

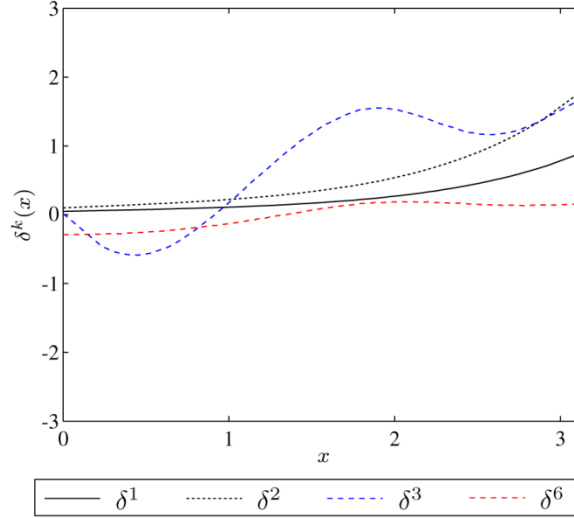
**Online Supplement for "Preposterior Analysis to Predict Identifiability in Experimental Calibration of Computer Models," by Paul D. Arendt, Daniel W. Apley, and Wei Chen**

**Sensitivity to the Actual Discrepancy Function and to the Choice of  $\sigma_\delta^2$**

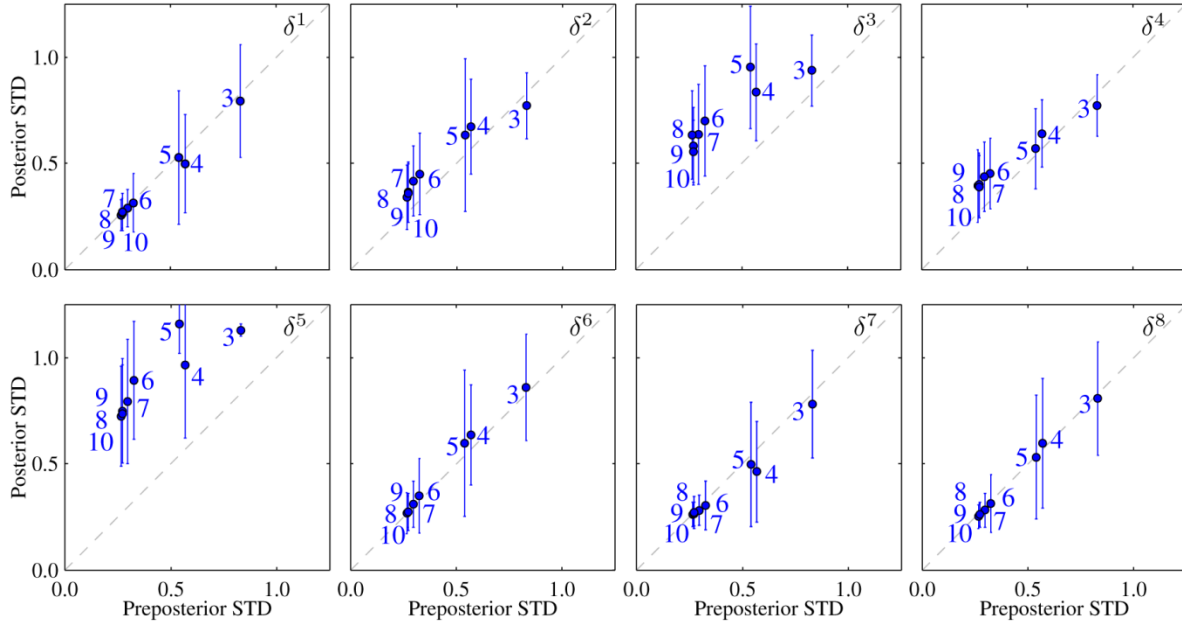
In the MC loop of the algorithm of Figure 1, the hypothetical discrepancy functions are generated from a GP model. In reality, however, the actual discrepancy functions can take whatever form the physics of the problem dictate, and we used non-GP discrepancy functions in the examples in Sections 5.1 and 5.2. Here we investigate the sensitivity of the posterior covariance to the actual discrepancy function for the sinusoidal example of Section 5.2. Specifically, we generate different sets of actual physical experimental data via Eq. (5) with the eight different discrepancy functions  $\{\delta^k(x): k = 1, 2, \dots, 8\}$  listed in Table S1, some of which are plotted in Figure S1. The first two discrepancy functions are mathematical functions with different magnitudes but similar shape, and the remaining discrepancy functions are realizations from GP models with the hyperparameters shown in Table S1. Figure S2 shows results analogous to Figure 7 for the eight different discrepancy functions. Most of the scatter plots in Figure S2 exhibit a strong linear trend, which indicates that the preposterior STD provides a reasonable indication of the relative improvement in identifiability that one can expect from running a larger physical experiment.

**Table S1.** Different discrepancy functions used in the sinusoidal example.

Discrepancy function	Description
$\delta^1(x)$	$0.25(0.1\exp\{x\} - 0.05x^2)$
$\delta^2(x)$	$0.1\exp\{x\} - 0.05x^2$
$\delta^3(x)$	Each $\delta^k(x)$ ( $k = 3,4,5$ ) is a realization of a GP model with $\{\beta^\delta, \sigma_\delta^2, \omega^\delta\} = \{0, 1, 9.26\}$
$\delta^4(x)$	
$\delta^5(x)$	
$\delta^6(x)$	Each $\delta^k(x)$ ( $k = 6,7,8$ ) is a realization of a GP model with $\{\beta^\delta, \sigma_\delta^2, \omega^\delta\} = \{0, 0.028, 9.26\}$
$\delta^7(x)$	
$\delta^8(x)$	



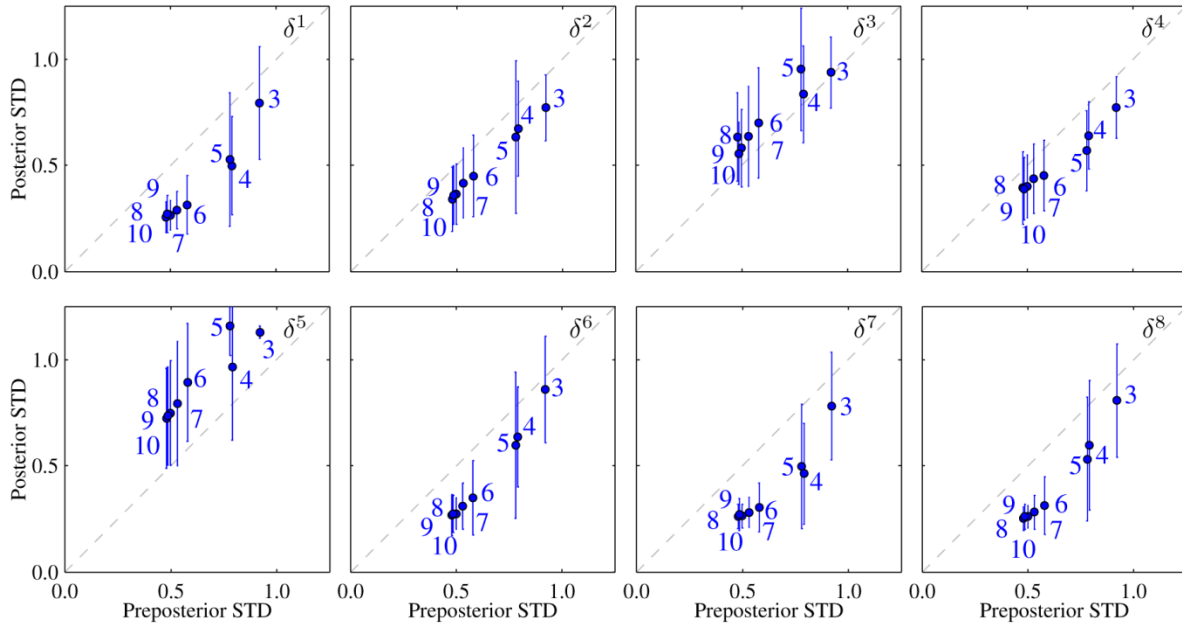
**Figure S1.** Plots of selected discrepancy functions listed in Table S1.



**Figure S2.** Plot of the posterior STD for the 8 different discrepancy functions versus the preposterior STD (both normalized by the prior STD). In Step 3 of the MC loop for calculating the preposterior STD,  $\sigma_\delta^2 = 0.028$  was used for all 8 cases. The numbers indicate  $N_e$ . In the vertical axis, the circles and error bars represent the mean posterior STD and  $\pm 1$  standard deviations of the posterior STD, with respect to the prior  $p(\theta)$ .

Because the parameter  $\sigma_\delta^2$  specified in Step 0c-i dictates the magnitude of the discrepancy functions that are generated when generating the hypothetical physical experimental data within the MC loop, the sensitivity of the preposterior STD to choice of  $\sigma_\delta^2$  is of interest. For calculating the preposterior

STDs in Figure S2, we used  $\sigma_\delta^2 = 0.028$  in Step 3 of the MC loop for all 8 cases (regardless of value of  $\sigma_\delta^2$  listed in Table S1 for generating the actual discrepancy function). Figure S3 shows analogous results but using a much larger  $\sigma_\delta^2 = 1.0$  in Step 3 of the MC loop. Relative to Figure S2, the primary difference in Figure S3 is that the points are all shifted to the right. This is consistent with what one would expect, since using a larger value of  $\sigma_\delta^2$  in Step 3 should result in a larger preposterior STD. Notice that there is still a strong linear trend in most of the scatter plots in Figure S3.



**Figure S3.** Plot of the posterior STD for the 8 different discrepancy functions versus the preposterior STD (both normalized by the prior STD). In Step 3 of the MC loop for calculating the preposterior STD,  $\sigma_\delta^2 = 1.0$  was used for all 8 cases.

### An Additional Example of Using the Preposterior Analysis to Select $N_e$

The beam and sinusoidal examples were used in Section 5.3 to illustrate how to use the preposterior analysis to select the number  $N_e$  of experimental runs. These examples involved only a single input variable  $x$  and used uniformly spaced grid designs. Here we consider an example that uses Latin hypercube designs with two input variable and a calibration parameter. Suppose the computer model is

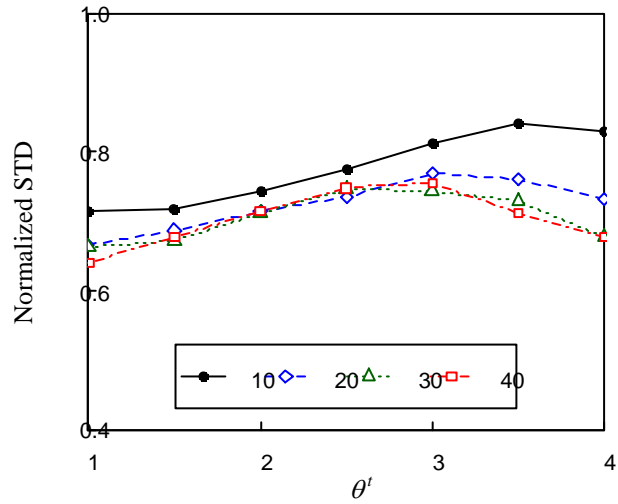
$$\begin{aligned}
y^m(x_1, x_2, \theta) &= \sin(\theta x_1) + x_1 x_2 \\
x_1 &\in [0, 1] \\
x_2 &\in [0, 1] \\
\theta &\in [0, 4]
\end{aligned}$$

and the physical experimental data are generated from the model

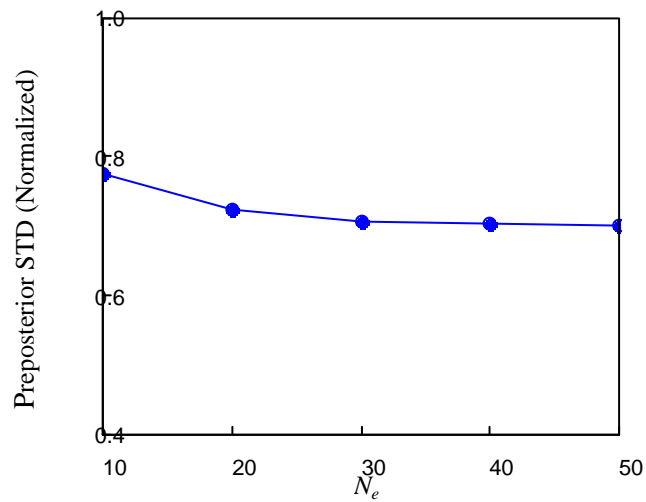
$$\begin{aligned}
y^e(x_1, x_2) &= y^m(x_1, x_2, \theta^*) + \delta(x_1, x_2) + \varepsilon \\
\varepsilon &\sim N(0, 0.01^2)
\end{aligned}$$

We began with a Latin hypercube design for the computer experiment with 40 points in three dimensions  $\{x_1, x_2, \theta\}$ . An optimal Latin hypercube design based on the maximin criterion was used. A Gaussian process model was fit to the computer experiment data via MLE in Step 0a of Figure 1. For the physical experimental design specified in Step 0b, we used maximin Latin hypercube designs of various size  $N_e = \{10, 20, 30, 40, 50\}$  over two dimensions  $\{x_1, x_2\}$ . Because of the random nature of the Latin hypercube designs, none of the physical experimental designs had  $\mathbf{x}$  settings that coincided with the computer experiment  $\mathbf{x}$  settings. Next, in Step 0c-i, we set the GP hyperparameters of the discrepancy function to be  $\beta^\delta = 0$ ,  $\omega^\delta = 3$ , and  $\sigma^\delta = 0.75$ , which corresponds to assigning them point mass priors. The relatively large value of  $\sigma^\delta$  results in a variety of relatively large discrepancy functions being generated within Step 3 of the preposterior algorithm. In step 0c-ii, we specified  $\lambda = 0.01^2$ . Lastly, in Step 0c-iii, we specified the prior for  $\theta$  to be uniform over the entire range  $[0, 4]$ .

From the preposterior analysis, Figure S4 shows the resulting fixed- $\theta$  preposterior standard deviation versus  $\theta^t$  for the designs with different  $N_e$ , and Figure S5 shows the (normalized) preposterior standard deviation as a function of  $N_e$ . From Figure S5, the preposterior standard deviation does not substantially decrease beyond  $N_e = 30$  roughly.



**Figure S4.** Fixed- $\theta$  preposterior STD (normalized by the prior STD of  $\theta$ ) versus  $\theta^t$  for Latin hypercube designs of various size  $N_e$  ( $N_e$  is specified in the legend).



**Figure S5.** Preposterior STD (normalized by the prior STD of  $\theta$ ) versus  $N_e$  for the Latin hypercube designs.