The Annals of Statistics
2015, Vol. 43, No. 6, 2331–2352
DOI: 10.1214/15-AOS1314

EFFICIENT CALIBRATION FOR IMPERFECT COMPUTER MODELS

BY RUI TUO[1] AND C. F. JEFF WU[2]

Chinese Academy of Sciences and Georgia Institute of Technology

Many computer models contain unknown parameters which need to be estimated using physical observations. Tuo and Wu (2014) show that the calibration method based on Gaussian process models proposed by Kennedy and O'Hagan [*J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** (2001) 425–464] may lead to an unreasonable estimate for imperfect computer models. In this work, we extend their study to calibration problems with stochastic physical data. We propose a novel method, called the L_2 calibration, and show its semiparametric efficiency. The conventional method of the ordinary least squares is also studied. Theoretical analysis shows that it is consistent but not efficient. Numerical examples show that the proposed method outperforms the existing ones.

1. Introduction. Computer simulations are widely used by researchers and engineers to understand, predict or control complex systems. Many physical phenomena and processes can be modeled with mathematical tools, like partial differential equations. These mathematical models are solved by numerical algorithms like the finite element method. For example, computer simulations can help predict the trajectory of a storm. In engineering, computer simulations have become more popular and sometimes indispensable in product and process designs. The design and analysis of experiments is a classic area of statistics. A new field has emerged, which considers the design and analysis for experiments in computer simulations, commonly referred to as "computer experiments." Unlike the physical experiments, computer experiments are usually deterministic. In addition, the input variables for computer experiments usually take real values, not discrete levels as in many physical experiments. Therefore, interpolation methods are widely used in computer experiments, while conventional methods, like ANOVA or regression models, are used much less often.

Received April 2014; revised January 2015.

[1]Supported in part by the Office of Advanced Scientific Computing Research; U.S. Department of Energy, project No. ERKJ259 "A mathematical environment for quantifying uncertainty: Integrated and optimized at the extreme scale." The work was performed at the Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC under Contract No. De-AC05-00OR22725. Supported by the National Center for Mathematics and Interdisciplinary Sciences, CAS and NSFC 11271355.

[2]Supported by NSF Grant DMS-13-08424 and DOE DE-SC0010548.

MSC2010 subject classifications. Primary 62P30, 62A01; secondary 62F12.

Key words and phrases. Computer experiments, uncertainty quantification, semiparametric efficiency, reproducing kernel Hilbert space.

In many computer experiments, some of the input parameters represent certain inherent attributes of the physical system. The true values of these variables are unknown because there may not be enough knowledge about the physical systems. For instance, in underground water simulations, the soil permeability is an important input parameter, but its true value is usually unknown. A standard approach to identify the unknown model parameters is known as *calibration*. To calibrate the unknown parameters, one needs to run the computer model under different model parameters, and run some physical experiments. The basic idea of calibration is to find the combination of the model parameters, under which the computer outputs match the physical responses.

One important topic in the calibration of computer models is to tackle the model uncertainty. Most physical models are built under certain assumptions or simplifications, which may not hold in reality. As a result, the computer output can rarely fit the physical response perfectly, even if the true values of the calibration parameters are known. We call such computer models *imperfect*. Kennedy and O'Hagan (2001) first discuss this model uncertainty problem and propose a Bayesian method, which models the discrepancy between the physical process and the computer output as a Gaussian process. Because of the importance of calibration for computer models, Kennedy–O'Hagan's approach has been widely used, including hydrology, radiological protection, cylinder implosion, spot welding, micro-cutting, climate prediction and cosmology. See Higdon et al. (2004, 2008, 2013), Bayarri et al. (2007a, 2007b), Joseph and Melkote (2009), Wang, Chen and Tsui (2009), Han, Santner and Rawlinson (2009), Goldstein and Rougier (2004), Murphy et al. (2007) and Goh et al. (2013).

Tuo and Wu (2014) study the asymptotic properties of the calibration parameter estimators given by Kennedy and O'Hagan (2001) and show that their method can lead to unreasonable estimate. The first theoretical framework for calibration problems of the Kennedy–O'Hagan type is established by Tuo and Wu (2014) under the assumption that the physical responses have no random error. This assumption is needed to make the mathematical analysis for a version of Kennedy–O'Hagan's approach feasible. Given the fact that the responses in physical experiments are rarely deterministic, it is necessary to extend the study to cases where the physical responses have measurement or observational errors. For convenience, we use the term "stochastic physical experiments" to denote physical responses with random errors.

In Tuo and Wu (2014), the theory of native spaces is used to derive the convergence rate for calibration with deterministic physical systems. Because of the random error in the current context, the interpolation theory fails to work. In this work, we will mainly use mathematical tools of weak convergence, including the limiting theory of empirical processes.

The main theme of this article is to propose a general framework for calibration and provide an efficient estimator for the calibration parameter. We utilize a nonparametric regression method to model the physical outputs. Similar models

are also considered in the literature of response surface methodology; see Myers (1999) and Anderson-Cook and Prewitt (2005). To estimate the calibration parameter, we extend the L_2 calibration method proposed by Tuo and Wu (2014) to the present context. This novel method is proven to be semiparametric efficient when the measurement error follows a normal distribution. A conventional method, namely, the ordinary least squares method, is also studied, and shown to be consistent but not efficient.

This paper is organized as follows. In Section 2, we extend the L_2 projection defined by Tuo and Wu (2014) to stochastic systems and propose the L_2 calibration method in the current context. In Section 3, the asymptotic behavior for L_2 calibration is studied. In Section 4, we consider the ordinary least squares method. The proposed method is illustrated and its performance studied in two numerical examples in Section 5. Concluding remarks are given in Section 6.

2. L_2 projection for systems with stochastic physical experiments. Let Ω denote the region of interest for the control variables, which is a convex and compact subset of $\mathbf{R}^d$. Let $x_1, \ldots, x_n$ be a set of points on Ω. Suppose the physical experiment is conducted once on each x_i, with the corresponding response denoted by y_i^p, for $i = 1, \ldots, n$, where the superscript p stands for "physical." In this work, we assume the physical system is stochastic, that is, the physical responses have random measurement or observational errors. To incorporate this randomness, we consider the following nonparametric model:

$$y_i = \zeta(x_i) + e_i, \tag{2.1}$$

where $\zeta(\cdot)$ is an unknown deterministic function and $\{e_i\}_{i=1}^n$ is a sequence of independent and identically distributed random variables with $Ee_i = 0$ and $Ee_i^2 = \sigma^2 < +\infty$. This model is also adopted by Kennedy and O'Hagan (2001), where $\zeta(\cdot)$ is called the *true process*. In addition, Kennedy and O'Hagan assumes that e_i's follow a normal distribution. Such a distribution assumption will be slightly relaxed in our theoretical analysis.

Let Θ be the parameter space for the calibration parameter θ. Suppose Θ is a compact region in $\mathbf{R}^q$. Denote the output of the deterministic computer code at $(x, \theta) \in \Omega \times \Theta$ by $y^s(x, \theta)$, where the superscript s stands for "simulation." In the frequentist framework of calibration established by Tuo and Wu (2014), the concept of L_2 *projection* plays a central role. Because the "true" calibration parameter [as stated in Kennedy and O'Hagan (2001)] is unidentifiable, Tuo and Wu (2014) define the purpose of calibration as that of finding the L_2 projection θ^*, which minimizes the L_2 distance between the physical response surface and the computer outputs as a function of the control variables. In the present context, the physical responses are observed with errors. A good definition of the "true" value of θ should exclude the uncertainty in y^p. Thus, we suggest the following definition for the L_2 projection using the true process $\zeta(\cdot)$:

$$\theta^* := \operatorname*{argmin}_{\theta \in \Theta} \left\| \zeta(\cdot) - y^s(\cdot, \theta) \right\|_{L_2(\Omega)}. \tag{2.2}$$

The main focus of this article is on the statistical inference for θ^*.

2.1. *L_2 calibration.* In this section, we will extend the L_2 calibration method proposed by Tuo and Wu (2014) to the present context. Since Tuo and Wu assume that the physical experiment is deterministic, they use the kernel interpolation method to approximate the physical response surface. Because of the existence of the random error in (2.1), the kernel interpolation can perform poorly because interpolation methods generally suffer from the problem of overfitting.

In spatial statistics, the effect of the random error is usually modeled with a white noise process, which is also referred to as a nugget term in the kriging modeling [Cressie (1993)]. In the computer experiment literature, it is also common to use the nugget term in Gaussian process modeling to tackle the numerical instability problems [Gramacy and Lee (2012), Peng and Wu (2014)].

Let $z(\cdot)$ be a Gaussian process with mean zero and covariance function $\Phi(\cdot,\cdot)$. Suppose $\{(x_i, y_i)\}_{i=1}^{y}$ are obtained, which satisfy $y_i = z(x_i) + \epsilon_i$ with ϵ_i's being i.i.d. and distributed as $N(0, \sigma^2)$. Then the predictive mean of $z(\cdot)$ is given by

$$\hat{z}(x) = \sum_{i=1}^{n} u_i \Phi(x_i, x), \tag{2.3}$$

where $u = (u_1, \ldots, u_n)^{\mathrm{T}}$ is the solution to the linear system

$$Y = (\mathbf{\Phi} + \sigma^2 \mathbf{I})u, \tag{2.4}$$

with $Y = (y_1, \ldots, y_n)^{\mathrm{T}}$ and $\mathbf{\Phi} = (\Phi(x_i, x_j))_{ij}$. By the representer theorem [Schölkopf, Herbrich and Smola (2001), Wahba (1990)], $\hat{z}(x)$ given by (2.3) and (2.4) is the solution of the following minimization problem with some $\lambda > 0$:

$$\operatorname*{argmin}_{f \in \mathcal{N}_\Phi(\Omega)} \frac{1}{n} \sum_{i=1}^{n} \big(y_i - f(x_i)\big)^2 + \lambda \|f\|^2_{\mathcal{N}_\Phi(\Omega)}, \tag{2.5}$$

where $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$ is the norm of the *reproducing kernel Hilbert space* $\mathcal{N}_\Phi(\Omega)$ generated by the kernel function Φ. We refer to Wendland (2005) and Wahba (1990) for detailed discussions about these spaces. The solution to (2.5) is referred to as the *nonparametric regressor* in the reproducing kernel Hilbert space [Berlinet and Thomas-Agnan (2004)].

Now we are ready to define the L_2 calibration method for systems with stochastic physical experiments. Suppose the physical experiment is conducted over a design set $\{x_1, \ldots, x_n\}$. Define

$$\hat{\zeta} := \operatorname*{argmin}_{f \in \mathcal{N}_\Phi(\Omega)} \frac{1}{n} \sum_{i=1}^{n} \big(y_i^p - f(x_i)\big)^2 + \lambda \|f\|^2_{\mathcal{N}_\Phi(\Omega)}, \tag{2.6}$$

where the smoothing parameter λ can be chosen using certain model selection criterion, for example, generalized cross validation (GCV); see Wahba (1990). We

define the L_2 calibration for θ as

$$\hat{\theta}^{L_2} := \operatorname*{argmin}_{\theta\in\Theta} \|\hat{\zeta}(\cdot) - \hat{y}^s(\cdot,\theta)\|_{L_2(\Omega)},$$

where $\hat{y}^s(\cdot,\cdot)$ is an *emulator* for the computer code $y^s(\cdot,\cdot)$. In this work, the emulator for the computer model can be constructed by any method provided that it approximates y^s well. For instance, $\hat{y}^s$ can be constructed by the radial basis function approximation [Wendland (2005)], Gaussian process models [Santner, Williams and Notz (2003)] or the polynomial chaos approximation [Xiu (2010)].

3. Asymptotic results for L_2 calibration. We now consider the asymptotic behavior of $\hat{\theta}^{L_2}$ as the sample size n becomes large. For mathematical rigor, we write

$$\hat{\zeta}_n = \operatorname*{argmin}_{f\in\mathcal{N}_\Phi(\Omega)} \frac{1}{n}\sum_{i=1}^{n}\big(y_i^p - f(x_i)\big)^2 + \lambda_n \|f\|^2_{\mathcal{N}_\Phi(\Omega)}, \tag{3.1}$$

for all sufficiently large n, where $\{\lambda_n\}_{i=1}^n$ is a prespecified sequence of positive values. For the ease of mathematical treatment, we assume the x_i's are a sequence of random samples rather than fixed design points. We also write the L_2 calibration estimator indexed by n as

$$\hat{\theta}_n^{L_2} := \operatorname*{argmin}_{\theta\in\Theta} \|\hat{\zeta}_n(\cdot) - \hat{y}_n^s(\cdot,\theta)\|_{L_2(\Omega)}, \tag{3.2}$$

where the emulator $\hat{y}^s$ is also indexed by n. We assume that $\hat{y}_n^s$ has increasing approximation power as n becomes large.

3.1. *Asymptotic results for $\hat{\zeta}_n$.* Before stating the asymptotic results for $\hat{\theta}_n^{L_2}$, we need first to show that $\hat{\zeta}_n$ tend to ζ. To study the convergence, we need some additional definitions from the theory of empirical processes [Kosorok (2008)]. For function space $\mathcal{F}$ over Ω, define the covering number $N(\delta,\mathcal{F},\|\cdot\|_{L_\infty(\Omega)})$ as the smallest value of N for which there exist functions $f_1,\dots,f_N$, such that for each $f\in\mathcal{F}$, $\|f-f_j\|_{L_\infty(\Omega)}\le\delta$ for some $j\in\{1,\dots,N\}$. The L_2 covering number with bracketing $N_{[]}(\delta,\mathcal{F},\|\cdot\|_{L_2(\Omega)})$ is the smallest value of N for which there exist L_2 functions $\{f_1^L, f_1^U,\dots,f_N^L,f_N^U\}$ with $\|f_j^U - f_j^L\|_{L_2(\Omega)}\le\delta$, $j=1,\dots,N$ such that for each $f\in\mathcal{F}$ there exists a j such that $f_j^L\le f\le f_j^U$.

We now state a result for general nonparametric regression. Suppose $\mathcal{F}$ is a space of functions over a compact region Ω equipped with a norm $\|\cdot\|$. Suppose the true model is

$$y_i = f_0(x_i) + \epsilon_i, \tag{3.3}$$

and x_i are i.i.d. from the uniform distribution $U(\Omega)$ over Ω. In addition, the sequences $\{x_i\}$ and $\{e_i\}$ are independent and e_i has zero mean. We use "$\preceq$" to denote

that the left-hand side is dominated by the right-hand side up to a constant. Let

$$\hat{f}_n = \operatorname*{argmin}_{f\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n}\bigl(y_i - f(x_i)\bigr)^2 + \lambda_n \|f\|^2, \tag{3.4}$$

for some $\lambda_n > 0$.

LEMMA 1. *Under the model* (3.3), *suppose* $f_0 \in \mathcal{F}$. *Let* $\mathcal{F}(\rho) := \{f \in \mathcal{F} : \|f\| \le \rho\}$. *Suppose there exists* $C_0 > 0$ *such that* $E[\exp(C_0|e_i|)] < \infty$. *Moreover, there exists* $0 < \tau < 2$ *such that*

$$\log N_{[]}\bigl(\delta, \mathcal{F}(\rho), \|\cdot\|_{L_2(\Omega)}\bigr) \preceq \rho^{\tau}\delta^{-\tau},$$

for all $\delta, \rho > 0$. *Then if* $\lambda_n^{-1} = O(n^{2/(2+\tau)})$, *the estimator* $\hat{f}_n$ *given by* (3.4) *satisfies*

$$\|\hat{f}_n\| = O_p(1) \quad \textit{and} \quad \|\hat{f}_n - f_0\|_{L_2(\Omega)} = O_p\bigl(\lambda_n^{1/2}\bigr).$$

PROOF. See van de Geer (2000). □

The covering numbers for some reproducing kernel Hilbert spaces have been calculated accurately in the literature. For instance, consider a Matérn kernel function given by

$$\Phi(s,t;\nu,\phi) = \frac{1}{\Gamma(\upsilon)2^{\nu-1}}\bigl(2\sqrt{\nu}\phi\|s-t\|\bigr)^{\nu}K_{\nu}\bigl(2\sqrt{\nu}\phi\|s-t\|\bigr), \tag{3.5}$$

with $\nu \ge 1$ [Santner, Williams and Notz (2003), Stein (1999)]. The reproducing kernel Hilbert space generated by this kernel function is equal to the (fractional) Sobolev space $H^{\nu+d/2}(\Omega)$, and $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$ and $\|\cdot\|_{H^{\nu+d/2}(\Omega)}$ are equivalent; see Corollary 1 of Tuo and Wu (2014). Let $H^{\mu}(\Omega,\rho) := \{f : \|f\|_{H^{\mu}(\Omega)} \le \rho\}$. Edmunds and Triebel (1996) prove that for $\mu > d/2$, the covering number of $H^{\mu}(\Omega,\rho)$ is bounded by

$$\log N\bigl(\delta, H^{\mu}(\Omega,\rho), \|\cdot\|_{L_\infty(\Omega)}\bigr) \le \left(\frac{C\rho}{\delta}\right)^{d/\mu},$$

where C is independent of ρ and δ. To calculate the L_2 metric entropy with bracketing, we note the fact that every $f, f' \in H^{\mu}(\Omega,\rho)$ with $\|f - f'\| \le \delta$ satisfy the inequality $f' - \delta \le f \le f' + \delta$. Thus, the union of the δ-balls centered at $f_1, \ldots, f_n$ is covered by the union of the "brackets" $[f_1 - \delta, f_1 + \delta], \ldots, [f_n - \delta, f_n + \delta]$, which together with the definition of the covering number and the L_2 covering number with bracketing, implies that

$$\log N_{[]}\bigl(2\delta\sqrt{\mathrm{Vol}(\Omega)}, H^{\mu}(\Omega,\rho), \|\cdot\|_{L_2(\Omega)}\bigr) \le \left(\frac{C\rho}{\delta}\right)^{d/\mu}, \tag{3.6}$$

where $\mathrm{Vol}(\Omega)$ denotes the volume of Ω, and $2\delta\sqrt{\mathrm{Vol}(\Omega)}$ is the $L_2(\Omega)$ norm of the function 2δ. Then by applying Lemma 1, the following result can be obtained after direct calculations.

PROPOSITION 1. *Under model* (3.3), *suppose* $f_0 \in \mathcal{F} = \mathcal{N}_\Phi(\Omega)$ *and* $\mathcal{N}_\Phi(\Omega)$ *can be embedded into* $H^\mu(\Omega)$ *with* $\mu > d/2$. *Choose* $\|\cdot\| = \|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$. *Then for* $\lambda_n^{-1} = O(n^{2\mu/(2\mu+d)})$, *the estimator* $\hat{f}_n$ *given by* (3.4) *satisfies*

$$\|\hat{f}_n\|_{\mathcal{N}_\Phi(\Omega)} = O_p(1) \quad \textit{and} \quad \|\hat{f}_n - f_0\|_{L_2(\Omega)} = O_p(\lambda_n^{1/2}).$$

In Proposition 1, one can choose $\lambda \asymp n^{-2\mu/(2\mu+d)}$ to obtain the best convergence rate $\|\hat{f}_n - f_0\|_{L_2(\Omega)} = O_p(n^{-\mu/(2\mu+d)})$, where "$\asymp$" denotes that its left-hand and the right-hand sides have the same order of magnitude. This rate is known to be optimal [Stone (1982)].

3.2. *Asymptotic normality.* The main purpose of calibration is to estimate the calibration parameter θ^*. In this section, we will prove some convergence properties of the L_2 calibration: its convergence rate is given by $\|\hat{\theta}_n^{L_2} - \theta^*\| = O_p(n^{-1/2})$ and the distribution of $\sqrt{n}(\hat{\theta}_n^{L_2} - \theta^*)$ tends to normal as $n \to \infty$ under certain conditions. This is a nontrivial result because the convergence rate for the nonparametric part $\|\hat{\zeta}_n(\cdot) - \zeta\|_{L_2(\Omega)}$ is generally slower than $O_p(n^{-1/2})$ (see Proposition 1).

We first list necessary conditions for the convergence result, which are grouped in three categories.

The first group consists of regularity conditions on the model. For any $\theta \in \Theta \subset \mathbf{R}^q$, write $\theta = (\theta_1, \ldots, \theta_q)$.

A1: The sequences $\{x_i\}$ and $\{e_i\}$ are independent; x_i's are i.i.d. from $U(\Omega)$; and $\{e_i\}$ is a sequence of i.i.d. random variables with zero mean and finite variance.

A2: θ^* is the unique solution to (2.2), and is an interior point of Θ.

A3: $\sup_{\theta\in\Theta} \|y^s(\cdot, \theta)\|_{L_2(\Omega)} < +\infty$.

A4: $V := E[\frac{\partial^2}{\partial\theta\,\partial\theta^{\mathrm{T}}}(\zeta(x_1) - y^s(x_1, \theta^*))^2]$ is invertible.

A5: There exists a neighborhood $U \subset \Theta$ of θ^*, such that

$$\sup_{\theta\in U} \left\| \frac{\partial y^s}{\partial \theta_j}(\cdot, \theta) \right\|_{\mathcal{N}_\Phi(\Omega)} < +\infty, \qquad \frac{\partial^2 y^s}{\partial\theta_j\,\partial\theta_k}(\cdot, \cdot) \in C(\Omega \times U),$$

for all $\theta \in U$ and $j, k = 1, \ldots, q$.

Next, we need some conditions on the nonparametric part.

B1: $\zeta \in \mathcal{N}_\Phi(\Omega)$ and $\mathcal{N}_\Phi(\Omega, \rho)$ is Donsker for all $\rho > 0$.

B2: $\|\hat{\zeta} - \zeta\|_{L_2(\Omega)} = o_p(1)$.

B3: $\|\hat{\zeta}\|_{\mathcal{N}_\Phi(\Omega)} = O_p(1)$.

B4: $\lambda_n = o_p(n^{-1/2})$.

The Donsker property is an important concept in the theory of empirical processes. For its definition and detailed discussion, we refer to van der Vaart and

Wellner (1996) and Kosorok (2008). One major result is that a class of functions over domain Ω, denoted as $\mathcal{F}$, is Donsker, if

$$\int_0^\infty \sqrt{\log N_{[]}(\delta, \mathcal{F}, L_2(\Omega))}\, d\delta < +\infty.$$

Thus, from (3.6) we can see that if $\mathcal{N}_\Phi(\Omega)$ can be embedded into $H^\mu(\Omega)$ for some $\mu > d/2$, $\mathcal{N}_\Phi(\Omega)$ is Donsker. Actually, if we further assume condition A1 and $E[\exp(C|e_i|)] < +\infty$ for some $C > 0$, the conditions of Proposition 1 are satisfied. Then by choosing a suitable sequence of $\{\lambda_n\}$, say $\lambda \asymp n^{-2\mu/(2\mu+d)}$, one can show that condition B4 holds and conditions B2 and B3 are ensured by Proposition 1.

Finally, we need to assume some convergence properties for the emulator. In this work, we assume that the approximation error caused by emulating the computer experiment is negligible compared to the estimation error caused by the measurement error in the physical experiment. Under this assumption, the asymptotic behavior of $\hat{\theta}_n^{L_2} - \theta^*$ is determined by the central limit theorem. Given that computer experiment is usually much cheaper to run than physical experiment, such an assumption is reasonable because the size of computer runs is in general much larger than the size of physical trials.

C1: $\|\hat{y}_n^s - y^s\|_{L_\infty(\Omega\times\Theta)} = o_p(n^{-1/2})$.
C2: $\|\frac{\partial \hat{y}^s}{\partial \theta_i} - \frac{\partial y^s}{\partial \theta_i}\|_{L_\infty(\Omega\times\Theta)} = o_p(n^{-1/2})$, for $i = 1, \ldots, q$.

Now we are ready to state the main theorem of this section on the asymptotic normality of the L_2 calibration.

THEOREM 1. *Under conditions* A1–A5, B1–B4 *and* C1–C2, *we have*

$$\hat{\theta}_n^{L_2} - \theta^* = -2V^{-1}\left\{\frac{1}{n}\sum_{i=1}^n e_i \frac{\partial y^s}{\partial \theta}(x_i, \theta^*)\right\} + o_p(n^{-1/2}), \tag{3.7}$$

where V is defined in condition A4.

PROOF. We first prove that $\hat{\theta}_n \to^p \theta^*$. From the definitions of θ^* and $\hat{\theta}_n^{L_2}$ in (2.2) and (3.2), it suffices to prove that $\|\hat{\zeta}_n(\cdot) - \hat{y}_n^s(\cdot, \theta)\|_{L_2(\Omega)}$ converges to $\|\zeta(\cdot) - y^s(\cdot, \theta)\|_{L_2(\Omega)}$ uniformly with respect to $\theta \in \Theta$ in probability, which is ensured by

$$\begin{aligned}
&\|\hat{\zeta}_n(\cdot) - \hat{y}^s(\cdot, \theta)\|_{L_2(\Omega)}^2 - \|\zeta(\cdot) - y^s(\cdot, \theta)\|_{L_2(\Omega)}^2 \\
&\quad = \int_\Omega \big(\hat{\zeta}_n(z) - \zeta(z) - \hat{y}^s(z, \theta) + y^s(z, \theta)\big) \\
&\qquad \times \big(\hat{\zeta}_n(z) + \zeta(z) - y^s(z, \theta) - \hat{y}^s(z, \theta)\big)\, dz \\
&\quad \le \big(\|\hat{\zeta}_n - \zeta\|_{L_2(\Omega)} + \|\hat{y}^s(\cdot, \theta) - y^s(\cdot, \theta)\|_{L_2(\Omega)}\big) \\
&\qquad \times \big(\|\hat{\zeta}_n(\cdot)\|_{L_2(\Omega)} + \|\zeta(\cdot)\|_{L_2(\Omega)} + \|y^s(\cdot, \theta)\|_{L_2(\Omega)} + \|\hat{y}^s(\cdot, \theta)\|_{L_2(\Omega)}\big),
\end{aligned} \tag{3.8}$$

where the inequality follows from the Schwarz inequality and the triangle inequality. Denote the volume of Ω by $\mathrm{Vol}(\Omega)$. It is easily seen that

$$\|f\|_{L_2(\Omega)} \le \mathrm{Vol}(\Omega)\|f\|_{L_\infty(\Omega)}$$

holds for all $f \in L_\infty(\Omega)$. Thus,

$$\begin{aligned}\|\hat{y}^s(\cdot,\theta) - y^s(\cdot,\theta)\|_{L_2(\Omega)} &\le \sqrt{\mathrm{Vol}(\Omega)}\|\hat{y}^s(\cdot,\theta) - y^s(\cdot,\theta)\|_{L_\infty(\Omega)} \\ &\le \sqrt{\mathrm{Vol}(\Omega)}\|\hat{y}^s - y^s\|_{L_\infty(\Omega\times\Theta)}.\end{aligned} \tag{3.9}$$

Additionally, we have

$$\begin{aligned}\|\hat{\zeta}_n\|_{L_2(\Omega)} &\le \mathrm{Vol}(\Omega)\|\hat{\zeta}_n\|_{L_\infty(\Omega)} \\ &= \mathrm{Vol}(\Omega)\sup_{x\in\Omega}\langle\hat{\zeta}_n, \Phi(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} \\ &\le \mathrm{Vol}(\Omega)\|\hat{\zeta}_n\|_{\mathcal{N}_\Phi(\Omega)}\sup_{x\in\Omega}\|\Phi(\cdot,x)\|_{\mathcal{N}_\Phi(\Omega)} = \mathrm{Vol}(\Omega)\|\hat{\zeta}_n\|_{\mathcal{N}_\Phi(\Omega)}.\end{aligned} \tag{3.10}$$

Combining (3.9), (3.10), B2 and C1, we have that (3.8) convergence to 0 uniformly with respect to $\theta \in \Theta$, which yields the consistency of $\hat{\theta}_n^{L_2}$.

Since $\hat{\theta}$ minimizes (3.2), following A1, A2 and A5 we have

$$\begin{aligned}0 &= \frac{\partial}{\partial\theta}\|\hat{\zeta}_n(\cdot) - \hat{y}^s(\cdot,\hat{\theta}_n^{L_2})\|_{L_2(\Omega)}^2 \\ &= 2\int_\Omega(\hat{\zeta}_n(z) - \hat{y}^s(z,\hat{\theta}_n^{L_2}))\frac{\partial\hat{y}^s}{\partial\theta}(z,\hat{\theta}_n^{L_2})\,dz,\end{aligned}$$

which, together with B2, C1 and C2, implies

$$\int_\Omega(\hat{\zeta}_n(z) - y^s(z,\hat{\theta}_n^{L_2}))\frac{\partial y^s}{\partial\theta}(z,\hat{\theta}_n^{L_2})\,dz = o_p(n^{-1/2}). \tag{3.11}$$

Let $L_n(f) = n^{-1}\sum_{i=1}^n(y_i^p - f(x_i))^2 + \lambda_n\|f\|_{\mathcal{N}_\Phi(\Omega)}^2$. By (2.6), $\hat{\zeta}_n$ minimizes L_n over $\mathcal{N}_\Phi(\Omega)$. Since $\hat{\theta}_n^{L_2}$ is consistent, by A5, $\frac{\partial y^s}{\partial\theta_j}(\cdot,\hat{\theta}_n^{L_2}) \in \mathcal{N}_\Phi(\Omega)$ for $j = 1,\dots,q$ and sufficiently large n. Thus, we have

$$\begin{aligned}0 &= \frac{\partial}{\partial t}L\left(\hat{\zeta}_n(\cdot) + t\frac{\partial y^s}{\partial\theta_j}(\cdot,\hat{\theta}_n^{L_2})\right)\Big|_{t=0} \\ &= \frac{2}{n}\sum_{i=1}^n\{\hat{\zeta}_n(x_i) - y_i^p\}\frac{\partial y^s}{\partial\theta_j}(x_i,\hat{\theta}_n^{L_2}) + 2\lambda_n\left\langle\hat{\zeta}_n, \frac{\partial y^s}{\partial\theta_j}(\cdot,\hat{\theta}_n^{L_2})\right\rangle_{\mathcal{N}_\Phi(\Omega)} \\ &= \frac{2}{n}\sum_{i=1}^n\{\hat{\zeta}_n(x_i) - \zeta(x_i)\}\frac{\partial y^s}{\partial\theta_j}(x_i,\hat{\theta}_n^{L_2}) - \frac{2}{n}\sum_{i=1}^n e_i\frac{\partial y^s}{\partial\theta_j}(x_i,\hat{\theta}_n^{L_2}) \\ &\quad + 2\lambda_n\left\langle\hat{\zeta}_n, \frac{\partial y^s}{\partial\theta_j}(\cdot,\hat{\theta}_n^{L_2})\right\rangle_{\mathcal{N}_\Phi(\Omega)} \\ &=: 2(C_n + D_n + E_n).\end{aligned} \tag{3.12}$$

First, we consider C_n. Let $A_i(g,\theta) = \{g(x_i) - \zeta(x_i)\}\frac{\partial y^s}{\partial \theta_j}(x_i,\theta)$ for $(g,\theta) \in \mathcal{N}_\Phi(\Omega,\rho) \times U$ with some $\rho > 0$ to be specified later. Then $E[A_i(g,\theta)] = \int_\Omega \{g(z) - \zeta(z)\}\frac{\partial y^s}{\partial \theta_j}(z,\theta)\,dz$. Define the empirical process

$$E_{1n}(g,\theta) = n^{-1/2}\sum_{i=1}^{n}\big\{A_i(g,\theta) - E\big[A_i(g,\theta)\big]\big\}.$$

By B1, $\mathcal{N}_\Phi(\Omega,k)$ is Donsker. Thus, $\mathcal{F}_1 = \{g - \zeta : g \in \mathcal{N}_\Phi(\Omega,\rho)\}$ is also Donsker. Condition A5 implies that $\mathcal{F}_2 = \{\frac{\partial y^s}{\partial \theta_j}(\cdot,\theta) : \theta \in U\}$ is Donsker. Since both $\mathcal{F}_1$ and $\mathcal{F}_2$ are uniformly bounded, the product class $\mathcal{F}_1 \times \mathcal{F}_2$ is also Donsker. For theorems on Donsker classes, we refer to Kosorok (2008) and the references therein. Thus, the asymptotic equicontinuity property holds, which suggests that [see Theorem 2.4 of Mammen and van de Geer (1997)] for any $\xi > 0$ there exists a $\delta > 0$ such that

$$\limsup_{n\to\infty} P\left(\sup_{f\in\mathcal{F}_1\times\mathcal{F}_2,\|f\|\le\delta}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(f(x_i) - E\big(f(x_i)\big)\big)\right| > \xi\right) < \xi,$$

where $\|\cdot\|$ is defined as $\|f\|^2 := E[f(x_i)]^2$. This implies that for all $\xi > 0$ there exists a $\delta > 0$ such that

$$\limsup_{n\to\infty} P\Big(\sup_{g\in\mathcal{N}_\Phi(\Omega,\rho),\theta\in U,\|g-\zeta\|_{L_2(\Omega)}\le\delta}\big|E_{1n}(g,\theta)\big| > \xi\Big) < \xi. \tag{3.13}$$

Now fix $\varepsilon > 0$. Condition B3 implies that there exists $\rho_0 > 0$, such that $P(\|\hat{\zeta}_n\|_{\mathcal{N}_\Phi(\Omega)} > \rho_0) \le \varepsilon/3$. Choose δ_0 to be a possible value of δ satisfying (3.13) with $\rho = \rho_0$ and $\xi = \varepsilon/3$. Define

$$\hat{\zeta}_n^\circ := \begin{cases} \hat{\zeta}_n, & \text{if } \|\hat{\zeta}_n\|_{\mathcal{N}_\Phi(\Omega)} \le \rho_0 \text{ and } \|\hat{\zeta}_n - \zeta\|_{L_2(\Omega)} \le \delta_0, \\ \zeta, & \text{elsewise.} \end{cases}$$

Therefore, for sufficiently large n we have

$$\begin{aligned} &P\big(\big|E_{1n}\big(\hat{\zeta}_n,\hat{\theta}_n^{L_2}\big)\big| > \varepsilon\big) \\ &\quad\le P\big(\big|E_{1n}\big(\hat{\zeta}_n^\circ,\hat{\theta}_n^{L_2}\big)\big| > \varepsilon\big) + P\big(\|\hat{\zeta}_n\|_{\mathcal{N}_\Phi(\Omega)} > \rho_0\big) \\ &\qquad + P\big(\|\hat{\zeta}_n - \zeta\|_{L_2(\Omega)} > \delta_0\big) \\ &\quad\le P\big(\big|E_{1n}\big(\hat{\zeta}_n^\circ,\hat{\theta}_n^{L_2}\big)\big| > \varepsilon/3\big) + \varepsilon/3 + \varepsilon/3 \\ &\quad\le P\Big(\sup_{g\in\mathcal{N}(\Omega,\rho_0),\theta\in U,\|g-\zeta\|_{L_2(\Omega)}\le\delta_0}\big|E_{1n}(g,\theta)\big| > \varepsilon/3\Big) + \varepsilon/3 + \varepsilon/3 \\ &\quad\le \varepsilon, \end{aligned}$$

where the first and the third inequalities follow from the definition of $\hat{\zeta}_n^{\circ}$; the second inequality follows from B2; the last inequality follows from (3.13). This implies that $E_{1n}(\hat{\zeta}_n, \hat{\theta}_n^{L_2})$ tends to zero in probability. Thus, we have

$$\begin{aligned} o_p(1) = E_{1n}\bigl(\hat{\zeta}_n, \hat{\theta}_n^{L_2}\bigr) &= n^{-1/2} \sum_{i=1}^{n} \bigl\{\hat{\zeta}_n(x_i) - \zeta(x_i)\bigr\} \frac{\partial y^s}{\partial \theta_j}\bigl(x_i, \hat{\theta}_n^{L_2}\bigr) \\ &\quad - n^{1/2} \int_{\Omega} \bigl\{\hat{\zeta}_n(z) - \zeta(z)\bigr\} \frac{\partial y^s}{\partial \theta_j}\bigl(z, \hat{\theta}_n^{L_2}\bigr)\, dz \\ &= n^{1/2} C_n - n^{1/2} \int_{\Omega} \bigl\{\hat{\zeta}_n(z) - \zeta(z)\bigr\} \frac{\partial y^s}{\partial \theta_j}\bigl(z, \hat{\theta}_n^{L_2}\bigr)\, dz, \end{aligned}$$

which implies

$$C_n = \int_{\Omega} \bigl\{\hat{\zeta}_n(z) - \zeta(z)\bigr\} \frac{\partial y^s}{\partial \theta_j}\bigl(z, \hat{\theta}_n^{L_2}\bigr)\, dz + o_p\bigl(n^{-1/2}\bigr). \tag{3.14}$$

By substituting (3.11) to (3.14) and using A2, we can apply the Taylor expansion to (3.14) at θ^* and obtain

$$\begin{aligned} C_n &= \int_{\Omega} \bigl\{y^s\bigl(z, \hat{\theta}_n^{L_2}\bigr) - \zeta(z)\bigr\} \frac{\partial y^s}{\partial \theta_j}\bigl(z, \hat{\theta}_n^{L_2}\bigr)\, dz + o_p\bigl(n^{-1/2}\bigr) \\ &= \left\{\frac{1}{2} \int_{\Omega} \frac{\partial^2}{\partial \theta^{\mathrm{T}} \partial \theta_j} \bigl(y^s(z, \tilde{\theta}_n) - \zeta(z)\bigr)^2\, dz\right\} \bigl(\hat{\theta}_n^{L_2} - \theta^*\bigr) + o_p\bigl(n^{-1/2}\bigr), \end{aligned} \tag{3.15}$$

where $\tilde{\theta}_n$ lies between $\hat{\theta}_n$ and θ^*. By the consistency of $\hat{\theta}_n^{L_2}$, we have $\tilde{\theta}_n \to^p \theta^*$. This implies that

$$\begin{aligned} &\int_{\Omega} \frac{\partial^2}{\partial \theta^{\mathrm{T}} \partial \theta} \bigl(y^s(z, \tilde{\theta}_n) - \zeta(z)\bigr)^2\, dz \\ &\quad \xrightarrow{p} \int_{\Omega} \frac{\partial^2}{\partial \theta^{\mathrm{T}} \partial \theta} \bigl(y^s(z, \theta^*) - \zeta(z)\bigr)^2\, dz \\ &\quad = V. \end{aligned} \tag{3.16}$$

Now we consider D_n. Define the empirical process

$$\begin{aligned} E_{2n}(\theta) &= n^{-1/2} \sum_{i=1}^{n} \left\{ e_i \frac{\partial y^s}{\partial \theta_j}(x_i, \theta) - e_i \frac{\partial y^s}{\partial \theta_j}(x_i, \theta^*) \right. \\ &\quad \left. - E\left[e_i \frac{\partial y^s}{\partial \theta_j}(x_i, \theta) - e_i \frac{\partial y^s}{\partial \theta_j}(x_i, \theta^*) \right] \right\} \\ &= n^{-1/2} \sum_{i=1}^{n} \left\{ e_i \frac{\partial y^s}{\partial \theta_j}(x_i, \theta) - e_i \frac{\partial y^s}{\partial \theta_j}(x_i, \theta^*) \right\}, \end{aligned}$$

where $\theta \in U$. By A5, the set $\{f_\theta \in C(\mathbf{R}\times\Omega) : f_\theta(e,x) = e\frac{\partial y^s}{\partial\theta_j}(x,\theta) - e\frac{\partial y^s}{\partial\theta_j}(x,\theta^*), \theta \in U\}$ is a Donsker class. This ensures that $E_{2n}(\cdot)$ weakly converges in $L_\infty(U)$ to a tight Guassian process, denoted by $G(\cdot)$. Without loss of generality, we assume that $G(\cdot)$ has continuous sample paths. We note that $G(\theta^*) = 0$ because $E_{2n}(\theta^*)$ for all n. Then as a consequence of the consistency of $\hat{\theta}_n^{L_2}$ and the continuous mapping theorem [van der Vaart (1998)], $E_{2n}(\hat{\theta}_n^{L_2}) \to^p G(\theta^*) = 0$, which gives

$$D_n = \frac{1}{n}\sum_{i=1}^{n} e_i \frac{\partial y^s}{\partial\theta_j}(x_i,\theta^*) + o_p(n^{-1/2}). \tag{3.17}$$

Finally, we estimate E_n. Applying A5, B3, B4, we have

$$E_n \le \lambda_n \|\hat{\zeta}\|_{\mathcal{N}_\Phi(\Omega)} \left\| \frac{\partial y^s}{\partial\theta_j}(\cdot,\hat{\theta}) \right\|_{\mathcal{N}_\Phi(\Omega)} = o_p(n^{-1/2}). \tag{3.18}$$

By combining (3.12), (3.15), (3.16), (3.17) and (3.18), we prove the desired result. □

Theorem 1 implies the asymptotic normality of $\sqrt{n}(\hat{\theta}_n^{L_2} - \theta^*)$, provided that

$$W := E\left[\frac{\partial y^s}{\partial\theta}(x_i,\theta^*)\frac{\partial y^s}{\partial\theta^{\mathrm{T}}}(x_i,\theta^*)\right] \tag{3.19}$$

is positive definite. Specifically,

$$\sqrt{n}(\hat{\theta}_n^{L_2} - \theta^*) \xrightarrow{d} N(0, 4\sigma^2 V^{-1} W V^{-1}). \tag{3.20}$$

3.3. *Semiparametric efficiency.* In this section, we discuss the efficiency of the proposed L_2 calibration. It will be shown that, as a semiparametric method, the L_2 calibration method reaches the highest possible efficiency if the measurement errors follow a normal distribution.

In statistics, a parametric model is one whose parameter space is finite dimensional, while a nonparametric model is one with an infinite dimensional parameter space. The definition of semiparametric models is, nevertheless, more complicated. Refer to Bickel et al. (1993), Groeneboom and Wellner (1992) for details. In simple terms, a semiparametric problem has an infinite dimensional parameter space but the parameter of interest in this problem is only finite dimensional. The calibration problems under consideration are semiparametric. To see this, consider the calibration model given by (2.1) and (2.2). The parameter space of model (2.1) contains an infinite dimensional function space which covers ζ. On the other hand, the parameter of interest is θ^* in (2.2), which is q-dimensional.

Now we briefly review the estimation efficiency in semiparametric problems. For details, we refer to Bickel et al. (1993), Kosorok (2008). Let Ξ be an infinite

dimensional parameter space whose true value is denoted by ξ_0. Denote the feature of interest as $\nu(\xi_0)$ with a known map $\nu : \Xi \mapsto \mathbf{R}^d$. Suppose T_n is an estimator for $\nu(\xi_0)$ based on n independent samples and that $\sqrt{n}(T_n - \nu(\xi_0))$ is asymptotically normal. Now let Ξ_0 be an arbitrary finite dimensional subset of Ξ satisfying $\xi_0 \in \Xi_0$. We consider the statistical estimation problem with the same observed data but with the parameter space Ξ_0. Under this parametric assumption and some other regularity conditions, an efficient estimator can be obtained by using the maximum likelihood (ML) method, denoted by $S_n^{\Xi_0}$. Since the construction of $S_n^{\Xi_0}$ uses more assumptions than T_n, the asymptotic variance of $S_n^{\Xi_0}$ should be less than or equal to that of T_n. We call T_n *semiparametric efficient* if there exists a Ξ_0 such that $S_n^{\Xi_0}$ has the same asymptotic variance as T_n.

For the calibration problem given by (2.1) and (2.2), consider the following q-dimensional parametric model indexed by γ:

$$\zeta_\gamma(\cdot) = \zeta(\cdot) + \gamma^{\mathrm{T}} \frac{\partial y^s}{\partial \theta}(\cdot, \theta^*), \tag{3.21}$$

with $\gamma \in \mathbf{R}^q$. Then (2.1) and (3.21) form a linear regression model. Regarding (2.1), the true value of γ is $\gamma_0 = 0$. Suppose that e_i in (2.1) follows $N(0, \sigma^2)$ with an unknown σ^2. Under the regularity conditions of Theorem 1, the ML estimator for observations $\{(x_i, y_i)\}_{i=1}^n$ is the least squares estimator, with the asymptotic expression

$$\hat{\gamma}_n = \frac{1}{n} W^{-1} \sum_{i=1}^n e_i \frac{\partial y^s}{\partial \theta}(x_i, \theta^*) + o_p(n^{-1/2}), \tag{3.22}$$

where W is defined in (3.19). Then a natural estimator for θ^* in (2.2) is

$$\hat{\theta}_n = \operatorname*{argmin}_{\theta \in \Theta} \| \zeta_{\hat{\gamma}_n}(\cdot) - y^s(\cdot, \theta) \|_{L_2(\Omega)}. \tag{3.23}$$

Again, we simplify the problem in (3.23) by assuming that y^s is a known function. The asymptotic variance of $\hat{\theta}_n$ can be obtained by the delta method. As in A1, assume that x_i follows the uniform distribution over Ω. Then $\| f \|_{L_2(\Omega)}^2 = E f^2(x_i)$ for all f. Define

$$\theta(t) = \operatorname*{argmin}_{\theta \in \Theta} E\big[\zeta_t(x_i) - y^s(x_i, \theta)\big]^2, \tag{3.24}$$

for each t near 0. Let

$$\begin{aligned} \Psi(\theta, t) &= \frac{\partial}{\partial \theta} E\big[\zeta_t(x_i) - y^s(x_i, \theta)\big]^2 \\ &= \frac{\partial}{\partial \theta} E\left[\zeta(x_i) - t^{\mathrm{T}} \frac{\partial y^s}{\partial \theta}(x_i, \theta^*) - y^s(x_i, \theta)\right]^2. \end{aligned}$$

Then (3.24) implies $\Psi(\theta(t), t) = 0$ for all t near 0. From the implicit function theorem, we have

$$\begin{aligned}
&\frac{\partial \theta(t)}{\partial t^{\mathrm{T}}}\bigg|_{t=0} \\
&\quad = -\left(\frac{\partial \Psi}{\partial \theta^{\mathrm{T}}}(\theta^*, 0)\right)^{-1} \frac{\partial \Psi}{\partial t^{\mathrm{T}}}(\theta^*, 0) \\
&\quad = -\left(E\frac{\partial^2}{\partial \theta\, \partial \theta^{\mathrm{T}}}\left[\zeta(x_i) - y^s(x_i, \theta^*)\right]^2\right)^{-1} 2E\left[\frac{\partial y^s}{\partial \theta^{\mathrm{T}}}(x_i, \theta^*)\frac{\partial y^s}{\partial \theta}(x_i, \theta^*)\right] \\
&\quad = -2V^{-1}W.
\end{aligned} \tag{3.25}$$

By the delta method,

$$\hat{\theta}_n - \theta^* = \theta(\hat{\gamma}_n) - \theta(0) = \frac{\partial \theta(t)}{\partial t^{\mathrm{T}}}\bigg|_{t=0} \hat{\gamma}_n + o_p\big(n^{-1/2}\big), \tag{3.26}$$

which, together with (3.22) and (3.25), yields

$$\hat{\theta}_n - \theta^* = -2V^{-1} \sum_{i=1}^{n} e_i \frac{\partial y^s}{\partial \theta}(x_i, \theta^*) + o_p\big(n^{-1/2}\big). \tag{3.27}$$

Noting that the asymptotic expression of the L_2 calibration given by (3.7) has the same form as the ML estimator for the parametric model in (3.27), we obtain the following theorem.

THEOREM 2. *Under the assumptions of Theorem 1, if e_i in (2.1) follows a normal distribution, then the L_2 calibration (3.2) is semiparametric efficient.*

Since the normal distribution is commonly used to model the random error in physical experiments [see, e.g., Wu and Hamada (2009)] and the calibration for computer experiments [see, e.g., Kennedy and O'Hagan (2001)]. Theorem 2 suggests that the proposed method is efficient for many practical problems. For non-normal error distributions, the ML estimator does not agree with the least squares estimator. Thus, the ML estimator cannot be expressed by (3.22). Consequently, the L_2 calibration defined by (3.1) and (3.2) is not semiparametric efficient. However, if the random error is from a parametric model, the proposed L_2 calibration can be modified to achieve the semiparametric efficiency. Denote the likelihood function of e_i by $l(\beta; e_i)$ with $\beta \in \mathcal{B}$, that is, e_i has a density $l(\beta_0; \cdot)$ for some unknown $\beta_0 \in \mathcal{B}$. Suppose $L(\cdot; x) := \log l(\cdot; x)$ is convex for all x. Then the penalized ML estimator for ζ is

$$\hat{\zeta}^{\mathrm{ML}} := \operatorname*{argmin}_{\beta \in \mathcal{B},\, f \in \mathcal{N}_\Phi(\Omega)} \frac{1}{n} \sum_{i=1}^{n} L\big(\beta; y_i - f(x_i)\big) + \lambda \|f\|^2_{\mathcal{N}_\Phi(\Omega)}. \tag{3.28}$$

Then define the modified L_2 calibration as

$$\hat{\theta}^{\mathrm{ML}} := \operatorname*{argmin}_{\theta\in\Theta} \|\hat{\zeta}^{\mathrm{ML}}(\cdot) - \hat{y}^s(\cdot,\theta)\|_{L_2(\Omega)}. \tag{3.29}$$

By using similar arguments, it can be proved that, under some regularity conditions, $\hat{\theta}^{\mathrm{ML}}$ is semiparametric efficient. For a related discussion, we refer to Shen (1997).

4. Ordinary least squares. In this section, we will study an alternative method, namely, the ordinary least squares (OLS) calibration. There are several versions of the OLS method discussed in statistics and applied mathematics for calibration problems and inverse problems [e.g., Evans and Stark (2002), Joseph and Melkote (2009)]. Here, we consider a general form, which is apparently new but covers the existing versions. As before, let $\hat{y}_n^s$ be a sequence of surrogate models for y^s. Define the OLS estimator for the calibration parameter as

$$\hat{\theta}_n^{\mathrm{OLS}} = \operatorname*{argmin}_{\theta\in\Theta} \sum_{i=1}^{n} \bigl(y_i^p - \hat{y}_n^s(x_i,\theta)\bigr)^2, \tag{4.1}$$

where x_i's and y_i's are from model (2.1).

Obviously, the OLS calibration is a natural choice when there is no difference between the true process and the optimal computer output, that is, $\zeta(\cdot) = y^s(\cdot,\theta^*)$. However, we are particularly interested in the asymptotic behavior of the OLS calibration when $\zeta(\cdot)$ and $y^s(\cdot,\theta^*)$ are different.

Analogous to Theorem 1 for the L_2 calibration, we have the following theorem on the asymptotic behavior of the OLS calibration.

THEOREM 3. *In addition to conditions* A1–A4 *and* C1–C2, *suppose that there exists a neighborhood* U *of* θ^*, *such that* $y^s(x,\cdot) \in C^{2,1}(U)$ *for all* $x \in \Omega$, *where* $C^{2,1}$ *denotes the space of functions whose second derivatives are Lipschitz. Then*

$$\hat{\theta}_n^{\mathrm{OLS}} - \theta^* = V^{-1}\left\{\frac{1}{n}\sum_{i=1}^{n} \frac{\partial}{\partial\theta}\bigl(y_i^p - y^s(x_i,\theta^*)\bigr)^2\right\} + o_p(n^{-1/2}).$$

PROOF. First, we prove $\hat{\theta}_n^{\mathrm{OLS}} \to^p \theta^*$. By condition A2, it suffices to show that

$$\begin{aligned}&\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{n}\bigl(y_i^p - \hat{y}_n^s(x_i,\theta)\bigr)^2 - \bigl(\|\zeta(\cdot) - y^s(\cdot,\theta)\|_{L_2(\Omega)}^2 + \sigma^2\bigr)\right| \\ &\quad\xrightarrow{p} 0.\end{aligned} \tag{4.2}$$

Note that

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n}\bigl(y_i^p-\hat{y}_n^s(x_i,\theta)\bigr)^2-\frac{1}{n}\sum_{i=1}^{n}\bigl(y_i^p-y^s(x_i,\theta)\bigr)^2\right| \\
&\quad=\left|\frac{1}{n}\sum_{i=1}^{n}\bigl(y^s(x_i,\theta)-\hat{y}^s(x_i,\theta)\bigr)\bigl(2y_i^p-y^s(x_i,\theta)-\hat{y}^s(x_i,\theta)\bigr)\right| \qquad (4.3)\\
&\quad\le\|y^s-\hat{y}^s\|_{L_\infty(\Omega)}\left(\frac{1}{n}\sum_{i=1}^{n}2\bigl(y_i^p-y^s(x_i,\theta)\bigr)+\|y^s-\hat{y}^s\|_{L_\infty(\Omega)}\right).
\end{aligned}$$

Since Θ is compact, the uniform law of large numbers [van der Vaart and Wellner (1996)] implies

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{n}\bigl(y_i^p-y^s(x_i,\theta)\bigr)-E\bigl[y_i^p-y^s(x_i,\theta)\bigr]\right|\xrightarrow{p}0 \qquad (4.4)$$

and

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{n}\bigl(y_i^p-y^s(x_i,\theta)\bigr)^2-E\bigl[y_i^p-y^s(x_i,\theta)\bigr]^2\right|\xrightarrow{p}0. \qquad (4.5)$$

Direct calculations give

$$E\bigl[y_i^p-y^s(x_i,\theta)\bigr]^2=\int_\Omega\bigl(\zeta(z)-y^s(z,\theta)\bigr)^2\,dz+\sigma^2, \qquad (4.6)$$

which, together with (4.3), (4.4), (4.5) and (4.6), proves (4.2).

By definition (4.1), condition A2 and the consistency of $\hat{\theta}_n^{\mathrm{OLS}}$, we have

$$\begin{aligned}
0&=\frac{\partial}{\partial\theta}\left\{\frac{1}{n}\sum_{i=1}^{n}\bigl(y_i^p-\hat{y}_n^s(x_i,\hat{\theta}_n^{\mathrm{OLS}})\bigr)^2\right\}\\
&=\frac{2}{n}\sum_{i=1}^{n}\frac{\partial\hat{y}_n^s}{\partial\theta}(x_i,\hat{\theta}_n^{\mathrm{OLS}})\{\hat{y}_n^s(x_i,\hat{\theta}_n^{\mathrm{OLS}})-y_i^p\},
\end{aligned}$$

which, together with the law of large numbers and conditions C1 and C2, yield

$$\begin{aligned}
o_p(n^{-1/2})&=\frac{2}{n}\sum_{i=1}^{n}\frac{\partial y^s}{\partial\theta}(x_i,\hat{\theta}_n^{\mathrm{OLS}})\{y^s(x_i,\hat{\theta}_n^{\mathrm{OLS}})-y_i^p\}\\
&=\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\bigl(y_i^p-y^s(x_i,\hat{\theta}_n^{\mathrm{OLS}})\bigr)^2.
\end{aligned}$$

The remainder of the proof follows from some direct calculations using the standard asymptotic theory for Z-estimators [van der Vaart (1998)]. $\square$

Theorem 3 shows that the OLS calibration is consistent even if the computer code is imperfect. Compared with the L_2 calibration, the OLS calibration is computationally more efficient. Also, the OLS calibration does not require tuning, while in the L_2 calibration the value of the tuning parameter λ in (2.6) needs to be determined. However, according to Theorem 3, the asymptotic variance of the OLS calibration does not reach the semiparametric lower bound given by (3.27).

We now study the conditions under which the L_2 calibration and the OLS calibration are asymptotically equivalent. Let $\Sigma_1 = 4\sigma^2 W$. Then the asymptotic variance of the L_2 calibration given by (3.20) is $V^{-1}\Sigma_1 V^{-1}$. Let

$$
\begin{aligned}
\Sigma_2 &= E\left[\frac{\partial}{\partial\theta}\bigl(y_i^p - y^s(x_i,\theta^*)\bigr)^2\right]^2 \\
&= 4E\left[\bigl(e_i + \zeta(x_i) - y^s(x_i,\theta^*)\bigr)^2 \frac{\partial y^s}{\partial\theta}(x_i,\theta^*)\frac{\partial y^s}{\partial\theta^{\mathrm{T}}}(x_i,\theta^*)\right] \qquad (4.7)\\
&= 4\sigma^2 W + 4E\left[\bigl(\zeta(x_i) - y^s(x_i,\theta^*)\bigr)^2 \frac{\partial y^s}{\partial\theta}(x_i,\theta^*)\frac{\partial y^s}{\partial\theta^{\mathrm{T}}}(x_i,\theta^*)\right].
\end{aligned}
$$

Then Theorem 3 shows that the asymptotic variance for the OLS calibration is $V^{-1}\Sigma_2 V^{-1}$. From (4.7), it is seen that $\Sigma_2 - \Sigma_1 \geq 0$. Additionally, $\Sigma_1 = \Sigma_2$ if and only if

$$
E\left[\bigl(\zeta(x_i) - y^s(x_i,\theta^*)\bigr)^2 \frac{\partial y^s}{\partial\theta}(x_i,\theta^*)\frac{\partial y^s}{\partial\theta^{\mathrm{T}}}(x_i,\theta^*)\right] = 0. \qquad (4.8)
$$

Suppose $\frac{\partial y^s}{\partial\theta}(x,\theta^*) \neq 0$ for all $x \in \Omega$. Then (4.8) holds only if $\zeta(x) = y^s(x,\theta^*)$ for almost every $x \in \Omega$, that is, there exists a perfect computer model. In this case, the OLS calibration has the same asymptotic distribution as the L_2 calibration. However, as suggested by Kennedy and O'Hagan (2001), the bias between $y^s(\cdot,\theta^*)$ and $\zeta(\cdot)$ can be large in practical situations. Thus, in general the OLS calibration is less efficient than the L_2 calibration.

5. Numerical studies. In this section, we compare the numerical behaviors of three methods for the estimation of the calibration parameters: the L_2 calibration, the OLS calibration and a version of the method proposed by Kennedy and O'Hagan (2001). The original version of the Kennedy–O'Hagan (abbreviated as KO) method is a Bayesian approach. In order to compare with the proposed frequentist methods, we consider the frequentist version of the KO method stated in Tuo and Wu (2014), where the maximum likelihood estimation is used.

5.1. *Example* 1: *Perfect computer model*. Suppose the true process is

$$
\zeta(x) = \exp(x/10)\sin x, \qquad (5.1)
$$

for $x \in \Omega = (0, 2\pi)$. The physical observations are given by

$$
y_i^p = \zeta(x_i) + e_i, \qquad (5.2)
$$

TABLE 1
Numerical comparison for perfect computer model. MSE = mean square error

True value:	$\sigma^2 = 0.1$		$\sigma^2 = 1$	
	−1		−1	
	Mean	**MSE**	**Mean**	**MSE**
L_2	−0.9990	6.497×10^{-5}	−0.8876	0.0906
OLS	−0.9999	1.160×10^{-4}	−0.9306	0.0908
KO	−0.9993	8.065×10^{-5}	−0.9325	0.0468

with

$$x_i = 2\pi i/50, \qquad e_i \sim N(0, \sigma^2) \qquad \text{for } i = 0, \ldots, 50. \tag{5.3}$$

We will consider two levels of σ^2 (with $\sigma^2 = 0.1$ and $\sigma^2 = 1$) so that the numerical stability of the methods with different noise levels is investigated.

Suppose the computer output is

$$y^s(x, \theta) = \zeta(x) - |\theta + 1|(\sin\theta x + \cos\theta x). \tag{5.4}$$

Then we have $\zeta(\cdot) = y^s(\cdot, -1)$. Thus, $\theta^* = -1$. And there is no discrepancy between $\zeta(\cdot)$ and $y^s(\cdot, \theta^*)$, that is, the computer model is perfect. For simplicity, we suppose that (5.4) is a known function so that we do not need an emulator for it.

We conducted 1000 random simulations to examine the performance of the L_2 calibration, the OLS calibration and the KO calibration for $\sigma^2 = 0.1$ and $\sigma^2 = 1$, respectively. For the L_2 calibration and the KO calibration, the Gaussian correlation family $\Phi(x_1, x_2) = \exp\{-\phi(x_1 - x_2)^2\}$ is used with the model parameter ϕ chosen by the cross-validation method [Rasmussen and Williams (2006), Santner, Williams and Notz (2003)]. The tuning parameters in the nonparametric regression is selected by the generalized cross validation [Wahba (1990)].

Table 1 shows the simulation results. The results for $\sigma^2 = 0.1$ and $\sigma^2 = 1$ are given in columns 2–3 and 4–5, respectively. The true values of θ^* are given in the second row. The last three rows give the mean value and the mean square error (MSE) over 1000 random simulations for the three methods.

It can be seen from Table 1 that all three methods give good estimation results in this example. The good performance of the KO method is not surprising because the computer model here is perfect. In their theoretical study on the KO method with deterministic physical experiments, Tuo and Wu (2014) obtained the limiting value of the KO method under certain conditions. Using Theorem 1 of Tuo and Wu (2014), it can be seen that, for deterministic physical experiments, the Kennedy–O'Hagan method would be consistent if the computer model is perfect. The simulation results in this example suggest that this statement may also hold for stochastic physical systems.

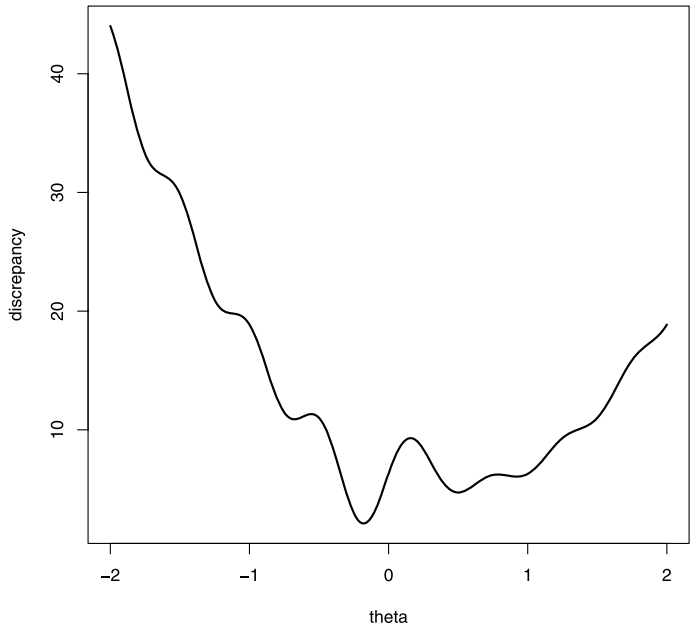

FIG. 1. *L_2 discrepancy function in Example* 2.

5.2. *Example* 2: *Imperfect computer model*. Now we consider an example with an imperfect computer model. Suppose the true process and the physical observations are the same as in Example 1, given by (5.1), (5.2) and (5.3). Suppose the computer model is

$$y^s(x,\theta) = \zeta(x) - \sqrt{\theta^2 - \theta + 1}(\sin\theta x + \cos\theta x). \tag{5.5}$$

As in Example 1, we suppose y^s is known. From (5.5), it can be seen that there does not exist a real number θ satisfying $y^s(\cdot,\theta) = \zeta(\cdot)$, because the quadratic function $\theta^2 - \theta + 1$ is always positive. Thus, this computer model is imperfect.

The L_2 discrepancy between the computer model and the physical model has an explicit form:

$$\|\zeta - y^s(\cdot,\theta)\|^2_{L_2(\Omega)} = (\theta^2 - \theta + 1)\left(2\pi - \frac{\cos(4\pi\theta) - 1}{2\theta}\right), \tag{5.6}$$

with a continuous extension at $\theta = 0$. Figure 1 plots the function (5.6) with $-2 < \theta < 2$. Numerical optimization shows that the minimizer of (5.6) is $\theta^* \approx -0.1789$.

As in Example 1, we conducted 1000 random simulations to compare the L_2 calibration, the OLS calibration and the KO calibration. We keep the remaining setup of this experiment the same as in Example 1. The mean value and standard deviation (SD) over 1000 simulations are shown in Table 2.

It can be seen from Table 2 that the L_2 calibration and the OLS calibration outperform the KO calibration. Furthermore, the mean value of the KO estimator

TABLE 2
Numerical comparison for imperfect computer model. SD = standard deviation

True value:	$\sigma^2 = 0.1$		$\sigma^2 = 1$	
	−0.1789		**−0.1789**	
	Mean	**SD**	**Mean**	**SD**
L_2	−0.1792	2.665×10^{-3}	−0.1773	0.0711
OLS	−0.1770	2.674×10^{-3}	−0.1684	0.1060
KO	−0.1224	7.162×10^{-3}	0.0034	0.3244

changes a lot as σ^2 changes. This is undesirable because a good estimator should not be sensitive to random error for large samples. Table 2 also shows that the standard deviation of the L_2 calibration is smaller than that of the OLS calibration. This agrees with our theoretical analysis, which shows that the L_2 calibration is more efficient than the OLS calibration for imperfect computer models. Overall, the KO calibration underperforms the L_2 calibration or the OLS calibration.

6. Concluding remarks and further discussions. In this work, we extend the framework established in Tuo and Wu (2014) to stochastic physical systems. We propose a novel method, called the L_2 calibration, and prove its asymptotic normality and semiparametric efficiency. We also study the OLS method and prove that it is consistent but not efficient. Although the OLS calibration is computationally less costly, the L_2 calibration should be seriously considered because of its high estimation efficiency. By using a more efficient estimator, fewer physical trials are needed to achieve the same estimation efficiency. In most practical problems, physical experiments are more expensive to run. Therefore, it would be worthwhile to save the physical runs by doing more computation. Thus, we recommend using the L_2 calibration over the OLS calibration.

Because of the identifiability problem in calibration, we define the purpose of calibration as that of finding the L_2 projection, that is, the parameter value which minimizes the discrepancy between the true process and the computer output under the L_2 norm. Noting that the "true" value of the calibration parameter in our framework depends on the choice of the norm, one may also consider the asymptotic results for calibration under a different norm. After some calculations, it can be shown that the main results of this work still hold if the new norm is equivalent to the L_2 norm. However, if a norm that is not equivalent to the L_2 norm, such as the L_∞ norm, is used, the idea in the proof of Theorem 1 will not work. We believe that, for those norms, there do not exist estimators with convergence rate $O(n^{-1/2})$. This will require further work.

We have reported the asymptotic properties of the L_2 calibration under the random design, that is, x_i are sampled independently from the uniform distribution.

Given the fact that many physical experiments are conducted under fixed designs [see books by Box, Hunter and Hunter (2005) and Wu and Hamada (2009)], the results for calibration for fixed designs need further investigation.

Acknowledgements. The authors are grateful to the Associate Editor and the referees for helpful comments.

REFERENCES

ANDERSON-COOK, C. M. and PREWITT, K. (2005). Some guidelines for using nonparametric methods for modeling data from response surface designs. *Journal of Modern Applied Statistical Methods* **4** 106–119.

BAYARRI, M. J., BERGER, J. O., CAFEO, J., GARCIA-DONATO, G., LIU, F., PALOMO, J., PARTHASARATHY, R. J., PAULO, R., SACKS, J. and WALSH, D. (2007a). Computer model validation with functional output. *Ann. Statist.* **35** 1874–1906. MR2363956

BAYARRI, M. J., BERGER, J. O., PAULO, R., SACKS, J., CAFEO, J. A., CAVENDISH, J., LIN, C.-H. and TU, J. (2007b). A framework for validation of computer models. *Technometrics* **49** 138–154. MR2380530

BERLINET, A. and THOMAS-AGNAN, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics* Kluwer, Boston, MA. MR2239907

BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. and WELLNER, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Univ. Press, Baltimore, MD. MR1245941

BOX, G. E. P., HUNTER, J. S. and HUNTER, W. G. (2005). *Statistics for Experimenters*: *Design*, *Innovation*, *and Discovery*, 2nd ed. Wiley, Hoboken, NJ. MR2140250

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. Wiley, New York. MR1239641

EDMUNDS, D. E. and TRIEBEL, H. (1996). *Function Spaces*, *Entropy Numbers*, *Differential Operators*. Cambridge Univ. Press, Cambridge. MR1410258

EVANS, S. N. and STARK, P. B. (2002). Inverse problems as statistics. *Inverse Probl.* **18** R55–R97. MR1929274

GOH, J., BINGHAM, D., HOLLOWAY, J. P., GROSSKOPF, M. J., KURANZ, C. C. and RUTTER, E. (2013). Prediction and computer model calibration using outputs from multifidelity simulators. *Technometrics* **55** 501–512. MR3176554

GOLDSTEIN, M. and ROUGIER, J. (2004). Probabilistic formulations for transferring inferences from mathematical models to physical systems. *SIAM J. Sci. Comput.* **26** 467–487 (electronic). MR2116356

GRAMACY, R. B. and LEE, H. K. H. (2012). Cases for the nugget in modeling computer experiments. *Stat. Comput.* **22** 713–722. MR2909617

GROENEBOOM, P. and WELLNER, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation* **19**. Birkhäuser, Basel. MR1180321

HAN, G., SANTNER, T. J. and RAWLINSON, J. J. (2009). Simultaneous determination of tuning and calibration parameters for computer experiments. *Technometrics* **51** 464–474. MR2756481

HIGDON, D., KENNEDY, M., CAVENDISH, J. C., CAFEO, J. A. and RYNE, R. D. (2004). Combining field data and computer simulations for calibration and prediction. *SIAM J. Sci. Comput.* **26** 448–466. MR2116355

HIGDON, D., GATTIKER, J., WILLIAMS, B. and RIGHTLEY, M. (2008). Computer model calibration using high-dimensional output. *J. Amer. Statist. Assoc.* **103** 570–583. MR2523994

HIGDON, D., GATTIKER, J., LAWRENCE, E., JACKSON, C., TOBIS, M., PRATOLA, M., HABIB, S., HEITMANN, K. and PRICE, S. (2013). Computer model calibration using the ensemble Kalman filter. *Technometrics* **55** 488–500. MR3176553

JOSEPH, V. R. and MELKOTE, S. N. (2009). Statistical adjustments to engineering models. *Journal of Quality Technology* **41** 362–375.

KENNEDY, M. C. and O'HAGAN, A. (2001). Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 425–464. MR1858398

KOSOROK, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York. MR2724368

MAMMEN, E. and VAN DE GEER, S. (1997). Penalized quasi-likelihood estimation in partial linear models. *Ann. Statist.* **25** 1014–1035. MR1447739

MURPHY, J. M., BOOTH, B. B. B., COLLINS, M., HARRIS, G. R., SEXTON, D. M. H. and WEBB, M. J. (2007). A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **365** 1993–2028.

MYERS, R. H. (1999). Response surface methodlogy: Current status and future directions. *Journal of Quality Technology* **31** 30–44.

PENG, C.-Y. and WU, C. F. J. (2014). On the choice of nugget in kriging modeling for deterministic computer experiments. *J. Comput. Graph. Statist.* **23** 151–168. MR3173765

RASMUSSEN, C. E. and WILLIAMS, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA. MR2514435

SANTNER, T. J., WILLIAMS, B. J. and NOTZ, W. I. (2003). *The Design and Analysis of Computer Experiments*. Springer, New York. MR2160708

SCHÖLKOPF, B., HERBRICH, R. and SMOLA, A. J. (2001). A generalized representer theorem. In *Computational Learning Theory* (*Amsterdam*, 2001). *Lecture Notes in Comput. Sci.* **2111** 416–426. Springer, Berlin. MR2042050

SHEN, X. (1997). On methods of sieves and penalization. *Ann. Statist.* **25** 2555–2591. MR1604416

STEIN, M. L. (1999). *Interpolation of Spatial Data*: *Some Theory for Kriging*. Springer, New York. MR1697409

STONE, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053. MR0673642

TUO, R. and WU, C. F. J. (2014). A theoretical framework for calibration in computer models: Parametrization, estimation and convergence properties. Technical report, Chinese Acad. Sci. and Georgia Inst. Technol.

VAN DE GEER, S. (2000). *Empirical Processes in M-Estimation* **45**. Cambridge Univ. Press, Cambridge.

VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*: *With Applications to Statistics*. Springer, New York. MR1385671

WAHBA, G. (1990). *Spline Models for Observational Data* **59**. SIAM, Philadelphia, PA. MR1045442

WANG, S., CHEN, W. and TSUI, K.-L. (2009). Bayesian validation of computer models. *Technometrics* **51** 439–451. MR2756479

WENDLAND, H. (2005). *Scattered Data Approximation*. *Cambridge Monographs on Applied and Computational Mathematics* **17**. Cambridge Univ. Press, Cambridge. MR2131724

WU, C. J. and HAMADA, M. S. (2009). *Experiments*: *Planning*, *Analysis*, *and Optimization* 2nd ed. Wiley, New York. MR2583259

XIU, D. (2010). *Numerical Methods for Stochastic Computations*: *A Spectral Method Approach*. Princeton Univ. Press, Princeton, NJ. MR2723020

ACADEMY OF MATHEMATICS
AND SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
BEIJING
CHINA 100190
E-MAIL: tuorui@amss.ac.cn

SCHOOL OF INDUSTRIAL
AND SYSTEMS ENGINEERING
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332-0205
USA
E-MAIL: jeffwu@isye.gatech.edu