SUPPLEMENTARY MATERIAL: PREDICTION BASED ON THE KENNEDY-O'HAGAN CALIBRATION MODEL: ASYMPTOTIC CONSISTENCY AND OTHER PROPERTIES

Chinese Academy of Sciences and Georgia Institute of Technology

Supplementary Material

In this supplementary material, we present the proofs for Lemma 1, Lemma 2, Lemma 4 and Theorem 2 in the main article.

S1 Technical Proofs

Proof of Lemma 1. We first assume $f \in F_\Phi$. If $f = s_{f,\mathbf{x}}$, there is nothing to prove. If $f \neq s_{f,\mathbf{x}}$, without loss of generality, we write

$$f(x) = \sum_{i=1}^{n+m} \alpha_i \Phi(x, x_i),$$

for an extra set of distinct points $\{x_{n+1}, \ldots, x_{n+m}\} \subset \Omega$. Now partition $(A_{i,j}) = \Phi(x_i, x_j), 1 \leq i, j \leq n + m$ into

$$A = \begin{pmatrix} (A_1)_{n\times n} & (A_2)_{n\times m} \\ (A_3)_{m\times n} & (A_4)_{m\times m} \end{pmatrix},$$

where $A_3 = A_2^{\mathrm{T}}$ because Φ is symmetric.

Let $\mathbf{y} = (f(x_1), \ldots, f(x_n))^{\mathrm{T}}, \mathbf{a}_1 = (\alpha_1, \ldots, \alpha_n)^{\mathrm{T}}, \mathbf{a}_2 = (\alpha_{n+1}, \ldots, \alpha_{n+m})^{\mathrm{T}}$. Clearly, $\mathbf{y} = A_1\mathbf{a}_1 + A_2\mathbf{a}_2$. By the definition of $s_{f,\mathbf{x}}$, we have

$$s_{f,\mathbf{x}}(x) = \sum_{i=1}^{n} u_i \Phi(x, x_i),$$

with $\mathbf{u} = (u_1, \ldots, u_n)^{\mathrm{T}}$ satisfying $\mathbf{y} = A_1\mathbf{u}$. Then from (**??**) we obtain

$$\begin{aligned}
& \langle s_{f,\mathbf{x}}, f - s_{f,\mathbf{x}} \rangle_{\mathcal{N}_\Phi(\Omega)} \\
= & \left\langle \sum_{i=1}^{n} u_i \Phi(x, x_i), \sum_{i=1}^{n} (\alpha_i - u_i)\Phi(x, x_i) + \sum_{i=n+1}^{n+m} \alpha_i \Phi(x, x_i) \right\rangle_{\mathcal{N}_\Phi(\Omega)} \\
= & \begin{pmatrix} \mathbf{u}^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 - \mathbf{u} \\ \mathbf{a}_2 \end{pmatrix} \\
= & \mathbf{u}^{\mathrm{T}} (A_1\mathbf{a}_1 + A_2\mathbf{a}_2 - A_1\mathbf{u}) \\
= & \mathbf{u}^{\mathrm{T}} (\mathbf{y} - \mathbf{y}) = 0.
\end{aligned} \tag{S1.1}$$

For a general $f \in \mathcal{N}_\Phi(\Omega)$, we can find a sequence $f_n \in F_\Phi$ with $f_n \to f$ in $\mathcal{N}_\Phi(\Omega)$ as $n \to \infty$. The desired result then follows from a limiting form of (S1.1). □

Proof of Lemma 2. For any $g \in L_2(\mathbf{R}^d) \cap C(\mathbf{R}^d)$, its native norm admits the representation

$$\|g\|^2_{\mathcal{N}_\Phi(\mathbf{R}^d)} = (2\pi)^{-d/2} \int_{\mathbf{R}^d} \frac{|\tilde{g}(\omega)|^2}{\tilde{\Phi}(\omega)} d\omega, \tag{S1.2}$$

where $\tilde{g}$ and $\tilde{\Phi}$ denote the Fourier transforms of g and Φ respectively. See Theorem 10.12 of Wendland [2005]. The (fractional) Sobolev norms have a similar representation

$$\|g\|^2_{H^s(\mathbf{R}^d)} = (2\pi)^{-d/2} \int_{\mathbf{R}^d} |\tilde{g}(\omega)|^2 (1 + \|\omega\|^2)^s d\omega. \tag{S1.3}$$

See Adams and Fournier [2003] for details. Tuo and Wu [2016] show that

$$\tilde{C}_{\upsilon,\gamma}(\omega) = 2^{d/2}(4\upsilon\gamma^2)^{\upsilon} \frac{\Gamma(\upsilon + d/2)}{\Gamma(\nu)} (4\upsilon\gamma^2 + \|\omega\|^2)^{-(\upsilon+d/2)}.$$

Using the inequality

$$(1 + b)\min(1, a) \le a + b \le (1 + b)\max(1, a),$$

for $a, b \ge 0$, we obtain

$$\begin{aligned}
\tilde{C}_{\upsilon,\gamma}(\omega) &\le 2^{d/2}(4\upsilon\gamma^2)^{\upsilon} \frac{\Gamma(\upsilon + d/2)}{\Gamma(\upsilon)} \max\left\{1, (4\upsilon\gamma^2)^{-(\upsilon+d/2)}\right\} (1 + \|\omega\|^2)^{-(\upsilon+d/2)} \\
&\le 2^{d/2} \frac{\Gamma(\upsilon + d/2)}{\Gamma(\upsilon)} \max\left\{(4\upsilon\gamma_2^2)^{\upsilon}, (4\upsilon\gamma_1^2)^{-d/2}\right\} (1 + \|\omega\|^2)^{-(\upsilon+d/2)} \\
&=: C_1(1 + \|\omega\|^2)^{-(\upsilon+d/2)}, \qquad \text{(S1.4)}
\end{aligned}$$

and

$$\begin{aligned}
\tilde{C}_{\upsilon,\gamma}(\omega) &\ge 2^{d/2}(4\upsilon\gamma^2)^{\upsilon} \frac{\Gamma(\upsilon + d/2)}{\Gamma(\upsilon)} \min\left\{1, (4\upsilon\gamma^2)^{-(\upsilon+d/2)}\right\} (1 + \|\omega\|^2)^{-(\upsilon+d/2)} \\
&\ge 2^{d/2} \frac{\Gamma(\upsilon + d/2)}{\Gamma(\upsilon)} \min\left\{(4\upsilon\gamma_1^2)^{\upsilon}, (4\upsilon\gamma_2^2)^{-d/2}\right\} (1 + \|\omega\|^2)^{-(\upsilon+d/2)} \\
&=: C_2(1 + \|\omega\|^2)^{-(\upsilon+d/2)}, \qquad \text{(S1.5)}
\end{aligned}$$

hold for all $\omega \in \mathbf{R}^d$.

Now we apply the extension theorem of the native spaces (Theorem 10.46 of Wendland, 2005) to obtain a function $f^E \in \mathcal{N}_{C_{\upsilon,\gamma}}(\mathbf{R}^d)$ such that $f^E|_\Omega = f$ and $\|f\|_{\mathcal{N}_{C_{\upsilon,\gamma}}(\Omega)} = \|f^E\|_{\mathcal{N}_{C_{\upsilon,\gamma}}(\mathbf{R}^d)}$ for each $\gamma \in [\gamma_1, \gamma_2]$. We use (S1.2)-(S1.4) to obtain

$$\begin{aligned}
\|f\|^2_{\mathcal{N}_{C_{\upsilon,\gamma}}(\Omega)} &= \|f^E\|^2_{\mathcal{N}_{C_{\upsilon,\gamma}}(\mathbf{R}^d)} = (2\pi)^{-d/2}\int_{\mathbf{R}^d}\frac{|\tilde{f}^E(\omega)|^2}{\tilde{C}_{\upsilon,\gamma}(\omega)}d\omega \\
&\geq C_1^{-1}(2\pi)^{-d/2}\int_{\mathbf{R}^d}|\tilde{f}^E(\omega)|^2(1+\|\omega\|^2)^{\upsilon+d/2}d\omega \\
&= C_1^{-1}\|f^E\|^2_{H^{\upsilon+d/2}(\mathbf{R}^d)} \geq C_1^{-1}\|f\|^2_{H^{\upsilon+d/2}(\Omega)}, \qquad \text{(S1.6)}
\end{aligned}$$

where the last inequality follows from the fact that $f^E|_\Omega = f$. On the other hand, because Ω is convex, f has an extension $f_E \in H^{\upsilon+d/2}(\mathbf{R}^d)$ satisfying $\|f_E\|_{H^k(\mathbf{R}^d)} \leq c\|f\|_{H^k(\Omega)}$ for some constant c independent of f. Then we use (S1.2), (S1.3) and (S1.5) to obtain

$$\begin{aligned}
\|f_E\|^2_{H^k(\Omega)} &\geq c^{-2}\|f\|^2_{H^k(\Omega)} \\
&= c^{-2}(2\pi)^{-d/2}\int_{\mathbf{R}^d}|\tilde{f}^E(\omega)|^2(1+\|\omega\|^2)^{\upsilon+d/2}d\omega \\
&\geq c^{-2}C_2(2\pi)^{-d/2}\int_{\mathbf{R}^d}\frac{|\tilde{f}^E(\omega)|^2}{\tilde{C}_{\upsilon,\gamma}(\omega)}d\omega \\
&= c^{-2}C_2\|f_E\|^2_{\mathcal{N}_{C_{\upsilon,\gamma}}(\mathbf{R}^d)} \geq c^{-2}C_2\|f\|^2_{\mathcal{N}_{C_{\upsilon,\gamma}}(\Omega)},
\end{aligned}$$

where the last inequality follows from the restriction theorem of the native space, which states that the restriction $f = f_E|_\Omega$ is contained in $\mathcal{N}_{C_{\upsilon,\gamma}}(\Omega)$ with a norm that is less than or equal to the norm $\|f_E\|_{\mathcal{N}_{C_{\upsilon,\gamma}}(\mathbf{R}^d)}$. See Theorem 10.47 of Wendland [2005]. The desired result is proved by combining

(S1.6) and (S1.7). □

Proof of Lemma 4. For $f \in \mathcal{N}_\Phi(\Omega)$, define

$$M(f) = L(f(x_1), \dots, f(x_n)) + \|f\|^2_{\mathcal{N}_\Phi(\Omega)}.$$

Now consider $s_{\hat{f},X}$, i.e., the interpolant of $\hat{f}$ over $X = \{x_1, \dots, x_n\}$ using the kernel function Φ. Because $\hat{f}(x_i) = s_{\hat{f},X}(x_i)$ for $i = 1, \dots, n$, we have

$$L(f(x_1), \dots, f(x_n)) = L(s_{\hat{f},X}(x_1), \dots, s_{\hat{f},X}(x_n)). \tag{S1.7}$$

In addition, it is easily seen from Lemma 1, (9) and (10) in the main article that

$$\|s_{\hat{f},X}\|^2_{\mathcal{N}_\Phi(\Omega)} \leq \|\hat{f}\|^2_{\mathcal{N}_\Phi(\Omega)}, \tag{S1.8}$$

and the equality holds if and only if $s_{\hat{f},X} = \hat{f}$. By combining (S1.7) and (S1.8) we obtain

$$M(s_{\hat{f},X}) \leq M(\hat{f}). \tag{S1.9}$$

Because $\hat{f}$ minimizes $M(f)$, the reverse of (S1.9) also holds. Hence we deduce $s_{\hat{f},X} = \hat{f}$, which proves the theorem according to the definition of the interpolant. □

Proof of Theorem 2. We first rewrite the minimization problem (16) in the main article as the following iterated form

$$\begin{aligned} & \min_{\substack{\theta\in\Theta \\ f\in\mathcal{N}_{C_{v,\gamma}}(\Omega)}} \sum_{i=1}^{n}(y_i^p - y^s(x_i,\theta) - f(x_i))^2 + \frac{\sigma^2}{\tau^2}\|f\|^2_{\mathcal{N}_{C_{v,\gamma}}(\Omega)} \\ = & \min_{\theta\in\Theta} \min_{f\in\mathcal{N}_{C_{v,\gamma}}(\Omega)} \sum_{i=1}^{n}(y_i^p - y^s(x_i,\theta) - f(x_i))^2 + \frac{\sigma^2}{\tau^2}\|f\|^2_{\mathcal{N}_{C_{v,\gamma}}(\Omega)} \end{aligned} \tag{S1.10}$$

Now we apply Lemma 4 to the inner minimization problem in (S1.10) and obtain the following representation for $\hat{\Delta}$:

$$\hat{\Delta} = \sum_{i=1}^{n} \alpha_i C_{v,\gamma}(x_i,\cdot),$$

with an undetermined vector of coefficients $\alpha = (\alpha_1,\ldots,\alpha_n)^{\mathrm{T}}$. Using the definition $\Sigma_\gamma = (C_{v,\gamma}(x_i,x_j))_{ij}$, clearly we have the matrix representation

$$\hat{\Delta}(\mathbf{x}) = \Sigma_\gamma \alpha. \tag{S1.11}$$

Now using (7) in the main article we have

$$\|\hat{\Delta}\|^2_{\mathcal{N}_{C_{v,\gamma}}(\Omega)} = \left\langle \sum_{i=1}^{n}\alpha_i C_{v,\gamma}(x_i,\cdot), \sum_{i=1}^{n}\alpha_i C_{v,\gamma}(x_i,\cdot)\right\rangle_{\mathcal{N}_{C_{v,\gamma}}(\Omega)} = \alpha^{\mathrm{T}}\Sigma_\gamma\alpha.$$

The minimization problem (16) in the main article then reduces to

$$\operatorname*{argmin}_{\substack{\theta\in\Theta \\ \alpha\in\mathbf{R}^n}} \|\mathbf{y}^p - y^s(\mathbf{x},\theta) - \alpha\Sigma_\gamma\|^2 + \frac{\sigma^2}{\tau^2}\alpha^{\mathrm{T}}\Sigma_\gamma\alpha.$$

Applying a change-of-variable argument using (S1.11) we obtain the following optimization formula

$$\operatorname*{argmin}_{\substack{\theta\in\Theta \\ \Delta(\mathbf{x})\in\mathbf{R}^n}} \|\mathbf{y}^p - y^s(\mathbf{x},\theta) - \Delta(\mathbf{x})\|^2 + \frac{\sigma^2}{\tau^2}\Delta(\mathbf{x})^{\mathrm{T}}\Sigma_\gamma^{-1}\Delta(\mathbf{x}).$$

Elementary calculations show its equivalence to the definition of $(\hat{\theta}_{KO}, \hat{\delta}(\mathbf{x}))$.

$\square$

Bibliography

Robert A Adams and John JF Fournier. *Sobolev Spaces*, volume 140. Access Online via Elsevier, 2003.

R. Tuo and C. F. J. Wu. A theoretical framework for calibration in computer models: parametrization, estimation and convergence properties. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):767–795, 2016.

H. Wendland. *Scattered Data Approximation*. Cambridge University Press, 2005.