Bayesian Model Calibration on Active Subspaces*

Allison L. Lewis[1], Ralph C. Smith[2], and Brian J. Williams[3]

***Abstract*— For many applications, the calibration of model parameters is complicated by the fact that parameters are often unidentifiable in the sense that they are not uniquely determined by the data. Moreover, for problems with moderate parameter dimensions – e.g., 8-50 parameters – deterministic, frequentist, and Bayesian model calibration techniques will often stall or fail to converge. To address both problems, we consider model calibration on active subspaces comprised of linear combinations of parameters. We first discuss the construction of active subspaces for cases when gradients may or may not be available and note how one can compute global sensitivity indices. We then employ a Delayed Rejection Adaptive Metropolis algorithm to infer parameter distributions on the active subspace and then map these distributions back to the full space. We illustrate the technique for problems in which analytic solutions are available, discretized elliptic PDEs, and a closure relation employed in reactor design.**

I. INTRODUCTION

The inference of calibration parameters in ordinary and partial differential equation (PDE) models is often challenging for various reasons. First, the numerical complexity of PDEs often dictates that a limited number of model realizations can be achieved with computational budgets. Secondly, inputs comprised of parameters, initial and boundary conditions, or exogenous forces are often unidentifiable in the sense that they are not uniquely determined by measured data. This can occur for input dimensions as low as two or three and the difficulty is compounded if parameters are linearly or nonlinearly related. For some problems, this can be mitigated if tight prior densities are employed for Bayesian inference. Finally, input dimensions can be moderate to high – e.g., 30 to over 1000 – which can limit the effectiveness of deterministic optimization routines or frequentist or Bayesian statistical inference algorithms.

In this paper, we illustrate how Bayesian inference on active subspaces of inputs can be used to address these difficulties for certain applications. As detailed in [2], and summarized in Section II, active subspace techniques isolate the subspace of identifiable inputs, which often has significantly lower dimension than the original input set. Inference on this subspace is thus well-posed and often substantially more efficient due to the reduced dimensionality. In conjunction, one can construct a surrogate model on the active subspace that is significantly more efficient to implement than the original PDE. This addresses the challenge associated with solving computationally intensive PDEs.

Since one typically desires point estimates or distributions for the original physical inputs, rather then the transformed set, we map inference results back to the physical space. This will not yield unique results for unidentifiable parameter sets. To address this, we employ the active subspace algorithms to construct global sensitivity indices that specify which of the physical parameters will be informed during Bayesian inference. For noninfluential parameters, the prior distribution will not be informed by the likelihood comprised of the data and model.

We note that Bayesian inference on active subspaces has previously been illustrated in [5]. In that work, the authors define the active subspaces based on the gradient of the negative log-likelihood and implement a random walk Metropolis-Hastings algorithm based on the data misfit. We employ a Delayed Rejection Adaptive Metropolis (DRAM) algorithm [7], [12], which achieves more effective sampling and mixing for moderate-dimensional parameter spaces. However, the primary novelty of this paper is the use of global sensitivity indices to specify those parameters that will be informed when inference results are mapped back to the physical space.

II. ACTIVE SUBSPACE CONSTRUCTION AND GLOBAL SENSITIVITY INDICES

We illustrate active subspace construction for a general function

$$f = f(\mathbf{q}),\ \mathbf{q} = [q_1, ..., q_m],$$

where the random variable $\mathbf{q}$ has an associated probability density function $\rho : \mathbb{R}^m \to \mathbb{R}_+$. We denote the gradient of f by $\nabla_{\mathbf{q}} f(\mathbf{q}) = [\frac{\partial f}{\partial q_1} \cdots \frac{\partial f}{\partial q_m}]^T$ and use it to construct the matrix

$$\mathbf{C} = \int (\nabla_{\mathbf{q}} f)(\nabla_{\mathbf{q}} f)^T \rho \, d\mathbf{q}. \tag{1}$$

Since $\mathbf{C}$ is symmetric and positive semi-definite, it admits an eigenvalue decomposition

$$\mathbf{C} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^T, \quad \mathbf{\Lambda} = \text{diag}(\lambda_1, ..., \lambda_m),$$

where $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$ and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_m]$. We partition the eigenvalues and eigenvectors according to their split

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \\ & \mathbf{\Lambda}_2 \end{bmatrix},\ \mathbf{W} = [\mathbf{W}_1\ \mathbf{W}_2], \tag{2}$$

*This research was supported by the Consortium for Advanced Simulation of Light Water Reactors (http://www. casl.gov), an Energy Innovation Hub (http://www.energy.gov/hubs) for Modeling and Simulation of Nuclear Reactors under U.S. Department of Energy Contract No. DE-AC05-00OR22725.

[1]Department of Mathematics, North Carolina State University, Raleigh, NC 27695 lewis.allison10@gmail.com

[2]Department of Mathematics, North Carolina State University, Raleigh, NC 27695 rsmith@ncsu.edu

[3]Los Alamos National Laboratory, Los Alamos, NM 87545 brianw@lanl.gov

should one occur [2]. We then define rotated coordinates $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^{m-n}$ by

$$\mathbf{y} = \mathbf{W}_1^T \mathbf{q}, \quad \mathbf{z} = \mathbf{W}_2^T \mathbf{q},$$

noting that f varies more along $\mathbf{y}$ directions than along $\mathbf{z}$.

To implement the method, one computes the gradient vector

$$\nabla_{\mathbf{q}} f^j = \nabla_{\mathbf{q}} f(\mathbf{q}^j), \quad j = 1, ..., M,$$

at each sample point $\mathbf{q}^j$ from ρ and constructs the matrix

$$\mathbf{G} = \frac{1}{\sqrt{M}} \left[\nabla_{\mathbf{q}} f^1 \cdots \nabla_{\mathbf{q}} f^M \right]. \tag{3}$$

We note that $\mathbf{G}\mathbf{G}^T \approx \mathbf{C}$. The matrix $\mathbf{G} \in \mathbb{R}^{m \times M}$ admits a singular value decomposition $\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$. This forms the basis for the gradient-based method employed in [1]–[3]. Readers are referred to [10] for details regarding a gradient-free algorithm, which can be employed when gradient or adjoint capabilities are not available.

To perform Bayesian inference on the active subspace, we fit a polynomial surface $g(\mathbf{y})$ in the manner detailed in [2]. We summarize active subspace-based Bayesian inference in Algorithm 1. Note that throughout this investigation, we employ Gaussian prior distributions to avoid the issues associated with rotating hypercubes, which may result in marginal and conditional distributions that are no longer uniform; see [2] for details. The extension to more general prior distributions constitutes future work. We employ a Delayed Rejection Adaptive Metropolis (DRAM) algorithm [7], [12] to compute marginal and pairwise posterior distributions.

Finally, we summarize the manner in which active subspaces can be used to quantify the global sensitivity of the parameters $q_1, \cdots, q_m$. For some problems, the components $\mathbf{U}(:, 1)$ can be used to rank the significance of each component q_i of $\mathbf{q}$ since it quantifies the directions in the parameter space where perturbations have the most influence. It is illustrated in [6], however, that the activity scores $\alpha(n)$, defined by

$$\alpha(n) = \sum_{j=1}^{n} \lambda_j \mathbf{w}_j^2, \tag{4}$$

provide a more comprehensive sensitivity index. Here $\mathbf{w}_j$ is the j^{th} column of the eigenvector matrix $\mathbf{W}$ and $\mathbf{w}_j^2$ should be interpreted as squaring each component of $\mathbf{w}_j$. Those parameters that play a larger role in the important directions specified by the active subspace will be assigned larger scores, whereas parameters that are relatively insignificant will have scores close to zero.

Algorithm 1 Bayesian Inference on Active Subspaces

(1) Compute a set of sample gradient vectors $\nabla_{\mathbf{q}} f^i = \nabla_{\mathbf{q}} f(\mathbf{q}^i)$, for $i = 1, ..., M$.

(2) Compute the SVD of the gradient matrix

$$\mathbf{G} = \frac{1}{\sqrt{M}} \left[\nabla_{\mathbf{q}} f^1 \quad ... \quad \nabla_{\mathbf{q}} f^M \right] = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$$

(3) Identify the dimension n of the active subspace, either based on gaps in the singular values or through the use of the activity scores (4).

(4) Project a set of training points $\mathbf{q}^j$, $j = 1, ..., N$, into the active subspace

$$\mathbf{y}^j = \mathbf{U}(:, 1:n)^T \mathbf{q}^j.$$

Fit a multivariate polynomial response surface $g(\mathbf{y})$ to the training data $\{\mathbf{y}^j, f(\mathbf{q}^j)\}$.

(5) Calibrate the active variables $\mathbf{y}$ using DRAM. Note that we minimize the sum of squares comparing the response surface $g(\mathbf{y})$ to a set of simulated data, $\{f(\mathbf{q}_{\text{nominal}}) + \varepsilon^k\}$. Here we take $\varepsilon^k \sim N(0, 0.05 \cdot f(\mathbf{q}_{\text{nominal}}))$, unless otherwise noted, for k ranging over the number of data points.

(6) Transform the resulting chain of active variable samples back into the full space

$$\mathbf{q}^p = \mathbf{U}(:, 1:n)\mathbf{y}^p + \mathbf{U}(:, n+1:m)\mathbf{z}^p,$$

for p ranging over the length of the DRAM chain. Here the 'inactive' variables $\mathbf{z}_j, j = 1, \ldots, N$ are iid with each component sampled from the prior distribution $\mathcal{N}(0, 1)$.

III. Examples

In this section, we provide four examples to illustrate aspects of the Bayesian inference algorithm on active subspaces. The first focuses on a simple algebraic example to illustrate the algorithm performance when the gradient can be computed analytically. Examples 2 and 3 illustrate the algorithm for discretized elliptic PDEs. In Example 4, we illustrate the algorithm for the Dittus-Boelter relation, which comprises a closure relation employed in two-phase nuclear thermal-hydraulic equations.

A. Example 1: Two-Parameter Exponential Relation

Consider the function

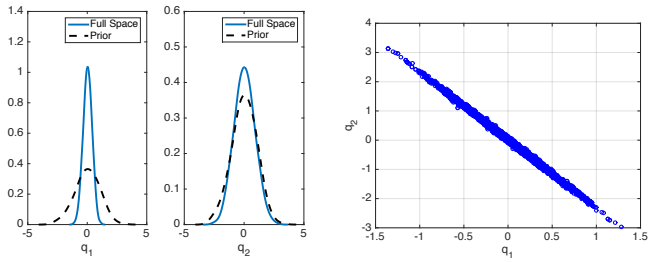
$$f(\mathbf{q}) = \exp(0.7q_1 + 0.3q_2), \tag{5}$$

with the parameter vector $\mathbf{q} = [q_1, q_2]$. The gradient vector is

$$\nabla_{\mathbf{q}} f = \begin{bmatrix} \frac{\partial f}{\partial q_1} & \frac{\partial f}{\partial q_2} \end{bmatrix}^T = \exp(0.7q_1 + 0.3q_2) \cdot [0.7 \quad 0.3]^T.$$

Computing the gradient vector repeatedly at samples $\{\mathbf{q}^i\}_{i=1}^M$ yields a gradient matrix $\mathbf{G}_{2 \times M}$ of rank one. To obtain a basis for the active subspace, we compute the singular value decomposition $\mathbf{G} = \mathbf{U}\Lambda\mathbf{V}^T$. The rank of the matrix $\mathbf{G}$ and the spectrum of singular values indicate a one-dimensional active subspace, for which $\mathbf{U}(:, 1) = [0.9191 \quad 0.3939]^T$ is a basis.

We first calibrate the model in the full two-dimensional parameter space, using synthetic data generated from the full model evaluated at the nominal parameter value $\mathbf{q} = [0, 0]$, with errors $\varepsilon_i \sim \mathcal{N}(0, 0.05)$. We employ Gaussian prior densities, $q_1, q_2 \sim \mathcal{N}(0, 1)$, when running DRAM. We note



(a) (b)

Fig. 1. (a) Posterior densities in the original parameter space, and (b) joint density for the function $f(\mathbf{q}) = \exp(0.7q_1 + 0.3q_2)$.

in Figure 1(a) that the prior density for the more influential parameter q_1 is informed by the likelihood whereas the prior density for the less significant q_2 is minimally informed. The fact that q_1 and q_2 are not jointly identifiable is illustrated by the nearly single-valued joint density in Figure1(b), which aligns with the inactive subspace.

We next employ Algorithm 1 to perform calibration in the active subspace thus reducing the dimension of our problem from two parameters to one. We plot the third-order response surface $g(\mathbf{y})$, computed in Step (4) of Algorithm 1, in Figure 2 along with test points. In Figure 3, we plot the prior densities and posterior densities computed on the 1-D active subspaces and transformed back to the physical space. We also compare the posterior densities computed on the full space. For the more significant parameter q_1, the prior is informed by the data model and two posterior densities are qualitatively close. For the less significant q_2, the prior density is minimally informed by the likelihood.

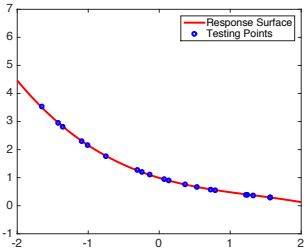


Fig. 2. Third-order univariate response surface $y(\mathbf{y})$ and test points.

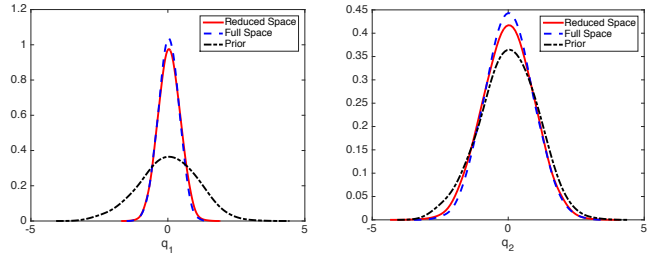


Fig. 3. Posterior densities computed in the full and reduced spaces.

B. Example 2: Elliptic PDE

Our next example is a variation of the elliptic PDE from [2]. We note that Bayesian inference on active subspaces is illustrated for this example in [5].

Let $u = u(\mathbf{s}, \mathbf{q})$ satisfy the elliptic PDE

$$-\nabla_{\mathbf{s}} \cdot (a(\mathbf{s}, \mathbf{q})\nabla_{\mathbf{s}} u(\mathbf{s}, a(\mathbf{s}, \mathbf{q}))) = 1, \ \mathbf{s} \in [0,1]^2, \tag{6}$$

with homogeneous Dirichlet boundary conditions $u = 0$ on the left, top, and bottom of the spatial domain, denoted by Γ_1, and a homogenous Neumann boundary condition $\frac{\partial u}{\partial s_1} = 0$ on the right side, denoted by Γ_2. The coefficient $a(\mathbf{s}, \mathbf{q})$ is taken to be a log-Gaussian second-order random field with mean zero and covariance function

$$\mathcal{C}(\mathbf{s}, \mathbf{s}') = \exp(\beta^{-1}||\mathbf{s} - \mathbf{s}'||_1). \tag{7}$$

The random field can be expressed in terms of the eigenvalues γ_i and orthonormal eigenfunctions ϕ_i of $\mathcal{C}$ by using a truncated Karhunen-Loeve (KL) expansion,

$$\log(a(\mathbf{s}, \mathbf{q})) = \sum_{i=1}^{m} q_i \gamma_i \phi_i. \tag{8}$$

Here q_i are independent and identically distributed (iid) standard normal random variables. For this example, we use $m = 20$; note that $m = 100$ parameters are used in the original problem.

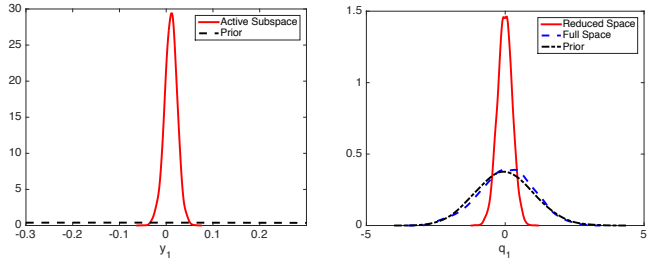
To obtain function evaluations at a given set of input parameters $\mathbf{q}$, the elliptic problem is discretized using a standard finite element method with a mesh containing 34,320 triangles and 17,361 nodes. The eigenfunctions $\phi_i = \phi_i(\mathbf{q})$ are approximated on this mesh for $i = 1, ..., N$ by solving the matrix equation $\mathbf{Ku} = \mathbf{f}$ for $\mathbf{u} = \mathbf{u}(\mathbf{q})$ at the mesh nodes. Here the stiffness matrix has elements $[\mathbf{K}]_{ij} = \int_\Omega a \nabla_{\mathbf{s}} \phi_i(\mathbf{s}) \cdot \nabla_{\mathbf{s}} \phi_j(\mathbf{s}) d\mathbf{s}$, $\mathbf{u} = [u_1, ..., u_N]^T$ and $[\mathbf{f}]_i = \int_\Omega \phi_i(\mathbf{s}) d\mathbf{s}$. The scalar response is then approximated by

$$f(\mathbf{q}) = \mathbf{c^T M u}(\mathbf{q}) \approx \frac{1}{|\Gamma_2|} \int_{\Gamma_2} u(\mathbf{s}, \mathbf{q}) d\mathbf{s}, \tag{9}$$

where $[\mathbf{M}]_{ij} = \frac{1}{|\Gamma_2|} \int_{\Gamma_2} \phi_i(\mathbf{s}) \phi_j(\mathbf{s}) d\mathbf{s}$ and the components of $\mathbf{c}$ are equal to one where they correspond to Γ_2, and zero elsewhere.

A singular value decomposition $\mathbf{G} = \mathbf{U}\Lambda\mathbf{V}^T$ of the gradient matrix yields a distinct one-dimensional active subspace, where $\mathbf{U}(:, 1)$ is a basis for the space. We generate synthetic data from the original model for use in calibration, and then calibrate the physical model in the full space and for a 2nd-order multivariate polynomial response surface in the reduced active subspace. The model calibration in the full 20-dimensional space takes approximately 38 minutes whereas the reduced space calibration required only 9.52 seconds. This demonstrates the advantage of constructing a lower-dimensional active subspace for model calibration and uncertainty quantification (UQ). We plot the posterior density for the active variable y_1 in Figure 4(a).

Based on the singular values, the parameter q_1 is considerably more significant than the other nineteen parameters. Therefore, after calibrating in the reduced space and transforming back into the original parameter space, we see that this parameter exhibits the greatest change from prior to posterior density whereas the posterior densities for all other parameters remain relatively close to their standard normal prior densities.



(a) (b)

Fig. 4. (a) Posterior density of active variable y_1. (b) Comparison of the posterior densities for q_1 for the full and reduced space calibrations.

In the full space calibration, we do not observe changes from prior to posterior density for any of the 20 variables. This is due to unidentifiability in the full parameter space and hence we are unable to inform any of the prior densities. In Figure 4(b), we display the posterior densities for q_1 for both the full and reduced space calibrations. The densities for the complete parameter set are plotted in Figure 5. This example illustrates the necessity of performing inference in the active subspace and transforming distributions back to the physical space for an unidentifiable parameter set.

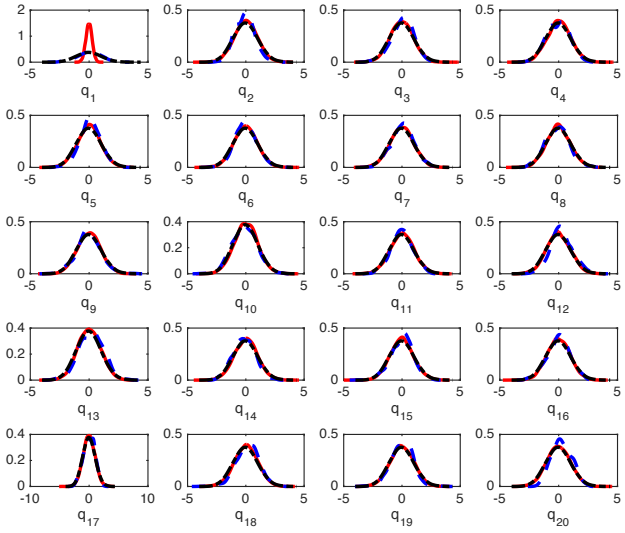


Fig. 5. Posterior densities for all 20 parameters. The prior densities are plotted in black dashed-dot lines, while the full space calibration is shown in blue dashed and the reduced active subspace calibration in solid red.

C. Example 3: Multiple Elliptic PDEs

We now modify the previous problem to illustrate calibration with an active subspace having multiple dimensions, as presented in [8]. We consider a family of PDEs

$$\begin{aligned} -\nabla_{\mathbf{s}} \cdot (a(\mathbf{s},\mathbf{q},w)\nabla_{\mathbf{s}} u(\mathbf{s},\mathbf{q},w)) &= 1, \\ \mathbf{s} \in [0,1]^2, \;\; w &= 1,...,W \end{aligned} \tag{10}$$

based upon the original problem from [2]. We employ identical boundary conditions for each PDE as in the original problem (6).

Each random field $a(\mathbf{s},\mathbf{q},w)$ is log-Gaussian with mean zero and corresponding covariance function $\mathcal{C}_w$, and is expressed by the Karhunen-Loevre expansion (8) with m_w eigenvalues and eigenvectors of the covariance matrix and m_w standard Gaussian random variables. We employ the quantity of interest

$$\begin{aligned} f(\mathbf{q}_1,...,\mathbf{q}_W) &= \sum_{w=1}^{W} \mathbf{c}^{\mathbf{T}}\mathbf{M}\mathbf{u}_w(\mathbf{x}_w) \\ &\approx \sum_{w=1}^{W} \frac{1}{|\Gamma_2|}\int_{\Gamma_2} u\left(a(\mathbf{s},\mathbf{q}_w,w)\right) d\mathbf{s}. \end{aligned} \tag{11}$$

We note that the quantity of interest is dependent upon $\sum_{w=1}^{W} m_w$ total parameters and should vary primarily along W directions. Using $W = 3$, we define the covariance functions

$$\mathcal{C} = \left(1 + \frac{||\mathbf{s}-\mathbf{s}'||_2^2}{2\beta}\right)^{\beta} \tag{12}$$

for values $\beta = \left\{\frac{2}{5}, \frac{4}{5}, \frac{6}{5}\right\}$. To compute values of f, the solution to each of the 3 PDEs is approximated on a finite element mesh of $N = 727$ nodes using the code from [3], [4], [9], [11]. Discarding eigenvalues that are less than 10^{-4} in magnitude, we obtain a total parameter space dimension of $\sum_{w=1}^{3} m_w = 91$. We note that the active subspace has dimension three due to the construction of the problem, which is supported by the singular value plot in Figure 6.

Figure 7 illustrates the comparison between the observed model responses and those predicted via a 3rd-order multivariate polynomial response surface for active subspace dimensions 1, 2, and 3 at 500 test points. We note that by

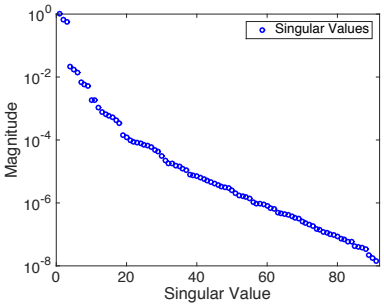


Fig. 6. Normalized singular values for the elliptic PDE (10). Note the gap between singular values 3 and 4, indicating a 3-dimensional active subspace.

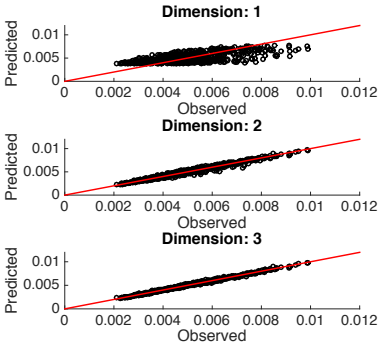


Fig. 7. Observed versus predicted outputs comparing the true quantity of interest to that predicted via the constructed response surface.

dimension 3, we observe very close agreement between the true and predicted values, once again supporting the choice of a 3-dimensional active subspace.

By construction, the 91 physical parameters are not identifiable since the model has a 3-D active subspace. Here we demonstrate how the activity scores $\alpha(n)$, defined in (4), can be used to designate which of the prior distributions for the 91 parameters will be informed via calibration on the active subspace. We also demonstrate that inference on the full 91-D space yields spurious posterior densities.

Using (4), we compute the activity scores plotted in Figure 8. The largest scores are for q_{38}, q_{66}, and q_1 with respective values of 2.84, 1.66, and 1.65. The next largest score is for q_{40} and is nearly two orders of magnitude smaller with a value of 0.04. Hence we expect that prior distributions for the sensitive parameters q_{38}, q_{66}, and q_1 will be informed during Bayesian calibration and that prior distributions for the remaining parameters will be minimally informed.

Calibration in the full 91-D spaces requires approximately 15 hours of computation time whereas calibration in the 3-D active subspace takes only 20 seconds. We plot the DRAM chains for the three most significant parameters in Figure 9. The chains constructed on the active subspace are visually well burned-in as required to construct reliable posterior densities. In contrast, the full-space chains are wandering and not converging in distribution, as is typical for unidentifiable parameters.

In Figure 10, we plot the prior and posterior densities for the three most sensitive parameters along with those for the next three most sensitive parameters q_{40}, q_{68} and q_3. As

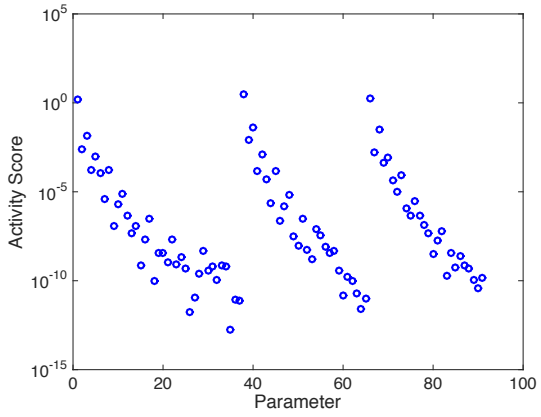
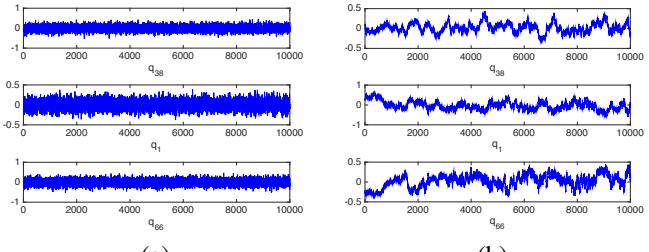


Fig. 8. Activity scores (4) for the 91 physical parameters.



(a) (b)

Fig. 9. DRAM chains for the three most significant parameters q_{38}, q_{66} and q_1 for the (a) active subspace calibration and (b) full space calibration.

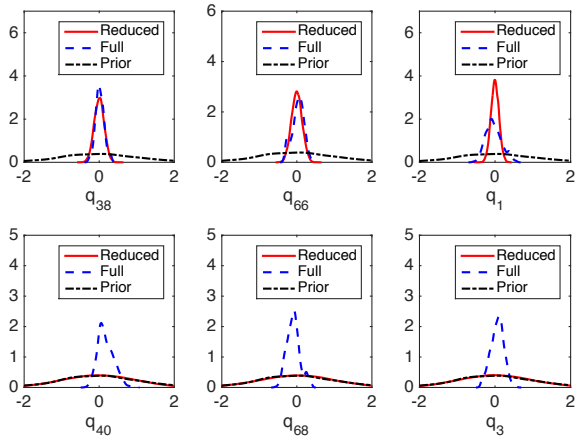


Fig. 10. Posterior densities for the six most significant parameters.

expected, calibration on the active subspace informs the prior distributions for q_{38}, q_{66}, and q_1, producing significantly narrower posterior densities but does not inform the prior distributions for the remaining 88 parameters.

In contrast, the full space calibration appears to inform all of the prior distributions. However, this contradicts the constructed dimension of the active subspace and is due to the fact that chains are not adequately converging for the full unidentifiable parameter set. This illustrates the necessity of using sensitivity analysis, based on activity scores, to ascertain which parameters will be informed and performing Bayesian inference on the active or identifiable subspace.

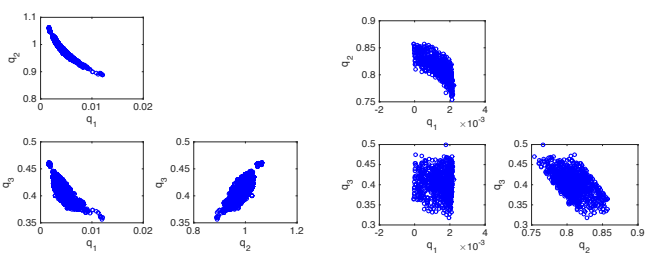
D. Example 4: Dittus-Boelter Equation

Here we consider the Dittus-Boelter equation

$$\mathrm{Nu} = q_1 \mathrm{Re}^{q_2} \mathrm{Pr}^{q_3}, \tag{13}$$

where Nu, Re and Pr respectively denote the Nusselt, Reynolds and Prandtl numbers and the parameters are $\mathbf{q} = [q_1, q_2, q_3]$. This closure relation is employed in thermal-hydraulic models used to quantify two- and three-phase flow in nuclear reactors. We employ experimental data with a sample size of $n = 56$ points for use in calibration. For calibration on the full space, we employ uniform distributions $U(0, 2\times\text{nominal})$ for the nominal values $[0.023, 0.8, 0.4]$. We employ Gaussian prior distributions when calibrating on the active subspace.

Performing calibration via the DRAM algorithm on the full parameter space reveals that the full parameter set is



(a) (b)

Fig. 11. Joint densities for the (a) full and (b) reduced space calibration.

unidentifiable; see Figure 11(a) for a joint density plot of each of the parameter pairs. We observe that a unique optimal parameter set does not exist due to the single-valued nature of the joint density plot for q_1 and q_2. This results in DRAM chains that never achieve burn-in without tight prior densities. For this reason, we exploit the active subspace for calibration.

We compute the gradient matrix, noting that each column of $\mathbf{G}$ is of the form

$$\mathbf{G}(:,j) = \begin{bmatrix} \mathrm{Re}^{q_2^j}\mathrm{Pr}^{q_3^j} \\ q_1^j \mathrm{Re}^{q_2^j}\mathrm{Pr}^{q_3^j}\log(\mathrm{Re}) \\ q_1^j \mathrm{Re}^{q_2^j}\mathrm{Pr}^{q_3^j}\log(\mathrm{Pr}) \end{bmatrix},$$

for the parameter sample $\mathbf{q}^j = [q_1^j, q_2^j, q_3^j]$. Computing the singular value decomposition of this matrix yields normalized singular values $\lambda_1 = 1$, $\lambda_2 = 0.1793$, and $\lambda_3 = 0.0060$, suggesting that a two-dimensional active subspace may be exploited for calibration.

Using a two-dimensional active subspace, we fit a 4th-order multivariate polynomial to a set of training data as illustrated in Figure 12. Performing calibration in the active subspace yields posterior densities for the active variables y_1 and y_2. The chain values for the inactive variable y_3 are sampled from a standard Gaussian prior distribution and all three variables are then transformed back into the original parameter space. Figure 13 shows a comparison of the posterior densities between the full and reduced space calibrations. The final computed mean parameter values obtained via the reduced space calibration were $\mathbf{q} = [0.0004, 0.8070, 0.4002]$ as compared to $\mathbf{q} = [0.0043, 0.9828, 0.4084]$ obtained via the full space calibration. See Figure 13 for a comparison of the posterior densities for both methods. With the exception of the first parameter, the mean values obtained via the active subspace are qualitatively very close to the nominal values $\mathbf{q} = [0.023, 0.8, 0.4]$. We also note that calibrating in the reduced space has resulted in an identifiable parameter set, as illustrated in Figure 11(b).

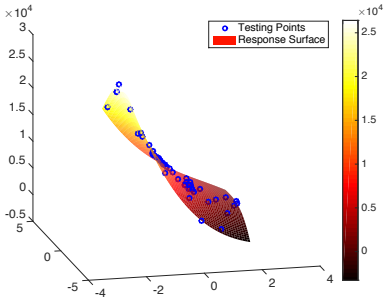


Fig. 12. 4th-order multivariate polynomial response surface and test points.

IV. Concluding Remarks

The results in Examples 2-4 demonstrate that Bayesian inference, with diffuse prior distributions, can produce erroneous posterior distributions when implemented on the full space of unidentifiable parameters. This is due to the fact that chains do not adequately mix or sample the full parameter space, as illustrated in Figure 9(b). For Example 2, this was manifested by the fact that the likelihood did not significantly influence any of the marginal distributions whereas for Example 3, it produced incorrect marginal distributions for unidentifiable parameters. Inference on the active subspace, followed by mapping back to the physical space, provides a well-posed and highly efficient technique to infer influential parameter distributions.

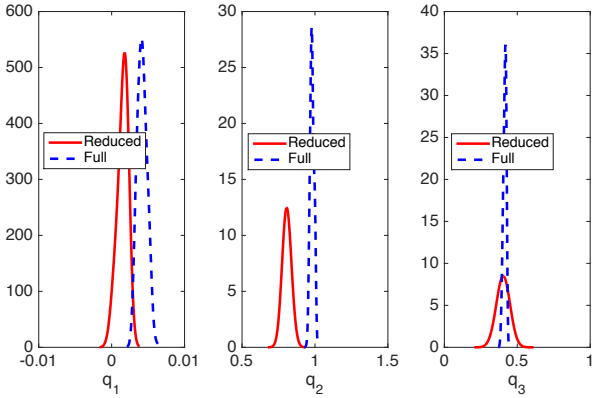


Fig. 13. Posterior densities for the full and reduced space calibrations.

As detailed in [5], Bayesian inference on the active subspace can introduce a bias and the authors provide a bound on the Hellinger distance between the true posterior distribution and that computed on the active subspace. In future work, we will investigate the use of the negative log-likelihood to determine the active subspace and employ the theory in [5] to quantify potential biases. We will also illustrate the use of these inference techniques for motivating applications arising in nuclear engineering.

References

[1] Y. Bang, H.S. Abdel-Khalik, and J.M. Hite, "Hybrid reduced order modeling applied to nonlinear models," *International Journal for Numerical Methods in Engineering*, 91, pp. 929-949, 2012.

[2] P.G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, SIAM Spotlights, Philadelphia, PA, 2015.

[3] P.G. Constantine, MATLAB code, https://bitbucket.org/paulcon/active-subspace-methods-in-theory-and-practice.

[4] P.G. Constantine, "Random Field Simulation," MATLAB code, http://www.mathworks.com/matlabcentral/fileexchange/27613-random-field-simulation.

[5] P.G. Constantine, C. Kent and T. Bui-Thanh, "Accelerating MCMC with active subspaces," *SIAM Journal on Scientific Computing*, 38(5), A2779–A2805, 2016.

[6] P. Diaz and P.G. Constantine, "Global sensitivity metrics from active subspaces," arXiv:1510.04361v2.

[7] H. Haario, M. Laine, A. Mira and E. Saksman, "DRAM: Efficient adaptive MCMC," *Statistics and Computing*, 16(4), pp. 339–354, 2006.

[8] J.T. Holodnak, *Topics in Randomized Algorithms for Numerical Linear Algebra*, Ph.D. Thesis, 2015, http://www.lib.ncsu.edu/resolver/1840.16/10327.

[9] A. Klimke, MATLAB Sparse Grid Interpolation Toolbox, http://www.ians.uni-stuttgart.de/ spinterp/.

[10] A. Lewis, R.C. Smith, and B. Williams, "Gradient-free active subspace construction using Morris screening elementary effects," *Computers and Mathematics with Applications*, 72, pp. 1603–1615, 2016.

[11] Mathworks, MATLAB Partial Differential Equations Toolbox, http://www.math works.com/ products/pde/.

[12] R.C. Smith, *Uncertainty Quantification: Theory, Implementation and Applications*, SIAM, Philadelphia, 2014.