w}$. To weight the spatiotemporal observations, the matrix $\mathbf{W}$ is defined as

$$\mathbf{W} = \sum_{i=1}^{n_s} w_i \mathbf{E}_i \quad \text{with} \quad \mathbf{E}_i = \mathbf{I}_{n_t} \otimes \mathbf{e}_i \mathbf{e}_i^{\top},$$

where $\otimes$ is the Kronecker product. Therefore, $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ is expressed as follows:

$$\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}) = \boldsymbol{\mathcal{F}}^{\top} \mathbf{W}^{1/2} \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{W}^{1/2} \boldsymbol{\mathcal{F}},$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We refer to [2] for details. This results in the $\boldsymbol{w}$ dependent posterior covariance operator:

$$\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w}) = \boldsymbol{\Gamma}_{\text{pr}}^{1/2}\mathbf{M}^{-1/2}\left(\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}) + \mathbf{I}\right)^{-1}\mathbf{M}^{1/2}\boldsymbol{\Gamma}_{\text{pr}}^{1/2}.$$

In our formulation, we assume uncorrelated observations across sensor locations and time which implies $\boldsymbol{\Gamma}_{\text{noise}}$ is a $n_s n_t \times n_s n_t$ block diagonal matrix, with n_t blocks of the form $\mathsf{diag}\,(\sigma_1^2, \ldots, \sigma_{n_s}^2)$; here σ_i^2, $i = 1, \ldots, n_s$, denote the measurement noise at individual sensor locations. Using this structure for $\boldsymbol{\Gamma}_{\text{noise}}$, we define

$$\mathbf{W}^{\text{noise}} \equiv \mathbf{W}^{1/2}\boldsymbol{\Gamma}_{\text{noise}}^{-1}\mathbf{W}^{1/2} = \sum_{i=1}^{n_{\text{sens}}} w_i \mathbf{E}_i^{\text{noise}} \tag{6}$$

with $\mathbf{E}_i^{\text{noise}} = \sigma_i^{-2}\mathbf{I}_{n_{\text{time}}} \otimes \mathbf{e}_i\mathbf{e}_i^\top$. Thus, we have $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}) = \boldsymbol{\mathcal{F}}^\top \mathbf{W}^{\text{noise}}\boldsymbol{\mathcal{F}}$ and the A-optimal criterion can be written as

$$\begin{aligned}\Phi_{\text{aopt}}(\boldsymbol{w}) &= \mathsf{tr}\,\left(\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\mathbf{M}^{-1/2}\left[\left(\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}) + \mathbf{I}\right)^{-1} - \mathbf{I}\right]\mathbf{M}^{1/2}\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\right)\\ &= \mathsf{tr}\,\left(\left[\left(\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}) + \mathbf{I}\right)^{-1} - \mathbf{I}\right]\mathbf{Z}\right)\end{aligned} \tag{7}$$

with

$$\mathbf{Z} \equiv \mathbf{M}^{1/2}\boldsymbol{\Gamma}_{\text{pr}}\mathbf{M}^{-1/2}. \tag{8}$$

Anticipating that we will use a gradient-based solver for solving (10), we also need the gradient of $\Phi_{\text{aopt}}(\boldsymbol{w})$ which we now derive. Using Theorems B.17 and B.19 in [22], the partial derivatives of (7) with respect to w_j, $j = 1, \ldots, n_s$, are

$$\partial_j \Phi_{\text{aopt}}(\boldsymbol{w}) = -\mathsf{tr}\,\left((\mathbf{I} + \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\partial_j \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})(\mathbf{I} + \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{Z}\right). \tag{9}$$

(We have used the notation ∂_j to denote $\frac{\partial}{\partial w_j}$.) Note that using the definition of $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$, we have $\partial_j \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}) = \boldsymbol{\mathcal{F}}^\top \mathbf{E}_j^{\text{noise}}\boldsymbol{\mathcal{F}}$, $j = 1, \ldots, n_s$.

The optimization problem for finding an A-optimal design. We now specialize the optimization problem (1) to the case of A-optimal sensor placement for linear inverse problems governed by time-dependent PDEs:

$$\min_{\boldsymbol{w} \in [0,1]^{n_s}} \Phi_{\text{aopt}}(\boldsymbol{w}) + \gamma P(\boldsymbol{w}). \tag{10}$$

As explained before, to enable efficient solution methods for the above optimization problem we need (i) a numerical method for fast computation of $\Phi_{\text{aopt}}(\boldsymbol{w})$ and its gradient, and (ii) a choice of penalty function that promotes sparse and binary weights. The former is facilitated by the randomized subspace iteration approach outlined earlier (see section 3), and for the latter we present an approach based on reweighted ℓ_1 minimization (see section 4).

3. Efficient computation of A-optimal criterion and its gradient. The computational cost of solving (10) is dominated by the PDE solves required in OED objective and gradient evaluations; these operations need to be performed repeatedly when using an optimization algorithm for solving (10). Therefore, to enable computing A-optimal designs for large-scale applications, efficient methods for objective and gradient computations are needed. In this section, we derive efficient and accurate

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Algorithm 1. Randomized subspace iteration.

Input: $\mathbf{A} \in \mathbb{R}^{n\times n}$ with target rank k, oversampling parameter $p \geq 2$, with $\ell \equiv k + p \leq n$, and $q \geq 1$ (number of subspace iterations).
Output: $\mathbf{Q} \in \mathbb{R}^{n\times \ell}, \mathbf{T} \in \mathbb{R}^{\ell\times\ell}$.
Draw a standard Gaussian random matrix $\mathbf{\Omega} \in \mathbb{R}^{n\times\ell}$.
Compute $\mathbf{Y} = \mathbf{A}^q\mathbf{\Omega}$.
Compute thin QR decomposition $\mathbf{Y} = \mathbf{QR}$.
Compute $\mathbf{T} = \mathbf{Q}^\top\mathbf{AQ}$.

randomized estimators for (7) and (9). The proposed estimators are matrix-free—they require only applications of the (prior-preconditioned) forward operator $\boldsymbol{\mathcal{F}}$ and its adjoint on vectors. Moreover, the computational cost of computing these estimators does not increase with the discretized parameter dimension. This is due to the fact that our estimators exploit the low-rank structure of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$, a problem property that is independent of the choice of discretization.

We introduce our proposed randomized estimators for the A-optimal design criterion $\Phi_{\mathrm{aopt}}(\boldsymbol{w})$ and its gradient $\nabla\Phi_{\mathrm{aopt}}(\boldsymbol{w})$ in subsection 3.1. Additionally, we present a detailed computational method for computing the proposed estimators. We analyze the errors associated with our proposed estimators in subsection 3.2.

3.1. Randomized estimators for $\Phi_{\mathrm{aopt}}(\boldsymbol{w})$ and its gradient. Consider the low-rank approximation of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$ given by $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w}) = \mathbf{Q}(\boldsymbol{w})\mathbf{T}(\boldsymbol{w})\mathbf{Q}^\top(\boldsymbol{w})$ with $\mathbf{Q}(\boldsymbol{w})$ and $\mathbf{T}(\boldsymbol{w})$ computed using Algorithm 1. Replacing $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$ by its approximation and using the cyclic property of the trace, we obtain the estimator for the A-optimal criterion (7):

$$\widehat{\Phi}_{\mathrm{aopt}}(\boldsymbol{w};\ell) = \mathsf{tr}\left(\left((\mathbf{I} + \widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w}))^{-1} - \mathbf{I}\right)\mathbf{Z}\right), \tag{11}$$

where $\mathbf{Z}$ is as in (8).

To derive an estimator for the gradient, once again, we replace $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$ with its low-rank approximation $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})$ in (9) to obtain

$$\widehat{\partial_j\Phi}_{\mathrm{aopt}}(\boldsymbol{w};\ell) = -\mathsf{tr}\left((\mathbf{I} + \widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w}))^{-1}\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\mathrm{noise}}\boldsymbol{\mathcal{F}}(\mathbf{I} + \widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w}))^{-1}\mathbf{Z}\right) \tag{12}$$

for $j = 1, \ldots n_s$.

Computational procedure. First, we discuss computation of the A-optimal criterion estimator using Algorithm 1. Typically, the algorithm can be used with $q = 1$, due to rapid decay of eigenvalues of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$. In this case, Algorithm 1 requires 2ℓ applications of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$. Since each application of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$ requires one $\boldsymbol{\mathcal{F}}$ apply (forward solve) and one $\boldsymbol{\mathcal{F}}^\top$ apply (adjoint solve), computing $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})$ requires 4ℓ PDE solves. Letting the spectral decomposition of the $\ell\times\ell$ matrix $\mathbf{T}(\boldsymbol{w})$ be given by $\mathbf{T}(\boldsymbol{w}) = \mathbf{U}(\boldsymbol{w})\mathbf{\Lambda}_{\mathbf{T}}(\boldsymbol{w})\mathbf{U}(\boldsymbol{w})^\top$ and denoting $\mathbf{V}(\boldsymbol{w}) \equiv \mathbf{Q}(\boldsymbol{w})\mathbf{U}(\boldsymbol{w})$, we have $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w}) = \mathbf{V}(\boldsymbol{w})\mathbf{\Lambda}_{\mathbf{T}}(\boldsymbol{w})\mathbf{V}^\top(\boldsymbol{w})$. Applying the Sherman–Morrison–Woodbury formula [19] and the cyclic property of the trace to (11), we obtain

$$\widehat{\Phi}_{\mathrm{aopt}}(\boldsymbol{w}) = -\mathsf{tr}\left(\mathbf{D}_{\mathbf{T}}(\boldsymbol{w})\mathbf{V}^\top(\boldsymbol{w})\mathbf{Z}\mathbf{V}(\boldsymbol{w})\right), \tag{13}$$

where $\mathbf{D}_{\mathbf{T}}(\boldsymbol{w}) = \mathbf{\Lambda}_{\mathbf{T}}(\boldsymbol{w})(\mathbf{I} + \mathbf{\Lambda}_{\mathbf{T}}(\boldsymbol{w}))^{-1}$. To simplify notation, the dependence of $\boldsymbol{w}$ is suppressed for the operators used in computing the estimators for the remainder of the article; however, the notation is retained for $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$ and $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Next, we describe computation of the gradient estimator (12). Here we assume $n_s n_t \leq n$; the extension to the case $n_s n_t > n$ is straightforward and is omitted. Again, using the Woodbury formula and cyclic property of the trace, we rewrite (12) as

$$\widehat{\partial_j \Phi}_{\text{aopt}}(\boldsymbol{w};\ell) = -\mathsf{tr}\left(\boldsymbol{\mathcal{F}}\left(\mathbf{I} - \mathbf{V}\mathbf{D}_{\mathbf{T}}\mathbf{V}^\top\right)\mathbf{Z}\left(\mathbf{I} - \mathbf{V}\mathbf{D}_{\mathbf{T}}\mathbf{V}^\top\right)\boldsymbol{\mathcal{F}}^\top \mathbf{E}_j^{\text{noise}}\right) \tag{14}$$

for $j = 1, \ldots, n_s$. Expanding this expression, we obtain

$$\begin{aligned}\widehat{\partial_j \Phi}_{\text{aopt}}(\boldsymbol{w}) = &-\mathsf{tr}\left(\mathbf{Z}\boldsymbol{\mathcal{F}}^\top \mathbf{E}_j^{\text{noise}}\boldsymbol{\mathcal{F}}\right) + 2\mathsf{tr}\left(\boldsymbol{\mathcal{F}}\mathbf{V}\mathbf{D}_{\mathbf{T}}\mathbf{V}^\top\mathbf{Z}\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\text{noise}}\right)\\ &-\mathsf{tr}\left(\boldsymbol{\mathcal{F}}\mathbf{V}\mathbf{D}_{\mathbf{T}}\mathbf{V}^\top\mathbf{Z}\mathbf{V}\mathbf{D}_{\mathbf{T}}\mathbf{V}^\top\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\text{noise}}\right).\end{aligned} \tag{15}$$

Note that the first term $s_j = -\mathsf{tr}\left(\mathbf{Z}\boldsymbol{\mathcal{F}}^\top \mathbf{E}_j^{\text{noise}}\boldsymbol{\mathcal{F}}\right)$ in (15) does not depend on the design $\boldsymbol{w}$, for $j = 1, \ldots, n_s$. As a result, this term can be precomputed and used in subsequent function evaluations. We expand s_j to be

$$\begin{aligned}s_j &= -\mathsf{tr}\left(\mathbf{Z}\boldsymbol{\mathcal{F}}^\top \mathbf{E}_j^{\text{noise}}\boldsymbol{\mathcal{F}}\right)\\ &= -\sum_{k=1}^{n_s n_t}\left(\boldsymbol{\mathcal{F}}^\top(\mathbf{E}_j^{\text{noise}})^{1/2}\widehat{\boldsymbol{e}}_k\right)^\top \mathbf{Z}\left(\boldsymbol{\mathcal{F}}^\top(\mathbf{E}_j^{\text{noise}})^{1/2}\widehat{\boldsymbol{e}}_k\right),\end{aligned}$$

where $\widehat{\boldsymbol{e}}_k$ is the kth column of the identity matrix of size $n_s n_t$. Because there are only n_t columns of $\mathbf{E}_j^{\text{noise}}$ with nonzero entries, the total cost to precompute s_j for $j = 1, \ldots, n_s$ is $n_s n_t$ PDE solves.

To compute the remaining terms in (15), we exploit the fact that $\mathbf{V}$ has ℓ columns. Notice all the other occurrences of $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{F}}^\top$ occur as a combination of $\boldsymbol{\mathcal{F}}\mathbf{V}$ and $\boldsymbol{\mathcal{F}}\mathbf{Z}\mathbf{V}$ (or of their transposes). Both of these terms require ℓ PDE solves to compute. As a result, the total cost to evaluate $\widehat{\Phi}_{\text{aopt}}(\boldsymbol{w})$ and $\widehat{\nabla\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$ is 4ℓ PDE solves to apply Algorithm 1 and 2ℓ PDE solves to compute $\boldsymbol{\mathcal{F}}\mathbf{V}$ and $\boldsymbol{\mathcal{F}}\mathbf{Z}\mathbf{V}$. We detail the steps for computing our estimators for A-optimal criterion and its gradient in Algorithm 2.

Alternative approaches and summary of computational cost. A closely related variation of Algorithm 2 is obtained by replacing step 1 of the algorithm (i.e., randomized subspace iteration) by the solution of an eigenvalue problem to compute the dominant eigenvalues of $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ "exactly." We refer to this method as Eig-k, where k is the target rank of $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$. This idea was explored for computing Bayesian D-optimal designs in [3]. The resulting cost is similar to that of the randomized method: it would cost $\mathcal{O}(k)$ PDE solves per iteration to compute the spectral decomposition of $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$, plus $\min\{n_s n_t, n\}$ PDE solves to precompute s_j for $j = 1, \ldots, n_s$. While both the randomized and Eig-k methods provide a viable scheme for computing the A-optimal criteria, our randomized method can exploit parallelism to lower computational costs. Each matrix-vector application with $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ in Algorithm 1 can be computed in parallel. However, if accurate eigenpairs of $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ are of importance to the problem, one can choose to use the Eig-k approach at the cost of computing a more challenging problem.

Another possibility suitable for problems where the forward model does not depend on the design (as is the case in the present work) is to precompute a low-rank SVD of $\boldsymbol{\mathcal{F}}$, which can then be applied as necessary to compute the A-optimal criterion and its gradient. This *frozen forward operator approach* has been explored in [2, 13] for the A-optimal criterion and in [3] for the D-optimal criterion. The resulting PDE cost of precomputing a low-rank approximation of $\boldsymbol{\mathcal{F}}$ is $\mathcal{O}(k)$, with k indicating the target rank. The frozen method is beneficial as no additional PDE solves are required

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

when applying Algorithm 2; however, this approach would not favor problems where $\boldsymbol{\mathcal{F}}$ depends on $\boldsymbol{w}$, nor can the modeling errors associated with the PDE be controlled in subsequent evaluations without the construction of another operator.

Finally, if the problem size is not too large (i.e., in small-scale applications), one could explicitly construct the forward operator $\boldsymbol{\mathcal{F}}$. This enables exact (excluding floating point errors) computation of the objective and its gradient. The total cost for evaluating $\boldsymbol{\mathcal{F}}$ involves an upfront cost of $\min\{n_s n_t, n\}$ PDE solves. We summarize the computational cost of Algorithm 2 along with the other alternatives mentioned in Table 1.

To summarize, the randomized methods for computing the OED objective and its gradient present several attractive features; our approach is well suited to large-scale applications; it is matrix-free, is simple to implement and parallelize, and exploits low-rank structure in the inverse problem. Moreover, as is the case with the Eig-k

Algorithm 2. Randomized method for computing $\widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$ and $\widehat{\nabla\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$.

Input: Target rank k, oversampling parameter $p \geq 0$, design $\boldsymbol{w}$, and s_j for $j = 1,\dots,n_s$.

Output: OED objective $\widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$ and gradient $\widehat{\nabla\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$.

Apply Algorithm 1 with $\ell = k+p$ and $q = 1$ to obtain $\mathbf{T} \in \mathbb{R}^{\ell\times\ell}$ and $\mathbf{Q} \in \mathbb{R}^{n\times\ell}$.
Compute eigendecomposition $[\mathbf{U}, \boldsymbol{\Lambda}_{\mathbf{T}}]$ of $\mathbf{T}$. Let $\mathbf{D}_{\mathbf{T}} = \boldsymbol{\Lambda}_{\mathbf{T}}(\mathbf{I} + \boldsymbol{\Lambda}_{\mathbf{T}})^{-1}$.
for $i = 1$ to ℓ **do**
  Compute $\boldsymbol{v_i} = \mathbf{Q}\boldsymbol{u_i}$, where $\boldsymbol{u_i}$ are the columns of $\mathbf{U}$.
end for
Compute

$$\widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell) = -\sum_{i=1}^{\ell} d_i \boldsymbol{v}_i^\top \mathbf{Z}\boldsymbol{v}_i,$$

where d_i is the ith diagonal of $\mathbf{D}_{\mathbf{T}}$.
for $i = 1$ to ℓ **do**
  Compute $\boldsymbol{a}_i = \boldsymbol{\mathcal{F}}\boldsymbol{v}_i$ and $\boldsymbol{b}_i = \boldsymbol{\mathcal{F}}\mathbf{Z}\boldsymbol{v}_i$.
end for
for $j = 1$ to n_s **do**
  Compute

$$\widehat{\partial_j\Phi}_{\text{aopt}}(\boldsymbol{w};\ell) = s_j + 2\sum_{i=1}^{\ell} d_i \boldsymbol{b}_i^\top \mathbf{E}_j^{\text{noise}} \boldsymbol{a}_i - \sum_{i=1}^{\ell}\sum_{k=1}^{\ell} d_i d_k \boldsymbol{a}_i^\top \mathbf{E}_j^{\text{noise}} \boldsymbol{a}_k \boldsymbol{v}_k^\top \mathbf{Z}\boldsymbol{v}_i.$$

end for
Return $\widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$ and $\widehat{\nabla\Phi}_{\text{aopt}}(\boldsymbol{w};\ell) = [\widehat{\partial_1\Phi}_{\text{aopt}}(\boldsymbol{w};\ell),\dots,\widehat{\partial_{n_s}\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)]^\top$.

TABLE 1
Computational cost measured in terms of PDE solves for different methods in computing $\Phi_{aopt}(\boldsymbol{w})$ and $\nabla\Phi_{aopt}(\boldsymbol{w})$. Typically, $q = 1$ in Algorithm 1 is sufficient.

Method	$\Phi_{\text{aopt}}(\boldsymbol{w})$ and $\nabla\Phi_{\text{aopt}}(\boldsymbol{w})$	Precompute	Storage cost
"Exact"	–	$\min\{n_s n_t, n\}$	$nn_s n_t$
Frozen	–	$\mathcal{O}(k)$	$(2n+1)k$
Eig-k	$\mathcal{O}(k)$	$\min\{n_s n_t, n\}$	n_s
Randomized	$2(q+2)(k+p)$	$\min\{n_s n_t, n\}$	n_s

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

approach, the performance of the randomized subspace iteration does not degrade as the dimension of the parameter increases due to mesh refinement. This is the case because the randomized subspace iteration relies on spectral properties of the prior-preconditioned data-misfit Hessian—a problem structure that is independent of discretization.

3.2. Error analysis. Here we analyze the error associated with our estimators computed using Algorithm 2. For fixed $\boldsymbol{w} \in [0,1]^{n_s}$, since $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w}) \in \mathbb{R}^{n\times n}$ is symmetric positive semidefinite, we can order its eigenvalues as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Suppose $\lambda_1, \ldots, \lambda_k$ are the dominant eigenvalues of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$, we define $\boldsymbol{\Lambda}_1 = \mathsf{diag}\,(\lambda_1, \ldots, \lambda_k)$ and $\boldsymbol{\Lambda}_2 = \mathsf{diag}\,(\lambda_{k+1}, \ldots, \lambda_n)$ and we assume that the eigenvalue ratio satisfies

$$\gamma_k = \|\boldsymbol{\Lambda}_2\|_2 \|\boldsymbol{\Lambda}_1^{-1}\|_2 = \frac{\lambda_{k+1}}{\lambda_k} < 1.$$

We now present the error bounds for the objective function and its gradient. To this end, we define the constant C as

$$C \equiv \frac{e^2(k+p)}{(p+1)^2}\left(\frac{1}{2\pi(p+1)}\right)^{\frac{2}{p+1}}\left(\mu+\sqrt{2}\right)^2\left(\frac{p+1}{p-1}\right) \tag{16}$$

with $r = \mathrm{rank}(\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w}))$ and $\mu \equiv \sqrt{r-k} + \sqrt{k+p}$.

THEOREM 3.1. *Let $\widehat{\Phi}_{aopt}(\boldsymbol{w};\ell)$ and $\widehat{\nabla\Phi}_{aopt}(\boldsymbol{w};\ell)$ be approximations of the A-optimal objective function $\Phi_{aopt}(\boldsymbol{w})$ and its gradient $\nabla\Phi_{aopt}(\boldsymbol{w})$, respectively, computed using Algorithm 2 for fixed $\boldsymbol{w} \in [0,1]^{n_s}$. Recall that k is the target rank of $\boldsymbol{\mathcal{H}}_m(\boldsymbol{w})$, $p \geq 2$ is the oversampling parameter such that $k+p \leq n$, and $q \geq 1$ is the number of subspace iterations. Assume that $\gamma_k < 1$. Then, with $f = x/(1+x)$*

$$\mathbb{E}\left[\left|\Phi_{aopt}(\boldsymbol{w}) - \widehat{\Phi}_{aopt}(\boldsymbol{w};\ell)\right|\right] \leq \|\mathbf{Z}\|_2\left(\mathsf{tr}\,\left(f(\boldsymbol{\Lambda}_2)\right) + \mathsf{tr}\,\left(f\left(\gamma_k^{2q-1}C\boldsymbol{\Lambda}_2\right)\right)\right). \tag{17}$$

Furthermore, with $\mathbf{P}_j = \boldsymbol{\mathcal{F}}^\top \mathbf{E}_j^{noise}\boldsymbol{\mathcal{F}}$, for $j = 1, \ldots, n_s$,

$$\mathbb{E}\left[\left|\partial_j\Phi_{aopt}(\boldsymbol{w}) - \partial_j\widehat{\Phi}_{aopt}(\boldsymbol{w};\ell)\right|\right] \leq 2\|\mathbf{Z}\|_2\|\mathbf{P}_j\|_2\left(\mathsf{tr}\,\left(f(\boldsymbol{\Lambda}_2)\right) + \mathsf{tr}\,\left(f\left(\gamma_k^{2q-1}C\boldsymbol{\Lambda}_2\right)\right)\right). \tag{18}$$

Proof. See Appendix A.2. ∎

In Theorem 3.1, the estimators are unbiased when the target rank equals the rank of $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$. If the eigenvalues decay rapidly, the bounds suggest that the estimators are accurate. Recall that $\mathrm{rank}(\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})) \leq \min\{n_s n_t, n\}$ is the number of nonzero eigenvalues. Consequently, it is seen that the bounds are independent of the dimension of the discretization n.

4. An optimization framework for finding binary designs. We seek A-optimal designs by solving an optimization problem of the form,

$$\min_{\boldsymbol{w}\in\mathcal{W}} \Phi(\boldsymbol{w}) + \gamma P(\boldsymbol{w}), \quad \mathcal{W} = [0,1]^{n_s}, \tag{19}$$

where the design criterion Φ is either the A-optimal criterion $\Phi_{\mathrm{aopt}}(\boldsymbol{w})$ or the modified A-optimal criterion $\Phi_{\mathrm{mod}}(\boldsymbol{w})$ (see section 5). In the previous sections, we laid out an efficient framework for computing accurate approximations to the A-optimal criterion

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and its gradient. We now discuss the choice of the penalty term P and the algorithm for solving the optimization problem.

The choice of the penalty term must satisfy two conditions: enforcing sparsity, measured by the number of nonzeros of the design vector, and binary designs, i.e., designs vectors whose entries are either 1 or 0. One possibility for the penalty function is the ℓ_0-"norm," $P_{\ell_0}(\boldsymbol{w}) = \|\boldsymbol{w}\|_0$, which measures the number of nonzero entries in the design. However, the resulting optimization problem is challenging to solve due to its combinatorial complexity. A common practice is to replace the ℓ_0-"norm" penalty by the ℓ_1-norm, $P_{\ell_1}(\boldsymbol{w}) = \|\boldsymbol{w}\|_1$. The penalty function P_{ℓ_1} has desirable features: it is a convex penalty function that promotes sparsity of the optimal design vector $\boldsymbol{w}$. However, the resulting design is sparse but not necessarily binary and additional postprocessing in the form of thresholding is necessary to enforce binary designs.

In what follows, we introduce a suitable penalty function that enforces both sparsity and binary designs and an algorithm for solving the OED optimization problem based on the MM approach. The resulting algorithm takes the form of a reweighted ℓ_1-minimization algorithm [7].

4.1. Penalty functions. We propose the penalty function

$$P_\epsilon(\boldsymbol{w}) = \sum_{i=1}^{n_s} \frac{|w_i|}{|w_i| + \epsilon}, \quad \boldsymbol{w} \in \mathbb{R}^{n_s}, \tag{20}$$

for a user-defined parameter $\epsilon > 0$. This penalty function approximates P_{ℓ_0} for small values of ϵ; however, as ϵ becomes smaller the corresponding optimization problem becomes harder. To illustrate the choice of penalty functions, in Figure 2, we plot $P_{0.05}$ along with P_{ℓ_0} and P_{ℓ_1}, with $n_s = 1$. Using P_ϵ in the OED problem leads to the optimization problem,

$$\min_{\boldsymbol{w} \in \mathcal{W}} \Phi(\boldsymbol{w}) + \gamma P_\epsilon(\boldsymbol{w}). \tag{21}$$

In (21), the absolute values in definition of P_ϵ can be dropped since we limit the search for optimal solutions in $\mathcal{W}$. Since $P_\epsilon(\boldsymbol{w})$ is concave, (21) is a nonconvex optimization problem. To tackle this, we adopt the MM approach.

4.2. MM approach and reweighted ℓ_1 algorithm. The idea behind the MM approach is to solve a sequence of optimization problems whose solutions converge to that of the original problem [16, 17]. This sequence is generated by a carefully constructed surrogate that satisfies two properties—the surrogate must majorize the

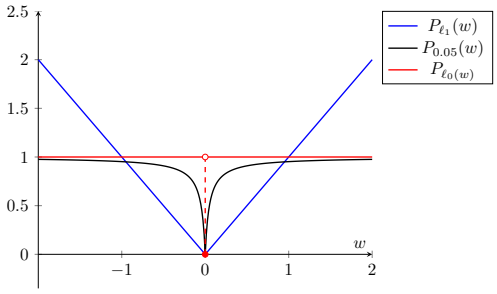

FIG. 2. *Different choices of penalty functions with* $n_s = 1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

objective function for all values, and the surrogate must match the objective function at the current iterate. More specifically, suppose

$$J(\boldsymbol{w}) = \Phi(\boldsymbol{w}) + \gamma \sum_{i=1}^{n_s} \frac{w_i}{w_i + \epsilon}.$$

Then the surrogate function $g(\boldsymbol{w}|\boldsymbol{w}^{(m)})$ at the current iterate $\boldsymbol{w}^{(m)}$ must satisfy

$$g\left(\boldsymbol{w}|\boldsymbol{w}^{(m)}\right) \geq J(\boldsymbol{w}) \quad \forall \boldsymbol{w} \in \mathcal{W},$$
$$g\left(\boldsymbol{w}^{(m)}|\boldsymbol{w}^{(m)}\right) = J(\boldsymbol{w}^{(m)}).$$

Granted the existence of this surrogate function, to find the next iterate $\boldsymbol{w}^{(m+1)}$ we solve the optimization problem

$$\boldsymbol{w}^{(m+1)} = \operatorname*{arg\,min}_{\boldsymbol{w}\in\mathcal{W}} g\left(\boldsymbol{w}|\boldsymbol{w}^{(m)}\right). \tag{22}$$

To show the objective function decreases at the next iterate, observe that the next iterate $\boldsymbol{w}^{(m+1)}$ stays within the feasible region and use the two properties of the surrogate function

$$J\left(\boldsymbol{w}^{(m+1)}\right) \leq g\left(\boldsymbol{w}^{(m+1)}|\boldsymbol{w}^{(m)}\right) \leq g\left(\boldsymbol{w}^{(m)}|\boldsymbol{w}^{(m)}\right) = J(\boldsymbol{w}^{(m)}).$$

To construct this surrogate function, we use the fact that a concave function is below its tangent [17, equation (4.7)]. Applying this to our concave penalty $P_\epsilon(\boldsymbol{w})$, we have

$$P_\epsilon(\boldsymbol{w}) \leq P_\epsilon\left(\boldsymbol{w}^{(m)}\right) + \left(\boldsymbol{w} - \boldsymbol{w}^{(m)}\right)^\top \nabla_{\boldsymbol{w}} P_\epsilon\left(\boldsymbol{w}^{(m)}\right) \quad \text{for all } \boldsymbol{w} \in \mathcal{W}.$$

With this majorization relation, we define the surrogate function to be

$$g\left(\boldsymbol{w}|\boldsymbol{w}^{(m)}\right) = \Phi(\boldsymbol{w}) + \gamma\left(P_\epsilon\left(\boldsymbol{w}^{(m)}\right) + \left(\boldsymbol{w} - \boldsymbol{w}^{(m)}\right)^\top \nabla_{\boldsymbol{w}} P_\epsilon\left(\boldsymbol{w}^{(m)}\right)\right).$$

By dropping the terms that do not depend on $\boldsymbol{w}$, it can be readily verified that (22) can be replaced by the equivalent problem

$$\begin{aligned} \boldsymbol{w}^{(m+1)} &= \operatorname*{arg\,min}_{\boldsymbol{w}\in\mathcal{W}} \Phi(\boldsymbol{w}) + \gamma \sum_{i=1}^{n_s} \frac{\epsilon w_i}{\left(w_i^{(m)} + \epsilon\right)^2} \\ &= \operatorname*{arg\,min}_{\boldsymbol{w}\in\mathcal{W}} \Phi(\boldsymbol{w}) + \gamma \|\boldsymbol{R}\left(\boldsymbol{w}^{(m)}\right)\boldsymbol{w}\|_1, \end{aligned} \tag{23}$$

where $\boldsymbol{R}(\boldsymbol{w}) = \operatorname{diag}(\frac{\epsilon}{(w_1+\epsilon)^2}, \ldots, \frac{\epsilon}{(w_{n_s}+\epsilon)^2})$. We see that (23) is of the form of a reweighted ℓ_1-optimization problem. The details of the optimization procedure are given in Algorithm 3. (We remark that, in Algorithm 3, other metrics measuring the difference between the successive weight vectors can be used in step 3.)

We conclude this section with a few remarks regarding this algorithm. In our application, $\Phi(\boldsymbol{w})$ is convex; therefore, each subproblem to update the design weights is also convex. To initialize the reweighted ℓ_1 algorithm, we start with the weights $r_i = 1$, $i = 1, \ldots, n_s$. This ensures that, in the first step, we are computing the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Algorithm 3. Reweighted ℓ_1 algorithm.

Input: Initial guess $\boldsymbol{w}^{(0)} \in \mathbb{R}^{n_s}$, stopping tolerance tol, penalty parameters $\gamma, \epsilon \geq 0$.
Output: Optimal design $\boldsymbol{w}^* \in \mathbb{R}^{n_s}$.

1: Initialize $m = 1$.
2: Compute

$$\boldsymbol{w}^{(1)} = \arg\min_{\boldsymbol{w} \in \mathcal{W}} \Phi(\boldsymbol{w}) + \gamma \|\boldsymbol{w}\|_1.$$

3: **while** $m < m_{max}$ and $\|\boldsymbol{w}^{(m)} - \boldsymbol{w}^{(m-1)}\|_2 >$ tol **do**
4: Update $m = m + 1$.
5: Compute

$$\boldsymbol{w}^{(m)} = \arg\min_{\boldsymbol{w} \in \mathcal{W}} \Phi(\boldsymbol{w}) + \gamma \sum_{i=1}^{n_s} r_i \cdot w_i.$$

6: Update

$$r_i = \frac{\epsilon}{\left(|w_i^{(m-1)}| + \epsilon\right)^2}, \quad i = 1, \ldots, n_s.$$

7: **end while**
8: Return $\boldsymbol{w}^{(m)} = \boldsymbol{w}^*$.

solution of the ℓ_1-penalized optimization problem. The subsequent reweighted ℓ_1 iterations further promote binary designs. To solve the subproblems in each reweighted ℓ_1 iteration we use an interior point algorithm; however, any solver for appropriate convex optimization may be used. In section 6, we will provide a discussion of our choice of ϵ for our application. It is also worth mentioning that besides the penalty function P_ϵ used above, another possible choice is $\sum_{i=1}^{n_s} \arctan(|w_i|/\epsilon)$ which yields the weights $\frac{1}{|w_i|^2+\epsilon^2}$.

5. Modifed A-optimal criterion. Motivated by reducing the computational cost of computing A-optimal designs, in this section, we introduce a *modified A-optimality* criterion. As mentioned in [8], we can consider a weighted A-optimal criterion $\mathsf{tr}\,(\boldsymbol{\Gamma}\boldsymbol{\Gamma}_{\text{post}}(\boldsymbol{w}))$, where $\boldsymbol{\Gamma}$ is a positive semidefinite weighting matrix. Similar to (5), we work with $\mathsf{tr}\,\left(\boldsymbol{\Gamma}(\boldsymbol{\Gamma}_{\text{post}} - \boldsymbol{\Gamma}_{\text{pr}})\right)$, since the term $\mathsf{tr}\,\left(\boldsymbol{\Gamma}(\boldsymbol{\Gamma}_{\text{pr}})\right)$ is independent of the weights $\boldsymbol{w}$. By choosing $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{\text{pr}}^{-1}$, we obtain the modified A-optimal criterion

$$\Phi_{\text{mod}}(\boldsymbol{w}) \equiv \mathsf{tr}\,\left((\mathbf{I} + \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1} - \mathbf{I}\right). \tag{24}$$

Note that the expression for $\Phi_{\text{mod}}(\boldsymbol{w})$ remains meaningful in the infinite-dimensional limit. This can be seen by noting that

$$\Phi_{\text{mod}}(\boldsymbol{w}) = \mathsf{tr}\,\left(\left(\mathbf{I} + \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})\right)^{-1} - \mathbf{I}\right) = -\mathsf{tr}\,\left(\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})\left(\mathbf{I} + \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})\right)^{-1}\right)$$

and using the fact that in the infinite-dimensional limit, for every $\boldsymbol{w} \in [0,1]^{n_s}$, $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ is trace-class and $\left(\mathbf{I} + \boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})\right)^{-1}$ is a bounded linear operator.

We show in our numerical results that the modified A-optimality criterion can be useful in practice, if a cheaper alternative to the A-optimal criterion is desired, and minimizing Φ_{mod} can provide designs that lead to small posterior uncertainty.

5.1. Derivation of estimators. Here we seek to improve the efficiency of the modified A-optimal criterion by computing a randomized estimator for the modified

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

A-optimal criterion and its gradient. As in the previous derivation of our estimators, we replace $\boldsymbol{\mathcal{H}}_{\mathrm{m}}(\boldsymbol{w})$ by its low-rank approximation to obtain the randomized estimator for the modified A-optimal criterion

$$\Phi_{\mathrm{mod}}(\boldsymbol{w}) \approx \mathsf{tr}\left(\left(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})\right)^{-1}-\mathbf{I}\right) \equiv \widehat{\Phi}_{\mathrm{mod}}(\boldsymbol{w};\ell). \tag{25}$$

Similarly, we use the same low-rank approximation $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})$ in the gradient of the modified A-optimal criterion to obtain the randomized estimator

$$\widehat{\partial_j \Phi}_{\mathrm{mod}}(\boldsymbol{w};\ell) = -\mathsf{tr}\left(\left(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})\right)^{-1}\boldsymbol{\mathcal{F}}^T\mathbf{E}_j^{\mathrm{noise}}\boldsymbol{\mathcal{F}}\left(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})\right)^{-1}\right) \tag{26}$$

for $j = 1, \ldots, n_s$.

5.2. Computational procedure and cost. Using techniques similar to those described in subsection 3.1, we can write the estimator for the modified A-optimal criterion in terms of the eigenvalues of $\mathbf{T}$:

$$\widehat{\Phi}_{\mathrm{mod}}(\boldsymbol{w};\ell) = -\mathsf{tr}\left(\mathbf{D}_{\mathbf{T}}\right). \tag{27}$$

Moreover,

$$\begin{aligned}\widehat{\partial_j \Phi}_{\mathrm{mod}}(\boldsymbol{w}) = &-\mathsf{tr}\left(\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\mathrm{noise}}\boldsymbol{\mathcal{F}}\right)+2\mathsf{tr}\left(\boldsymbol{\mathcal{F}}\mathbf{V}\mathbf{D}_{\mathbf{T}}\mathbf{V}^\top\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\mathrm{noise}}\right)\\ &-\mathsf{tr}\left(\boldsymbol{\mathcal{F}}\mathbf{V}\mathbf{D}_{\mathbf{T}}^2\mathbf{V}^\top\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\mathrm{noise}}\right)\end{aligned} \tag{28}$$

with $\mathbf{D}_{\mathbf{T}} = \boldsymbol{\Lambda}_{\mathbf{T}}(\mathbf{I}+\boldsymbol{\Lambda}_{\mathbf{T}})^{-1}$ and $\mathbf{V}$ defined as in subsection 3.1.

The procedure for computing the estimators for the modified A-optimal criterion follows the steps in Algorithm 2 closely. Instead of presenting an additional algorithm, we provide an overview of the computation of the estimators for the modified A-optimal criterion along with the associated computational cost in terms of the number of PDE solves.

To evaluate the estimators for the modified A-optimal criterion, the only precomputation we perform is to obtain

$$s_j = -\mathsf{tr}\left(\boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\mathrm{noise}}\boldsymbol{\mathcal{F}}\right), \quad j = 1, \ldots, n_s.$$

This term appears in the estimator for the gradient and accumulates a total cost of $n_s n_t$ PDE solves. The remaining terms in the estimators depend on a design $\boldsymbol{w}$ and, in particular, the eigenvalues and eigenvectors of $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})$. As with the estimators for the A-optimal criterion, the eigenvalues and eigenvectors of $\widehat{\boldsymbol{\mathcal{H}}}_{\mathrm{m}}(\boldsymbol{w})$ are obtained by Algorithm 1. Recall that the cost associated with Algorithm 1, with $q = 1$, is 4ℓ PDE solves. Once the eigenvalues and eigenvectors are computed, (25) can be evaluated without any additional PDE solves. The remaining computational effort occurs in the evaluation of the gradient. Because of our modification to the A-optimal criterion, the expression (28) is efficiently evaluated by computing $\boldsymbol{\mathcal{F}}\mathbf{V}$. Therefore, the total cost of evaluating the estimators for the modified A-optimal criterion and its gradient is 5ℓ PDE solves. From this computational cost analysis, we see the modified A-optimal estimators require ℓ less PDE solves than the A-optimal estimators.

5.3. Error analysis. We now quantify the absolute error of our estimators with the following theorem.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 5.1. *Let* $\widehat{\Phi}_{mod}(\boldsymbol{w};\ell)$ *and* $\widehat{\nabla\Phi}_{mod}(\boldsymbol{w};\ell)$ *be the randomized estimators approximating the modified A-optimal objective function* $\Phi_{mod}(\boldsymbol{w})$ *and its gradient* $\nabla\Phi_{mod}(\boldsymbol{w})$*, respectively. Using the notation and assumptions of Theorem* 3.1*, for fixed* $\boldsymbol{w} \in [0,1]^{n_s}$

$$\mathbb{E}\left[\left|\Phi_{mod}(\boldsymbol{w}) - \widehat{\Phi}_{mod}(\boldsymbol{w};\ell)\right|\right] \leq \mathsf{tr}\ (f(\boldsymbol{\Lambda}_2)) + \mathsf{tr}\ \left(f\left(\gamma_k^{2q-1} C\boldsymbol{\Lambda}_2\right)\right),$$

and for $j = 1, \ldots, n_s$,

$$\mathbb{E}\left[\left|\partial_j\Phi_{mod}(\boldsymbol{w}) - \widehat{\partial_j\Phi}_{mod}(\boldsymbol{w};\ell)\right|\right] \leq 2\|\mathbf{P}_j\|_2 \left(\mathsf{tr}\ (f(\boldsymbol{\Lambda}_2)) + \mathsf{tr}\ \left(f\left(\gamma_k^{2q-1} C\boldsymbol{\Lambda}_2\right)\right)\right),$$

where C *is defined in* (16).

Proof. See Appendix A.3. ∎

Notice the bounds presented in Theorems 3.1 and 5.1 differ by a factor of $\|\mathbf{Z}\|_2$. Since the modified A-optimal criterion removes one application of $\boldsymbol{\Gamma}_{\text{pr}}$ from the computation of the A-optimal criterion, the bounds related to the modified A-optimal criterion no longer have the factor of $\|\mathbf{Z}\|_2$.

6. Numerical results. In this section, we present numerical results that test various aspects of the proposed methods. We begin with a brief description of the inverse advection-diffusion problem used to illustrate the proposed OED methods, in subsection 6.1. The setup of our model problem is adapted from that in [2], where further details about the forward and inverse problems can be found. In subsection 6.1, we also describle the numerical methods used for solving the forward problem, as well as the optimization solver for the OED problem. In subsection 6.2, we test the accuracy of our randomized estimators and illustrate our error bounds. Then in subsection 6.3, we investigate the performance of our proposed reweighted ℓ_1-optimization approach. Next, we utilize the proposed optimization framework to compute A-optimal designs in subsection 6.4. Finally, in subsection 6.5, we compare A-optimal sensor placements with those computed by minimizing the modified A-optimal criterion.

6.1. Model problem and solvers. We consider a two-dimensional time-dependent advection-diffusion equation

$$\begin{aligned} u_t - \kappa\Delta u + \boldsymbol{v}\cdot\nabla u &= 0 && \text{in } \mathcal{D}\times(0,T),\\ u(\cdot,0) &= m && \text{in } \mathcal{D},\\ \kappa\nabla u\cdot\boldsymbol{n} &= 0 && \text{on } \partial\mathcal{D}\times(0,T), \end{aligned}$$

which models the transport of contaminants (e.g., in the atmosphere or the subsurface). Here κ is the diffusion coefficient and is taken to be $\kappa = 0.01$. The velocity field $\boldsymbol{v}$ is computed by solving a steady-state Navier–Stokes equations, as in [2]. The domain $\mathcal{D}$, depicted in Figure 3, is the unit square in $\mathbb{R}^2$ with the gray rectangles, modeling obstacles/buildings, removed. The boundary $\partial\mathcal{D}$ is the union of the outer boundary and the boundaries of the obstacles. The PDE is discretized using linear triangular continuous Galerkin finite elements in space and implicit Euler in time. We let the final simulation time be $T = 5$.

The inverse problem involves reconstructing the initial state m from space-time point measurements of $u(\boldsymbol{x},t)$. We consider $n_s = 109$ sensor candidate locations distributed throughout the domain, which are indicated by hollow squares in Figure 3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

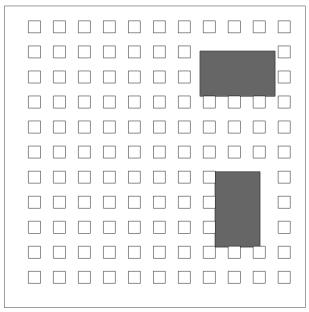

FIG. 3. *Domain with* 109 *candidate sensor locations.*

Measurement data is collected from a subset of these locations at three observation times—$t = 1, 2$, and 3.5. To simulate noisy observations, 2% noise is added to the simulated data.

Recall, in section 2, we define our prior covariance operator to be a Laplacian-like operator. Following [2], an application of the square root of the prior on $s \in L^2(\mathcal{D})$ is $v = \mathcal{A}^{-1}s$, which satisfies the following weak form:

$$\int_{\mathcal{D}} \theta \nabla v \cdot \nabla p + \alpha v p \, d\boldsymbol{x} = \int_{\mathcal{D}} sp \, d\boldsymbol{x} \qquad \text{for every } p \in H^1(\mathcal{D}). \tag{29}$$

Here θ and α control the variance and correlation length and are chosen to be $\alpha = 0.1$ and $\theta = 0.002$, respectively.

The optimization solver used for the OED problem is a quasi-Newton interior point method. Specifically, to solve each subproblem of Algorithm 3, we use the MATLAB interior point solver provided by the `fmincon` function; BFGS approximation to the Hessian is used for line search. We use a vector of all ones, $\mathbf{1} \in \mathbb{R}^{n_s}$, as the initial guess for the optimization solver.

For the numerical experiments presented in this section, the random matrix $\boldsymbol{\Omega}$ in Algorithm 1 is fixed during the optimization process. However, because of the randomness, the accuracy of the estimators, and the optimal design thus obtained, may vary with different realizations of $\boldsymbol{\Omega}$. By conducting additional numerical experiments (not reported here) the stochastic nature of the estimators resulted in minor variability (at most one or two sensor locations) in the OEDs for a modest value of ℓ. On the other hand, if ℓ is sufficiently large, we observed that the same design was obtained with different realizations of $\boldsymbol{\Omega}$

6.2. Accuracy of estimators. Here we examine the accuracy of the randomized estimators with respect to ℓ, the number of columns in the sampling matrix $\boldsymbol{\Omega}$ in the randomized subspace iteration algorithm. Specifically, we compute

$$e_1(\ell) = \frac{|\Phi_{\text{aopt}}(\boldsymbol{w}) - \widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)|}{|\Phi_{\text{aopt}}(\boldsymbol{w})|}, \quad e_2(\ell) = \frac{\|\nabla\Phi_{\text{aopt}}(\boldsymbol{w}) - \widehat{\nabla\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)\|_2}{\|\nabla\Phi_{\text{aopt}}(\boldsymbol{w})\|_2},$$
$$e_3(\ell) = \frac{|\Phi_{\text{mod}}(\boldsymbol{w}) - \widehat{\Phi}_{\text{mod}}(\boldsymbol{w};\ell)|}{|\Phi_{\text{mod}}(\boldsymbol{w})|}, \quad e_4(\ell) = \frac{\|\nabla\Phi_{\text{mod}}(\boldsymbol{w}) - \widehat{\nabla\Phi}_{\text{mod}}(\boldsymbol{w};\ell)\|_2}{\|\nabla\Phi_{\text{mod}}(\boldsymbol{w})\|_2},$$

with $\boldsymbol{w}$ taken to be a vector of all ones, that is, with all sensors activated. We let ℓ vary from 17 to 327, because the rank of $\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ is no larger than the number of

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

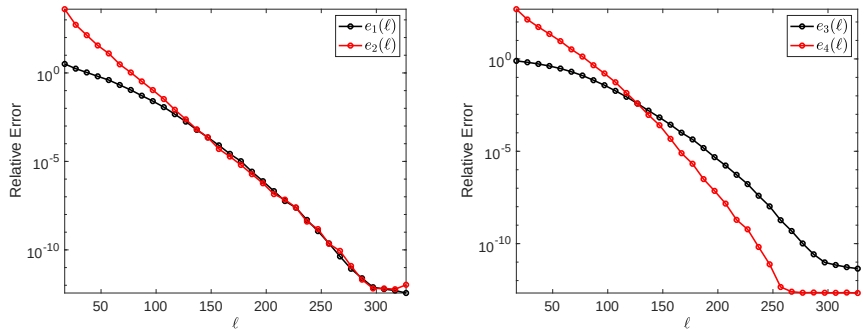

FIG. 4. *Relative error of the randomized estimators for the A-optimal criterion (left), and those corresponding to the modified A-optimal criterion (right) for varying ℓ.*

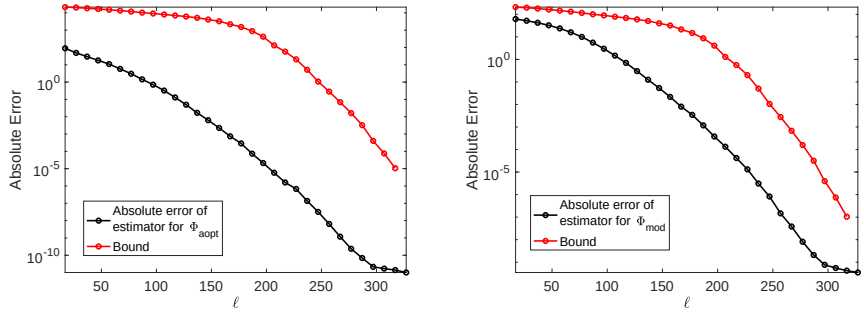

FIG. 5. *Absolute error bound for $\widehat{\Phi}_{aopt}(\boldsymbol{w})$ from Theorem 3.1 (left), and $\widehat{\Phi}_{mod}(\boldsymbol{w})$ from Theorem 5.1 (right) for varying ℓ.*

observations taken, $n_s n_t = 327$. Figure 4 illustrates the relative error in the estimators for the A-optimal criterion and its gradient (left) and the modified A-optimal criterion and its gradient (right), as ℓ is varied. We observe that the error decreases rapidly with increasing ℓ. This illustrates accuracy and efficiency of our estimators.

Next, we consider the absolute error in the estimators for the objective function and compare them with the theoretical bounds derived in Theorems 3.1 and 5.1. In Figure 5, we compare the absolute error in the estimators with bounds from Theorems 3.1 and 5.1. As before, we take $\boldsymbol{w} = [1, 1, \ldots, 1]^\top \in \mathbb{R}^{n_s}$. We observe that our error bound captures the general trend in the error. Moreover, the error bound for the modified A-optimality is better since it does not have the additional factor of $\|\mathbf{Z}\|_2$.

6.3. Performance of the reweighted ℓ_1 algorithm. We now consider solving (10) with Algorithms 2 and 3. We first consider the choice of the value of ϵ in (23). The user-defined parameter ϵ controls the steepness of the penalty function at the origin; see Figure 2. However, we observed that if the penalty function is too steep, the optimization solvers took more iterations without substantially altering the optimal designs. We found $\epsilon = 1/2^8$ to be sufficiently small in our numerical experiments, and we keep this fixed for the remainder of the numerical experiments.

We perform two numerical experiments examining the impact of changing ℓ and the penalty parameter γ (in (10)). In the first experiment, we fix the penalty coefficient at an experimentally determined value of $\gamma = 3$, and vary ℓ; the results are recorded

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

in Table 2. Notice when $\ell \geq 127$, the objective function value evaluated with the optimal solution, the number of active sensors, and number of subproblem solves do not change. This suggests that the randomized estimators are sufficiently accurate with $\ell = 127$ and this yields a substantial reduction in computational cost.

The second experiment involves varying γ, which indirectly controls the number of sensors in computed designs, with ℓ kept fixed. Here we fix $\ell = 207$, which corresponds to an accuracy on the order of 10^{-7} for the A-optimal criterion (cf. Figure 4). In Figure 6, we report the design weights sorted in descending order, as γ varies. This shows that the reweighted ℓ_1 algorithm indeed produces binary designs for a range of penalty parameters. We also notice in Figure 7 that as γ increases the sparsity increases (i.e., the number of active sensors ns_{active} decreases) and the number of function evaluations increases. The right panel compares the cost of solving an ℓ_1-penalized problem for the corresponding penalty parameter γ. Since this problem is the first iterate of the reweighted ℓ_1 algorithm we see that an additional cost is required to obtain binary designs and this cost increases with increasing γ (i.e., more sparse designs).

6.4. Computing optimal designs. In Figure 8 (left), we report an A-optimal sensor placement obtained using our optimization framework, with $\ell = 207$ and $\gamma = 5$; the resulting optimal sensor locations, with 18 active sensors, are superimposed on the posterior standard deviation field. While the design is computed to yield a minimal average variance of the posterior distribution, it is also important to consider the mean of this distribution. For completeness, in Figure 8 we show the "true" initial condition (middle panel), used to generate synthetic data, and the mean of the resulting posterior distribution for the 18 active sensor design (right panel).

TABLE 2

Number of subproblem solves (23)*, function evaluations, and active sensors for varying ℓ with the reweighted ℓ_1 algorithm and $n_s = 109$.*

ℓ	Subproblem solves	Function count	Function value	ns_{active}
57	10	395	39.0176	38
67	9	610	44.7219	30
77	9	790	44.5484	29
87	9	756	44.1142	30
127	9	1015	44.1139	30
207	9	978	44.1139	30
307	9	970	44.1139	30

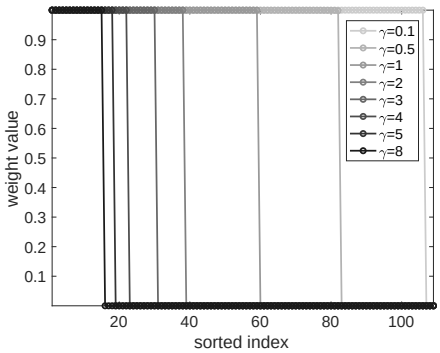

FIG. 6. *Optimal designs as a result of varying γ with the reweighted ℓ_1 algorithm for the A-optimal criterion. We set $\ell = 207$.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

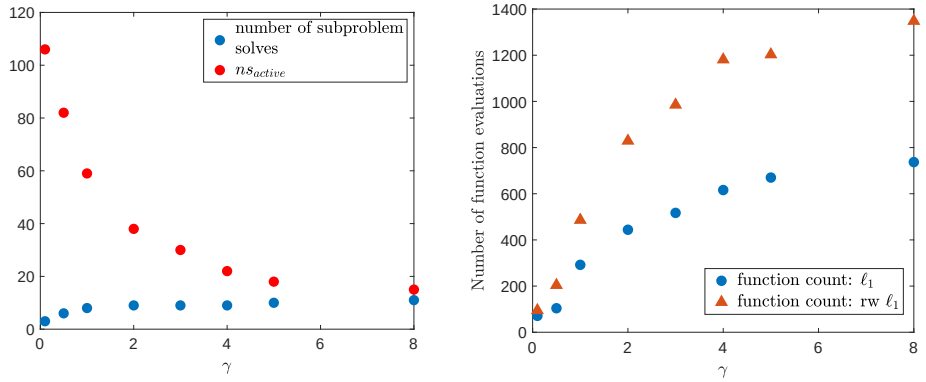

FIG. 7. *The effect of varying γ on the reweighted ℓ_1 algorithm for the A-optimal criterion. We set $\ell = 207$.*

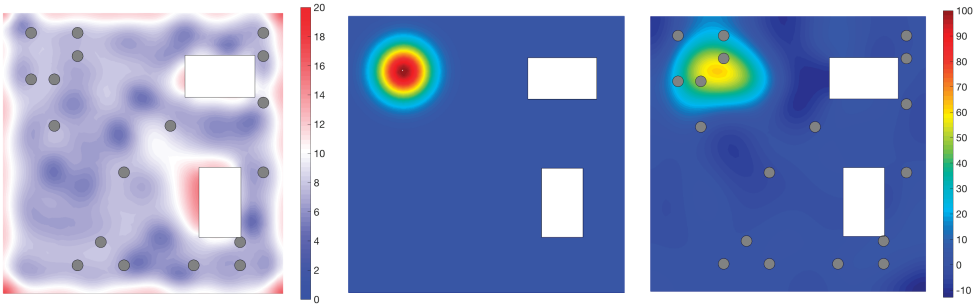

FIG. 8. *Standard deviation computed using the optimal design indicated by the gray circles (left). True initial condition (middle) and initial condition reconstruction (right). The optimal design was computed using $\ell = 207$, the reweighted ℓ_1 algorithm, and $\gamma = 5$.*

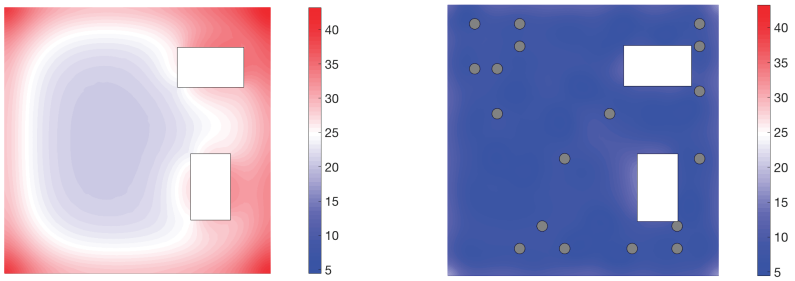

FIG. 9. *Comparison of the prior standard deviation field (left) with the posterior standard deviation field (right) computed using the optimal design indicated by the gray circles.*

With the 18 sensor design, we also illustrate the resulting uncertainty reduction by looking at the prior and posterior standard deviation fields; see Figure 9.

We now compare the designs obtained using the reweighted ℓ_1 algorithm and our estimators against designs chosen at random, illustrating the effectiveness of the proposed A-optimal design strategy. Recall that varying γ allows us to obtain optimal designs with different numbers of active sensors. For each value of γ, we use Algorithms 2 and 3 to compute an optimal design. We then draw 15 random designs

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

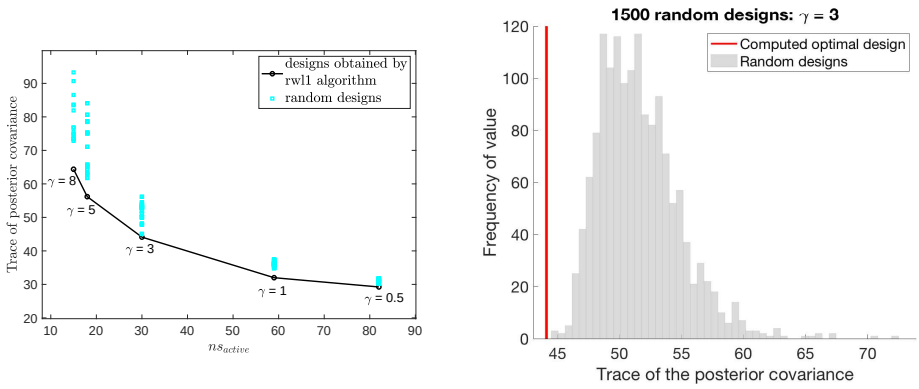

FIG. 10. *The true A-optimal criterion computed using the optimal and randomly generated designs. The optimal designs were computed using the A-optimal criterion and the reweighted ℓ_1 algorithm for different values of γ (left). Comparing the computed optimal design when $\gamma = 3$ to* 1500 *randomly generated designs (right).*

with the same number of active sensors as the optimal design obtained using our algorithms. To enable a consistent comparison, we evaluate the exact A-optimal criterion $\Phi_{\text{aopt}}(\boldsymbol{w})$ at the computed optimal designs and the random ones; the results are reported in the left panel of Figure 10. The values corresponding to the computed optimal designs are indicated as dots on the black solid line. The values obtained from the random designs are indicated by the squares. We note that the designs computed with the reweighted ℓ_1 algorithm consistently beat the random designs, as expected. This observation is emphasized in the right panel of Figure 10. Here, we compared the computed optimal design (when $\gamma = 3$) with 1500 randomly generated designs using the exact trace of the posterior covariance. Again, the computed design results in a lower true A-optimal value than the random designs.

6.5. Comparing A-optimal and modified A-optimal designs. Here we provide a quantitative comparison of sensor placements obtained by minimizing A-optimal and modified A-optimal criteria using our proposed algorithms. Specifically, for various values of γ, we solve (10) with both the A-optimal and modified A-optimal estimators to obtain two sets of designs. By varying γ, the resulting designs obtained with the A-optimal and modified A-optimal estimators have different number of active sensors. Using both sets of designs, we evaluate the exact A-optimal criterion $\Phi_{\text{aopt}}(\boldsymbol{w})$; these are displayed in Figure 11. Observe that in all cases the computed A-optimal and modified A-optimal designs lead to similar levels of average posterior variance. This suggests that the modified A-optimal criterion could be used as a surrogate for the A-optimal criterion. Using the modified A-optimal criterion decreases the overall number of PDE solves and yields designs that result in values of the average posterior variance close to those produced by A-optimal designs.

7. Conclusion. We have established an efficient and flexible computational framework for A-optimal design of experiments in large-scale Bayesian linear inverse problems. The proposed randomized estimators for the OED objective and its gradient are accurate, efficient, and simple to implement and parallelize. Specifically, the randomized estimators exploit the low-rank structure in the inverse problem; namely, the low-rank structure of the prior-preconditioned data-misfit Hessian—a common feature of ill-posed inverse problems. Our reweighted ℓ_1-minimization strategy is

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

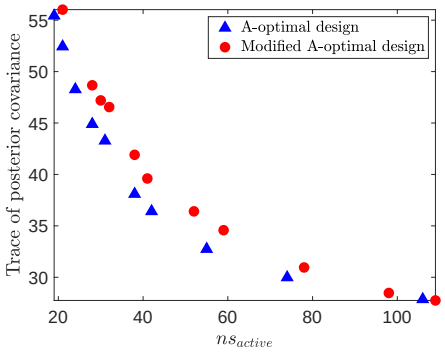

FIG. 11. *Comparison of designs obtained by minimizing the (approximate) A-optimal and modified A-optimal criteria. For each design, we report the exact trace of the corresponding posterior covariance operator.*

tailored to sensor placement problems, where finding binary optimal design vectors is desirable. We also presented the modified A-optimal criterion, which is more computationally efficient to compute and can provide designs that, while suboptimal if the goal is to compute A-optimal designs, provide a systematic means for obtaining sensor placements with small posterior uncertainty levels.

Open questions that we seek to explore in our future work include adaptive determination of the target rank k within the optimization algorithm, to further reduce computational costs, while ensuring sufficiently accurate estimates of the OED objective and gradient. Another possible line of inquiry is to use different low-rank approximations, such as Nyström's method, and extend the randomized estimators to approximate trace of matrix functions. We also seek to incorporate the randomized estimators in a suitable optimization framework for Bayesian nonlinear inverse problems, in our future work.

Appendix A. Proofs of bounds.

A.1. Trace of matrix function. In the proofs below, we use the Loewner partial ordering [15, Chapter 7.7]; we briefly recapitulate some main results that will be useful in our proof. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ be symmetric positive definite; then $\mathbf{A} \preceq \mathbf{B}$ means that $\mathbf{B} - \mathbf{A}$ is positive semidefinite. For any $\mathbf{S} \in \mathbb{R}^{n\times m}$, it also follows that $\mathbf{S}^\top \mathbf{A}\mathbf{S} \preceq \mathbf{S}^\top \mathbf{B}\mathbf{S}$. Let $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$ be the eigendecomposition of $\mathbf{A}$. Then, $f(\mathbf{A}) = \mathbf{U} f(\mathbf{\Lambda})\mathbf{U}^\top$ and $\mathsf{tr}\,(f(\mathbf{A})) = \sum_{i=1}^n f(\lambda_i)$. If f is monotonically increasing, then $\mathsf{tr}\,(f(\mathbf{A})) \leq \mathsf{tr}\,(f(\mathbf{B}))$ since $\mathbf{A} \preceq \mathbf{B}$ implies $\lambda_i(\mathbf{A}) \leq \lambda_i(\mathbf{B})$ for $i = 1, \ldots, n$.

The following bound allows us to bound the trace of a matrix function in terms of its diagonal subblocks.

LEMMA A.1. *Let*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22} \end{bmatrix}$$

be a symmetric positive definite matrix. Let f be a nonnegative concave function on $[0, \infty)$. Then

$$\mathsf{tr}\,(f(\mathbf{A})) \leq \mathsf{tr}\,(f(\mathbf{A}_{11})) + \mathsf{tr}\,(f(\mathbf{A}_{22}))\,.$$

Proof. See Theorem 2.1 and Remark 2.4 in [18].

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We are ready to state and prove our main result of this section, which is the key in proving Theorems 3.1 and 5.1.

THEOREM A.1. *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix with eigendecomposition*

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^\top = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda}_1 & \\ & \mathbf{\Lambda}_2\end{bmatrix}\begin{bmatrix}\mathbf{U}_1^\top \\ \mathbf{U}_2^\top\end{bmatrix},$$

where $\mathbf{\Lambda}_1 = \mathsf{diag}\,(\lambda_1, \ldots, \lambda_k)$ and $\mathbf{\Lambda}_2 = \mathsf{diag}\,(\lambda_{k+1}, \ldots, \lambda_n)$ contain the eigenvalues arranged in descending order. Assume that the eigenvalue ratio $\gamma_k \equiv \frac{\lambda_{k+1}}{\lambda_k} < 1$. Let k be the target rank, let $p \geq 2$ be the oversampling parameter such that $\ell \equiv k+p \leq n$, and let $q \geq 1$ be the number of subspace iterations. Furthermore, assume that $\mathbf{Q} \in \mathbb{R}^{n\times\ell}$ and $\mathbf{T} \in \mathbb{R}^{\ell\times\ell}$ are computed using Algorithm 1 and define $\widehat{\mathbf{A}} \equiv \mathbf{Q}\mathbf{T}\mathbf{Q}^\top$. Then

$$0 \leq \mathbb{E}\left[\mathsf{tr}\left(\left(\mathbf{I}+\widehat{\mathbf{A}}\right)^{-1}\right) - \mathsf{tr}\left((\mathbf{I}+\mathbf{A})^{-1}\right)\right] \leq \mathsf{tr}\left(f(\mathbf{\Lambda}_2)\right) + \mathsf{tr}\left(f\left(\gamma_k^{2q-1}C\mathbf{\Lambda}_2\right)\right),$$

where $f = x/(1+x)$, and the constant C is defined in (16).

Proof. Suppose $\mathrm{rank}(\mathbf{A}) = r$. Then, $\mathbf{A}$ has at most r nonzero eigenvalues, and thus, we can define $\mathbf{\Lambda}_{r-k} = \mathrm{diag}(\lambda_{k+1}, \ldots, \lambda_r)$ so that

$$\mathbf{\Lambda}_2 = \begin{bmatrix}\mathbf{\Lambda}_{r-k} & \\ & \mathbf{0}_{n-r-k}\end{bmatrix}. \tag{30}$$

We split this proof into several steps.

Step 0: *Lower bound.* Let $\tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_\ell \geq 0$ be the eigenvalues of $\mathbf{T}$ (and also $\widehat{\mathbf{A}}$). By the Cauchy interlacing theorem (see [20, Lemma 1] for the specific version of the argument), $\lambda_i \geq \tilde{\lambda}_i$ for $i = 1, \ldots, \ell$. Using properties of the trace operator

$$\begin{aligned}\mathsf{tr}\left((\mathbf{I}+\widehat{\mathbf{A}})^{-1}\right) - \mathsf{tr}\left((\mathbf{I}+\mathbf{A})^{-1}\right) &= \sum_{i=1}^{\ell}\frac{1}{1+\tilde{\lambda}_i} + (n-\ell) - \sum_{i=1}^{n}\frac{1}{1+\lambda_i} \\ &= \sum_{i=1}^{\ell}\frac{\lambda_i - \tilde{\lambda}_i}{(1+\lambda_i)(1+\tilde{\lambda}_i)} + \sum_{i=\ell+1}^{n}\frac{\lambda_i}{1+\lambda_i}.\end{aligned}$$

Since each term in the summation is nonnegative, the lower bound follows.

Step 1. *Trace of matrix function.* We first write $\widehat{\mathbf{A}} = \mathbf{Q}\mathbf{T}\mathbf{Q}^\top = \mathbf{Q}\mathbf{Q}^\top\mathbf{A}\mathbf{Q}\mathbf{Q}^\top = \mathbf{P}_\mathbf{Q}\mathbf{A}\mathbf{P}_\mathbf{Q}$, where $\mathbf{P}_\mathbf{Q} = \mathbf{Q}\mathbf{Q}^\top$ is an orthogonal projection matrix onto the range of $\mathbf{Q}$. Since $\widehat{\mathbf{A}}$ has the same eigenvalues as $\mathbf{R} \equiv \mathbf{A}^{1/2}\mathbf{P}_\mathbf{Q}\mathbf{A}^{1/2}$ [15, Theorem 1.3.22],

$$\mathsf{tr}\left((\mathbf{I}+\widehat{\mathbf{A}})^{-1}\right) = \mathsf{tr}\left(\left(\mathbf{I}+\mathbf{A}^{1/2}\mathbf{P}_\mathbf{Q}\mathbf{A}^{1/2}\right)^{-1}\right). \tag{31}$$

Also, since $\mathbf{P}_\mathbf{Q} \preceq \mathbf{I}$, it follows that $\mathbf{R} = \mathbf{A}^{1/2}\mathbf{P}_\mathbf{Q}\mathbf{A}^{1/2} \preceq \mathbf{A}$ and $\mathbf{0} \preceq \mathbf{A}-\mathbf{R}$. Therefore, from the proof of [5, Lemma X.1.4] and (31),

$$\mathsf{tr}\left((\mathbf{I}+\widehat{\mathbf{A}})^{-1}\right) - \mathsf{tr}\left((\mathbf{I}+\mathbf{A})^{-1}\right) \leq \mathsf{tr}\left(\left(\mathbf{I}-(\mathbf{I}+\mathbf{A}-\mathbf{R})^{-1}\right)\right) = \mathsf{tr}\left(f(\mathbf{A}-\mathbf{R})\right),$$

where $f(x)$ was defined in the statement of the theorem.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Step 2. *Reducing the dimensionality.* Let $\mathbf{F_S} \equiv \mathbf{\Lambda}_2^q \mathbf{\Omega}_2 \mathbf{\Omega}_1^\dagger \mathbf{\Lambda}_1^{-q}$. In Algorithm 1, we compute $\mathbf{Y} = \mathbf{A}^q \mathbf{\Omega}$ and let $\mathbf{Y} = \mathbf{Q}\mathbf{R_Y}$ be the thin QR factorization of $\mathbf{Y}$. Let $\mathbf{W_Q} = \mathbf{R_Y}\mathbf{\Omega}_1^\dagger \mathbf{\Lambda}_1^{-q}(\mathbf{I} + \mathbf{F_S}^\top \mathbf{F_S})^{-1/2} \in \mathbb{R}^{\ell \times k}$ be defined as in the proof of [20, Theorem 6]. It was also shown that $\mathbf{W_Q}$ has orthonormal columns, so that

$$\mathbf{Q}\mathbf{W_Q} = \mathbf{U}\begin{bmatrix}\mathbf{I}\\ \mathbf{F_S}\end{bmatrix}\left(\mathbf{I} + \mathbf{F_S}^\top \mathbf{F_S}\right)^{-1/2} \in \mathbb{R}^{n\times k}$$

has orthonormal columns. The following sequence of identities also hold: $\mathbf{W_Q}\mathbf{W_Q}^\top \preceq \mathbf{I}$, $\mathbf{Q}\mathbf{W_Q}\mathbf{W_Q}^\top\mathbf{Q}^\top \preceq \mathbf{Q}\mathbf{Q}^\top$, and

$$\mathbf{A} - \mathbf{R} \preceq \mathbf{A} - \mathbf{A}^{1/2}\mathbf{Q}\mathbf{W_Q}\mathbf{W_Q}^\top\mathbf{Q}^\top\mathbf{A}^{1/2} \equiv \mathbf{S}.$$

Since $f(x)$ is a monotonic increasing function, $\mathsf{tr}\ (f(\mathbf{A}-\mathbf{R})) \le \mathsf{tr}\ (f(\mathbf{S}))$.

Step 3. *Split into the diagonal blocks.* We can rewrite $\mathbf{S}$ as

$$\mathbf{S} = \mathbf{U}\begin{bmatrix}\mathbf{S}_1 & *\\ * & \mathbf{S}_2\end{bmatrix}\mathbf{U}^\top,$$

where $*$ represents blocks that can be ignored and

$$\mathbf{S}_1 \equiv \mathbf{\Lambda}_1^{1/2}\left(\mathbf{I} - \left(\mathbf{I} + \mathbf{F_S}^\top\mathbf{F_S}\right)^{-1}\right)\mathbf{\Lambda}_1^{1/2}, \qquad \mathbf{S}_2 = \mathbf{\Lambda}_2^{1/2}\left(\mathbf{I} - \mathbf{F_S}\left(\mathbf{I} + \mathbf{F_S}^\top\mathbf{F_S}\right)^{-1}\mathbf{F_S}^\top\right)\mathbf{\Lambda}_2^{1/2}.$$

We can invoke Lemma A.1, since $f(x) = x/(1+x)$ is concave and nonnegative on $[0,\infty)$. Therefore, we have

$$\mathsf{tr}\ (f(\mathbf{S})) \le \mathsf{tr}\ (f(\mathbf{S}_1)) + \mathsf{tr}\ (f(\mathbf{S}_2)).$$

Note that the matrix $\mathbf{U}$ disappears, because the trace is unitarily invariant.

Step 4. *Completing the structural bound.* Using an SVD-based argument it can be shown that $\mathbf{I} - (\mathbf{I} + \mathbf{F_S}^\top\mathbf{F_S})^{-1} \preceq \mathbf{F_S}^\top\mathbf{F_S}$, so that

$$\mathbf{S}_1 = \mathbf{\Lambda}_1^{1/2}\left(\mathbf{I} - \left(\mathbf{I} + \mathbf{F_S}^\top\mathbf{F_S}\right)^{-1}\right)\mathbf{\Lambda}_1^{1/2} \preceq \mathbf{\Lambda}_1^{1/2}\mathbf{F_S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}.$$

Therefore, since f is monotonically increasing

$$\mathsf{tr}\ (f(\mathbf{S}_1)) \le \mathsf{tr}\ \left(f\left(\mathbf{\Lambda}_1^{1/2}\mathbf{F_S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}\right)\right) = \sum_{j=1}^{k} f\left(\lambda_j\left[\mathbf{\Lambda}_1^{1/2}\mathbf{F_S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}\right]\right). \tag{32}$$

Note that $\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}$ is $(n-k)\times k$ has at most $\min\{n-k,k\}$ nonzero singular values. Looking more into the structure of $\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}$, and using (30), we can write

$$\begin{aligned}\mathbf{F_S}\mathbf{\Lambda}_1^{1/2} &= \mathbf{\Lambda}_2^{1/2}\begin{bmatrix}\mathbf{\Lambda}_{r-k}^{q-1/2} & \\ & \mathbf{0}\end{bmatrix}\mathbf{\Omega}_2\mathbf{\Omega}_1^\dagger\mathbf{\Lambda}_1^{-q+1/2}\\ &= \mathbf{\Lambda}_2^{1/2}\begin{bmatrix}\mathbf{\Lambda}_{r-k}^{q-1/2}\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\mathbf{\Lambda}_1^{-q+1/2}\\ \mathbf{0}\end{bmatrix},\end{aligned}$$

where $\widehat{\mathbf{\Omega}}_2 \in \mathbb{R}^{(r-k)\times(k+p)}$ such that $\mathbf{\Omega}_2 = \left[\begin{smallmatrix}\widehat{\mathbf{\Omega}}_2\\ *\end{smallmatrix}\right]$. Using the multiplicative singular value inequalities [15, equation (7.3.14)] and repeated use of the submultiplicative inequality give

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$$\sigma_j\left(\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}\right) \le \gamma_k^{q-1/2}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2\sigma_j\left(\mathbf{\Lambda}_2^{1/2}\right), \qquad j=1,\dots,\min\{k,n-k\}. \tag{33}$$

The analysis splits into two cases:

Case 1*:* $k \le n-k$. Since f is monotonically increasing, using (32) and (33)

$$\begin{aligned}\mathsf{tr}\,(f(\mathbf{S}_1)) &\le \sum_{j=1}^{k} f\left(\lambda_j\left[\mathbf{\Lambda}_1^{1/2}\mathbf{F}_\mathbf{S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}\right]\right) \le \sum_{j=1}^{k} f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\sigma_j^2\left(\mathbf{\Lambda}_2^{1/2}\right)\right)\\ &\le \sum_{j=1}^{n-k} f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\sigma_j^2\left(\mathbf{\Lambda}_2^{1/2}\right)\right) = \mathsf{tr}\left(f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\mathbf{\Lambda}_2\right)\right).\end{aligned}$$

Case 2*:* $k > n-k$. Since $\mathbf{\Lambda}_1^{1/2}\mathbf{F}_\mathbf{S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}$ has at most $n-k$ nonzero eigenvalues, use the fact that $f(0)=0$, along with (32) and (33), to obtain

$$\begin{aligned}\mathsf{tr}\,(f(\mathbf{S}_1)) &\le \sum_{j=1}^{k} f\left(\lambda_j\left[\mathbf{\Lambda}_1^{1/2}\mathbf{F}_\mathbf{S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}\right]\right) = \sum_{j=1}^{n-k} f\left(\lambda_j\left[\mathbf{\Lambda}_1^{1/2}\mathbf{F}_\mathbf{S}^\top\mathbf{F_S}\mathbf{\Lambda}_1^{1/2}\right]\right)\\ &\le \sum_{j=1}^{n-k} f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\sigma_j^2\left(\mathbf{\Lambda}_2^{1/2}\right)\right) = \mathsf{tr}\left(f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\mathbf{\Lambda}_2\right)\right).\end{aligned}$$

To summarize, in both cases $\mathsf{tr}\,(f(\mathbf{S}_1)) \le (f(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\mathbf{\Lambda}_2))$. Similarly, since $\mathbf{0} \preceq \mathbf{F_S}(\mathbf{I}+\mathbf{F}_\mathbf{S}^\top\mathbf{F_S})^{-1}\mathbf{F}_\mathbf{S}^\top$, we can show

$$\mathbf{S}_2 = \mathbf{\Lambda}_2^{1/2}\left(\mathbf{I}-\mathbf{F_S}\left(\mathbf{I}+\mathbf{F}_\mathbf{S}^\top\mathbf{F_S}\right)^{-1}\mathbf{F}_\mathbf{S}^\top\right)\mathbf{\Lambda}_2^{1/2} \preceq \mathbf{\Lambda}_2,$$

so that $\mathsf{tr}\,(f(\mathbf{S}_2)) \le \mathsf{tr}\,(f(\mathbf{\Lambda}_2))$. Combine with Step 3 to obtain

$$\mathsf{tr}\,(f(\mathbf{S})) \le \mathsf{tr}\,(f(\mathbf{\Lambda}_2)) + \mathsf{tr}\left(f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\mathbf{\Lambda}_2\right)\right).$$

Combine this with the results of Steps 1 and 2 to obtain the structural bound

$$\mathsf{tr}\left(\left(\mathbf{I}+\widehat{\mathbf{A}}\right)^{-1}\right) - \mathsf{tr}\left((\mathbf{I}+\mathbf{A})^{-1}\right) \le \mathsf{tr}\left(f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\mathbf{\Lambda}_2\right)\right) + \mathsf{tr}\,(f(\mathbf{\Lambda}_2)).$$

Step 5*. The expectation bound.* Note that $\widehat{\mathbf{\Omega}}_2 \in \mathbb{R}^{(r-k)\times(k+p)}$ and $\mathbf{\Omega}_1 \in \mathbb{R}^{k\times(k+p)}$. From the proof of [20, Theorem 1], we have $\mathbb{E}\,[\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2] \le C$, where C was defined in (16). By Jensen's inequality, using the fact that $f(x) = x/(1+x)$ is concave on $[0,\infty)$ we have

$$\begin{aligned}\mathbb{E}\left[\mathsf{tr}\left(\left(\mathbf{I}+\widehat{\mathbf{A}}\right)^{-1}\right) - \mathsf{tr}\left((\mathbf{I}+\mathbf{A})^{-1}\right)\right] &\le \mathsf{tr}\,(f(\mathbf{\Lambda}_2)) + \mathbb{E}\left[\mathsf{tr}\left(f\left(\gamma_k^{2q-1}\|\widehat{\mathbf{\Omega}}_2\mathbf{\Omega}_1^\dagger\|_2^2\mathbf{\Lambda}_2\right)\right)\right]\\ &\le \mathsf{tr}\,(f(\mathbf{\Lambda}_2)) + \mathsf{tr}\left(f\left(\gamma_k^{2q-1}C\mathbf{\Lambda}_2\right)\right).\end{aligned}$$

Combining this with the lower bound (Step 0) completes the proof. ∎

A.2. Proof of Theorem 3.1. For the remaining discussion, recall the notation from Theorem 3.1

$$\mathbf{P}_j = \boldsymbol{\mathcal{F}}^\top\mathbf{E}_j^{\text{noise}}\boldsymbol{\mathcal{F}}, \tag{34}$$

where $\boldsymbol{\mathcal{F}}$ and $\mathbf{E}_j^{\text{noise}}$ are defined in (3) and (6), respectively. We will also need

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

LEMMA A.2 (see [3]). *Let* $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ *and let* $\mathbf{B}$ *be a symmetric positive semi-definite matrix. Then, we have* $|\mathsf{tr}\,(\mathbf{AB})| \leq \|\mathbf{A}\|_2 \mathsf{tr}\,(\mathbf{B})$.

Theorem 3.1. Recall our estimator $\widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$ from (11). For fixed ℓ, using Lemma A.2 we have

$$\begin{aligned}\mathbb{E}|\Phi_{\text{aopt}}(\boldsymbol{w}) - \widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)| &= \mathbb{E}\left|\mathsf{tr}\,\left((\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{Z} - (\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{Z}\right)\right| \\ &\leq \|\mathbf{Z}\|_2 \mathbb{E}\left|\mathsf{tr}\,\left((\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1} - (\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\right)\right|.\end{aligned}$$

Applying Theorem A.1 establishes (17).

Next, we consider (18). Recall the estimator $\widehat{\partial_j \Phi}_{\text{aopt}}(\boldsymbol{w};\ell)$ from (12). We can write the absolute error as

$$\begin{aligned}&|\partial_j \Phi_{\text{aopt}}(\boldsymbol{w}) - \widehat{\partial_j \Phi}_{\text{aopt}}(\boldsymbol{w};\ell)| \\ &= \left|\mathsf{tr}\,\left(((\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{P}_j(\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1} - (\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{P}_j(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1})\mathbf{Z}\right)\right|,\end{aligned}$$

where $\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}) = \mathbf{Q}\mathbf{T}\mathbf{Q}^{\top}$. We use the decomposition

$$\begin{aligned}&\left((\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{P}_j(\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1} - (\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{P}_j(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\right)\mathbf{Z} \\ &\qquad = -\left(\mathbf{D}\mathbf{P}_j(\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1} + (\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\mathbf{P}_j\mathbf{D}\right)\mathbf{Z},\end{aligned}$$

where $\mathbf{D} \equiv (\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1} - (\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}$. Repeated application of Lemma A.2 gives

$$\begin{aligned}|\partial_j \Phi_{\text{aopt}}(\boldsymbol{w}) - \partial_j \widehat{\Phi}_{\text{aopt}}(\boldsymbol{w};\ell)| \leq \|\mathbf{P}_j\|_2\|\mathbf{Z}\|_2\Big(&\|(\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\|_2 \\ &+ \|(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\|_2\Big)\mathsf{tr}\,(\mathbf{D}).\end{aligned}$$

Since $\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w})$ and $\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w})$ have eigenvalues greater than or equal to one, $\|(\mathbf{I}+\boldsymbol{\mathcal{H}}_{\text{m}}(\boldsymbol{w}))^{-1}\|_2 + \|(\mathbf{I}+\widehat{\boldsymbol{\mathcal{H}}}_{\text{m}}(\boldsymbol{w}))^{-1}\|_2 \leq 2$. Finally, taking the expectation and applying Theorem A.1, we have the desired result. ∎

A.3. Proof of Theorem 5.1.

Proof. The proof follows along similar lines as the proof of Theorem 3.1 except the fact that in Theorem 5.1 we do not have $\mathbf{Z}$ in the expressions. ∎

Acknowledgment. We would like to thank Eric Chi for useful discussions regarding the MM algorithm.

REFERENCES

[1] A. ALEXANDERIAN, P. J. GLOOR, O. GHATTAS, ET AL., *On Bayesian A-and D-optimal experimental designs in infinite dimensions*, Bayesian Anal., 11 (2016), pp. 671–695.

[2] A. ALEXANDERIAN, N. PETRA, G. STADLER, AND O. GHATTAS, *A-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems with regularized ℓ_0-sparsification*, SIAM J. Sci. Comput., 36 (2014), pp. pp. A2122–A2148, https://doi.org/10.1137/130933381.

[3] A. ALEXANDERIAN AND A. K. SAIBABA, *Efficient D-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems*, SIAM J. Sci. Comput., 24 (2018), pp. 2956–2985.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[4] A. Attia, A. Alexanderian, and A. K. Saibaba, *Goal-oriented optimal design of experiments for large-scale Bayesian linear inverse problems*, Inverse Problems, 34 (2018), 095009.
[5] R. Bhatia, *Matrix Analysis*, Grad. Texts in Math., Springer, New York, 1997, https://doi.org/10.1007/978-1-4612-0653-8.
[6] T. Bui-Thanh, O. Ghattas, J. Martin, and G. Stadler, *A computational framework for infinite-dimensional Bayesian inverse problems Part* I*: The linearized case, with application to global seismic inversion*, SIAM J. Sci. Comput., 35 (2013), pp. A2494–A2523, https://doi.org/10.1137/12089586X.
[7] E. J. Candès, M. B. Wakin, and S. P. Boyd, *Enhancing sparsity by reweighted ℓ_1 minimization*, J. Fourier Anal. Appl., 14 (2008), pp. 877–905.
[8] K. Chaloner and I. Verdinelli, *Bayesian experimental design: A review*, Statist. Sci., 10 (1995), pp. 273–304.
[9] M. Dashti and A. M. Stuart, *The Bayesian approach to inverse problems*, in Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer, New York, 2017.
[10] J. Fohring and E. Haber, *Adaptive A-optimal experimental design for linear dynamical systems*, SIAM/ASA J. Uncertain. Quantif., 4 (2016), pp. 1138–1159, https://doi.org/10.1137/15M1034738.
[11] E. Haber, L. Horesh, and L. Tenorio, *Numerical methods for experimental design of large-scale linear ill-posed inverse problems*, Inverse Problems, 24 (2008), pp. 125–137.
[12] E. Haber, L. Horesh, and L. Tenorio, *Numerical methods for the design of large-scale nonlinear discrete ill-posed inverse problems*, Inverse Problems, 26 (2010), 025002, http://stacks.iop.org/0266-5611/26/i=2/a=025002.
[13] E. Haber, Z. Magnant, C. Lucero, and L. Tenorio, *Numerical methods for A-optimal designs with a sparsity constraint for ill-posed inverse problems*, Comput. Optim. Appl., 52 (2012), pp. 293–314.
[14] N. Halko, P.-G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288.
[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2013.
[16] D. R. Hunter and K. Lange, *A tutorial on MM algorithms*, Amer. Statist., 58 (2004), pp. 30–37, https://doi.org/10.1198/0003130042836.
[17] K. Lange, *MM Optimization Algorithms*, Other Titles Appl. Math. 147, SIAM, Philadelphia, 2016.
[18] E.-Y. Lee, *Extension of Rotfel'd theorem*, Linear Algebra Appl., 435 (2011), pp. 735–741.
[19] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Other Titles Appl. Math., 71, SIAM, Philadelphia, 2000, https://books.google.com/books?id=Brpp1Cvzs14C.
[20] A. K. Saibaba, A. Alexanderian, and I. C. F. Ipsen, *Randomized matrix-free trace and log-determinant estimators*, Numer. Math., 137 (2017), pp. 353–395.
[21] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numeri., 19 (2010), pp. 451–559.
[22] D. Uciński, *Optimal Measurement Methods for Distributed Parameter System Identification*, CRC Press, Boca Raton, FL, 2005.
[23] J. Yu, V. M. Zavala, and M. Anitescu, *A scalable design of experiments framework for optimal sensor placement*, J. Process Control, 67 (2018), pp. 44–55.

Downloaded 07/13/20 to 138.195.214.230. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.