# Supplementary material for *An analysis of Ermakov-Zolotukhin quadrature using kernels*

**Ayoub Belhadji**
Univ Lyon, ENS de Lyon
Inria, CNRS, UCBL
LIP UMR 5668, Lyon, France
ayoub.belhadji@ens-lyon.fr

## A Detailed proofs

### A.1 Proof of Proposition 2

By definition, we have

$$\forall x_1, x_2 \in \mathcal{X},\;\; \kappa_N^{\boldsymbol{\gamma}}(x_1,x_2) = \sum_{n \in [N]} \gamma_n \phi_n(x_1)\phi_n(x_2), \tag{1}$$

therefore

$$\forall x_1, x_2 \in \mathcal{X},\;\; \kappa_N^{\boldsymbol{\gamma}}(x_1,x_2) = \sum_{n \in [N]} \rho_n \phi_n(x_1)\tilde{\gamma}_n\phi_n(x_2), \tag{2}$$

where

$$\forall n \in [N],\;\; \rho_n = \frac{\gamma_n}{\tilde{\gamma}_n}. \tag{3}$$

Then

$$\forall \boldsymbol{x} \in \mathcal{X}^N,\;\; \boldsymbol{\kappa}_N^{\boldsymbol{\gamma}}(\boldsymbol{x}) = \boldsymbol{\Phi}_N^{\boldsymbol{\rho}}(\boldsymbol{x})^{\intercal}\boldsymbol{\Phi}_N^{\tilde{\boldsymbol{\gamma}}}(\boldsymbol{x}) \;\;, \tag{4}$$

where

$$\boldsymbol{\Phi}_N^{\boldsymbol{\rho}}(\boldsymbol{x}) = (\rho_n\phi_n(x_i))_{(n,i)\in[N]\times[N]} \in \mathbb{R}^{N\times N}, \tag{5}$$

and

$$\boldsymbol{\Phi}_N^{\tilde{\boldsymbol{\gamma}}}(\boldsymbol{x}) = (\tilde{\gamma}_n\phi_n(x_i))_{(n,i)\in[N]\times[N]} \in \mathbb{R}^{N\times N}. \tag{6}$$

Moreover, by definition of $\mu_g^{\boldsymbol{\gamma}}$, we have

$$\forall x \in \mathcal{X},\;\; \mu_g^{\boldsymbol{\gamma}}(x) = \sum_{n \in [N]} \gamma_n \langle g, \phi_n\rangle_\omega \phi_n(x), \tag{7}$$

therefore

$$\forall x \in \mathcal{X},\;\; \mu_g^{\boldsymbol{\gamma}}(x) = \sum_{n \in [N]} \tilde{\gamma}_n \langle g, \phi_n\rangle_\omega \rho_n\phi_n(x), \tag{8}$$

so that

$$\forall \boldsymbol{x} \in \mathcal{X}^N,\;\; \mu_g^{\boldsymbol{\gamma}}(\boldsymbol{x}) = \boldsymbol{\Phi}_N^{\boldsymbol{\rho}}(\boldsymbol{x})^{\intercal}\boldsymbol{\alpha}, \tag{9}$$

where

$$\boldsymbol{\alpha} = (\tilde{\gamma}_n)_{n\in[N]} \in \mathbb{R}^N. \tag{10}$$

35th Conference on Neural Information Processing Systems (NeurIPS 2021).

Combining (4) and (9), we prove that for any $\boldsymbol{x} \in \mathcal{X}^N$ such that $\operatorname{Det} \boldsymbol{\kappa}_N(\boldsymbol{x}) > 0$, we have

$$
\begin{aligned}
\hat{\boldsymbol{w}}^{\gamma,N,g}(\boldsymbol{x}) &= \boldsymbol{\kappa}_N^{\gamma}(\boldsymbol{x})^{-1} \mu_g^{\gamma}(\boldsymbol{x}) \\
&= \boldsymbol{\Phi}_N^{\tilde{\gamma}}(\boldsymbol{x})^{-1} \;\; \boldsymbol{\Phi}_N^{\boldsymbol{\rho}}(\boldsymbol{x})^{-1^{\intercal}} \mu_g^{\gamma}(\boldsymbol{x}) \\
&= \boldsymbol{\Phi}_N^{\tilde{\gamma}}(\boldsymbol{x})^{-1} \;\; \boldsymbol{\Phi}_N^{\boldsymbol{\rho}}(\boldsymbol{x})^{-1^{\intercal}} \boldsymbol{\Phi}_N^{\boldsymbol{\rho}}(\boldsymbol{x})^{\intercal} \boldsymbol{\alpha} \\
&= \boldsymbol{\Phi}_N^{\tilde{\gamma}}(\boldsymbol{x})^{-1} \;\; \boldsymbol{\alpha}. \qquad (11)
\end{aligned}
$$

### A.2 Useful results

We gather in this section some results that will be useful in the following proofs.

#### A.2.1 A useful lemma

We prove in the following a lemma that we will use in Section A.3.

**Lemma 1.** *Let $\boldsymbol{x} \in \mathcal{X}^N$ such that $\operatorname{Det} \boldsymbol{\kappa}_N(\boldsymbol{x}) > 0$. For $n, n' \in [N]$, define*

$$\tau_{n,n'}(\boldsymbol{x}) = \sqrt{\sigma_n}\sqrt{\sigma_{n'}}\phi_n(\boldsymbol{x})^{\intercal} \boldsymbol{K}_N(\boldsymbol{x})^{-1}\phi_{n'}(\boldsymbol{x}). \qquad (12)$$

*Then*

$$\forall n, n' \in [N], \;\; \tau_{n,n'}(\boldsymbol{x}) = \delta_{n,n'}. \qquad (13)$$

*Proof.* We have

$$\boldsymbol{K}_N(\boldsymbol{x}) = \boldsymbol{\Phi}_N^{\sqrt{\boldsymbol{\sigma}}}(\boldsymbol{x})^{\intercal} \boldsymbol{\Phi}_N^{\sqrt{\boldsymbol{\sigma}}}(\boldsymbol{x}), \qquad (14)$$

where

$$\boldsymbol{\Phi}_N^{\sqrt{\boldsymbol{\sigma}}}(\boldsymbol{x}) = (\sqrt{\sigma}_n \phi_n(x_i))_{(n,i)\in[N]\times[N]}. \qquad (15)$$

Let $n, n' \in [N]$. We have

$$\sqrt{\sigma_n}\phi_n(\boldsymbol{x})^{\intercal} = \boldsymbol{e}_n^{\intercal} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x}), \qquad (16)$$

and

$$\sqrt{\sigma_{n'}}\phi_{n'}(\boldsymbol{x}) = \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})^{\intercal} \boldsymbol{e}_{n'}. \qquad (17)$$

Therefore

$$
\begin{aligned}
\tau_{n,n'}(\boldsymbol{x}) &= \boldsymbol{e}_n^{\intercal} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x}) \big(\boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})^{\intercal} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})\big)^{-1} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})^{\intercal} \boldsymbol{e}_{n'} \qquad (18) \\
&= \boldsymbol{e}_n^{\intercal} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x}) \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})^{-1} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})^{\intercal^{-1}} \boldsymbol{\Phi}_N^{\boldsymbol{\sigma}}(\boldsymbol{x})^{\intercal} \boldsymbol{e}_{n'} \qquad (19) \\
&= \boldsymbol{e}_n^{\intercal} \boldsymbol{e}_{n'} \qquad (20) \\
&= \delta_{n,n'}. \qquad (21)
\end{aligned}
$$

□

#### A.2.2 A borrowed result

We recall in the following a result proven in [1] that we will use in Section A.4.

**Proposition 1.** *[Theorem 1 in [1]] Let $\boldsymbol{x}$ be a random subset of $\mathcal{X}$ that follows the distribution of DPP of kernel $\kappa_N$ and reference measure $\omega$. Let $f \in \mathbb{L}_2(\omega)$, and $n, n' \in [N]$ such that $n \neq n'$. Then*

$$\mathbb{C}\mathrm{ov}_{\mathrm{DPP}}(I^{\mathrm{EZ},n}(f), I^{\mathrm{EZ},n'}(f)) = 0. \qquad (22)$$

### A.3 Proof of Proposition 5

Let $\boldsymbol{x} \in \mathcal{X}^N$ such that the condition $\operatorname{Det} \boldsymbol{\kappa}_N(\boldsymbol{x}) > 0$ is satisfied, and let $g \in \mathcal{E}_N$. We start by the proof of (36). By definition of $\mu_g$, we have

$$\mu_g(x) = \sum_{n\in[N]} \sigma_n \langle g, \phi_n\rangle_{\omega} \phi_n(x), \qquad (23)$$

so that

$$\mu_g(\boldsymbol{x}) = \sum_{n\in[N]} \sigma_n \langle g, \phi_n\rangle_{\omega} \phi_n(\boldsymbol{x}). \qquad (24)$$

Proposition 2 yields

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g} = \boldsymbol{K}_N(\boldsymbol{x})^{-1}\mu_g(\boldsymbol{x}). \tag{25}$$

Therefore

$$\begin{aligned}
\hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\mu_g(\boldsymbol{x}) &= \mu_g(\boldsymbol{x})\boldsymbol{K}_N(\boldsymbol{x})^{-1}\mu_g(\boldsymbol{x}) \\
&= \sum_{n,n' \in [N]} \sigma_n \sigma_{n'} \langle g, \phi_n\rangle_\omega \langle g, \phi_{n'}\rangle_\omega \phi_n(\boldsymbol{x})^\intercal \boldsymbol{K}_N(\boldsymbol{x})^{-1}\phi_{n'}(\boldsymbol{x}) \\
&= \sum_{n \in [N]} \sigma_n \langle g, \phi_n\rangle_\omega^2 \tau_{n,n}(\boldsymbol{x}) \\
&+ \sum_{\substack{n,n' \in [N] \\ n \neq n'}} \sqrt{\sigma_n \sigma_{n'}} \langle g, \phi_n\rangle_\omega \langle g, \phi_{n'}\rangle_\omega \tau_{n,n'}(\boldsymbol{x}),
\end{aligned} \tag{26}$$

where $\tau_{n,n'}(\boldsymbol{x})$ is defined by

$$\tau_{n,n'}(\boldsymbol{x}) = \sqrt{\sigma_n}\sqrt{\sigma_{n'}}\phi_n(\boldsymbol{x})^\intercal \boldsymbol{K}_N(\boldsymbol{x})^{-1}\phi_{n'}(\boldsymbol{x}). \tag{27}$$

Now, Lemma 1 yields

$$\begin{aligned}
\sum_{n \in [N]} \sigma_n \langle g, \phi_n\rangle_\omega^2 \tau_{n,n}(\boldsymbol{x}) &= \sum_{n \in [N]} \sigma_n \langle g, \phi_n\rangle_\omega^2 \\
&= \|\mu_g\|_{\mathcal{F}}^2,
\end{aligned} \tag{28}$$

and

$$\sum_{\substack{n,n' \in [N] \\ n \neq n'}} \sqrt{\sigma_n}\sqrt{\sigma_{n'}} \langle g, \phi_n\rangle_\omega \langle g, \phi_{n'}\rangle_\omega \tau_{n,n'}(\boldsymbol{x}) = 0. \tag{29}$$

Combining (26), (28) and (29), we obtain

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\mu_g(\boldsymbol{x}) = \|\mu_g\|_{\mathcal{F}}^2. \tag{30}$$

We move now to the proof of (37). We have by the Mercer decomposition

$$\boldsymbol{K}(\boldsymbol{x}) = \sum_{m=1}^{+\infty} \sigma_m \phi_m(\boldsymbol{x})\phi_m(\boldsymbol{x})^\intercal \tag{31}$$

$$= \sum_{m=1}^{N} \sigma_m \phi_m(\boldsymbol{x})\phi_m(\boldsymbol{x})^\intercal + \sum_{m=N+1}^{+\infty} \sigma_m \phi_m(\boldsymbol{x})\phi_m(\boldsymbol{x})^\intercal. \tag{32}$$

Moreover, observe that

$$\boldsymbol{K}_N(\boldsymbol{x}) = \sum_{m=1}^{N} \sigma_m \phi_m(\boldsymbol{x})\phi_m(\boldsymbol{x})^\intercal, \tag{33}$$

and

$$\boldsymbol{K}_N^{\perp}(\boldsymbol{x}) = \sum_{m=N+1}^{+\infty} \sigma_m \phi_m(\boldsymbol{x})\phi_m(\boldsymbol{x})^\intercal. \tag{34}$$

Therefore

$$\boldsymbol{K}(\boldsymbol{x}) = \boldsymbol{K}_N(\boldsymbol{x}) + \boldsymbol{K}_N^{\perp}(\boldsymbol{x}), \tag{35}$$

so that

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\boldsymbol{K}(\boldsymbol{x})\hat{\boldsymbol{w}}^{\mathrm{EZ},g} = \hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\boldsymbol{K}_N(\boldsymbol{x})\hat{\boldsymbol{w}}^{\mathrm{EZ},g} + \hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\boldsymbol{K}_N^{\perp}(\boldsymbol{x})\hat{\boldsymbol{w}}^{\mathrm{EZ},g}. \tag{36}$$

In order to evaluate $\hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\boldsymbol{K}_N(\boldsymbol{x})\hat{\boldsymbol{w}}^{\mathrm{EZ},g}$, we use Proposition 2, and we get

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g} = \boldsymbol{K}_N(\boldsymbol{x})^{-1}\mu_g(\boldsymbol{x}), \tag{37}$$

so that

$$\begin{aligned}
\hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\boldsymbol{K}_N(\boldsymbol{x})\hat{\boldsymbol{w}}^{\mathrm{EZ},g} &= \hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\boldsymbol{K}_N(\boldsymbol{x})\boldsymbol{K}_N(\boldsymbol{x})^{-1}\mu_g(\boldsymbol{x}) \\
&= \hat{\boldsymbol{w}}^{\mathrm{EZ},g^\intercal}\mu_g(\boldsymbol{x}) \\
&= \|\mu_g\|_{\mathcal{F}}^2.
\end{aligned} \tag{38}$$

Finally, by definition

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g} = \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1}\boldsymbol{\epsilon}, \tag{39}$$

where $\boldsymbol{\epsilon} = \sum_{n\in[N]} \langle g, \phi_n\rangle_\omega \boldsymbol{e}_n$. Therefore

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g\intercal} \boldsymbol{K}_N(\boldsymbol{x})^{\perp} \hat{\boldsymbol{w}}^{\mathrm{EZ},g} = \boldsymbol{\epsilon}^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \boldsymbol{K}_N(\boldsymbol{x})^{\perp} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1}\boldsymbol{\epsilon}. \tag{40}$$

### A.4 Proof of Theorem 7

Let $m \in \mathbb{N}^*$ such that $m \geq N+1$. We prove that

$$\forall \boldsymbol{\epsilon}, \tilde{\boldsymbol{\epsilon}} \in \mathbb{R}^N,\ \mathbb{E}_{\mathrm{DPP}} \boldsymbol{\epsilon}^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} = \sum_{n \in N} \epsilon_n \tilde{\epsilon}_n. \tag{41}$$

For this purpose, let $\boldsymbol{\epsilon}, \tilde{\boldsymbol{\epsilon}} \in \mathbb{R}^N$, and observe that

$$\begin{aligned} \boldsymbol{\epsilon}^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \phi_m(\boldsymbol{x}) &= \sum_{n\in[N]} \epsilon_n \boldsymbol{e}_n^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \phi_m(\boldsymbol{x}) && (42)\\ &= \sum_{n\in[N]} \epsilon_n \hat{\boldsymbol{w}}^{\mathrm{EZ},n} \phi_m(\boldsymbol{x}) && (43)\\ &= \sum_{n\in[N]} \epsilon_n I^{\mathrm{EZ},n}(\phi_m). && (44) \end{aligned}$$

and

$$\phi_m(\boldsymbol{x})^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} = \sum_{n\in[N]} \tilde{\epsilon}_n I^{\mathrm{EZ},n}(\phi_m). \tag{45}$$

Therefore

$$\boldsymbol{\epsilon}^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} = \sum_{n\in[N]} \sum_{n'\in[N]} \epsilon_n \tilde{\epsilon}_{n'} I^{\mathrm{EZ},n}(\phi_m) I^{\mathrm{EZ},n'}(\phi_m), \tag{46}$$

then

$$\mathbb{E}_{\mathrm{DPP}} \boldsymbol{\epsilon}^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} = \sum_{n\in[N]} \sum_{n'\in[N]} \epsilon_n \tilde{\epsilon}_{n'} \mathbb{E}_{\mathrm{DPP}} I^{\mathrm{EZ},n}(\phi_m) I^{\mathrm{EZ},n'}(\phi_m). \tag{47}$$

Now, for $n, n' \in [N]$,

$$\mathbb{E}_{\mathrm{DPP}} I^{\mathrm{EZ},n}(\phi_m) = \int_{\mathcal{X}} \phi_m(x)\phi_n(x) \mathrm{d}\omega(x) = 0, \tag{48}$$

and

$$\mathbb{E}_{\mathrm{DPP}} I^{\mathrm{EZ},n'}(\phi_m) = \int_{\mathcal{X}} \phi_m(x)\phi_{n'}(x) \mathrm{d}\omega(x) = 0. \tag{49}$$

Therefore

$$\mathbb{E}_{\mathrm{DPP}} I^{\mathrm{EZ},n}(\phi_m) I^{\mathrm{EZ},n'}(\phi_m) = \mathbb{C}\mathrm{ov}_{\mathrm{DPP}}(I^{\mathrm{EZ},n}(\phi_m), I^{\mathrm{EZ},n'}(\phi_m)). \tag{50}$$

Now, by Proposition 1, we have $\mathbb{C}\mathrm{ov}_{\mathrm{DPP}}(I^{\mathrm{EZ},n}(\phi_m), I^{\mathrm{EZ},n'}(\phi_m)) = \delta_{n,n'}$, so that

$$\mathbb{E}_{\mathrm{DPP}} I^{\mathrm{EZ},n}(\phi_m) I^{\mathrm{EZ},n'}(\phi_m) = \delta_{n,n'}, \tag{51}$$

and

$$\begin{aligned} \mathbb{E}_{\mathrm{DPP}} \boldsymbol{\epsilon}^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1\intercal} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^\intercal \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} &= \sum_{n\in[N]} \sum_{n'\in[N]} \epsilon_n \tilde{\epsilon}_{n'} \mathbb{E}_{\mathrm{DPP}} I^{\mathrm{EZ},n}(\phi_m) I^{\mathrm{EZ},n'}(\phi_m)\\ &= \sum_{n\in[N]} \sum_{n'\in[N]} \epsilon_n \tilde{\epsilon}_{n'} \delta_{n,n'}\\ &= \sum_{n\in[N]} \epsilon_n \tilde{\epsilon}_n. && (52) \end{aligned}$$

Now, for $\boldsymbol{\epsilon} \in \mathbb{R}^N$ and $m \in \mathbb{N}^*$ define $Y_{\boldsymbol{\epsilon},m}$ by

$$Y_{\boldsymbol{\epsilon},m} = \sigma_m \boldsymbol{\epsilon}^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1^{\intercal}} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \boldsymbol{\epsilon}. \tag{53}$$

We have

$$\mathbb{E}_{\mathrm{DPP}} Y_{\boldsymbol{\epsilon},m} = \sigma_m \sum_{n \in [N]} \epsilon_n^2, \tag{54}$$

and the $Y_{\boldsymbol{\epsilon},m}$ are non-negative since

$$Y_{\boldsymbol{\epsilon},m} = \sigma_m (\boldsymbol{\epsilon}^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1^{\intercal}} \phi_m(\boldsymbol{x}))^2 \geq 0, \tag{55}$$

moreover,

$$\sum_{m=N+1}^{+\infty} \mathbb{E}_{\mathrm{DPP}} Y_{\boldsymbol{\epsilon},m} < +\infty. \tag{56}$$

Therefore, by Beppo Levi's lemma

$$\begin{aligned} \mathbb{E}_{\mathrm{DPP}} \sum_{m=N+1}^{+\infty} Y_{\boldsymbol{\epsilon},m} &= \sum_{m=N+1}^{+\infty} \mathbb{E}_{\mathrm{DPP}} Y_{\boldsymbol{\epsilon},m} \\ &= \sum_{n \in [N]} \epsilon_n^2 \sum_{m=N+1}^{+\infty} \sigma_m. \end{aligned} \tag{57}$$

Now, in general for $m \in \mathbb{N}^*$ such that $m \geq N+1$, we have

$$\sigma_m \boldsymbol{\epsilon}^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1^{\intercal}} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} \leq \frac{1}{2}(Y_{\boldsymbol{\epsilon},m} + Y_{\tilde{\boldsymbol{\epsilon}},m}), \tag{58}$$

so that for $M \geq N+1$, we have

$$\sum_{m=N+1}^{M} \sigma_m \boldsymbol{\epsilon}^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1^{\intercal}} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} \leq \frac{1}{2}( \sum_{m=N+1}^{+\infty} Y_{\boldsymbol{\epsilon},m} + \sum_{m=N+1}^{+\infty} Y_{\tilde{\boldsymbol{\epsilon}},m}). \tag{59}$$

Therefore, by dominated convergence theorem we conclude that

$$\mathbb{E}_{\mathrm{DPP}} \sum_{m=N+1}^{+\infty} \sigma_m \boldsymbol{\epsilon}^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1^{\intercal}} \phi_m(\boldsymbol{x}) \phi_m(\boldsymbol{x})^{\intercal} \boldsymbol{\Phi}_N(\boldsymbol{x})^{-1} \tilde{\boldsymbol{\epsilon}} = \sum_{m=N+1}^{+\infty} \sigma_m \sum_{n \in [N]} \epsilon_n \tilde{\epsilon}_n. \tag{60}$$

### A.5 Proof of Theorem 3

Let $g \in \mathcal{E}_N$, and denote $\boldsymbol{\epsilon} = \sum_{n \in [N]} \langle g, \phi_n \rangle_{\omega} \boldsymbol{e}_n$. Combining Theorem 6 and Theorem 7, we obtain

$$\mathbb{E}_{\mathrm{DPP}} \| \mu_g - \sum_{i \in [N]} \hat{w}_i^{\mathrm{EZ},g} k(x_i, .) \|_{\mathcal{F}}^2 = \sum_{m \geq N+1} \sigma_m \sum_{n \in [N]} \epsilon_n^2. \tag{61}$$

Now let $g \in \mathcal{L}_2(\omega)$, we have

$$\begin{aligned} \| \mu_g - \sum_{i \in [N]} \hat{w}_i^{\mathrm{EZ},g} k(x_i, .) \|_{\mathcal{F}}^2 &= \| \mu_g - \mu_{g_N} + \mu_{g_N} - \sum_{i \in [N]} \hat{w}_i^{\mathrm{EZ},g} k(x_i, .) \|_{\mathcal{F}}^2 && (62) \\ &\leq 2 \Big( \| \mu_g - \mu_{g_N} \|_{\mathcal{F}}^2 + \| \mu_{g_N} - \sum_{i \in [N]} \hat{w}_i^{\mathrm{EZ},g} k(x_i, .) \|_{\mathcal{F}}^2 \Big), && (63) \end{aligned}$$

where $g_N = \sum_{n \in [N]} \langle g, \phi_n \rangle_{\omega} \phi_n \in \mathcal{E}_N$.

Now, observe that

$$\mu_g^{\gamma,N} = \mu_{g_N}^{\gamma,N}, \tag{64}$$

so that

$$\hat{\boldsymbol{w}}^{\mathrm{EZ},g} = \hat{\boldsymbol{w}}^{\mathrm{EZ},g_N}. \tag{65}$$

Therefore

$$\|\mu_g - \sum_{i\in[N]} \hat{w}_i^{\mathrm{EZ},g} k(x_i,.)\|_{\mathcal{F}}^2 \leq 2\Big(\|\mu_g - \mu_{g_N}\|_{\mathcal{F}}^2 + \|\mu_{g_N} - \sum_{i\in[N]} \hat{w}_i^{\mathrm{EZ},g_N} k(x_i,.)\|_{\mathcal{F}}^2\Big). \tag{66}$$

Now, we have

$$\begin{aligned}\|\mu_g - \mu_{g_N}\|_{\mathcal{F}}^2 &= \sum_{m\geq N+1} \sigma_m \langle g,\phi_m\rangle_{\omega}^2 \\ &\leq \sigma_{N+1} \sum_{m\geq N+1} \langle g,\phi_m\rangle_{\omega}^2 \\ &\leq r_{N+1}\|g\|_{\omega}^2. \end{aligned} \tag{67}$$

Moreover, by (26) we have

$$\begin{aligned}\mathbb{E}_{\mathrm{DPP}}\|\mu_{g_N} - \sum_{i\in[N]} \hat{w}_i^{\mathrm{EZ},g_N} k(x_i,.)\|_{\mathcal{F}}^2 &= \sum_{n\in[N]} \langle g,\phi_n\rangle_{\omega}^2 r_{N+1} \\ &\leq \|g\|_{\omega}^2 r_{N+1}. \end{aligned} \tag{68}$$

Combining (66), (67) and (68), we obtain

$$\|\mu_g - \sum_{i\in[N]} \hat{w}_i^{\mathrm{EZ},g} k(x_i,.)\|_{\mathcal{F}}^2 \leq 4\|g\|_{\omega}^2 r_{N+1}. \tag{69}$$

## References

[1] G. Gautier, R. Bardenet, and M. Valko. On two ways to use determinantal point processes for monte carlo integration. In *Advances in Neural Information Processing Systems*, volume 32, 2019.