

## Supplementary Material

### A Large Dimensional Analysis of Least Squares Support Vector Machines

#### APPENDIX A PROOF OF THEOREM 1

Our key interest here is on the decision function of LS-SVM:  $g(\mathbf{x}) = \alpha^\top \mathbf{k}(\mathbf{x}) + b$  with  $(\alpha, b)$  given by

$$\begin{cases} \alpha &= \mathbf{S}^{-1} \left( \mathbf{I}_n - \frac{\mathbf{1}_n \mathbf{1}_n^\top \mathbf{S}^{-1}}{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n} \right) \mathbf{y} \\ b &= \frac{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{y}}{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n} \end{cases}$$

$$\text{and } \mathbf{S}^{-1} = \left( \mathbf{K} + \frac{n}{\gamma} \mathbf{I}_n \right)^{-1}.$$

Before going into the detailed proof, as we will frequently deal with random variables evolving as  $n, p$  grow large, we shall use the extension of the  $O(\cdot)$  notation introduced in [20]: for a random variable  $x \equiv x_n$  and  $u_n \geq 0$ , we write  $x = O(u_n)$  if for any  $\eta > 0$  and  $D > 0$ , we have  $n^D P(x \geq n^\eta u_n) \rightarrow 0$ . Note that under Assumption 1 it is equivalent to use either  $O(u_n)$  or  $O(u_p)$  since  $n, p$  scales linearly. In the following we shall use constantly  $O(u_n)$  for simplicity.

When multidimensional objects are concerned,  $\mathbf{v} = O(u_n)$  means the maximum entry of a vector (or a diagonal matrix)  $\mathbf{v}$  in absolute value is of order  $O(u_n)$  and  $\mathbf{M} = O(u_n)$  means that the operator norm of  $\mathbf{M}$  is of order  $O(u_n)$ . We refer the reader to [20] for more discussions on these practical definitions.

Under the growth rate settings of Assumption 1 from [20], the approximation of the kernel matrix  $\mathbf{K}$  is given by

$$\mathbf{K} = -2f'(\tau) (\mathbf{P}\mathbf{\Omega}^\top \mathbf{\Omega} \mathbf{P} + \mathbf{A}) + \beta \mathbf{I}_n + O(n^{-\frac{1}{2}}) \quad (12)$$

with  $\beta = f(0) - f(\tau) + \tau f'(\tau)$  and  $\mathbf{A} = \mathbf{A}_n + \mathbf{A}_{\sqrt{n}} + \mathbf{A}_1$ ,  $\mathbf{A}_n = -\frac{f(\tau)}{2f'(\tau)} \mathbf{1}_n \mathbf{1}_n^\top$  and  $\mathbf{A}_{\sqrt{n}}, \mathbf{A}_1$  given by (18) and (19) at the top of next page, where we denote

$$t_a \triangleq \frac{\text{tr}(\mathbf{C}_a - \mathbf{C}^\circ)}{\sqrt{p}} = O(1) \\ (\boldsymbol{\psi})^2 \triangleq [(\psi_1)^2, \dots, (\psi_n)^2]^\top.$$

We start with the term  $\mathbf{S}^{-1}$ . The terms of leading order in  $\mathbf{K}$ , i.e.,  $-2f'(\tau)\mathbf{A}_n$  and  $\frac{n}{\gamma}\mathbf{I}_n$  are both of operator norm  $O(n)$ . Therefore a Taylor expansion can be performed as

$$\begin{aligned} \mathbf{S}^{-1} &= \left( \mathbf{K} + \frac{n}{\gamma} \mathbf{I}_n \right)^{-1} = \frac{1}{n} \left[ \mathbf{L}^{-1} - \frac{2f'(\tau)}{n} \right. \\ &\quad \left. \left( \mathbf{A}_{\sqrt{n}} + \mathbf{A}_1 + \mathbf{P}\mathbf{\Omega}^\top \mathbf{\Omega} \mathbf{P} \right) + \frac{\beta \mathbf{I}_n}{n} + O(n^{-\frac{3}{2}}) \right]^{-1} \\ &= \frac{\mathbf{L}}{n} + \frac{2f'(\tau)}{n^2} \mathbf{L} \mathbf{A}_{\sqrt{n}} \mathbf{L} + \mathbf{L} \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{L} + O(n^{-\frac{5}{2}}) \end{aligned}$$

with  $\mathbf{L} = \left( f(\tau) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} + \frac{\mathbf{I}_n}{\gamma} \right)^{-1}$  of order  $O(1)$  and  $\mathbf{Q} = \frac{2f'(\tau)}{n^2} \left( \mathbf{A}_1 + \mathbf{P}\mathbf{\Omega}^\top \mathbf{\Omega} \mathbf{P} + \frac{2f'(\tau)}{n} \mathbf{A}_{\sqrt{n}} \mathbf{L} \mathbf{A}_{\sqrt{n}} \right)$ .

With the Sherman-Morrison formula we are able to compute explicitly  $\mathbf{L}$  as

$$\begin{aligned} \mathbf{L} &= \left( f(\tau) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} + \frac{\mathbf{I}_n}{\gamma} \right)^{-1} = \gamma \left( \mathbf{I}_n - \frac{\gamma f(\tau)}{1 + \gamma f(\tau)} \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right)^{-1} \\ &= \frac{\gamma}{1 + \gamma f(\tau)} \mathbf{I}_n + \frac{\gamma^2 f(\tau)}{1 + \gamma f(\tau)} \mathbf{P} = O(1). \end{aligned} \quad (13)$$

Writing  $\mathbf{L}$  as a linear combination of  $\mathbf{I}_n$  and  $\mathbf{P}$  is useful when computing  $\mathbf{L} \mathbf{1}_n$  or  $\mathbf{1}_n^\top \mathbf{L}$ , because by the definition of  $\mathbf{P} = \mathbf{I}_n - \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$ , we have  $\mathbf{1}_n^\top \mathbf{P} = \mathbf{P} \mathbf{1}_n = \mathbf{0}$ .

We shall start with the term  $\mathbf{1}_n^\top \mathbf{S}^{-1}$ , since it is the basis of several other terms appearing in  $\alpha$  and  $b$ ,

$$\begin{aligned} \mathbf{1}_n^\top \mathbf{S}^{-1} &= \frac{\gamma \mathbf{1}_n^\top}{1 + \gamma f(\tau)} \left[ \frac{\mathbf{I}_n}{n} + \frac{2f'(\tau)}{n^2} \mathbf{A}_{\sqrt{n}} \mathbf{L} + \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{L} \right] \\ &\quad + O(n^{-\frac{3}{2}}) \end{aligned}$$

since  $\mathbf{1}_n^\top \mathbf{L} = \frac{\gamma}{1 + \gamma f(\tau)} \mathbf{1}_n^\top$ .

With  $\mathbf{1}_n^\top \mathbf{S}^{-1}$  at hand, we next obtain,

$$\begin{aligned} \mathbf{1}_n \mathbf{1}_n^\top \mathbf{S}^{-1} &= \frac{\gamma}{1 + \gamma f(\tau)} \left[ \underbrace{\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}}_{O(1)} + \underbrace{\frac{2f'(\tau)}{n^2} \mathbf{1}_n \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \mathbf{L}}_{O(n^{-1/2})} \right. \\ &\quad \left. + \underbrace{\mathbf{1}_n \mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{L}}_{O(n^{-1})} \right] + O(n^{-\frac{3}{2}}) \end{aligned} \quad (14)$$

$$\begin{aligned} \mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{y} &= \frac{\gamma}{1 + \gamma f(\tau)} \left[ \underbrace{c_2 - c_1}_{O(1)} + \underbrace{\frac{2f'(\tau)}{n^2} \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \mathbf{L} \mathbf{y}}_{O(n^{-1/2})} \right. \\ &\quad \left. + \underbrace{\mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{L} \mathbf{y}}_{O(n^{-1})} \right] + O(n^{-\frac{3}{2}}) \end{aligned}$$

$$\begin{aligned} \mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n &= \frac{\gamma}{1 + \gamma f(\tau)} \left[ \underbrace{1}_{O(1)} + \underbrace{\frac{2f'(\tau)}{n^2} \frac{\gamma \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \mathbf{1}_n}{1 + \gamma f(\tau)}}_{O(n^{-1/2})} \right. \\ &\quad \left. + \underbrace{\frac{\gamma}{1 + \gamma f(\tau)} \mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{1}_n}_{O(n^{-1})} \right] + O(n^{-\frac{3}{2}}). \end{aligned}$$

The inverse of  $\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n$  can consequently be computed using a Taylor expansion around its leading order, allowing an error term of  $O(n^{-\frac{3}{2}})$  as

$$\begin{aligned} \frac{1}{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n} &= \frac{1 + \gamma f(\tau)}{\gamma} \left[ \underbrace{1}_{O(1)} - \underbrace{\frac{2f'(\tau)}{n^2} \frac{\gamma \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \mathbf{1}_n}{1 + \gamma f(\tau)}}_{O(n^{-1/2})} \right. \\ &\quad \left. - \underbrace{\frac{\gamma}{1 + \gamma f(\tau)} \mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{1}_n}_{O(n^{-1})} \right] + O(n^{-\frac{3}{2}}). \end{aligned} \quad (15)$$

$$\mathbf{A}_{\sqrt{n}} = -\frac{1}{2} \left[ \boldsymbol{\psi} \mathbf{1}_n^\top + \mathbf{1}_n \boldsymbol{\psi}^\top + \left\{ t_a \frac{\mathbf{1}_{n_a}}{\sqrt{p}} \right\}_{a=1}^2 \mathbf{1}_n^\top + \mathbf{1}_n \left\{ t_b \frac{\mathbf{1}_{n_b}}{\sqrt{p}} \right\}_{b=1}^2 \right] \quad (18)$$

$$\begin{aligned} \mathbf{A}_1 = & -\frac{1}{2} \left[ \left\{ \|\boldsymbol{\mu}_a - \boldsymbol{\mu}_b\|^2 \frac{\mathbf{1}_{n_a} \mathbf{1}_{n_b}^\top}{p} \right\}_{a,b=1}^2 + 2 \left\{ \frac{(\boldsymbol{\Omega} \mathbf{P})_a^\top (\boldsymbol{\mu}_b - \boldsymbol{\mu}_a) \mathbf{1}_{n_b}^\top}{\sqrt{p}} \right\}_{a,b=1}^2 - 2 \left\{ \frac{\mathbf{1}_{n_a} (\boldsymbol{\mu}_b - \boldsymbol{\mu}_a)^\top (\boldsymbol{\Omega} \mathbf{P})_b}{\sqrt{p}} \right\}_{a,b=1}^2 \right] \\ & - \frac{f''(\tau)}{4f'(\tau)} \left[ (\boldsymbol{\psi})^2 \mathbf{1}_n^\top + \mathbf{1}_n [(\boldsymbol{\psi})^2]^\top + \left\{ t_a^2 \frac{\mathbf{1}_{n_a}}{p} \right\}_{a=1}^2 \mathbf{1}_n^\top + \mathbf{1}_n \left\{ t_b^2 \frac{\mathbf{1}_{n_b}}{p} \right\}_{b=1}^2 + 2 \left\{ t_a t_b \frac{\mathbf{1}_{n_a} \mathbf{1}_{n_b}^\top}{p} \right\}_{a,b=1}^2 + 2\mathcal{D}\{t_a \mathbf{I}_{n_a}\}_{a=1}^2 \boldsymbol{\psi} \frac{\mathbf{1}_n^\top}{\sqrt{p}} \right. \\ & \left. + 2\boldsymbol{\psi} \left\{ t_b \frac{\mathbf{1}_{n_b}}{\sqrt{p}} \right\}_{b=1}^2 + 2 \frac{\mathbf{1}_n}{\sqrt{p}} (\boldsymbol{\psi})^\top \mathcal{D}\{t_a \mathbf{1}_{n_a}\}_{a=1}^2 + 2 \left\{ t_a \frac{\mathbf{1}_{n_a}}{\sqrt{p}} \right\}_{a=1}^2 (\boldsymbol{\psi})^\top + 4 \left\{ \text{tr}(\mathbf{C}_a \mathbf{C}_b) \frac{\mathbf{1}_{n_a} \mathbf{1}_{n_b}^\top}{p^2} \right\}_{a,b=1}^2 + 2\boldsymbol{\psi} (\boldsymbol{\psi})^\top \right] \quad (19) \end{aligned}$$

$$\begin{aligned} \tilde{\mathbf{k}}(\mathbf{x}) = & f'(\tau) \left[ \left\{ \frac{\|\boldsymbol{\mu}_b - \boldsymbol{\mu}_a\|^2}{p} \mathbf{1}_{n_b} \right\}_{b=1}^2 - \frac{2}{\sqrt{p}} \left\{ \mathbf{1}_{n_b} (\boldsymbol{\mu}_b - \boldsymbol{\mu}_a)^\top \right\}_{b=1}^2 \boldsymbol{\omega}_{\mathbf{x}} + \frac{2}{\sqrt{p}} \mathcal{D} \left( \left\{ \mathbf{1}_{n_b} (\boldsymbol{\mu}_b - \boldsymbol{\mu}_a)^\top \right\}_{b=1}^2 \boldsymbol{\Omega} \right) \right] \\ & + \frac{f''(\tau)}{2} \left[ \left\{ \frac{(t_a + t_b)^2}{p} \mathbf{1}_{n_b} \right\}_{b=1}^2 + 2\mathcal{D} \left( \left\{ \frac{t_a + t_b}{\sqrt{p}} \mathbf{1}_{n_b} \right\}_{b=1}^2 \right) \boldsymbol{\psi} + 2 \left\{ \frac{t_a + t_b}{\sqrt{p}} \mathbf{1}_{n_b} \right\}_{b=1}^2 \boldsymbol{\psi}_{\mathbf{x}} + (\boldsymbol{\psi})^2 + 2\boldsymbol{\psi}_{\mathbf{x}} \boldsymbol{\psi} + \boldsymbol{\psi}_{\mathbf{x}}^2 \mathbf{1}_n \right. \\ & \left. + \left\{ \frac{4}{p^2} \text{tr}(\mathbf{C}_a \mathbf{C}_b) \mathbf{1}_{n_b} \right\}_{b=1}^2 \right] \quad (20) \end{aligned}$$

Combing (14) with (15) we deduce

$$\begin{aligned} \frac{\mathbf{1}_n \mathbf{1}_n^\top \mathbf{S}^{-1}}{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n} = & \underbrace{\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}}_{O(1)} + \underbrace{\frac{2f'(\tau)}{n^2} \mathbf{1}_n \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \left[ \mathbf{L} - \frac{\gamma \mathbf{1}_n \mathbf{1}_n^\top}{1 + \gamma f(\tau)} \right]}_{O(n^{-1/2})} \\ & + \underbrace{\mathbf{1}_n \mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \left[ \mathbf{L} - \frac{\gamma \mathbf{1}_n \mathbf{1}_n^\top}{1 + \gamma f(\tau)} \right]}_{O(n^{-1})} + O(n^{-\frac{3}{2}}) \quad (16) \end{aligned}$$

and similarly the following approximation of  $b$  as

$$\begin{aligned} b = & \underbrace{c_2 - c_1}_{O(1)} - \underbrace{\frac{2\gamma}{\sqrt{p}} c_1 c_2 f'(\tau) (t_2 - t_1)}_{O(n^{-1/2})} - \underbrace{\frac{\gamma f'(\tau)}{n} \mathbf{y}^\top \mathbf{P} \boldsymbol{\psi}}_{O(n^{-1})} \\ & - \underbrace{\frac{\gamma f''(\tau)}{2n} \mathbf{y}^\top \mathbf{P} (\boldsymbol{\psi})^2 + \frac{4\gamma c_1 c_2}{p} [c_1 T_1 + (c_2 - c_1) D - c_2 T_2]}_{O(n^{-1})} \\ & + O(n^{-\frac{3}{2}}) \quad (17) \end{aligned}$$

where

$$\begin{aligned} D = & \frac{f'(\tau)}{2} \|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\|^2 + \frac{f''(\tau)}{4} (t_1 + t_2)^2 + f''(\tau) \frac{\text{tr} \mathbf{C}_1 \mathbf{C}_2}{p} \\ T_a = & f''(\tau) t_a^2 + f''(\tau) \frac{\text{tr} \mathbf{C}_1 \mathbf{C}_2}{p} \end{aligned}$$

which gives the asymptotic approximation of  $b$ .

Moving to  $\boldsymbol{\alpha}$ , note from (13) that  $\mathbf{L} - \frac{\gamma \mathbf{1}_n \mathbf{1}_n^\top}{1 + \gamma f(\tau)} = \gamma \mathbf{P}$ , and we can thus rewrite:

$$\begin{aligned} \frac{\mathbf{1}_n \mathbf{1}_n^\top \mathbf{S}^{-1}}{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n} = & \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} + \frac{2\gamma f'(\tau)}{n^2} \mathbf{1}_n \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \mathbf{P} \\ & + \gamma \mathbf{1}_n \mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{P} + O(n^{-\frac{3}{2}}). \end{aligned}$$

At this point, for  $\boldsymbol{\alpha} = \mathbf{S}^{-1} \left( \mathbf{I}_n - \frac{\mathbf{1}_n \mathbf{1}_n^\top \mathbf{S}^{-1}}{\mathbf{1}_n^\top \mathbf{S}^{-1} \mathbf{1}_n} \right) \mathbf{y}$ , we have

$$\begin{aligned} \boldsymbol{\alpha} = & \mathbf{S}^{-1} \left[ \mathbf{I}_n - \frac{2\gamma f'(\tau)}{n^2} \mathbf{1}_n \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \right. \\ & \left. - \gamma \mathbf{1}_n \mathbf{1}_n^\top \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \right] \mathbf{P} \mathbf{y} + O(n^{-\frac{5}{2}}). \end{aligned}$$

Here again, we use  $\mathbf{1}_n^\top \mathbf{L} = \frac{\gamma}{1 + \gamma f(\tau)} \mathbf{1}_n^\top$  and  $\mathbf{L} - \frac{\gamma}{1 + \gamma f(\tau)} \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} = \gamma \mathbf{P}$ , to eventually get

$$\begin{aligned} \boldsymbol{\alpha} = & \underbrace{\frac{\gamma}{n} \mathbf{P} \mathbf{y}}_{O(n^{-1})} + \underbrace{\gamma^2 \mathbf{P} \left( \mathbf{Q} - \frac{\beta}{n^2} \mathbf{I}_n \right) \mathbf{P} \mathbf{y}}_{O(n^{-2})} \\ & - \underbrace{\frac{\gamma^2}{1 + \gamma f(\tau)} \left( \frac{2f'(\tau)}{n^2} \right)^2 \mathbf{L} \mathbf{A}_{\sqrt{n}} \mathbf{1}_n \mathbf{1}_n^\top \mathbf{A}_{\sqrt{n}} \mathbf{P} \mathbf{y}}_{O(n^{-2})} + O(n^{-\frac{5}{2}}). \quad (21) \end{aligned}$$

Note here the absence of a term of order  $O(n^{-3/2})$  in the expression of  $\boldsymbol{\alpha}$  since  $\mathbf{P} \mathbf{A}_{\sqrt{n}} \mathbf{P} = 0$  from (18).

We shall now work on the vector  $\mathbf{k}(\mathbf{x})$  for a new datum  $\mathbf{x}$ , following the same analysis as in [20] for the kernel matrix  $\mathbf{K}$ , assuming that  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_a, \mathbf{C}_a)$  and recalling the random variables definitions,

$$\begin{aligned} \boldsymbol{\omega}_{\mathbf{x}} & \triangleq (\mathbf{x} - \boldsymbol{\mu}_a) / \sqrt{p} \\ \boldsymbol{\psi}_{\mathbf{x}} & \triangleq \|\boldsymbol{\omega}_{\mathbf{x}}\|^2 - \mathbb{E}\|\boldsymbol{\omega}_{\mathbf{x}}\|^2 \end{aligned}$$

we show that the  $j$ -th entry of  $\mathbf{k}(\mathbf{x})$  can be written as

$$\begin{aligned} [\mathbf{k}(\mathbf{x})]_j &= \underbrace{f(\tau)}_{O(1)} + \underbrace{f'(\tau) \left[ \frac{t_a + t_b}{\sqrt{p}} + \psi_x + \psi_j - 2(\boldsymbol{\omega}_x)^\top \boldsymbol{\omega}_j \right]}_{O(n^{-1/2})} \\ &+ \underbrace{\frac{\|\boldsymbol{\mu}_b - \boldsymbol{\mu}_a\|^2}{p} + \frac{2}{\sqrt{p}}(\boldsymbol{\mu}_b - \boldsymbol{\mu}_a)^\top (\boldsymbol{\omega}_j - \boldsymbol{\omega}_x)}_{O(n^{-1})} + \frac{f''(\tau)}{2} \\ &\underbrace{\left[ \left( \frac{t_a + t_b}{\sqrt{p}} + \psi_j + \psi_x \right)^2 + \frac{4}{p^2} \text{tr} \mathbf{C}_a \mathbf{C}_b \right]}_{O(n^{-1})} + O(n^{-\frac{3}{2}}). \end{aligned} \quad (22)$$

Combining (21) and (22), we deduce

$$\begin{aligned} \boldsymbol{\alpha}^\top \mathbf{k}(\mathbf{x}) &= \underbrace{\frac{2\gamma}{\sqrt{p}} c_1 c_2 f'(\tau) (t_2 - t_1)}_{O(n^{-1/2})} + \underbrace{\frac{\gamma}{n} \mathbf{y}^\top \mathbf{P} \tilde{\mathbf{k}}(\mathbf{x})}_{O(n^{-1})} \\ &+ \underbrace{\frac{\gamma f'(\tau)}{n} \mathbf{y}^\top \mathbf{P} (\boldsymbol{\psi} - 2\mathbf{P} \boldsymbol{\Omega}^\top \boldsymbol{\omega}_x)}_{O(n^{-1})} + O(n^{-\frac{3}{2}}) \end{aligned} \quad (23)$$

with  $\tilde{\mathbf{k}}(\mathbf{x})$  given in (20).

At this point, note that the term of order  $O(n^{-\frac{1}{2}})$  in the final object  $g(\mathbf{x}) = \boldsymbol{\alpha}^\top \mathbf{k}(\mathbf{x}) + b$  disappears because in both (17) and (23) the term of order  $O(n^{-1/2})$  is  $\frac{2\gamma}{\sqrt{p}} c_1 c_2 f'(\tau) (t_2 - t_1)$  but of opposite signs. Also, we see that the leading term  $c_2 - c_1$  in  $b$  will remain in  $g(\mathbf{x})$  as stated in Remark 2.

The development of  $\mathbf{y}^\top \mathbf{P} \tilde{\mathbf{k}}(\mathbf{x})$  induces many simplifications, since i)  $\mathbf{P} \mathbf{1}_n = \mathbf{0}$  and ii) random variables as  $\boldsymbol{\omega}_x$  and  $\boldsymbol{\psi}$  in  $\tilde{\mathbf{k}}(\mathbf{x})$ , once multiplied by  $\mathbf{y}^\top \mathbf{P}$ , thanks to probabilistic averaging of independent zero-mean terms, are of smaller order and thus become negligible. We thus get

$$\begin{aligned} \frac{\gamma}{n} \mathbf{y}^\top \mathbf{P} \tilde{\mathbf{k}}(\mathbf{x}) &= 2\gamma c_1 c_2 f'(\tau) \left[ \frac{\|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_a\|^2 - \|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_a\|^2}{p} \right. \\ &\quad \left. - 2(\boldsymbol{\omega}_x)^\top \frac{\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1}{\sqrt{p}} \right] + \frac{\gamma f''(\tau)}{2n} \mathbf{y}^\top \mathbf{P} (\boldsymbol{\psi})^2 + \gamma c_1 c_2 f''(\tau) \left[ \right. \\ &\quad \left. 2 \left( \frac{t_a}{\sqrt{p}} + \psi_x \right) \frac{t_2 - t_1}{\sqrt{p}} + \frac{t_2^2 - t_1^2}{p} + \frac{4}{p^2} \text{tr}(\mathbf{C}_a \mathbf{C}_2 - \mathbf{C}_a \mathbf{C}_1) \right] \\ &\quad + O(n^{-\frac{3}{2}}). \end{aligned} \quad (24)$$

This result, together with (23), completes the analysis of the term  $\boldsymbol{\alpha}^\top \mathbf{k}(\mathbf{x})$ . Combining (23)-(24) with (17) we conclude the proof of Theorem 1.

## APPENDIX B PROOF OF THEOREM 2

This section is dedicated to the proof of the central limit theorem for

$$\hat{g}(\mathbf{x}) = c_2 - c_1 + \gamma(\mathfrak{P} + c_x \mathfrak{D})$$

with the shortcut  $c_x = -2c_1 c_2^2$  for  $\mathbf{x} \in \mathcal{C}_1$  and  $c_x = 2c_1^2 c_2$  for  $\mathbf{x} \in \mathcal{C}_2$ , and  $\mathfrak{P}, \mathfrak{D}$  as defined in (7) and (8).

Our objective is to show that for  $a \in \{1, 2\}$ ,  $n(\hat{g}(\mathbf{x}) - G_a) \xrightarrow{d} 0$  with

$$G_a \sim \mathcal{N}(E_a, \text{Var}_a)$$

where  $E_a$  and  $\text{Var}_a$  are given in Theorem 2. We recall that  $\mathbf{x} = \boldsymbol{\mu}_a + \sqrt{p} \boldsymbol{\omega}_x$  with  $\boldsymbol{\omega}_x \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_a/p)$ .

Letting  $\mathbf{z}_x$  such that  $\boldsymbol{\omega}_x = \mathbf{C}_a^{1/2} \mathbf{z}_x / \sqrt{p}$ , we have  $\mathbf{z}_x \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$  and we can rewrite  $\hat{g}(\mathbf{x})$  in the following quadratic form (of  $\mathbf{z}_x$ ) as

$$\hat{g}(\mathbf{x}) = \mathbf{z}_x^\top \mathbf{A} \mathbf{z}_x + \mathbf{z}_x^\top \mathbf{b} + c$$

with

$$\begin{aligned} \mathbf{A} &= 2\gamma c_1 c_2 f''(\tau) \frac{\text{tr}(\mathbf{C}_2 - \mathbf{C}_1)}{p} \frac{\mathbf{C}_a}{p} \\ \mathbf{b} &= -\frac{2\gamma f'(\tau)}{n} \frac{(\mathbf{C}_a)^{\frac{1}{2}}}{\sqrt{p}} \boldsymbol{\Omega} \mathbf{P} \mathbf{y} - \frac{4c_1 c_2 \gamma f'(\tau)}{\sqrt{p}} \frac{(\mathbf{C}_a)^{\frac{1}{2}}}{\sqrt{p}} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) \\ c &= c_2 - c_1 + \gamma c_x \mathfrak{D} - 2\gamma c_1 c_2 f''(\tau) \frac{\text{tr}(\mathbf{C}_2 - \mathbf{C}_1)}{p} \frac{\text{tr} \mathbf{C}_a}{p}. \end{aligned}$$

Since  $\mathbf{z}_x$  is (standard) Gaussian and has the same distribution as  $\mathbf{U} \mathbf{z}_x$  for any orthogonal matrix  $\mathbf{U}$  (i.e., such that  $\mathbf{U}^\top \mathbf{U} = \mathbf{U} \mathbf{U}^\top = \mathbf{I}_n$ ), we choose  $\mathbf{U}$  that diagonalize  $\mathbf{A}$  such that  $\mathbf{A} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top$ , with  $\boldsymbol{\Lambda}$  diagonal so that  $\hat{g}(\mathbf{x})$  and  $\tilde{g}(\mathbf{x})$  have the same distribution where

$$\tilde{g}(\mathbf{x}) = \mathbf{z}_x^\top \boldsymbol{\Lambda} \mathbf{z}_x + \mathbf{z}_x^\top \tilde{\mathbf{b}} + c = \sum_{i=1}^n \left( z_i^2 \lambda_i + z_i \tilde{b}_i + \frac{c}{n} \right)$$

and  $\tilde{\mathbf{b}} = \mathbf{U}^\top \mathbf{b}$ ,  $\lambda_i$  the diagonal elements of  $\boldsymbol{\Lambda}$  and  $z_i$  the elements of  $\mathbf{z}_x$ .

Conditioning on  $\boldsymbol{\Omega}$ , we thus result in the sum of independent but not identically distributed random variables  $r_i = z_i^2 \lambda_i + z_i \tilde{b}_i + \frac{c}{n}$ . We then resort to the Lyapunov CLT [33, Theorem 27.3].

We begin by estimating the expectation and the variance

$$\begin{aligned} \mathbb{E}[r_i | \boldsymbol{\Omega}] &= \lambda_i + \frac{c}{n} \\ \text{Var}[r_i | \boldsymbol{\Omega}] &= \sigma_i^2 = 2\lambda_i^2 + \tilde{b}_i^2 \end{aligned}$$

of  $r_i$ , so that

$$\begin{aligned} \sum_{i=1}^n \mathbb{E}[r_i | \boldsymbol{\Omega}] &= c_2 - c_1 + \gamma c_x \mathfrak{D} = E_a \\ s_n^2 &= \sum_{i=1}^n \sigma_i^2 = 2 \text{tr}(\mathbf{A}^2) + \mathbf{b}^\top \mathbf{b} \\ &= 8\gamma^2 c_1^2 c_2^2 (f''(\tau))^2 \frac{(\text{tr}(\mathbf{C}_2 - \mathbf{C}_1))^2}{p^2} \frac{\text{tr} \mathbf{C}_a^2}{p^2} \\ &\quad + 4\gamma^2 \left( \frac{f'(\tau)}{n} \right)^2 \mathbf{y}^\top \mathbf{P} \boldsymbol{\Omega}^\top \frac{\mathbf{C}_a}{p} \boldsymbol{\Omega} \mathbf{P} \mathbf{y} \\ &\quad + \frac{16\gamma^2 c_1^2 c_2^2 (f'(\tau))^2}{p} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^\top \frac{\mathbf{C}_a}{p} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) \\ &\quad + O(n^{-\frac{5}{2}}). \end{aligned}$$

We shall rewrite  $\boldsymbol{\Omega}$  into two blocks as:

$$\boldsymbol{\Omega} = \left[ \frac{(\mathbf{C}_1)^{\frac{1}{2}}}{\sqrt{p}} \mathbf{Z}_1, \quad \frac{(\mathbf{C}_2)^{\frac{1}{2}}}{\sqrt{p}} \mathbf{Z}_2 \right]$$

where  $\mathbf{Z}_1 \in \mathbb{R}^{p \times n_1}$  and  $\mathbf{Z}_2 \in \mathbb{R}^{p \times n_2}$  with i.i.d. Gaussian entries with zero mean and unit variance. Then

$$\boldsymbol{\Omega}^\top \frac{\mathbf{C}_a}{p} \boldsymbol{\Omega} = \frac{1}{p^2} \begin{bmatrix} \mathbf{Z}_1^\top (\mathbf{C}_1)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_1)^{\frac{1}{2}} \mathbf{Z}_1 & \mathbf{Z}_1^\top (\mathbf{C}_1)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{Z}_2 \\ \mathbf{Z}_2^\top (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_1)^{\frac{1}{2}} \mathbf{Z}_1 & \mathbf{Z}_2^\top (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{Z}_2 \end{bmatrix}$$

and with  $\mathbf{P}\mathbf{y} = \mathbf{y} - (c_2 - c_1)\mathbf{1}_n$ , we deduce

$$\begin{aligned} \mathbf{y}^\top \mathbf{P} \boldsymbol{\Omega}^\top \frac{\mathbf{C}_a}{p} \boldsymbol{\Omega} \mathbf{P} \mathbf{y} &= \frac{4}{p^2} \left( c_1^2 \mathbf{1}_{n_1}^\top \mathbf{Z}_1^\top (\mathbf{C}_1)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_1)^{\frac{1}{2}} b \mathbf{Z}_1 \mathbf{1}_{n_1} \right. \\ &\quad \left. - 2c_1 c_2 \mathbf{1}_{n_1}^\top \mathbf{Z}_1^\top (\mathbf{C}_1)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{Z}_2 \mathbf{1}_{n_2} \right. \\ &\quad \left. + c_2^2 \mathbf{1}_{n_1}^\top \mathbf{Z}_2^\top (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{C}_a (\mathbf{C}_2)^{\frac{1}{2}} \mathbf{Z}_2 \mathbf{1}_{n_2} \right). \end{aligned}$$

Since  $\mathbf{Z}_i \mathbf{1}_{n_i} \sim \mathcal{N}(\mathbf{0}, n_i \mathbf{I}_{n_i})$ , by applying the trace lemma [39] Lemma B.26] we get

$$\mathbf{y}^\top \mathbf{P} \boldsymbol{\Omega}^\top \frac{\mathbf{C}_a}{p} \boldsymbol{\Omega} \mathbf{P} \mathbf{y} - \frac{4nc_1^2 c_2^2}{p^2} \left( \frac{\text{tr } \mathbf{C}_1 \mathbf{C}_a}{c_1} + \frac{\text{tr } \mathbf{C}_2 \mathbf{C}_a}{c_2} \right) \xrightarrow{\text{a.s.}} 0. \quad (25)$$

Consider now the events

$$\begin{aligned} E &= \left\{ \left| \mathbf{y}^\top \mathbf{P} \boldsymbol{\Omega}^\top \frac{\mathbf{C}_a}{p} \boldsymbol{\Omega} \mathbf{P} \mathbf{y} - \rho \right| < \epsilon \right\} \\ \bar{E} &= \left\{ \left| \mathbf{y}^\top \mathbf{P} \boldsymbol{\Omega}^\top \frac{\mathbf{C}_a}{p} \boldsymbol{\Omega} \mathbf{P} \mathbf{y} - \rho \right| > \epsilon \right\} \end{aligned}$$

for any fixed  $\epsilon$  with  $\rho = \frac{4nc_1^2 c_2^2}{p^2} \left( \frac{\text{tr } \mathbf{C}_1 \mathbf{C}_a}{c_1} + \frac{\text{tr } \mathbf{C}_2 \mathbf{C}_a}{c_2} \right)$  and write

$$\begin{aligned} \mathbb{E} \left[ \exp \left( iun \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n} \right) \right] &= \mathbb{E} \left[ \exp \left( iun \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n} \right) \middle| E \right] \\ \mathbb{P}(E) + \mathbb{E} \left[ \exp \left( iun \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n} \right) \middle| \bar{E} \right] \mathbb{P}(\bar{E}) & \quad (26) \end{aligned}$$

We start with the variable  $\tilde{g}(\mathbf{x})|E$  and check that Lyapunov's condition for  $\tilde{r}_i = r_i - \mathbb{E}[r_i]$ , conditioning on  $E$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^4} \sum_{i=1}^n \mathbb{E}[|\tilde{r}_i|^4] = 0$$

holds by rewriting

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^4} \sum_{i=1}^n \mathbb{E}[|\tilde{r}_i|^4] = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{60\lambda_i^4 + 12\lambda_i^2 \tilde{b}_i^2 + 3\tilde{b}_i^4}{s_n^4} = 0$$

since both  $\lambda_i$  and  $\tilde{b}_i$  are of order  $O(n^{-3/2})$ .

As a consequence of the above, we have the CLT for the random variable  $\tilde{g}(\mathbf{x})|E$ , thus

$$\mathbb{E} \left[ \exp \left( iun \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n} \right) \middle| E \right] \rightarrow \exp(-\frac{u^2}{2}).$$

Next, we see that the second term in (26) goes to zero because  $|\mathbb{E}[\exp(iun \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n}) | \bar{E}]| \leq 1$  and  $\mathbb{P}(\bar{E}) \rightarrow 0$  from (25) and we eventually deduce

$$\mathbb{E} \left[ \exp \left( iun \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n} \right) \right] \rightarrow \exp(-\frac{u^2}{2}).$$

With the help of Lévy's continuity theorem, we thus prove the CLT of the variable  $n \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{s_n}$ . Since  $s_n^2 \rightarrow \text{Var}_a$ , with Slutsky's theorem, we have the CLT for  $n \frac{\tilde{g}(\mathbf{x}) - \mathbb{E}_a}{\sqrt{\text{Var}_a}}$  (thus for  $n \frac{\hat{g}(\mathbf{x}) - \mathbb{E}_a}{\sqrt{\text{Var}_a}}$ ), and eventually for  $n \frac{g(\mathbf{x}) - \mathbb{E}_a}{\sqrt{\text{Var}_a}}$  by Theorem 1 which completes the proof.