A. Proof of Proposition 1

We first introduce the following lemmas.

Lemma 1 (Liu et al. 2016, Proposition 3.5)**.** *Let $\mathcal{H}$ denote the Reproducing Kernel Hilbert Space (RKHS) induced by kernel k. If $k(\cdot,\cdot)$ has continuous second order partial derivatives, and both $k(\mathbf{x},\cdot)$ and $k(\cdot,\mathbf{x})$ satisfy the boundary condition in eq. (7), then $\forall f \in \mathcal{H}$, f satisfies the same boundary condition.*

Lemma 2 (Mercer's theorem)**.** *Let k be a continuous kernel on compact metric space $\mathcal{X}$. q is a finite Borel measure on $\mathcal{X}$. Then for $\{\psi_j\}_{j\geq 1}$ that satisfies eq. (1), $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$:*

$$k(\mathbf{x}, \mathbf{y}) = \sum_j \mu_j \psi_j(\mathbf{x}) \psi_j(\mathbf{y}).$$

Proof. See Sejdinovic & Gretton, Theorem 50. □

Lemma 3 (Sejdinovic & Gretton, Theorem 51)**.** *Let $\mathcal{X}$ be a compact metric space and k: $\mathcal{X} \times \mathcal{X} \to \mathcal{R}$ a continuous kernel, Define:*

$$\mathcal{H} = \left\{ f = \sum_i a_i \psi_i : \left\{ \frac{a_i}{\sqrt{\mu_i}} \right\} \in \ell^2 \right\}.$$

Then $\mathcal{H}$ is the same space as the RKHS induced by k.

Then we prove Proposition 1.

Proof. In Lemma 3 we set $a_j = 1, a_i = 0 (i \neq j)$, then we have $\psi_j \in \mathcal{H}$. Then according to Lemma 1, we can conclude that ψ_j satisfies the boundary condition. □

B. Error Bound of SSGE

Define

$$g_i(\mathbf{x}) = \sum_{j=1}^{\infty} \beta_{ij} \psi_j(\mathbf{x}), \quad g_{i,J}(\mathbf{x}) = \sum_{j=1}^{J} \beta_{ij} \psi_j(\mathbf{x}), \quad \tilde{g}_{i,J}(\mathbf{x}) = \sum_{j=1}^{J} \beta_{ij} \hat{\psi}_j(\mathbf{x}), \quad \hat{g}_{i,J}(\mathbf{x}) = \sum_{j=1}^{J} \hat{\beta}_{ij} \hat{\psi}_j(\mathbf{x}), \tag{33}$$

which correspond to the major approximations in each step.

Lemma 4 (Izbicki et al. 2014)**.** *For all $1 \leq j \leq J$,*

$$\int \left(\hat{\psi}_j(\mathbf{x}) - \psi_j(\mathbf{x}) \right)^2 \, dq = O_q \left(\frac{1}{\mu_j \delta_j^2 M} \right), \tag{34}$$

where $\delta_j = \mu_j - \mu_{j+1}$.

Lemma 5 (Izbicki et al. 2014)**.** *For all $1 \leq j \leq J$,*

$$\int \hat{\psi}_j(\mathbf{x})^2 \, dq = O_q \left(\frac{1}{\mu_j \Delta_J^2 M} \right) + 1, \tag{35}$$

and for all $1 \leq i \leq J, i \neq j$,

$$\int \hat{\psi}_i(\mathbf{x}) \hat{\psi}_j(\mathbf{x}) \, dq = O_q \left(\left(\frac{1}{\sqrt{\mu_i}} + \frac{1}{\sqrt{\mu_j}} \right) \frac{1}{\Delta_J \sqrt{M}} \right), \tag{36}$$

where $\Delta_J = \min_{1 \leq j \leq J} \delta_j$.

Lemma 6.

$$\int \left| \tilde{g}_{i,J}(\mathbf{x}) - g_{i,J}(\mathbf{x}) \right|^2 \, dq = J O_q \left(\frac{1}{\mu_J \Delta_J^2 M} \right). \tag{37}$$

Proof. By Cauchy-Schwartz inequality, Assumption 2 and Lemma 4:

$$\begin{aligned}\int \left|\tilde{g}_{i,J}(\mathbf{x}) - g_{i,J}(\mathbf{x})\right|^2 \, dq &= \int \left|\sum_{j=1}^{J} \beta_{ij} \left(\psi_j(\mathbf{x}) - \hat{\psi}_j(\mathbf{x})\right)\right|^2 \, dq \\ &\leq \left(\sum_{j=1}^{J} \beta_{ij}^2\right) \left(\sum_{j=1}^{J} \int \left(\psi_j(\mathbf{x}) - \hat{\psi}_j(\mathbf{x})\right)^2 \, dq\right) \\ &= J O_q\left(\frac{1}{\mu_J \Delta_J^2 M}\right).\end{aligned} \tag{38}$$

□

Lemma 7. *For all* $1 \leq j \leq J$*,*

$$\left(\int \left(\nabla_{x_i}\psi_j(\mathbf{x}) - \nabla_{x_i}\hat{\psi}_j(\mathbf{x})\right) \, dq\right)^2 = O_q\left(\frac{C}{\mu_j \delta_j^2 M}\right). \tag{39}$$

Proof. Denote $\delta(\mathbf{x}) = \psi_j(\mathbf{x}) - \hat{\psi}_j(\mathbf{x})$. According to Assumption 1, it is easy to see that $\hat{\psi}_j(\mathbf{x})$ satisfies the boundary condition:

$$\int \nabla_{\mathbf{x}}[\hat{\psi}_j(\mathbf{x})q(\mathbf{x})]d\mathbf{x} = \mathbf{0}. \tag{40}$$

And from the proof of Proposition 1, we know $\psi_j(\mathbf{x})$ satisfies the boundary condition. Combining the two, we have:

$$\int \nabla_{\mathbf{x}}[\delta(\mathbf{x})q(\mathbf{x})] \, d\mathbf{x} = \mathbf{0}. \tag{41}$$

By eq. (41), Lemma 4 and Assumption 2, we have

$$\begin{aligned}\left(\int \nabla_{x_i}\delta(\mathbf{x})dq\right)^2 &= \left(\int \nabla_{x_i}[\delta(\mathbf{x})q(\mathbf{x})] - \delta(\mathbf{x})\nabla_{x_i}q(\mathbf{x}) \, d\mathbf{x}\right)^2 \\ &= \left(\int \delta(\mathbf{x})\nabla_{x_i}\log q(\mathbf{x}) \, dq\right)^2 \\ &\leq \left(\int \delta(\mathbf{x})^2 \, dq\right)\left(\int g_i(\mathbf{x})^2 \, dq\right) \\ &= O_q\left(\frac{C}{\mu_j \delta_j^2 M}\right).\end{aligned} \tag{42}$$

□

Lemma 8. *For all* $1 \leq j \leq J$*,*

$$(\beta_{ij} - \hat{\beta}_{ij})^2 = O_q\left(\frac{1}{M}\right) + O_q\left(\frac{C}{\mu_j \delta_j^2 M}\right). \tag{43}$$

Proof.

$$\begin{aligned}\frac{1}{2}(\beta_{ij}-\hat{\beta}_{ij})^2 &\le \left(\beta_{ij}-\frac{1}{M}\sum_{m=1}^{M}\nabla_{x_i}\psi_j(\mathbf{x}^m)\right)^2+\left(\frac{1}{M}\sum_{m=1}^{M}\left(\nabla_{x_i}\psi_j(\mathbf{x}^m)-\nabla_{x_i}\hat{\psi}_j(\mathbf{x}^m)\right)\right)^2\\ &\le O_q\left(\frac{1}{M}\right)+2\left[\frac{1}{M}\sum_{m=1}^{M}\left(\nabla_{x_i}\psi_j(\mathbf{x}^m)-\nabla_{x_i}\hat{\psi}_j(\mathbf{x}^m)\right)-\int\left(\nabla_{x_i}\psi_j(\mathbf{x})-\nabla_{x_i}\hat{\psi}_j(\mathbf{x})\right)\,dq\right]^2\\ &\quad+2\left[\int\left(\nabla_{x_i}\psi_j(\mathbf{x})-\nabla_{x_i}\hat{\psi}_j(\mathbf{x})\right)\,dq\right]^2\\ &=O_q\left(\frac{1}{M}\right)+2O_q\left(\frac{1}{M}\right)+2\left(\int\left(\nabla_{x_i}\psi_j(\mathbf{x})-\nabla_{x_i}\hat{\psi}_j(\mathbf{x})\right)dq\right)^2.\end{aligned} \tag{44}$$

Therefore, by Lemma 7 we have

$$(\beta_{ij}-\hat{\beta}_{ij})^2=O_q\left(\frac{1}{M}\right)+O_q\left(\frac{C}{\mu_j\delta_j^2M}\right). \tag{45}$$

□

Lemma 9.

$$\int|\tilde{g}_{i,J}(\mathbf{x})-\hat{g}_{i,J}(\mathbf{x})|^2\,dq=J^2\left(O_q\left(\frac{1}{M}\right)+O_q\left(\frac{C}{\mu_j\Delta_j^2M}\right)\right) \tag{46}$$

Proof. By applying Minkowski inequality, Cauchy-Schwartz inequality, Lemma 8 and Lemma 5, we have

$$\begin{aligned}&\int|\tilde{g}_{i,J}(\mathbf{x})-\hat{g}_{i,J}(\mathbf{x})|^2\,dq=\int\left|\sum_{j=1}^{J}\beta_{ij}\hat{\psi}_j(\mathbf{x})-\sum_{j=1}^{J}\hat{\beta}_{ij}\hat{\psi}_j(\mathbf{x})\right|^2dq=\int\left|\sum_{j=1}^{J}(\beta_{ij}-\hat{\beta}_{ij})\hat{\psi}_j(\mathbf{x})\right|^2dq\\ &\le\left\{\sum_{j=1}^{J}\left[\int\left|(\beta_{ij}-\hat{\beta}_{ij})\hat{\psi}_j(\mathbf{x})\right|^2dq\right]^{\frac{1}{2}}\right\}^2\le\left\{\sum_{j=1}^{J}\left[\int\left|(\beta_{ij}-\hat{\beta}_{ij})\right|^2dq\int\hat{\psi}_j^2(\mathbf{x})\,dq\right]^{\frac{1}{2}}\right\}^2\\ &=\left\{\sum_{j=1}^{J}\left[O_q\left(\frac{1}{M}\right)+O_q\left(\frac{C}{\mu_j\delta_j^2M}\right)\right]^{\frac{1}{2}}\left[O_q\left(\frac{1}{\mu_j\Delta_J^2M}\right)+1\right]^{\frac{1}{2}}\right\}^2\\ &=J^2\left(O_q\left(\frac{1}{M}\right)+O_q\left(\frac{C}{\mu_J\Delta_J^2M}\right)\right)\end{aligned} \tag{47}$$

□

Theorem 3 (Estimation Error)**.**

$$\int|\hat{g}_{i,J}(\mathbf{x})-g_{i,J}(\mathbf{x})|^2\;dq=J^2\left(O_q\left(\frac{1}{M}\right)+O_q\left(\frac{C}{\mu_J\Delta_J^2M}\right)\right)+JO_q\left(\frac{1}{\mu_J\Delta_J^2M}\right) \tag{48}$$

Proof. By lemma 6 and lemma 9.

$$\begin{aligned}\int|\hat{g}_{i,J}(\mathbf{x})-g_{i,J}(\mathbf{x})|^2\;dq&\le\int|\tilde{g}_{i,J}(\mathbf{x})-g_{i,J}(\mathbf{x})|^2\;dq+\int|\tilde{g}_{i,J}(\mathbf{x})-\hat{g}_{i,J}(\mathbf{x})|^2\,dq\\ &=J^2\left(O_q\left(\frac{1}{M}\right)+O_q\left(\frac{C}{\mu_J\Delta_J^2M}\right)\right)+JO_q\left(\frac{1}{\mu_J\Delta_J^2M}\right)\end{aligned} \tag{49}$$

□

Theorem 4 (Truncation Error).

$$\int |g_{i,J}(\mathbf{x}) - g_i(\mathbf{x})|^2 \, dq = \|g_i\|_{\mathcal{H}}^2 O(\mu_J) \tag{50}$$

Proof.

$$\int |g_{i,J}(\mathbf{x}) - g_i(\mathbf{x})|^2 \, dq = \sum_{j>J} \beta_{ij}^2 = \sum_{j>J} \frac{\beta_{ij}^2}{\mu_j} \mu_j \le \mu_J \sum_{j>J} \frac{\beta_{ij}^2}{\mu_j} = \mu_J \|g_i\|_{\mathcal{H}}^2 \tag{51}$$

□

Theorem 5 (Error Bound of SSGE).

$$\int \left(\hat{g}_{i,J}(\mathbf{x}) - g_i(\mathbf{x})\right)^2 \, dq = J^2 \left(O_q\left(\frac{1}{M}\right) + O_q\left(\frac{C}{\mu_J \Delta_J^2 M}\right)\right) + J O_q\left(\frac{1}{\mu_J \Delta_J^2 M}\right) + \|g_i\|_{\mathcal{H}}^2 O(\mu_J). \tag{52}$$

Proof. By theorem 3 and theorem 4, we have

$$\begin{aligned} \int \left(\hat{g}_{i,J}(\mathbf{x}) - g_i(\mathbf{x})\right)^2 \, dq &\le \int |\hat{g}_{i,J}(\mathbf{x}) - g_{i,J}(\mathbf{x})|^2 \, dq + \int |g_{i,J}(\mathbf{x}) - g_i(\mathbf{x})|^2 \, dq \\ &= J^2 \left(O_q\left(\frac{1}{M}\right) + O_q\left(\frac{C}{\mu_J \Delta_J^2 M}\right)\right) + J O_q\left(\frac{1}{\mu_J \Delta_J^2 M}\right) + \|g_i\|_{\mathcal{H}}^2 O(\mu_J) \end{aligned} \tag{53}$$

□