SUPPLEMENTARY MATERIAL TO 'STATISTICAL AND COMPUTATIONAL TRADE-OFFS IN ESTIMATION OF SPARSE PRINCIPAL COMPONENTS'

By Tengyao Wang[*] Quentin Berthet[*,†] Richard J. Samworth[*]

University of Cambridge[*]
California Institute of Technology[†]

1. Ancillary results. We collect here various results used in the proofs in Appendices A, B and C in the main document Wang, Berthet and Samworth (2015).

Proposition 1. *Let $P \in \mathrm{RCC}_p(n, \ell, C)$ and suppose that $\ell \log p \leq n$. Then*

$$\mathbb{E} \sup_{u \in B_0(\ell)} |\hat{V}(u) - V(u)| \leq \biggl(1 + \frac{1}{\log p}\biggr) C \sqrt{\frac{\ell \log p}{n}}.$$

Proof. By setting $\delta = p^{1-t}$ in the RCC condition, we find that

$$\mathbb{P}\biggl(\sup_{u \in B_0(\ell)} |\hat{V}(u) - V(u)| \geq C \max\biggl\{\sqrt{\frac{t\ell \log p}{n}}, \frac{t\ell \log p}{n}\biggr\}\biggr) \leq \min(1, p^{1-t})$$

for all $t \geq 0$. It follows that

$$\begin{aligned}
\mathbb{E} \sup_{u \in B_0(\ell)} |\hat{V}(u) - V(u)| &= \int_0^\infty \mathbb{P}\biggl(\sup_{u \in B_0(\ell)} |\hat{V}(u) - V(u)| \geq s\biggr)\, ds \\
&\leq C\sqrt{\frac{\ell \log p}{n}} + C\sqrt{\frac{\ell \log p}{n}} \int_1^{\frac{n}{\ell \log p}} \frac{1}{2} p^{1-t} t^{-1/2}\, dt + C\frac{\ell \log p}{n} \int_{\frac{n}{\ell \log p}}^\infty p^{1-t}\, dt \\
&\leq C\sqrt{\frac{\ell \log p}{n}} \biggl\{1 + \int_1^\infty p^{1-t}\, dt\biggr\} = \biggl(1 + \frac{1}{\log p}\biggr) C\sqrt{\frac{\ell \log p}{n}},
\end{aligned}$$

as required. ☐

Lemma 2. *Let $\epsilon \in (0, 1/2)$, let $\ell \in \{1, \ldots, p\}$ and let $A \in \mathbb{R}^{p \times p}$ be a symmetric matrix. Then there exists $\mathcal{N}_\epsilon \subseteq B_0(\ell)$ with cardinality at most $\binom{p}{\ell} \pi \ell^{1/2} (1 - \epsilon^2/16)^{-(\ell-1)/2} (2/\epsilon)^{\ell-1}$ such that*

$$\sup_{u \in B_0(\ell)} |u^\top A u| \leq (1 - 2\epsilon)^{-1} \max_{u \in \mathcal{N}_\epsilon} |u^\top A u|.$$

PROOF. Let $\mathcal{I}_\ell := \bigl\{I \subseteq \{1,\ldots,p\} : |I| = \ell\bigr\}$, and for $I \in \mathcal{I}_\ell$, let $B_I := \{u \in B_0(\ell) : u_{I^c} = 0\}$. Thus

$$B_0(\ell) = \bigcup_{I \in \mathcal{I}_\ell} B_I.$$

For each $I \in \mathcal{I}_\ell$, by Lemma 10 of Kim and Samworth (2014), there exists $\mathcal{N}_{I,\epsilon} \subseteq B_I$ such that $|\mathcal{N}_{I,\epsilon}| \leq \pi \ell^{1/2}(1-\epsilon^2/16)^{-(\ell-1)/2}(2/\epsilon)^{\ell-1}$ and such that for any $x \in B_I$, there exists $x' \in \mathcal{N}_{I,\epsilon}$ with $\|x - x'\| \leq \epsilon$. Let $u_I \in \mathrm{argmax}_{u \in B_I} |u^\top A u|$ and find $v_I \in \mathcal{N}_{I,\epsilon}$ such that $\|u_I - v_I\| \leq \epsilon$. Then

$$\begin{aligned}|u_I^\top A u_I| &\leq |v_I^\top A v_I| + |(u_I - v_I)^\top A v_I| + |u_I^\top A(u_I - v_I)| \\ &\leq \max_{u \in \mathcal{N}_{I,\epsilon}} |u^\top A u| + 2\epsilon |u_I^\top A u_I|.\end{aligned}$$

Writing $\mathcal{N}_\epsilon := \cup_{I \in \mathcal{I}_\ell} \mathcal{N}_{I,\epsilon}$, we note that $\mathcal{N}_\epsilon$ has cardinality no larger than $\binom{p}{\ell}\pi \ell^{1/2}(1-\epsilon^2/16)^{-(\ell-1)/2}(2/\epsilon)^{\ell-1}$ and that

$$\begin{aligned}\sup_{u \in B_0(\ell)} |u^\top A u| = \max_{I \in \mathcal{I}_\ell} \sup_{u \in B_I} |u^\top A u| &\leq (1-2\epsilon)^{-1} \max_{I \in \mathcal{I}_\ell} \max_{u \in \mathcal{N}_{I,\epsilon}} |u^\top A u| \\ &= (1-2\epsilon)^{-1} \max_{u \in \mathcal{N}_\epsilon} |u^\top A u|,\end{aligned}$$

as required. □

LEMMA 3 (Variant of the Gilbert–Varshamov Lemma). *Let $\alpha, \beta \in (0,1)$ and $k, p \in \mathbb{N}$ be such that $k \leq \alpha\beta p$. Writing $\mathcal{S} := \bigl\{x = (x_1,\ldots,x_p)^\top \in \{0,1\}^p : \sum_{j=1}^p x_j = k\bigr\}$, there exists a subset $\mathcal{S}_0$ of $\mathcal{S}$ such that for all distinct $x = (x_1,\ldots,x_p)^\top, y = (y_1,\ldots,y_p)^\top \in \mathcal{S}_0$, we have $\sum_{j=1}^p \mathbb{1}_{\{x_j \neq y_j\}} \geq 2(1-\alpha)k$ and such that*

$$\log|\mathcal{S}_0| \geq \rho k \log(p/k),$$

where $\rho := \frac{\alpha}{-\log(\alpha\beta)}(-\log\beta + \beta - 1)$.

PROOF. See Massart (2007, Lemma 4.10). □

Let P and Q be two probability measures on a measurable space $(\mathcal{X}, \mathcal{B})$. Recall that if P is absolutely continuous with respect to Q, then the Kullback–Leibler divergence between P and Q is $D(P\|Q) := \int_{\mathcal{X}} \log(dP/dQ)\, dP$, where dP/dQ denotes the Radon–Nikodym derivative of P with respect to Q. If P is not absolutely continuous with respect to Q, we set $D(P\|Q) := \infty$.

LEMMA 4 (Generalised Fano's Lemma). *Let $P_1, \ldots, P_M$ be probability distributions on a measurable space $(\mathcal{X}, \mathcal{B})$, and assume that $D(P_i \| P_j) \leq \beta$ for all $i \neq j$. Then any measurable function $\hat{\psi} : \mathcal{X} \to \{1, \ldots, M\}$ satisfies*

$$\max_{1 \leq i \leq M} P_i(\hat{\psi} \neq i) \geq 1 - \frac{\beta + \log 2}{\log M}.$$

PROOF. See Yu (1997, Lemma 3). □

LEMMA 5. *Suppose that $P \in \mathcal{P}$ and that $X_1, \ldots, X_n \overset{iid}{\sim} P$. Let $\Sigma := \int_{\mathbb{R}^p} xx^\top \, dP(x)$ and $\hat{\Sigma} := n^{-1} \sum_{i=1}^n X_i X_i^\top$. If $V(u) := \mathbb{E}\{(u^\top X_1)^2\}$ and $\hat{V}(u) := n^{-1} \sum_{i=1}^n (u^\top X_i)^2$ for $u \in B_0(2)$, then*

$$\|\hat{\Sigma} - \Sigma\|_\infty \leq 2 \sup_{u \in B_0(2)} \bigl|\hat{V}(u) - V(u)\bigr|.$$

PROOF. Let e_r denote the rth standard basis vector in $\mathbb{R}^p$ and write $X_i = (X_{i,1}, \ldots, X_{i,p})^\top$. Then

$$\begin{aligned}
\|\hat{\Sigma} - \Sigma\|_\infty &= \max_{r,s \in \{1,\ldots,p\}} \biggl| \frac{1}{n} \sum_{i=1}^n (X_{i,r} X_{i,s}) - \mathbb{E}(X_{1,r} X_{1,s}) \biggr| \\
&\leq \max_{r,s \in \{1,\ldots,p\}} \biggl| \frac{1}{n} \sum_{i=1}^n \Bigl\{ \Bigl(\frac{1}{2} e_r + \frac{1}{2} e_s\Bigr)^\top X_i \Bigr\}^2 - \mathbb{E}\Bigl[\Bigl\{ \Bigl(\frac{1}{2} e_r + \frac{1}{2} e_s\Bigr)^\top X_1 \Bigr\}^2 \Bigr] \biggr| \\
&\quad + \max_{r,s \in \{1,\ldots,p\}} \biggl| \frac{1}{n} \sum_{i=1}^n \Bigl\{ \Bigl(\frac{1}{2} e_r - \frac{1}{2} e_s\Bigr)^\top X_i \Bigr\}^2 - \mathbb{E}\Bigl[\Bigl\{ \Bigl(\frac{1}{2} e_r - \frac{1}{2} e_s\Bigr)^\top X_1 \Bigr\}^2 \Bigr] \biggr| \\
&\leq 2 \sup_{u \in B_0(2)} \bigl| \hat{V}(u) - V(u) \bigr|,
\end{aligned}$$

as required. □

Recall the definition of the Graph Vector distribution $\mathrm{GV}_p^g(\pi_0)$ from the proof of Theorem 6 in the main document Wang, Berthet and Samworth (2015).

LEMMA 6. *Let $g = (g_1, \ldots, g_p)^\top \in \{0,1\}^p$, and let $Y_1, \ldots, Y_n$ be independent random vectors, each distributed as $\mathrm{GV}_p^g(\pi_0)$ for some $\pi_0 \in (0, 1/2]$. For any $u \in B_0(\ell)$, let $V(u) := \mathbb{E}\{(u^\top Y_1)^2\}$ and $\hat{V}(u) := n^{-1} \sum_{i=1}^n (u^\top Y_i)^2$. Then for every $1 \leq \ell \leq 2/\pi_0$, every $n \in \mathbb{N}$ and every $\delta > 0$,*

$$\mathbb{P}\biggl[\sup_{u \in B_0(\ell)} |\hat{V}(u) - V(u)| \geq 750 \max\biggl\{ \sqrt{\frac{\ell \log(p/\delta)}{n}}, \frac{\ell \log(p/\delta)}{n} \biggr\} \biggr] \leq \delta.$$

In other words, $\mathrm{GV}_p^g(\pi_0) \in \mathrm{RCC}_p(\ell, 750)$ for all $\pi_0 \in (0, 1/2]$ and $\ell \leq 2/\pi_0$.

PROOF. We can write

$$Y_i = \xi_i\big\{(1-\epsilon_i)R_i + \epsilon_i(g+\tilde{R}_i)\big\},$$

where ξ_i, ϵ_i and R_i are independent, where ξ_i is a Rademacher random variable, where $\epsilon_i \sim \mathrm{Bern}(\pi_0)$, where $R_i = (r_{i1},\ldots,r_{ip})^\top$ has independent Rademacher coordinates, and where $\tilde{R}_i = (\tilde{r}_{i1},\ldots,\tilde{r}_{ip})^\top$ with $\tilde{r}_{ij} := (1-g_j)r_{ij}$. Thus, for any $u \in B_0(\ell)$, we have

$$(u^\top Y_i)^2 = (1-\epsilon_i)(u^\top R_i)^2 + \epsilon_i(u^\top g)^2 + \epsilon_i(u^\top \tilde{R}_i)^2 + 2\epsilon_i(u^\top\tilde{R}_i)(u^\top g).$$

Hence, writing $S := \{j : g_j = 1\}$,

$$\begin{aligned}|\hat{V}(u) - V(u)| &\leq \biggl|\frac{1}{n}\sum_{i=1}^n (1-\epsilon_i)(u^\top R_i)^2 - (1-\pi_0)\biggr| + \frac{(u^\top g)^2}{n}\biggl|\sum_{i=1}^n(\epsilon_i - \pi_0)\biggr| \\ &\quad + \biggl|\frac{1}{n}\sum_{i=1}^n \epsilon_i(u^\top\tilde{R}_i)^2 - \pi_0\|u_{S^c}\|_2^2\biggr| + \biggl|\frac{2u^\top g}{n}\sum_{i=1}^n\epsilon_i(u^\top\tilde{R}_i)\biggr| \\ &\leq \biggl|\frac{1}{n}\sum_{i=1}^n(1-\epsilon_i)\big\{(u^\top R_i)^2 - 1\big\}\biggr| + \frac{1+(u^\top g)^2 + \|u_{S^c}\|_2^2}{n}\biggl|\sum_{i=1}^n(\epsilon_i-\pi_0)\biggr| \\ &\quad + \biggl|\frac{1}{n}\sum_{i=1}^n\epsilon_i\big\{(u^\top\tilde{R}_i)^2 - \|u_{S^c}\|_2^2\big\}\biggr| + \biggl|\frac{2u^\top g}{n}\sum_{i=1}^n\epsilon_i(u^\top\tilde{R}_i)\biggr|. \qquad (1)\end{aligned}$$

We now control the four terms on the right-hand side of (1) separately. For the first term, note that the distribution of R_i is subgaussian with parameter 1. Writing $N_{\boldsymbol\epsilon} := \sum_{i=1}^n \epsilon_i$, it follows by the same argument as in the proof of Proposition 1(i) in Wang, Berthet and Samworth (2015) that for any $s > 0$,

$$\begin{aligned}&\mathbb{P}\biggl(\sup_{u\in B_0(\ell)}\biggl|\frac{1}{n}\sum_{i=1}^n(1-\epsilon_i)\big\{(u^\top R_i)^2-1\big\}\biggr| \geq 2s\biggr) \\ &\qquad = \mathbb{E}\biggl\{\mathbb{P}\biggl(\sup_{u\in B_0(\ell)}\biggl|\frac{1}{n-N_{\boldsymbol\epsilon}}\sum_{i:\epsilon_i=0}\big\{(u^\top R_i)^2-1\big\}\biggr| \geq \frac{2ns}{n-N_{\boldsymbol\epsilon}}\biggm| N_{\boldsymbol\epsilon}\biggr)\biggr\} \\ &\qquad \leq e^9 p^\ell \mathbb{E}\biggl[\exp\biggl\{-\frac{n(\frac{ns}{n-N_{\boldsymbol\epsilon}})^2}{4(\frac{ns}{n-N_{\boldsymbol\epsilon}})+32}\biggr\}\biggr] \leq e^9 p^\ell \exp\biggl(-\frac{ns^2}{4s+32}\biggr).\end{aligned}$$

We deduce that for any $\delta > 0$,

$$\begin{aligned}\mathbb{P}\biggl(\sup_{u\in B_0(\ell)}\biggl|\frac{1}{n}\sum_{i=1}^n(1-\epsilon_i)\big\{(u^\top R_i)^2-1\big\}\biggr| \geq 16\max\biggl\{\sqrt{\frac{\ell\log(p/\delta)}{n}}, \frac{\ell\log(p/\delta)}{n}\biggr\}\biggr) \\ \leq e^9\delta. \qquad (2)\end{aligned}$$

For the second term on the right-hand side of (1), note first that for any $u \in B_0(\ell)$, we have by Cauchy–Schwarz that

$$(u^\top g)^2 \leq \|u_S\|_0 \|u_S\|_2^2 \leq \|u_S\|_0 \leq \ell.$$

We deduce using Bernstein's inequality for Binomial random variables (e.g. Shorack and Wellner, 1986, p. 855) that for any $s > 0$,

$$\begin{aligned}
&\mathbb{P}\biggl\{\sup_{u \in B_0(\ell)} \frac{1 + (u^\top g)^2 + \|u_{S^c}\|_2^2}{n} \biggl|\sum_{i=1}^n (\epsilon_i - \pi_0)\biggr| \geq s\biggr\} \\
&\qquad \leq \mathbb{P}\biggl\{\frac{1}{n}\biggl|\sum_{i=1}^n (\epsilon_i - \pi_0)\biggr| \geq \frac{s}{3\ell}\biggr\} \leq 2\exp\biggl(-\frac{ns^2}{18\ell^2\pi_0 + 2s\ell}\biggr) \\
&\qquad \leq 2\max\biggl\{\exp\biggl(-\frac{ns^2}{(19+\sqrt{37})\ell^2\pi_0}\biggr), \exp\biggl(-\frac{ns}{(1+\sqrt{37})\ell}\biggr)\biggr\}.
\end{aligned}$$

By assumption, $\ell\pi_0 \leq 2$. Hence, for any $\delta > 0$,

$$\begin{aligned}
&\mathbb{P}\biggl\{\sup_{u \in B_0(\ell)} \frac{1 + (u^\top g)^2 + \|u_{S^c}\|_2^2}{n} \biggl|\sum_{i=1}^n (\epsilon_i - \pi_0)\biggr| \geq \\
&\qquad (1+\sqrt{37})\max\biggl(\sqrt{\frac{\ell\log(1/\delta)}{n}}, \frac{\ell\log(1/\delta)}{n}\biggr)\biggr\} \leq 2\delta. \qquad (3)
\end{aligned}$$

The third term on the right-hand side of (1) can be handled in a very similar way to the first. We find that for every $\delta > 0$,

$$\begin{aligned}
&\mathbb{P}\biggl(\sup_{u \in B_0(\ell)} \biggl|\frac{1}{n}\sum_{i=1}^n \epsilon_i\bigl\{(u^\top \tilde{R}_i)^2 - \|u_{S^c}\|_2^2\bigr\}\biggr| \geq \\
&\qquad 16\max\biggl\{\sqrt{\frac{\ell\log(p/\delta)}{n}}, \frac{\ell\log(p/\delta)}{n}\biggr\}\biggr) \leq e^9\delta. \qquad (4)
\end{aligned}$$

For the final term, by definition of $\tilde{R}_i$, we have for any $u \in B_0(\ell)$ that

$$\biggl|\frac{2u^\top g}{n}\sum_{i=1}^n \epsilon_i(u^\top \tilde{R}_i)\biggr| \leq \frac{2\ell^{1/2}}{n}\biggl|\sum_{j:g_j=0} u_j \sum_{i:\epsilon_i=1} r_{ij}\biggr| \leq \frac{2\ell}{n}\max_{j:g_j=0}\biggl|\sum_{i:\epsilon_i=1} r_{ij}\biggr|.$$

Hence by Hoeffding's inequality, for any $s > 0$,

$$\mathbb{P}\biggl\{\sup_{u\in B_0(\ell)}\biggl|\frac{2u^\top g}{n}\sum_{i=1}^n \epsilon_i(u^\top \tilde{R}_i)\biggr| \geq s\biggr\} \leq \mathbb{E}\biggl\{\mathbb{P}\biggl(\max_{1\leq j\leq p}\biggl|\sum_{i:\epsilon_i=1} r_{ij}\biggr| \geq \frac{ns}{2\ell}\biggm| N_{\boldsymbol{\epsilon}}\biggr)\biggr\}$$
$$\leq 2p\mathbb{E}\biggl\{\exp\biggl(-\frac{n^2s^2}{8\ell^2 N_{\boldsymbol{\epsilon}}}\biggr)\biggr\} \leq 2p\inf_{t>0}\biggl\{\exp\biggl(-\frac{n^2s^2}{8\ell^2 t}\biggr) + \mathbb{P}(N_{\boldsymbol{\epsilon}} > t)\biggr\}$$
$$\leq 2p\inf_{t>0}\biggl\{\exp\biggl(-\frac{n^2s^2}{8\ell^2 t}\biggr) + \exp\biggl(-t\log\frac{t}{n\pi_0} + t - n\pi_0\biggr)\biggr\},$$

where the final line follows by Bennett's inequality (e.g. Shorack and Wellner, 1986, p. 440). Choosing $t = \max\bigl(e^2 n\pi_0, \frac{ns}{2^{3/2}\ell}\bigr)$, we find

$$\mathbb{P}\biggl\{\sup_{u\in B_0(\ell)}\biggl|\frac{2u^\top g}{n}\sum_{i=1}^n \epsilon_i(u^\top \tilde{R}_i)\biggr| \geq s\biggr\}$$
$$\leq 2p\max\biggl\{\exp\biggl(-\frac{ns^2}{8e^2\ell^2\pi_0}\biggr) + \exp\biggl(-\frac{ns}{2^{3/2}\ell}\biggr), 2\exp\biggl(-\frac{ns}{2^{3/2}\ell}\biggr)\biggr\}$$
$$\leq 4p\max\biggl\{\exp\biggl(-\frac{ns^2}{16e^2\ell}\biggr), \exp\biggl(-\frac{ns}{2^{3/2}\ell}\biggr)\biggr\}.$$

We deduce that for any $\delta > 0$,

(5)
$$\mathbb{P}\biggl[\sup_{u\in B_0(\ell)}\biggl|\frac{2u^\top g}{n}\sum_{i=1}^n \epsilon_i(u^\top \tilde{R}_i)\biggr| \geq 4e\max\biggl\{\sqrt{\frac{\ell\log(p/\delta)}{n}}, \frac{\ell\log(p/\delta)}{n}\biggr\}\biggr] \leq 4\delta.$$

We conclude from (1), (2), (3), (4) and (5) that for any $\delta > 0$,

$$\mathbb{P}\biggl[\sup_{u\in B_0(\ell)}|\hat{V}(u) - V(u)| \geq 750\max\biggl\{\sqrt{\frac{\ell\log(p/\delta)}{n}}, \frac{\ell\log(p/\delta)}{n}\biggr\}\biggr] \leq \delta,$$

as required. □

LEMMA 7. *Let $v = (v_1, \ldots, v_p)^\top \in B_0(k)$ and let $\hat{v} = (\hat{v}_1, \ldots, \hat{v}_p)^\top \in \mathbb{R}^p$ be such that $\|\hat{v}\|_2 = 1$. Let $S := \{j \in \{1, \ldots, p\} : v_j \neq 0\}$. Then for any $\hat{S} \in \mathrm{argmax}_{1\leq j_1<\ldots<j_k\leq p}\sum_{r=1}^k |\hat{v}_{j_r}|$, we have*

$$L(\hat{v}, v)^2 \geq \frac{1}{2}\sum_{j\in S\setminus\hat{S}} v_j^2.$$

PROOF. By the Cauchy–Schwarz inequality, and then by definition of $\hat{S}$,

$$
\begin{aligned}
1 - L(\hat{v}, v)^2 &= \biggl(\sum_{j \in S \setminus \hat{S}} \hat{v}_j v_j + \sum_{j \in S \cap \hat{S}} \hat{v}_j v_j\biggr)^2 \\
&\leq \biggl(2 \sum_{j \in S \setminus \hat{S}} \hat{v}_j^2 + \sum_{j \in S \cap \hat{S}} \hat{v}_j^2\biggr)\biggl(\frac{1}{2} \sum_{j \in S \setminus \hat{S}} v_j^2 + \sum_{j \in S \cap \hat{S}} v_j^2\biggr) \\
&\leq \biggl(\sum_{j \in \hat{S} \setminus S} \hat{v}_j^2 + \sum_{j \in S \setminus \hat{S}} \hat{v}_j^2 + \sum_{j \in S \cap \hat{S}} \hat{v}_j^2\biggr)\biggl(1 - \frac{1}{2} \sum_{j \in S \setminus \hat{S}} v_j^2\biggr) \leq 1 - \frac{1}{2} \sum_{j \in S \setminus \hat{S}} v_j^2,
\end{aligned}
$$

as required. □

LEMMA 8. *Let $A \in \mathbb{R}^{d \times d}$ be a symmetric matrix. Let $A^{(r)}$ be the principal submatrix of A obtained by deleting the rth row and rth column of A. If A has a unique (up to sign) leading eigenvector v, then*

$$\lambda_2(A) \leq \lambda_1(A^{(r)}) \leq \lambda_1(A) - v_{1,r}^2(\lambda_1(A) - \lambda_2(A))$$

PROOF. The first inequality in the lemma is implied by Cauchy's Interlacing Theorem (see, e.g. Horn and Johnson (2012, Theorem 4.3.17)). It remains to show the second inequality. Let $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_d$ be eigenvalues of A (counting multiplicities), and $v_1, \ldots, v_d$ be unit-length eigenvectors of A such that $Av_i = \lambda_i v_i$ and $v_i^\top v_j = 0$ for all $i \neq j$. We have

$$
\begin{aligned}
\lambda_1(A^{(r)}) &= \max_{\substack{\|u\|_2=1 \\ u_r=0}} u^\top A u = \max_{\substack{\|u\|_2=1 \\ u_r=0}} u^\top \biggl(\sum_{i=1}^d \lambda_i v_i v_i^\top\biggr) u \\
&\leq \max_{\substack{\|u\|_2=1 \\ u_r=0}} \biggl\{(\lambda_1 - \lambda_2) u^\top v_1 v_1^\top u + \lambda_2 u^\top \biggl(\sum_{i=1}^d v_i v_i^\top\biggr) u\biggr\} \\
&\leq \max_{\substack{\|u\|_2=1 \\ u_r=0}} (\lambda_1 - \lambda_2)|u^\top v_1|^2 + \lambda_2 \\
&\leq (\lambda_1 - \lambda_2)(1 - v_{1,r}^2) + \lambda_2 \\
&= \lambda_1 - v_{1,r}^2(\lambda_1 - \lambda_2),
\end{aligned}
$$

where we used Cauchy–Schwarz inequality in the penultimate line. □

Recall the definition of the total variation distance d_{TV} given in the proof of Theorem 6 in the main document Wang, Berthet and Samworth (2015).

LEMMA 9. *Let X and Y be random elements taking values in a measurable space $(F, \mathcal{F})$, and let $(G, \mathcal{G})$ be another measurable space.*

(a) If $\phi : F \to G$ is measurable, then

$$d_{\mathrm{TV}}\big(\mathcal{L}(\phi(X)), \mathcal{L}(\phi(Y))\big) \leq d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big).$$

(b) Let Z be a random element taking values in $(G, \mathcal{G})$, and suppose that Z is independent of (X, Y). Then

$$d_{\mathrm{TV}}\big(\mathcal{L}(X, Z), \mathcal{L}(Y, Z)\big) = d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big).$$

PROOF. (a) For any $A \in \mathcal{G}$, we have

$$\begin{aligned} |\mathbb{P}\{\phi(X) \in A\} - \mathbb{P}\{\phi(Y) \in A\}| &= |\mathbb{P}\{X \in \phi^{-1}(A)\} - \mathbb{P}\{Y \in \phi^{-1}(A)\}| \\ &\leq d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big). \end{aligned}$$

Since $A \in \mathcal{G}$ was arbitrary, the result follows.

(b) Define $\phi : F \times G \to F$ by $\phi(w, z) := w$. Then ϕ is measurable, and using the result of part (a),

$$\begin{aligned} d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big) &= d_{\mathrm{TV}}\big(\mathcal{L}(\phi(X, Z)), \mathcal{L}(\phi(Y, Z))\big) \\ &\leq d_{\mathrm{TV}}\big(\mathcal{L}(X, Z), \mathcal{L}(Y, Z)\big). \end{aligned}$$

For the other inequality, let $\mathcal{A}$ denote the set of subsets A of $\mathcal{F} \otimes \mathcal{G}$ with the property that given $\epsilon > 0$, there exist sets $B_{1,F}, \ldots, B_{n,F} \in \mathcal{F}$ and disjoint sets $B_{1,G}, \ldots, B_{n,G} \in \mathcal{G}$ such that, writing $B := \cup_{i=1}^n (B_{i,F} \times B_{i,G})$, we have $\mathbb{P}\big((X, Z) \in A \triangle B\big) < \epsilon$ and $\mathbb{P}\big((Y, Z) \in A \triangle B\big) < \epsilon$. Here, the binary operator $\triangle$ denotes the symmetric difference of two sets, so that $A \triangle B := (A \cap B^c) \cup (A^c \cap B)$. Note that $\mathcal{F} \times \mathcal{G} \subseteq \mathcal{A}$. Now suppose $A \in \mathcal{A}$ so that, given $\epsilon > 0$, we can find sets $B_{1,F}, \ldots, B_{n,F} \in \mathcal{F}$ and disjoint sets $B_{1,G}, \ldots, B_{n,G} \in \mathcal{G}$ with the properties above. Observe that we can write

$$B^c = \bigcup_{I \subseteq \{1, \ldots, n\}} \biggl(\bigcap_{i \in I} B_{i,F}^c \times \bigcap_{i \in I} B_{i,G} \cap \bigcap_{i \in I^c} B_{i,G}^c \biggr).$$

For each $I \subseteq \{1, \ldots, n\}$, the sets $\cap_{i \in I} B_{i,F}^c$ belong to $\mathcal{F}$, and $\big\{\cap_{i \in I} B_{i,G} \cap \cap_{i \in I^c} B_{i,G}^c : I \subseteq \{1, \ldots, n\}\big\}$ is a family of disjoint sets in $\mathcal{G}$. Moreover,

$$\mathbb{P}\big((X, Z) \in A^c \triangle B^c\big) = \mathbb{P}\big((X, Z) \in A \triangle B\big) < \epsilon,$$

and similarly $\mathbb{P}\big((Y, Z) \in A^c \triangle B^c\big) < \epsilon$. We deduce that $A^c \in \mathcal{A}$. Finally, if (A_n) is a disjoint sequence in $\mathcal{A}$, then let $A := \cup_{n=1}^\infty A_n$, and given $\epsilon > 0$, find

$m \in \mathbb{N}$ such that $\mathbb{P}\big((X,Z) \in A \backslash \cup_{i=1}^m A_i\big) < \epsilon/2$ and $\mathbb{P}\big((Y,Z) \in A \backslash \cup_{i=1}^m A_i\big) < \epsilon/2$. Now, for each $i = 1,\ldots,m$, find sets $B_{i1,F},\ldots,B_{in_i,F} \in \mathcal{F}$ and disjoint sets $B_{i1,G},\ldots,B_{in_i,G} \in \mathcal{G}$ such that, writing $B_i := \cup_{j=1}^{n_i}(B_{ij,F} \times B_{ij,G})$, we have $\mathbb{P}\big((X,Z) \in A_i \triangle B_i\big) < \epsilon/(2m)$ and $\mathbb{P}\big((Y,Z) \in A_i \triangle B_i\big) < \epsilon/(2m)$. It is convenient to relabel the sets $\{(B_{ij,F}, B_{ij,G}) : i = 1,\ldots,m, j = 1,\ldots,n_i\}$ as $\{(C_{1,F}, C_{1,G}),\ldots,(C_{N,F}, C_{N,G})\}$, where $N := \sum_{i=1}^m n_i$. This means that we can write

$$\bigcup_{i=1}^m B_i = \bigcup_{k=1}^N (C_{k,F} \times C_{k,G}) = \bigcup_{K \subseteq \{1,\ldots,N\}, K \neq \emptyset} \biggl(\bigcup_{k\in K} C_{k,F} \times \bigcap_{k \in K} C_{k,G} \cap \bigcap_{k \in K^c} C_{k,G}^c\biggr).$$

Now, for each non-empty subset K of $\{1,\ldots,N\}$, the set $\cup_{k\in K} C_{k,F}$ belongs to $\mathcal{F}$, and $\bigl\{\cap_{k\in K} C_{k,G} \cap \cap_{k\in K^c} C_{k,G}^c : K \subseteq \{1,\ldots,N\}, K \neq \emptyset\bigr\}$ is a family of disjoint sets in $\mathcal{G}$. Moreover,

$$\mathbb{P}\big((X,Z) \in A \triangle \cup_{i=1}^m B_i\big) \leq \sum_{i=1}^m \mathbb{P}\big((X,Z) \in A_i \triangle B_i\big) + \frac{\epsilon}{2} < \epsilon,$$

and similarly, $\mathbb{P}\big((Y,Z) \in A \triangle \cup_{i=1}^m B_i\big) < \epsilon$. We deduce that $A \in \mathcal{A}$, so $\mathcal{A}$ is a σ-algebra containing $\mathcal{F} \times \mathcal{G}$, so $\mathcal{A}$ contains $\mathcal{F} \otimes \mathcal{G}$.

Now suppose that $A \in \mathcal{F} \otimes \mathcal{G}$. By the argument above, given $\epsilon > 0$, there exist sets $B_{1,F},\ldots,B_{n,F} \in \mathcal{F}$ and disjoint sets $B_{1,G},\ldots,B_{n,G} \in \mathcal{G}$ such that $\mathbb{P}\big((X,Z) \in A \triangle \cup_{i=1}^m (B_{i,F} \times B_{i,G})\big) < \epsilon/2$ and $\mathbb{P}\big((Y,Z) \in A \triangle \cup_{i=1}^m (B_{i,F} \times B_{i,G})\big) < \epsilon/2$. It follows that

$$\begin{aligned}
&\big|\mathbb{P}\big((X,Z) \in A\big) - \mathbb{P}\big((Y,Z) \in A\big)\big| \\
&\leq \sum_{i=1}^m \big|\mathbb{P}\big(X \in B_{i,F}, Z \in B_{i,G}\big) - \mathbb{P}\big(Y \in B_{i,F}, Z \in B_{i,G}\big)\big| + \epsilon \\
&= \sum_{i=1}^m \mathbb{P}(Z \in B_{i,G})\big|\mathbb{P}(X \in B_{i,F}) - \mathbb{P}(Y \in B_{i,F})\big| + \epsilon \leq d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big) + \epsilon.
\end{aligned}$$

Since $A \in \mathcal{A}$ and $\epsilon > 0$ were arbitrary, we conclude that

$$d_{\mathrm{TV}}\big(\mathcal{L}(X,Z), \mathcal{L}(Y,Z)\big) \leq d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big),$$

as required. □

2. A brief introduction to computational complexity theory.

The following is intended to give a short introduction to notions in computational complexity theory referred to in Wang, Berthet and Samworth

(2015). A good reference for further information is Arora and Barak (2009), from which much of the following is inspired.

A *computational problem* is the task of generating a desired output based on a given input. Formally, defining $\{0,1\}^* := \cup_{k=1}^{\infty}\{0,1\}^k$ to be the set of all finite strings of zeros and ones, we can view a computational problem as a function $F : \{0,1\}^* \to \mathcal{P}\bigl(\{0,1\}^*\bigr)$, where $\mathcal{P}(A)$ denotes the power set of a set A. The interpretation is that $F(s)$ describes the set of acceptable output strings (solutions) for a particular input string s.

Loosely speaking, an *algorithm* is a collection of instructions for performing a task. Despite the widespread use of algorithms in mathematics throughout history, it was not until 1936 that Alonzo Church and Alan Turing formalised the notion by defining notational systems called the λ-calculus and Turing machines respectively (Church, 1936; Turing, 1936). Here we define an algorithm to be a *Turing machine*:

DEFINITION 1. *A* Turing machine M *is a pair* (Q, δ), *where*

- Q *is a finite set of states, among which are two distinguished states* q_{start} *and* q_{halt}.
- δ *is a 'transition' function from* $Q \times \{0,1,\text{␣}\}$ *to* $Q \times \{0,1,\text{␣}\} \times \{\mathrm{L}, \mathrm{R}\}$.

A Turing Machine can be thought of as having a reading head that can access a tape consisting of a countably infinite number of squares, labelled $0, 1, 2, \ldots$. When the Turing machine is given an input $s \in \{0,1\}^*$, the tape is initialised with the components of s in its first $|s|$ tape squares (where $|\cdot|$ denotes the length of a string in $\{0,1\}^*$) and with 'blank symbols' ␣ in its remaining squares. The Turing machine starts in the state $q_{\mathrm{start}} \in Q$ with its head on the 0th square and operates according to its transition function δ. When the machine is in state $q \in Q$ with its head over the ith tape square that contains the symbol $a \in \{0,1,\text{␣}\}$, and if $\delta(q,a) = (q', a', \mathrm{L})$, the machine overwrites a with a', updates its state to q', and moves to square $i-1$ (or to square $i+1$ if the third component of the transition function is R instead of L). The Turing machine stops if it reaches state $q_{\mathrm{halt}} \in Q$ and outputs the vector of symbols on the tape before the first blank symbol. If the Turing machine M terminates (in finitely many steps) with input s, we write $M(s)$ for its output.

We say an algorithm (Turing machine) M *solves a computational problem* F if M terminates for every input $s \in \{0,1\}^*$, and $M(s) \in F(s)$. A computational problem is *solvable* if there exists a Turing machine that solves it. It turns out that other notions of an algorithm (including Church's λ-calculus and modern computer programming languages) are equivalent in the sense that the set of solvable problems is the same.

A *polynomial time algorithm* is a Turing machine M for which there exist $a, b > 0$ such that for all input strings $s \in \{0,1\}^*$, M terminates after at most $a|s|^b$ transitions. We say a problem F is *polynomial time solvable*, written $F \in \mathsf{P}$, if there exists a polynomial time algorithm that solves it[1].

A *nondeterministic Turing machine* has the same definition as that for a Turing machine except that the transition function δ becomes a set-valued function $\delta : Q \times \{0, 1, \llcorner\!\lrcorner\} \to \mathcal{P}\big(Q \times \{0, 1, \llcorner\!\lrcorner\} \times \{\mathrm{L}, \mathrm{R}\}\big)$. The idea is that, while in state q with its head over symbol a, a nondeterministic Turing machine replicates $|\delta(q, a)|$ copies of itself (and its tape) in the current configuration, each exploring a different possible future configuration in the set $\delta(q, a)$. Each replicate branches to further replicates in the next step. The process continues until one of its replicates reaches the state q_{halt}. At that point, the Turing machine replicate that has halted outputs its tape content and all replicates stop computation. A *nondeterministic polynomial time algorithm* is a nondeterministic Turing machine M_{nd} for which there exist $a, b > 0$ such that for all input strings $s \in \{0,1\}^*$, M_{nd} terminates after at most $a|s|^b$ steps. (We count all replicates of M_{nd} making one parallel transition as one step.) We say a computational problem F is *nondeterministically polynomial time solvable*, written $F \in \mathsf{NP}$, if there exists a nondeterministic polynomial time algorithm that solves it[2].

Clearly $\mathsf{P} \subseteq \mathsf{NP}$, but it is not currently known if these classes are equal. It is widely believed that $\mathsf{P} \neq \mathsf{NP}$, and many computational lower bounds for particular computational problems have been proved conditional under this assumption. Working under this hypothesis, a common strategy is to relate the algorithmic complexity of one computational problem to another. We say a computational problem F is *polynomial time reducible* to another problem G, written as $F \le_{\mathsf{P}} G$, if there exist polynomial time algorithms M_{in} and M_{out} such that $M_{\mathrm{out}} \circ G \circ M_{\mathrm{in}}(s) \subseteq F(s)$. In other words, $F \le_{\mathsf{P}} G$ if we can convert an input of F to an input of G through M_{in}, and translate every solution of G back to a solution for F through M_{out}.

DEFINITION 2. *A computational problem G is NP-*hard *if $F \le_{\mathsf{P}} G$ for all $F \in \mathsf{NP}$. It is NP-*complete *if it is in NP and is NP-hard.*

[1] In fact, some authors write FP (short for 'Functional Polynomial Time') for the class we have denoted as P here. The notation P is then reserved for the subset of computational problems consisting of so-called *decision problems* F, where $F(s) \in \big\{\{0\}, \{1\}\big\}$ for all $s \in \{0,1\}^*$.

[2] Again, some authors write FNP for the class we have denoted as NP here.

Karp (1972) showed that a large number of natural computational problems are NP-complete, including the Clique problem mentioned in Section 4. The Turing machines and nondeterministic Turing machines introduced above are both non-random. In some situations (e.g. statistical problems), it is useful to consider random procedures:

DEFINITION 3. *A* probabilistic Turing machine M_{pr} *is a triple* (Q, δ, X)*, where*

- Q *is a finite set of states, among which are two distinguished states* q_{start} *and* q_{halt}*.*
- δ *is a transition function from* $Q \times \{0, 1, \llcorner\!\lrcorner\} \times \{0, 1\}$ *to* $Q \times \{0, 1, \llcorner\!\lrcorner\} \times \{\mathrm{L}, \mathrm{R}\}$*.*
- $X = (X_1, X_2, \ldots)$ *is an infinite sequence of independent Bern*$(1/2)$ *random variables.*

In its tth step, if a probabilistic Turing machine M_{pr} is in state q with its reading head over symbol a, and $\delta(q, a, X_t) = (q', a', L)$, then M_{pr} overwrites a with a', updates its state to q' and moves its reading head to the left (or to the right if $\delta(q, a, X_t) = (q', a', R)$). A *randomised polynomial time algorithm* is a probabilistic Turing machine M_{pr} for which there exist $a, b > 0$ such that for any $s \in \{0, 1\}^*$, M_{pr} terminates in at most $a|s|^b$ steps. We say a computational problem F is *solvable in randomised polynomial time*, written as $F \in \mathsf{BPP}$, if, given $\epsilon > 0$, there exists a randomised polynomial time algorithm $M_{\mathrm{pr},\epsilon}$ such that $\mathbb{P}\bigl(M_{\mathrm{pr},\epsilon}(s) \in F(s)\bigr) \geq 1 - \epsilon$.

In the above discussion, the classes P, NP, BPP are all defined through worst-case performance of an algorithm, since we require the time bound to hold for every input string s. However, in many statistical applications, the input string s is drawn from some distribution $\mathcal{D}$ on $\{0, 1\}^*$, and it is the average performance of the algorithm, rather than the worst case scenario, that is of more interest. We say such a random problem is solvable in randomised polynomial time if, given $\epsilon > 0$, there exists a randomised polynomial time algorithm $M_{\mathrm{pr},\epsilon}$ such that, when $s \sim \mathcal{D}$, independent of X, we have $\mathbb{P}\bigl(M_{\mathrm{pr}}(s) \in F(s)\bigr) \geq 1 - \epsilon$. Note that the probability here is taken over both the randomness in s and the randomness in X. Similar to the non-random cases, we can talk about randomised polynomial time reduction. If M_F is a randomised polynomial time algorithm for a computational problem F, then $M_{\mathrm{out}} \circ M_F \circ M_{\mathrm{in}}$ is a potential randomised polynomial time algorithm for another problem G for suitably constructed randomised polynomial time algorithms M_{in} and M_{out}. One such construction is the key to the proof of Theorem 6 in the main document Wang, Berthet and Samworth (2015).

References.

Arora, S. and Barak, B. (2009) *Computational Complexity: A Modern Approach.* Cambridge University Press, Cambridge.

Church, A. (1936) An unsolvable problem of elementary number theory. *Amer. J. Math.*, **58**, 345–363.

Horn, R. A. and Johnson, C. R. (2012) *Matrix Analysis*. Cambridge University Press.

Karp, R. M. (1972) Reducibility among combinatorial problems. In R. E. Miller et al. (Eds.), *Complexity of Computer Computations*, 85–103. Springer, New York.

Kim, A. K.-H. and Samworth R. J. (2014) Global rates of convergence in log-concave density estimation. Available at `http://arxiv.org/abs/1404.2298`.

Massart, P. (2007) *Concentration Inequalities and Model Selection: Ecole d'Eté de Probabilités de Saint-Flour XXXIII - 2003*. Springer, Berlin/Heidelberg.

Shorack, G. R. and Wellner, J. A. (1986) *Empirical Processes with Applications to Statistics*. Wiley, New York.

Turing, A. (1936) On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc.*, **2**, 230–265.

Wang, T., Berthet, Q. and Samworth, R. J. (2015) Statistical and computational trade-offs in estimation of sparse principal components. *Submitted*.

Yu, B. (1997) Assouad, Fano and Le Cam. In Pollard, D., Torgersen, E. and Yang G. L. (Eds.) *Festschrift for Lucien Le Cam: Research Papers in Probability and Statistics*, 423–435. Springer, New York.

Statistical Laboratory
Wilberforce Road
Cambridge, CB3 0WB
United Kingdom
E-mail: r.samworth@statslab.cam.ac.uk
E-mail: t.wang@statslab.cam.ac.uk
E-mail: q.berthet@statslab.cam.ac.uk
URL: http://www.statslab.cam.ac.uk/~rjs57
URL: http://www.statslab.cam.ac.uk/~tw389
URL: http://www.statslab.cam.ac.uk/~qb204