Supplementary Appendix

Venkat Chandrasekaran and Michael I. Jordan

Proof of Proposition 5

As with the proof of Proposition 4, we condition on $\mathbf{z} = \tilde{\mathbf{z}}$. Setting $\boldsymbol{\delta} = \mathbf{x} - \tilde{\mathbf{x}}$ and setting $\hat{\boldsymbol{\delta}}_n(\mathcal{C}) = \hat{\mathbf{x}}_n(\mathcal{C})|_{\mathbf{z}=\tilde{\mathbf{z}}} - \tilde{\mathbf{x}}$, we can rewrite the estimation problem [**2**] from the main paper as follows:

$$\hat{\boldsymbol{\delta}}_n(\mathcal{C}) = \arg\min_{\boldsymbol{\delta}\in\mathbb{R}^p} \quad \frac{1}{2}\left\|(\mathbf{x}^* - \tilde{\mathbf{x}}) + \tfrac{\sigma}{\sqrt{n}}\tilde{\mathbf{z}} - \boldsymbol{\delta}\right\|_{\ell_2}^2 \quad \text{s.t.} \quad \boldsymbol{\delta} \in \mathcal{C} - \tilde{\mathbf{x}}.$$

Letting R_1 and R_2 denote orthogonal subspaces that contain Q_1 and Q_2, i.e., $Q_1 \subseteq R_1$ and $Q_2 \subseteq R_2$, and letting $\boldsymbol{\delta}^{(1)} = \mathcal{P}_{R_1}(\boldsymbol{\delta}), \boldsymbol{\delta}^{(2)} = \mathcal{P}_{R_2}(\boldsymbol{\delta}), \hat{\boldsymbol{\delta}}_n^{(1)}(\mathcal{C}) = \mathcal{P}_{R_1}(\hat{\boldsymbol{\delta}}_n(\mathcal{C})), \hat{\boldsymbol{\delta}}_n^{(2)}(\mathcal{C}) = \mathcal{P}_{R_2}(\hat{\boldsymbol{\delta}}_n(\mathcal{C}))$ denote the projections of $\boldsymbol{\delta}, \hat{\boldsymbol{\delta}}_n(\mathcal{C})$ onto R_1, R_2, we can rewrite the above reformulated optimization problem as:

$$\begin{aligned}\left[\hat{\boldsymbol{\delta}}_n^{(1)}(\mathcal{C}), \hat{\boldsymbol{\delta}}_n^{(2)}(\mathcal{C})\right] = \arg\min_{\boldsymbol{\delta}^{(1)}\in Q_1, \boldsymbol{\delta}^{(2)}\in Q_2} \quad & \frac{1}{2}\left\|\mathcal{P}_{R_1}\left[(\mathbf{x}^* - \tilde{\mathbf{x}}) + \tfrac{\sigma}{\sqrt{n}}\tilde{\mathbf{z}}\right] - \boldsymbol{\delta}^{(1)}\right\|_{\ell_2}^2 \\ & + \frac{1}{2}\left\|\mathcal{P}_{R_2}\left[(\mathbf{x}^* - \tilde{\mathbf{x}}) + \tfrac{\sigma}{\sqrt{n}}\tilde{\mathbf{z}}\right] - \boldsymbol{\delta}^{(2)}\right\|_{\ell_2}^2.\end{aligned}$$

As the sets Q_1, Q_2 live in orthogonal subspaces, the two variables $\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}$ in this problem can be optimized separately. Consequently, we have that $\|\hat{\boldsymbol{\delta}}_n^{(2)}(\mathcal{C})\|_{\ell_2} \leq \alpha$ and that

$$\|\hat{\boldsymbol{\delta}}_n^{(1)}(\mathcal{C})\|_{\ell_2} \leq \sup_{\bar{\boldsymbol{\delta}}\in\mathrm{cone}(Q_1)\cap B_{\ell_2}^p} \langle \bar{\boldsymbol{\delta}}, \tfrac{\sigma}{\sqrt{n}}\tilde{\mathbf{z}} + (\mathbf{x}^* - \tilde{\mathbf{x}})\rangle.$$

This bound can be established following the same sequence of steps as in the proof of Proposition 4. Combining the two bounds on $\hat{\boldsymbol{\delta}}_n^{(1)}(\mathcal{C})$ and $\hat{\boldsymbol{\delta}}_n^{(2)}(\mathcal{C})$, one can then check that

$$\|\hat{\boldsymbol{\delta}}_n^{(1)}(\mathcal{C})\|_{\ell_2}^2 + \|\hat{\boldsymbol{\delta}}_n^{(2)}(\mathcal{C})\|_{\ell_2}^2 \leq 2\left[\tfrac{\sigma^2}{n} g(\mathrm{cone}(Q_1) \cap B_{\ell_2}^p) + \|\mathbf{x}^* - \tilde{\mathbf{x}}\|_{\ell_2}^2\right] \ + \ \alpha^2.$$

To obtain a bound on $\|\hat{\mathbf{x}}_n(\mathcal{C})|_{\mathbf{z}=\tilde{\mathbf{z}}} - \mathbf{x}^*\|_{\ell_2}^2$ we note that

$$\begin{aligned}\|\hat{\mathbf{x}}_n(\mathcal{C})|_{\mathbf{z}=\tilde{\mathbf{z}}} - \mathbf{x}^*\|_{\ell_2}^2 &\leq 2\left[\|\hat{\mathbf{x}}_n(\mathcal{C})|_{\mathbf{z}=\tilde{\mathbf{z}}} - \tilde{\mathbf{x}}\|_{\ell_2}^2 + \|\mathbf{x}^* - \tilde{\mathbf{x}}\|_{\ell_2}^2\right] \\ &\leq 2\|\hat{\boldsymbol{\delta}}_n^{(1)}(\mathcal{C})\|_{\ell_2}^2 + 2\|\hat{\boldsymbol{\delta}}_n^{(2)}(\mathcal{C})\|_{\ell_2}^2 + 2\|\mathbf{x}^* - \tilde{\mathbf{x}}\|_{\ell_2}^2.\end{aligned}$$

Taking expectations concludes the proof. □

Proof of Proposition 9

The main steps of this proof follow the steps of a similar result in [1], with the principal difference being that we wish to bound Gaussian squared-complexity rather than Gaussian complexity. A central theme in this proof is the appeal to Gaussian isoperimetry. Let $\mathbb{S}^{p-1}$ denote the sphere in p dimensions. Then in bounding the expected squared-distance to the dual cone $\mathcal{K}^*$ with $\mathcal{K}^* \cap \mathbb{S}^{p-1}$ having a volume of μ, we need only consider the extremal case of a spherical cap in $\mathbb{S}^{p-1}$ having a volume of μ. The manner in which this is made precise will become clear in the proof. Before proceeding with the main proof, we state and derive a result on the solid angle subtended by a spherical cap in $\mathbb{S}^{p-1}$ to which we will need to appeal repeatedly:

Lemma 2 *Let $\psi(\mu)$ denote the solid angle subtended by a spherical cap in $\mathbb{S}^{p-1}$ with volume $\mu \in \left(\frac{1}{4}\exp\{-\frac{p}{20}\}, \frac{1}{4e^2}\right)$. Then*

$$\psi(\mu) \geq \frac{\pi}{2}\left(1 - \sqrt{\frac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}\right).$$

Proof of Lemma 2: Consider the following definition of a spherical cap, parametrized by height h:

$$J = \{\mathbf{a} \in \mathbb{S}^{p-1} \mid \mathbf{a}_1 \geq h\}.$$

Here $\mathbf{a}_1$ denotes the first coordinate of $\mathbf{a} \in \mathbb{R}^p$. Given a spherical cap of height $h \in [0,1]$, the solid angle ψ is given by:

$$\psi = \frac{\pi}{2} - \sin^{-1}(h). \tag{10}$$

We can thus obtain bounds on the solid angle of a spherical cap via bounds on its height. The following result from [2] relates the volume of a spherical cap to its height:

Lemma 3 *[2] For $\frac{2}{\sqrt{p}} \leq h \leq 1$ the volume $\tilde{\mu}(p,h)$ of a spherical cap of height h in $\mathbb{S}^{p-1}$ is bounded as*

$$\frac{1}{10h\sqrt{p}}(1-h^2)^{\frac{p-1}{2}} \leq \tilde{\mu}(p,h) \leq \frac{1}{2h\sqrt{p}}(1-h^2)^{\frac{p-1}{2}}.$$

Continuing with the proof of Lemma 2, note that for $\frac{2}{\sqrt{p}} \leq h \leq 1$

$$\frac{1}{2h\sqrt{p}}(1-h^2)^{\frac{p-1}{2}} \leq \frac{1}{4}(1-h^2)^{\frac{p-1}{2}} \leq \frac{1}{4}\exp\left(-\tfrac{p-1}{2}h^2\right).$$

Choosing $h = \sqrt{\frac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}$ we have $\frac{2}{\sqrt{p}} \leq h \leq 1$ based on the assumption $\mu \in \left(\frac{1}{4}\exp\{-p/20\}, \frac{1}{4e^2}\right)$. Consequently, we can apply Lemma 3 with this value of h combined with (10) to conclude that

$$\tilde{\mu}\left(p, \sqrt{\frac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}\right) \leq \mu.$$

Hence the solid angle $\psi\left(\tilde{\mu}\left(p, \sqrt{\frac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}\right)\right)$ is less than the solid angle $\psi(\mu)$. Consequently, we use (10) to conclude that

$$\psi(\mu) \geq \frac{\pi}{2} - \sin^{-1}\left(\sqrt{\frac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}\right).$$

Using the bound $\sin^{-1}(h) \leq \frac{\pi}{2}h$, we obtain the desired bound. □

Proof of Proposition 9: We bound the Gaussian squared-complexity of $\mathcal{K}$ by bounding the expected squared-distance to the polar cone $\mathcal{K}^*$. Let $\bar{\mu}(U;t)$ for $U \subseteq \mathbb{S}^{p-1}$ and $t > 0$ denote the volume of the set of points in $\mathbb{S}^{p-1}$ that are within a Euclidean distance of at most t from U (recall that the volume of this set is equivalent to the measure of the set with respect to the normalized Haar measure on $\mathbb{S}^{p-1}$). We have the following sequence of relations by appealing to the independence of the direction $\mathbf{g}/\|\mathbf{g}\|_{\ell_2}$ and of the length $\|\mathbf{g}\|_{\ell_2}$ of a standard normal vector $\mathbf{g}$:

$$\begin{aligned}
\mathbb{E}[\mathrm{dist}(\mathbf{g},\mathcal{K}^*)^2] &= \mathbb{E}[\|\mathbf{g}\|_{\ell_2}^2 \mathrm{dist}(\mathbf{g}/\|\mathbf{g}\|_{\ell_2},\mathcal{K}^*)^2] \\
&= p\ \mathbb{E}[\mathrm{dist}(\mathbf{g}/\|\mathbf{g}\|_{\ell_2},\mathcal{K}^*)^2] \\
&\leq p\ \mathbb{E}[\mathrm{dist}(\mathbf{g}/\|\mathbf{g}\|_{\ell_2},\mathcal{K}^*\cap\mathbb{S}^{p-1})^2] \\
&= p\int_0^\infty \mathbb{P}[\mathrm{dist}(\mathbf{g}/\|\mathbf{g}\|_{\ell_2},\mathcal{K}^*\cap\mathbb{S}^{p-1})^2 > t]dt \\
&= p\int_0^\infty \mathbb{P}[\mathrm{dist}(\mathbf{g}/\|\mathbf{g}\|_{\ell_2},\mathcal{K}^*\cap\mathbb{S}^{p-1}) > \sqrt{t}]dt \\
&= 2p\int_0^\infty s\mathbb{P}[\mathrm{dist}(\mathbf{g}/\|\mathbf{g}\|_{\ell_2},\mathcal{K}^*\cap\mathbb{S}^{p-1}) > s]ds \\
&= 2p\int_0^\infty s[1-\bar{\mu}(\mathcal{K}^*\cap\mathbb{S}^{p-1};s)]ds.
\end{aligned}$$

Here the third equality follows based on the integral version of the expected value. Let $V \subseteq \mathbb{S}^{p-1}$ denote a spherical cap with the same volume μ as $\mathcal{K}^*\cap\mathbb{S}^{p-1}$. Then we have by spherical isoperimetry that $\bar{\mu}(V;s) \geq \bar{\mu}(\mathcal{K}^*\cap\mathbb{S}^{p-1};s)$ for all $s \geq 0$ [3]. Thus

$$\mathbb{E}[\mathrm{dist}(\mathbf{g},\mathcal{K}^*)^2] \leq 2p\int_0^\infty s[1-\bar{\mu}(V;s)]ds. \tag{11}$$

From here onward, we focus exclusively on bounding the integral.

Let $\tau(\psi)$ denote the volume of a spherical cap subtending a solid angle of ψ radians. Recall that ψ is a quantity between 0 and π. As in Lemma 2 let $\psi(\mu)$ denote the solid angle of a spherical cone subtending a solid angle of μ. Since the Euclidean distance between points on a sphere is always smaller than the geodesic distance, we have that $\bar{\mu}(V;s) \geq \tau(\psi(\mu)+s)$. Further, we have the following explicit formula for $\tau(\psi)$ [4]:

$$\tau(\psi) = \omega_p^{-1}\int_0^\psi \sin^{p-1}(v)dv,$$

where $\omega_p = \int_0^\pi \sin^{p-1}(v)dv$ is the normalization constant. Combining these latter two observations, we can bound the integral in (11) as:

$$\begin{aligned}
\int_0^\infty s[1-\bar{\mu}(V;s)]ds &\leq \int_0^\infty s[1-\tau(\psi(\mu)+s)]ds \\
&= \int_0^{\pi-\psi(\mu)} s[1-\tau(\psi(\mu)+s)]ds \\
&= \frac{(\pi-\psi(\mu))^2}{2} - \int_0^{\pi-\psi(\mu)} s\tau(\psi(\mu)+s)ds \\
&= \frac{(\pi-\psi(\mu))^2}{2} - \omega_p^{-1}\int_0^{\pi-\psi(\mu)}\int_0^{\psi(\mu)+s} s\sin^{p-1}(v)dvds
\end{aligned}$$

Next we change the order of integration to obtain:

$$\begin{aligned}
\int_0^\infty s[1-\bar{\mu}(V;s)]ds &\leq \frac{(\pi-\psi(\mu))^2}{2} - \omega_p^{-1}\int_0^\pi \int_{\max\{v-\psi(\mu),0\}}^{\pi-\psi(\mu)} \sin^{p-1}(v) s ds dv \\
&= \frac{(\pi-\psi(\mu))^2}{2} - \omega_p^{-1}\int_0^\pi \frac{1}{2}\left[(\pi-\psi(\mu))^2 - (\max\{v-\psi(\mu),0\})^2\right]\sin^{p-1}(v)dv \\
&= \frac{\omega_p^{-1}}{2}\int_0^\pi (\max\{v-\psi(\mu),0\})^2 \sin^{p-1}(v)dv \\
&= \frac{\omega_p^{-1}}{2}\int_{\psi(\mu)}^\pi (v-\psi(\mu))^2 \sin^{p-1}(v)dv.
\end{aligned}$$

We now appeal to the inequalities $\omega_p^{-1} \leq \sqrt{p-1}/2$ and $\sin(x) \leq \exp(-(x-\frac{\pi}{2})^2/2)$ for $x \in [0,\pi]$ to obtain

$$\int_0^\infty s[1-\bar{\mu}(V;s)]ds \leq \frac{\sqrt{p-1}}{2}\int_{\psi(\mu)}^\pi (v-\psi(\mu))^2 \exp\left[-\tfrac{p-1}{2}(v-\tfrac{\pi}{2})^2\right]dv.$$

Performing a change of variables with $a = \sqrt{p-1}(v-\frac{\pi}{2})$, we have

$$\begin{aligned}
\int_0^\infty s[1-\bar{\mu}(V;s)]ds &\leq \frac{1}{2}\int_{\sqrt{p-1}(\psi(\mu)-\pi/2)}^{\sqrt{p-1}\pi/2} (\tfrac{a}{\sqrt{p-1}} + (\tfrac{\pi}{2}-\psi(\mu)))^2 \exp[-\tfrac{a^2}{2}]da \\
= \ & \frac{1}{2}\int_{\sqrt{p-1}(\psi(\mu)-\pi/2)}^{\sqrt{p-1}\pi/2} \left[\tfrac{a^2}{p-1} + (\tfrac{\pi}{2}-\psi(\mu))^2 + \tfrac{2a}{\sqrt{p-1}}(\tfrac{\pi}{2}-\psi(\mu))\right]\exp[-\tfrac{a^2}{2}]da \\
\leq \ & \frac{1}{2}\left[\int_{-\infty}^\infty \tfrac{a^2}{p-1}\exp[-\tfrac{a^2}{2}]da + \int_{-\infty}^\infty (\tfrac{\pi}{2}-\psi(\mu))^2\exp[-\tfrac{a^2}{2}]da + \int_0^\infty \tfrac{2a}{\sqrt{p-1}}(\tfrac{\pi}{2}-\psi(\mu))\exp[-\tfrac{a^2}{2}]da\right] \\
= \ & \frac{1}{2}\left[\tfrac{\sqrt{2\pi}}{p-1} + \sqrt{2\pi}(\tfrac{\pi}{2}-\psi(\mu))^2 + \tfrac{2}{\sqrt{p-1}}(\tfrac{\pi}{2}-\psi(\mu))\cdot(-\exp[-\tfrac{a^2}{2}])|_0^\infty\right] \\
= \ & \frac{1}{2}\left[\tfrac{\sqrt{2\pi}}{p-1} + \sqrt{2\pi}(\tfrac{\pi}{2}-\psi(\mu))^2 + \tfrac{2}{\sqrt{p-1}}(\tfrac{\pi}{2}-\psi(\mu))\right]
\end{aligned}$$

Here the inequality was obtained by suitably changing the limits of integration. We now employ Lemma 2 to obtain the final bound:

$$\begin{aligned}
g(\mathcal{K}\cap B_{\ell_2}^p) &\leq p\left[\tfrac{\sqrt{2\pi}}{p-1} + \sqrt{2\pi}\left(\tfrac{\pi}{2}\sqrt{\tfrac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}\right)^2 + \tfrac{2}{\sqrt{p-1}}\left(\tfrac{\pi}{2}\sqrt{\tfrac{2\log\left(\frac{1}{4\mu}\right)}{p-1}}\right)\right] \\
&= \tfrac{p\sqrt{2\pi}}{p-1}\left[1 + \pi\log\left(\tfrac{1}{4\mu}\right) + \sqrt{\pi}\sqrt{\log\left(\tfrac{1}{4\mu}\right)}\right] \\
&\leq 20\log\left(\tfrac{1}{4\mu}\right).
\end{aligned}$$

Here the final bound holds because $\mu < 1/4e^2$ and $p \geq 12$. □

References

[1] Chandrasekaran V, Recht B, Parrilo P, Willsky A (2012) The convex geometry of linear inverse problems. *Foundations of Computational Mathematics* 12:805–849.

[2] Brieden A, et al. (1998) *Approximation of diameters: randomization doesn't help In Proceedings of the 39th Annual Symposium on Foundations of Computer Science* pp 244–251.

[3] Ledoux M (2000) *The Concentration of Measure Phenomenon* (American Mathematical Society).

[4] Klain D, Rota G (1997) *Introduction to Geometric Probability* (Cambridge University Press).