Submitted to the Annals of Statistics
arXiv: math.PR/0000000

SUPPLEMENT TO 'NEEDLES AND STRAWS IN A HAYSTACK: POSTERIOR CONCENTRATION FOR POSSIBLY SPARSE SEQUENCES'

By Ismaël Castillo and Aad van der Vaart

This Supplement complements the paper [1]. It contains the proofs of Theorems 2.5, 2.7 and 2.9 as well as the proofs of some technical lemmas used in [1]. It also contains proofs of the results on the posterior coordinate-wise median.

1. Proof of technical lemmas in [1].

Proof of Lemma 4.2. The numbers $b_p := \binom{n-p_n}{p}/\binom{n}{p+k}$ satisfy

$$\frac{b_p}{b_{p-1}} = \frac{n-p_n-p+1}{n-p-k+1}\frac{p+k}{p} = \Big(1-\frac{p_n-k}{n-p-k+1}\Big)\Big(1+\frac{k}{p}\Big).$$

Hence $b_p \le b_{p-1}(1+C_1^{-1})$ for $p \ge C_1 p_n$ and any $k \le p_n$, whence the numbers $a_p = \pi_n(p+k)b_p$ satisfy

$$a_p \le a_{p-1}D(1+C_1^{-1}) \le a_{C_1p_n}F^{p-C_1p_n},$$

for $p \ge C_1p_n$ and $F = D(1+C_1^{-1})$, provided C_1 is larger than the constant C in the assumptions. Because $D < 1$, there exists suitable C_1 such that $F < 1$. Then

$$\frac{\sum_{p=C_1p_n}^{n-p_n} pa_p}{\sum_{p=0}^{n-p_n} a_p} \le C_1p_n + \frac{\sum_{p=C_1p_n}^{n-p_n}(p-C_1p_n)a_{C_1p_n}F^{p-C_1p_n}}{a_{C_1p_n}} \lesssim C_1p_n + 1.$$

The missing initial part of the normalized sum in the left side also contributes at most C_1p_n. This concludes the proof for the bound on ν_k.

For the final assertion we must take the dependence of $a_p = a_{p,k}$ on k into account. The preceding argument shows that $a_{p,k} \le a_{Cp_n,k}F^{p-Cp_n}$ for constants C, F that do not depend on k. Therefore $a_{p,k}/\sum_p a_{p,k} \le F^{p-Cp_n}$ for every $p \ge Cp_n$ and every k. The sum in the lemma is thus bounded above by $\sum_{p\ge P_n} F^{p-Cp_n}e^{m_2D_1p_n}$, for $P_n \ge Cp_n$, where $F < 1$. □

AMS 2000 subject classifications: Primary 62G05, 62G20
Keywords and phrases: Bayesian estimators, Sparsity, Gaussian sequence model, Mixture priors, Asymptotics, Contraction.

PROOF OF LEMMA 5.1. For simplicity of notation we can choose $\theta_0 = 0$. If $\|\theta - \theta_1\| \leq \|\theta_1\|/2$, then $\|\theta\| \geq \|\theta_1\|/2$ and hence $\langle\theta, \theta_1\rangle = \big(\|\theta\|^2 + \|\theta_1\|^2 - \|\theta - \theta_1\|^2\big)/2 \geq \|\theta_1\|^2/2$. Therefore, the test $\phi = 1_{\theta_1^T X > D\|\theta_1\|}$ satisfies

$$\begin{aligned} P_{n,\theta_0}\phi &= 1 - \Phi(D), \\ P_{n,\theta}(1-\phi) &= \Phi\big((D\|\theta_1\| - \langle\theta, \theta_1\rangle)/\|\theta_1\|\big) \leq \Phi(D - \rho), \end{aligned}$$

for $\rho = \|\theta_1\|/2$. The infimum over D of $\alpha\big(1 - \Phi(D)\big) + \beta\Phi(D - \rho)$ is attained for $D = \rho^{-1}\log(\alpha/\beta) + \rho/2$, for which $D - \rho = \rho^{-1}\log(\alpha/\beta) - \rho/2$, which leads to the first inequality.

If $D - \rho \leq 0 \leq D$, then the bound $1 - \Phi(x) \leq e^{-x^2/2}$ for $x \geq 0$ gives that the infimum is bounded above by

$$\alpha e^{-D^2/2} + \beta e^{-(D-\rho)^2/2} = 2\sqrt{\alpha\beta}\, e^{-(\log(\alpha/\beta)^2/(2\rho^2)} e^{-\rho^2/8}.$$

If $0 < D - \rho < D$, then the term $\alpha\big(1 - \Phi(D)\big)$ can be bounded as before. For the second term we use that $\beta < \alpha e^{-\rho^2/2}$, so that $\beta\Phi(D - \rho) \leq \beta \leq \sqrt{\alpha\beta}\, e^{-\rho^2/4}$. If $D - \rho \leq D < 0$, then we similarly treat the first term differently, by bounding this by α. □

PROOF OF LEMMA 5.2. To prove the first assertion, we apply Jensen's inequality to see that

$$\log \int \frac{p_{n,\theta}}{p_{n,\theta_0}}(X)\, \frac{d\Pi(\theta)}{\|\tilde\Pi\|} \geq \int \log \frac{p_{n,\theta}}{p_{n,\theta_0}}(X)\, \frac{d\tilde\Pi(\theta)}{\|\tilde\Pi\|} = \tilde\mu^T(X - \theta_0) - \tilde\sigma^2/2.$$

To prove the second assertion we apply the first with $\tilde\Pi$ equal to Π restricted to the ball $\{\theta : \|\theta - \theta_0\| \leq r\}$. The relevant characteristics corresponding to this measure satisfy $\|\tilde\mu\| \leq r$ and $\tilde\sigma^2 \leq r^2$. Under P_{n,θ_0} the variable $\tilde\mu^T(X - \theta_0)$ is distributed as $Z\|\tilde\mu\|$, for a standard normal variable Z. The assertion follows from the inequality $\Pr(Zr \leq -r^2 + r^2/2) \leq \exp(-r^2/8)$. □

PROOF OF LEMMA 5.3. This follows from the explicit formula $v_p = \pi^{p/2}/\Gamma(p/2 + 1)$, and Stirling's formula with bounds $\sqrt{2\pi} < \Gamma(x + 1)/\big((x/e)^x\sqrt{x}\big) < \sqrt{2\pi}e^{1/(12x)}$, which holds for any $x > 0$, see e.g. [5]. □

PROOF OF LEMMA 7.1. By Lemma 5.2 in [1], there exist events $\mathcal{A}_n$ with $\Pr(A_n^c) \leq e^{-r_n^2/8} \to 0$ such that $\int p_{n,\theta}/p_{n,\theta_0}\, d\Pi_n(\theta) \geq e^{-r_n^2}\Pi_n\big(\theta : \|\theta - \theta_0\| < r_n\big)$. By the definition of the posterior as a quotient, it follows that

$$\begin{aligned} P_{n,\theta_0}\Pi_n\big(\theta : \|\theta - \theta_0\| < s_n |\, X\big)1_{\mathcal{A}_n} &\leq \frac{P_{n,\theta_0}\int_{\theta:\|\theta-\theta_0\|<s_n} p_{n,\theta}/p_{n,\theta_0}\, d\Pi_n(\theta)}{e^{-r_n^2}\Pi_n\big(\theta : \|\theta - \theta_0\| < r_n\big)} \\ &\leq \frac{\Pi_n\big(\theta : \|\theta - \theta_0\| < s_n\big)}{e^{-r_n^2}\Pi_n\big(\theta : \|\theta - \theta_0\| < r_n\big)}, \end{aligned}$$

by Fubini's theorem. The right side tends to zero by assumption. □

PROOF OF LEMMA 7.2. We use the notations $p_n = |S_{\theta_0}|$ and $p = |S \cap S_{\theta_0}|$, so that $|S_{\theta_0} \setminus S| = p_n - p$. Using Stirling's formula, we obtain that for large enough c_1 it holds

$$\begin{aligned} \frac{v_p}{v_{p_n}} &= \pi^{\frac{p-p_n}{2}} \frac{\Gamma(p_n/2+1)}{\Gamma(p/2+1)} \le c_1 \sqrt{\frac{2\pi(p_n/2)}{2\pi(p/2)}} \frac{(p_n/2)^{p_n/2}}{(p/2)^{p/2}} \\ &\le c_1 \exp\left(\frac{p_n}{2}\log\frac{p_n}{2} - \frac{p}{2}\log\frac{p}{2}\right). \end{aligned}$$

Thus for any $1 \le p \le p_n$,

$$\frac{v_p}{v_{p_n}} \frac{1}{r_n^{p_n-p}} \le c_1 \exp\left(\frac{p_n-p}{2}\log\frac{p_n}{2r_n^2} + \frac{p}{2}\log\frac{p_n}{p}\right).$$

Since the function $p \mapsto (p/2)\log(p_n/p)$ is inferior to $p_n/(2e)$ for any $1 \le p \le p_n$, we conclude, using the assumption on α_n, that the last display is bounded from above by $c_1 \exp(C_2 p_n)$. This concludes the proof. □

2. Proof of Theorem 2.5 in [1]. The proof is based on refinements of the general scheme to obtain posterior rates presented in [2, 3] that uses tests. Different from the proofs of Theorems 2.2 and 2.4 in [1], which gives separate proofs for low and high-dimensional models, the form of the prior π_n enables the use of tests for any dimension p between 0 and n. We start with a series of lemmas and next state an explicit bound in Proposition 2.1 below.

LEMMA 2.1. *For any $p \in \mathbb{N}$ and $\alpha_p, \beta_p > 0$, there exists a test ϕ_p based on $X \sim N_n(\theta, I)$ with, for any $r > 1$ and every integer $j \ge 1$,*

$$P_{n,\theta_0}\phi_p \le 66\sqrt{\frac{\beta_p}{\alpha_p}}\binom{n}{p}48^p \exp(-r^2/32), \tag{2.1}$$

$$\sup_{\theta\in\mathbb{R}^n:|S_\theta|\le p, \|\theta-\theta_0\|\ge jr} P_{n,\theta}(1-\phi_p) \le 2\sqrt{\frac{\alpha_p}{\beta_p}}\exp(-j^2r^2/32). \tag{2.2}$$

PROOF. The set $\Theta_p = \{\theta \in \mathbb{R}^n : |S_\theta| = p\}$ can be partitioned into the shells

$$\mathcal{C}_{j,p}(r) = \{\theta \in \Theta_p : jr < \|\theta - \theta_0\| \le (j+1)r\}. \tag{2.3}$$

Similarly as in the proof of Corollary 1 in [4], we cover each of these shells by a minimal collection of balls of radius $jr/2$ with centers inside the shells, and next construct ϕ_p as the maximum of all the tests as in Lemma 5.1 in [1] attached to one of the centres, for some shell with $j \geq 1$. Every θ in such a ball with center θ_1 satisfies $\|\theta - \theta_1\| \leq jr/2 \leq \|\theta_0 - \theta_1\|/2$, since $\theta_1 \in \mathcal{C}_{j,p}(r)$. Hence each test satisfies the inequalities of Lemma 5.1 in [1].

If $S_\theta \subset S$, then $\|\theta - \theta_0\|^2 = \|\pi_S\theta - \pi_S\theta_0\|^2 + \|\pi_{S^c}\theta_0\|^2$, and hence the number of centres is bounded by

$$N\big(jr/4, \{\theta \in \Theta_p : \|\theta - \theta_0\| \leq 2jr\}, \|\cdot\|\big) \leq \sum_{S:|S|=p} N\big(jr/4, \{\theta \in \mathbb{R}^S : \|\theta - \pi_S\theta_0\| \leq 2jr\}, \|\cdot\|\big) 1_{\|\pi_{S^c}\theta_0\| \leq 2jr}.$$

(The covering number is taken at $jr/4$ rather than at $jr/2$ to account for the fact the centers of the balls are to be inside the shell.) This is bounded above by the number of sets $\binom{n}{p}$ of size p times $(8 \times 6)^p$, the entropy term being bounded using Lemma 4.1 in [6].

We conclude the proof as the proof of Corollary 1 in [4], noting that

$$\begin{aligned} P_{n,\theta_0}\phi_p &\leq 2\sqrt{\frac{\beta_p}{\alpha_p}}\binom{n}{p} 48^p \frac{e^{-r^2/32}}{1 - e^{-r^2/32}}. \\ &\leq \frac{2}{1 - e^{-1/32}}\sqrt{\frac{\beta_p}{\alpha_p}}\binom{n}{p} 48^p e^{-r^2/32}. \square \end{aligned}$$

LEMMA 2.2. *Under conditions (2.5)-(2.6) in [1] for all* $\theta \in \mathbb{R}^S$ *with* $\|(\theta, 0_{S^c}) - \theta_0\| < r$,

$$e^{-2c_1|S| - c_1|S_{\theta_0}| - r^2/64} \leq \frac{g_S(\theta)}{g_{S_{\theta_0}}(\pi_{S_{\theta_0}}\theta_0)} \leq e^{2c_1|S| + c_1|S_{\theta_0}| + r^2/64}.$$

PROOF. We first use condition (2.5) in [1] to see that

$$\Big|\log \frac{g_S(\theta)}{g_S(\pi_S\theta_0)}\Big| \leq c_1|S| + \|\theta - \pi_S\theta_0\|^2/64.$$

Next we use condition (2.6) in [1] twice, first with $S' = S_{\theta_0} \cap S \subset S$ and next with $S' = S_{\theta_0} \cap S \subset S_{\theta_0}$, to see that

$$\begin{aligned} &\Big|\log \frac{g_S(\pi_S\theta_0)}{g_{S_{\theta_0}\cap S}(\pi_{S_{\theta_0}\cap S}\theta_0)}\Big| \leq c_1|S|, \\ &\Big|\log \frac{g_{S_{\theta_0}\cap S}(\pi_{S_{\theta_0}\cap S}\theta_0)}{g_{S_{\theta_0}}(\pi_{S_{\theta_0}}\theta_0)}\Big| \leq c_1|S_{\theta_0}| + \|\pi_{S_{\theta_0}\setminus S}\theta_0\|^2/64. \end{aligned}$$

The lemma follows upon combining these three inequalities with the observation that $\|\theta - \pi_S\theta_0\|^2 + \|\pi_{S_{\theta_0}\setminus S}\theta_0\|^2 = \|(\theta, 0_{S^c}) - \theta_0\|^2$. $\square$

LEMMA 2.3. *Under conditions (2.5) and (2.6) in [1], setting $d = \sqrt{2e\pi}$ and $d_1 = 1/\sqrt{\pi}$, we have that, for every $r > 0$, and $p_n = |S_{\theta_0}|$,*

$$\begin{aligned}\Pi_n(\theta \in \mathbb{R}^n : \|\theta - \theta_0\| < r) \\ \geq d_1 \sum_{p=p_n}^{n} \frac{\binom{n-p_n}{p-p_n}}{\binom{n}{p}} \pi_n(p)\, p^{-p/2} d^p r^p e^{-r^2/64} e^{-c_1 p_n} g_{S_{\theta_0}}(\pi_{S_{\theta_0}} \theta_0).\end{aligned}$$

PROOF. By definition the left side of the lemma is equal to

$$\sum_{p=1}^{n} \pi_n(p) \frac{1}{\binom{n}{p}} \sum_{|S|=p} G_S(\theta \in \mathbb{R}^S : \|\theta - \pi_S\theta_0\|^2 + \|\pi_{S_{\theta_0} \setminus S}\theta_0\|^2 < r^2).$$

If $S \supset S_{\theta_0}$, then $\|\pi_{S_{\theta_0} \setminus S}\theta_0\| = 0$. Hence the preceding display is at least

$$\sum_{p=p_n}^{n} \pi_n(p) \frac{1}{\binom{n}{p}} \sum_{|S|=p, S \supset S_{\theta_0}} G_S(\theta \in \mathbb{R}^S : \|\theta - \pi_S\theta_0\| < r).$$

For v_p the volume of the p-dimensional Euclidean unit ball, the measure $G_S(\theta \in \mathbb{R}^p : \|\theta - \pi_S\theta_0\| < r)$ in this expression can be bounded below by (with $p = |S|$)

$$\inf_{\theta \in \mathbb{R}^S : \|\theta - \pi_S\theta_0\| < r} g_S(\theta) r^p v_p \geq g_{S_{\theta_0}}(\pi_{S_{\theta_0}}\theta_0) d_1 e^{-2c_1 p - c_1 p_n - r^2/64} r^p d^p p^{-p/2},$$

by Lemma 2.2, since $\|(\theta, 0_{S^c}) - \theta_0\| = \|\theta - \pi_S\theta_0\|$. The resulting lower bounds for the terms in the sum do depend on S only through $p = |S|$. For each $p \geq p_n$ there exist $\binom{n-p_n}{p-p_n}$ subsets of size $|S| = p$ that contain S_{θ_0}. □

LEMMA 2.4. *For any integer $q \in \mathbb{N}$, for any $C \geq 1$,*

$$I_C(q) = \int_C^{+\infty} x^q e^{-x} dx \leq (q+1)(C \vee q)^q e^{-C}.$$

For a real $q > 0$, one extends this result using the fact that $I_C(q) \leq I_C(\lceil q \rceil)/C^{\lceil q \rceil - q}$.

PROOF. By partial integration the function $I_C(q)$, for fixed $C > 0$, can be seen to satisfy $I_C(q) = C^q e^{-C} + qI_C(q-1)$. For $q \in \mathbb{N}$ repeated application of this recursion gives that

$$I_C(q) = e^{-C}(C^q + qC^{q-1} + q(q-1)C^{q-2} + \cdots + q!).$$

It follows that $I_C(q) \leq e^{-C}(q+1)(q \vee C)^q$ for any $q \in \mathbb{N}$. □

LEMMA 2.5. *For any $p \in \mathbb{N}$ and any constants $D \geq 1$ and $M \geq 1$,*

$$\sum_{j \geq M} (jD)^p e^{-j^2 D^2} \leq 2(2p+1) e^{-\left(\frac{p}{2} \vee M^2 D^2\right)} \left(\frac{p}{2} \vee M^2 D^2\right)^{p/2}.$$

PROOF. The function $\psi_p : u \to (uD)^p e^{-u^2 D^2}$ attains its maximum on the positive reals at $u = \sqrt{p/2}/D =: u^*$ and the corresponding value is $\bar{\psi}_p := (p/2)^{p/2} e^{-p/2}$.

If $M > u^*$, the terms of the sum are decreasing in j and thus

$$\sum_{j \geq M} (jD)^p e^{-j^2 D^2} \leq (MD)^p e^{-M^2 D^2} + \int_M^\infty (xD)^p e^{-x^2 D^2} dx.$$

It suffices to control the integral, with $C = M^2 D^2$

$$\int_M^\infty (xD)^p e^{-x^2 D^2}\, dx = \frac{1}{2D} I_C\left(\frac{p-1}{2}\right).$$

In view of Lemma 2.4 and the facts that $D \geq 1$ and $M > u^*$, the latter quantity is bounded by $e^{-C} C^{(p-1)/2}(p+1)/4$ if p is odd, and by $e^{-C} C^{p/2}(1+p/2)/2\sqrt{C}$ if p is even. In both cases, using that $C \geq 1$, it is bounded by $e^{-C} C^{p/2}(p/4 + 1/2)$.

If $M \leq u^*$, we split the sum in two parts corresponding to $M \leq j \leq \lfloor u^* \rfloor$ and $j \geq \lfloor u^* \rfloor + 1$. The second part is bounded as above while for the first part, we bound each term by the maximum value $\bar{\psi}_p$. Thus in this case

$$\sum_{j \geq M} (jD)^p e^{-j^2 D^2} \leq (\lfloor u^* \rfloor - M + 2)\bar{\psi}_p + \int_{u^*}^\infty (xD)^p e^{-x^2 D^2} dx.$$

The integral is bounded by $(p/4 + 1/2)\bar{\psi}_p$ using Lemma 2.4 as above while the first term is bounded by $3\sqrt{p/2}\bar{\psi}_p \leq 3p\bar{\psi}_p$. □

Given any prior π_n on dimension, p_n an integer between 0 and n, setting $h = e^{5+2c_1}$ and $d = \sqrt{2e\pi}$, define

$$C_n(r; \pi_n, p_n) = 20 e^{c_1 p_n} \frac{\sum_{p=0}^n \sqrt{\pi_n(p)}\, (hn/p)^{p/2} (1 \vee r^2/p)^{p/4}}{\left(\sum_{p=p_n}^n \frac{\binom{n-p_n}{p-p_n}}{\binom{n}{p}} \pi_n(p) d^p (r^2/p)^{p/2}\right)^{1/2}}.$$

PROPOSITION 2.1. *If the densities g_S have a finite second moment and satisfy (2.5)-(2.6) in [1], then for any $1 \leq p_n \leq n$ and any $r \geq 1$,*

$$\sup_{\theta_0 \in \ell_0[p_n]} P_{n,\theta_0} \Pi_n(\theta : \|\theta - \theta_0\| > 10r |\, X) \leq e^{-r^2/9}\big(C_n(r, \pi_n, p_n) + 1\big).$$

PROOF. For any $p \in \mathbb{N}$ fix a test ϕ_p as given in Lemma 2.1, with α_p, β_p to be chosen later. Then, for $\Theta_p = \{\theta \in \mathbb{R}^n : |S_\theta| = p\}$ and any events $\mathcal{A}_n$,

$$\begin{aligned}
&P_{n,\theta_0}\Pi_n(\theta : \|\theta - \theta_0\| > 10r|\, X) \\
&\quad = P_{n,\theta_0}\sum_{p=1}^{n}\Pi_n(\theta \in \Theta_p : \|\theta - \theta_0\| > 10r|\, X) \\
&\quad = \sum_{p=0}^{n} P_{n,\theta_0}\phi_p + \sum_{p=1}^{n} P_{n,\theta_0}\Pi_n(\theta \in \Theta_p : \|\theta - \theta_0\| > 10r|\, X)(1-\phi_p)1_{\mathcal{A}_n} \\
&\qquad\qquad + P_{n,\theta_0}(\mathcal{A}_n^c) \\
&\quad =: A + B + C.
\end{aligned}$$

Using (2.1) and that $\binom{n}{p} \le \sum_{j=0}^{p}\binom{n}{j} \le (ne/p)^p$, we obtain, for $10r > 1$,

$$A \le a\sum_{p=1}^{n}\sqrt{\frac{\beta_p}{\alpha_p}}\binom{n}{p}48^p e^{-r^2/32} \le a\sum_{p=0}^{n}\sqrt{\frac{\beta_p}{\alpha_p}}e^{5p+p\log n/p-r^2/32},$$

with $a = 66$. In view of the second assertion of Lemma 5.2 in [1] applied with r equal to ηr for a small number η, to be determined later, there exist events $\mathcal{A}_n$ such that

$$C = P_{n,\theta_0}(\mathcal{A}_n^c) \le e^{-\eta^2 r^2/8}, \tag{2.4}$$

while on the event $\mathcal{A}_n$,

$$\int \frac{p_{n,\theta}}{p_{n,\theta_0}}\,d\Pi_n(\theta) \ge e^{-\eta^2 r^2}\Pi_n(\theta : \|\theta - \theta_0\| < \eta r).$$

Therefore, for $\mathcal{C}_{j,p}(r)$ as defined in (2.3), in view of Fubini's theorem,

$$\begin{aligned}
B &= \sum_{p=0}^{n}\sum_{j\ge 10} P_{n,\theta_0}\left[(1-\phi_p)1_{\mathcal{A}_n}\frac{\int_{\mathcal{C}_{j,p}(r)} p_{n,\theta}/p_{n,\theta_0}\,d\Pi_n(\theta)}{\int p_{n,\theta}/p_{n,\theta_0}\,d\Pi_n(\theta)}\right] \\
&\le e^{\eta^2 r^2}\sum_{p=0}^{n}\sum_{j\ge 10}\frac{\int_{\mathcal{C}_{j,p}(r)} P_{n,\theta}(1-\phi_p)\,d\Pi_n(\theta)}{\Pi_n(\theta : \|\theta - \theta_0\| < \eta r)} \\
&\le 2e^{\eta^2 r^2}\sum_{p=0}^{n}\sqrt{\frac{\alpha_p}{\beta_p}}\sum_{j\ge 10}\frac{\Pi_n(\mathcal{C}_{j,p}(r))}{\Pi_n(\theta : \|\theta - \theta_0\| < \eta r)}e^{-j^2 r^2/32},
\end{aligned}$$

where the last inequality follows from inequality (2.2). Here, with v_p the volume of the p-dimensional unit ball,

$$\begin{aligned}\Pi_n(\mathcal{C}_{j,p}(r)) &= \frac{\pi_n(p)}{\binom{n}{p}} \sum_{S:|S|=p} G_S(\theta\in\mathbb{R}^S : jr < \|(\theta,0_{S^c})-\theta_0\| \le (j+1)r) \\ &\le \frac{\pi_n(p)}{\binom{n}{p}} \sum_{S:|S|=p} v_p\,(j+1)^p r^p \max_{\theta\in\mathbb{R}^S:\|(\theta,0_{S^c})-\theta_0\|\le (j+1)r} g_S(\theta) \\ &\le \pi_n(p) v_p\,(2jr)^p g_{S_{\theta_0}}(\pi_{S_{\theta_0}}\theta_0) e^{2c_1p+c_1p_n+(j+1)^2r^2/64},\end{aligned}$$

by Lemma 2.2. Because $j^2/32-(j+1)^2/64 \ge j^2/100$ for $j\ge 10$, we obtain

$$\begin{aligned}&\sum_{j\ge 10} \Pi_n(\mathcal{C}_{j,p}(r)) e^{-j^2r^2/32} \\ &\quad\le \sum_{j\ge 10} \pi_n(p) v_p\,(jr/10)^p 20^p g_{S_{\theta_0}}(\pi_{S_{\theta_0}}\theta_0) e^{2c_1p+c_1p_n-j^2r^2/100} \\ &\quad\le 4d_2\pi_n(p)p^{-p/2} e^{c_1p_n+2c_1p} g_{S_{\theta_0}}(\pi_{S_{\theta_0}}\theta_0) F^p e^{-r^2}\Big(\frac{p}{2}\vee r^2\Big)^{p/2},\end{aligned}$$

where $F=20\sqrt{2e\pi}\le e^5$, by Lemmas 5.2 in [1] and 2.5. Combining this with the result of Lemma 2.3, and choosing $\eta^2=8/9$ so that $1-\eta^2-\eta^2/64>1/9$, we obtain

$$B \le \frac{\sum_{p=0}^n \sqrt{\alpha_p/\beta_p}\,\pi_n(p) e^{2c_1(p_n+p)} F^p \big(1\vee (r^2/p)\big)^{p/2}}{\sum_{p=p_n}^n \frac{\binom{n-p_n}{p-p_n}}{\binom{n}{p}} \pi_n(p) d^p (r^2/p)^{p/2}}\, e^{-r^2/9} b,$$

where $b=4d_2/d_1\le 5$. We now balance A and B by choosing

$$\begin{aligned}\alpha_p &= e^{p\log n/p} a, \\ \beta_p &= \frac{\pi_n(p) e^{2c_1(p_n+p)} F^p \big(1\vee (r^2/p)\big)^{p/2}}{\sum_{p=p_n}^n \frac{\binom{n-p_n}{p-p_n}}{\binom{n}{p}} \pi_n(p)(d^2r^2/p)^{p/2}} b.\end{aligned}$$

With this choice, for $h=e^{5+2c_1}$,

$$A+B \le e^{c_1p_n} \frac{\sum_{p=0}^n \sqrt{\pi_n(p)}\,(hn/p)^{p/2}\big(1\vee (r^2/p)\big)^{p/4}}{\Big(\sum_{p=p_n}^n \frac{\binom{n-p_n}{p-p_n}}{\binom{n}{p}} \pi_n(p)(d^2r^2/p)^{p/2}\Big)^{1/2}} e^{-r^2/9}\sqrt{ab}.$$

Together with the bound (2.4) this concludes the proof of the first assertion. □

Finally we are ready for the proof of Theorem 2.5. For $\pi_n(p)$ satisfying (2.10) in [1] with $a \geq 1$ we have $\pi_n(p)(hn/p)^p \leq e^{-p\log(b/h)}$. In view of the inequalities $(1 \vee u)^p \leq 1 + u^p$ and $p\log(t/p) \leq t/e$ for any positive reals u, t, the sum in the numerator of $C_n(r; \pi_n, p_n)$ is bounded above, for $b \geq e^2 h$, by

$$\sum_{p=0}^{n} e^{-(p/2)\log(b/h)} + \sum_{p=0}^{n} e^{-p/4} e^{(p/4)\log(r^2 e^{-3}/p)} \lesssim e^{r^2 e^{-4}/4}.$$

This last bound is less than $e^{r^2/90}$. The sum in the denominator of $C_n(r; \pi_n, p_n)$ is bounded below by its p_nth term, which for $r^2 \geq r_n^2 \geq p_n$ is bounded below by $\pi_n(p_n)/\binom{n}{p_n} \gtrsim e^{-2r_n^2}$. Thus we obtain that, for $r \geq r_n$,

$$e^{-r^2/9}\big(C_n(r; \pi_n, p_n) + 1\big) \lesssim e^{-r^2/10 + 2r_n^2 + c_1 p_n}.$$

For r replaced by $r + 4.5 r_n$ this is bounded above by $e^{-r^2/10}$.

3. Proof of Theorem 2.7 in [1]. The second assertion of Theorem 2.7 in [1] can be obtained from the first and Theorem 2.6 in [1] using interpolation of the d_q and Euclidean distances, following a similar method as for $\ell_0[p_n]$-classes in the last paragraph of Section 5 in [1].

Denoting again the projection of θ_0 into $\ell_0[p_n^*]$ by θ_1, we have, for $|S_\theta| \leq (A-1)p_n^*$,

$$\begin{aligned} d_q(\theta, \theta_0) &\lesssim d_q(\theta, \theta_1) + d_q(\theta_1, \theta_0) \lesssim (Ap_n^*)^{1-q/2}\|\theta - \theta_1\|^q + d_q(\theta_1, \theta_0) \\ &\lesssim (Ap_n^*)^{1-q/2}\|\theta - \theta_0\|^q + (Ap_n^*)^{1-q/2}\|\theta_0 - \theta_1\|^q + d_q(\theta_1, \theta_0). \end{aligned}$$

For $\theta_0 \in m_s[p_n]$ for $q > s$ and p_n^* given by (2.10) in [1] the second and third terms can be seen to be of the order $\mu_{n,s,q}^*$, where we use (6.1) in [1] to bound the second term and the d_q-analogon of (6.1) in [1] for the third. By Theorem 2.6 in [1] the first term on the right is of the order $(p_n^*)^{1-q/2}\mu_{n,s,2}^*$, which is $\mu_{n,s,q}^*$.

For the first assertion of Theorem 2.7 in [1], we follow the proof of Theorem 2.1 in [1] until (4.2), where we take S_0 to be the set of indices of the p_n^* largest coordinates of θ_0 in absolute value. We take $B = \{\theta : |S_\theta \cap S_0^c| \geq Rp_n^*\}$. The denominator in (4.2) in [1] is bounded from below as below (4.2) in [1],

$$\Pi_n(B|\, X, \theta_{S_0} = \bar\theta_1) \leq \int_B \frac{p_{\bar n_2, \bar\theta_2}}{p_{\bar n_2, 0_{S_0^c}}}(X_{S_0^c}) \frac{d\Pi_n(\bar\theta_2|\, \bar\theta_1)}{\Pi_n(\bar\theta_2 = 0|\, \bar\theta_1)}, \tag{3.1}$$

where now $\bar n_2 = n - p_n^*$. If S_2 are the indices of the nonzero coordinates of $\bar\theta_2 \in \mathbb{R}^{S_0^c}$, θ_2 the vector of their values and $n_2 = |S_2|$, then

$$\frac{p_{\bar n_2, \bar\theta_2}}{p_{\bar n_2, 0_{S_0^c}}}(X_{S_0^c}) = \frac{p_{n_2, \theta_2}}{p_{n_2, \theta_{0,S_2}}}(X_{S_2}) \frac{p_{n_2, \theta_{0,S_2}}}{p_{n_2, 0_{S_2}}}(X_{S_2}),$$

where the last ratio can be written as $\exp(\|\theta_{0,S_2}\|^2/2+(X_{S_2}-\theta_{0,S_2})^T\theta_{0,S_2})$. Let us consider the event, with $E_n^2(S_2)=3|S_2|\log((n-p_n^*)e/|S_2|)$,

$$\mathcal{A}=\bigcap_{p=Rp_n^*}^{n-p_n^*}\bigcap_{|S_2|=p}\{(X_{S_2}-\theta_{0,S_2})^T\theta_{0,S_2}\le\|\theta_{0,S_2}\|\,E_n(S_2)\}.$$

From equation (6.1) in [1] and the definition of S_2, we have that

$$\|\theta_{0,S_2}\|^2\le\|\theta_{0,S_0^c}\|^2\lesssim\mu^*_{n,s,2}.$$

Using the inequality $ab\le b^2+a^2/4$, deduce that on the event $\mathcal{A}$,

$$\frac{p_{n_2,\theta_{0,S_2}}}{p_{n_2,0_{S_2}}}(X_{S_2})\le e^{C\mu^*_{n,s,2}+\frac{E_n^2(S_2)}{4}}.$$

It follows that on the event $\mathcal{A}$ the right side of (3.1) is bounded above by

$$\sum_{p=Rp_n^*}^{n-p_n^*}\sum_{|S_2|=p}\frac{\pi_{n,k}(p)}{\pi_{n,k}(0)}\frac{m_1^{p+p_n}}{\binom{n-p_n^*}{p}}e^{C\mu^*_{n,s,2}+\frac{E_n^2(S_2)}{4}}\int\frac{p_{n_2,\theta_2}}{p_{n_2,\theta_{0,S_2}}}(X_{S_2})\gamma_{S_2}(\theta_2)\,d\theta_2.$$

Taking the maximum over k inside and next the expectation under P_{n,θ_0},

$$P_{n,\theta_0}\Pi_n(B|\,X)1_{\mathcal{A}}\le e^{C\mu^*_{n,s,2}}\sum_{p=Rp_n^*}^{n-p_n^*}\max_{0\le k\le p_n^*}\frac{\pi_{n,k}(p)}{\pi_{n,k}(0)}C^{p+p_n}e^{\frac{3}{4}p\log\frac{(n-p_n^*)e}{p}}.$$

Using the explicit form of the complexity prior (2.8) and proceeding as in the proof of Lemma 4.1 in [1], one obtains, for any k between 0 and p_n^*,

$$\frac{\pi_{n,k}(p)}{\pi_{n,k}(0)}\le e^{ak\log\left(1+\frac{p}{k}\right)-ap\log\left(\frac{bn}{p+k}\right)}\le e^{-ap\log\left(\frac{en}{p+k}\right)},$$

where the last inequality holds when p is greater than a large enough multiple of k, which is achieved by taking the constant R large enough. Deduce that, for a large enough and some positive constant c,

$$P_{n,\theta_0}\Pi_n(B|\,X)1_{\mathcal{A}}\le e^{C\mu^*_{n,s,2}}\sum_{p=Rp_n^*}^{n-p_n^*}e^{-\frac{a}{2}p\log\left(\frac{en}{p+p_n^*}\right)}\lesssim e^{-c\mu^*_{n,s,2}}.$$

The expectation over the complement of $\mathcal{A}$ is bounded above by

$$\begin{aligned}P_{n,\theta_0}(\mathcal{A}^c)&\le\sum_{p=Rp_n^*}^{n-p_n^*}\binom{n-p_n^*}{p}e^{-3p\log((n-p_n^*)e/p)/2}\\&\le\sum_{p=Rp_n^*}^{n-p_n^*}e^{-[3/2-1]p\log((n-p_n^*)e/p)}\lesssim((n-p_n^*)e/p_n^*)^{-\frac{Rp_n^*}{4}}.\quad\square\end{aligned}$$

4. Proof of Theorem 2.9 in [1]. Let us define $\alpha_n = c\underline{\varepsilon_n}$ and $\gamma_n = d\underline{\varepsilon_n}$ with $c < d$ small enough constants to be defined later. Due to Lemma 7.1 in [1], it suffices to prove that for some $\theta_0 \in \mathbb{R}^n$, it holds

$$Q_n = \frac{\Pi_n(\theta : \ \|\theta - \theta_0\|^2 \leq n\alpha_n^2)}{\Pi_n(\theta : \ \|\theta - \theta_0\|^2 \leq n\gamma_n^2)} \leq \exp(-2n\gamma_n^2).$$

Let us define θ_0 by $\theta_{0,k} = 0$ for any $k > d_{3,n} = (3d_{2,n} - d_{1,n})/2$ and, with M some sufficiently large constant to be defined later,

$$\theta_{0,k} = \begin{cases} Mn\underline{\varepsilon_n}^2 & \text{if } k \in \{1, \dots, d_{1,n}\} = S_1 \\ 4n\alpha_n^2/(d_{2,n} - d_{1,n}) & \text{if } k \in \{d_{1,n} + 1, \dots, d_{3,n}\} \end{cases}$$

Note that if the support of θ is S and if $\|\theta - \theta_0\|^2 \leq n\alpha_n^2$, then S belongs to the set $\mathcal{Q}$ of supports defined by

$$\mathcal{Q} = \{S : \ S \supset S_1, \ |S \cap \{d_{1,n} + 1, \dots, d_{3,n}\}| \geq d_{2,n} - d_{1,n}\}.$$

In particular, for any $S \in \mathcal{Q}$, we have $|S| \geq d_{2,n}$. We also have the inclusion

$$\{\theta, \ |S_\theta| = S_1, \ \|\theta - \pi_{S_1}\theta_0\|^2 \leq n\gamma_n^2/2\} \subset \{\theta \in \mathbb{R}^n, \ \|\theta - \theta_0\|^2 \leq n\gamma_n^2\},$$

as long as we impose $12c \leq d$. Thus

$$\begin{aligned} Q_n &\leq \sum_{k=d_{2,n}}^{n} \sum_{S \in \mathcal{Q}, \ |S|=k} \frac{\Pi_n(\|\theta - \theta_0\|^2 \leq n\alpha_n^2 | S)}{\Pi_n(\|\theta - \theta_0\|^2 \leq n\gamma_n^2)} \frac{\pi_n(k)}{\binom{n}{k}} \\ &\leq \sum_{k=d_{2,n}}^{n} \sum_{S \in \mathcal{Q}, \ |S|=k} \frac{\Pi_n(\|\theta - \theta_0\|^2 \leq n\alpha_n^2 | S)}{\Pi_n(\|\theta - \pi_{S_1}\theta_0\|^2 \leq n\gamma_n^2/2 | S_1)} \frac{\pi_n(k)\binom{n}{d_{1,n}}}{\pi_n(d_{1,n})\binom{n}{k}} \end{aligned}$$

Now note that if $\gamma_n^2 \geq 2\alpha_n^2$ (which amounts to impose $d \geq \sqrt{2}c$), the first ratio of probabilities in the last display is bounded above by 1. Using the monotonicity of the prior π_n, one obtains that $\pi_n(k) \leq \pi_n(d_{2,n})$ for any $k \geq d_{2,n}$. Thus $Q_n \leq \exp(-n\underline{\varepsilon_n}^2) \leq \exp(-2n\gamma_n^2)$, as long as $d \leq C/\sqrt{2}$, which implies the result due to Lemma 7.1 in [1]. □

5. Proof for Example 2.6 in [1] on weakly-mixing priors. Let us first check that these priors satisfy (2.5)-(2.6) in [1]. Let $S' \subset S \subset \{1, \dots, n\}$ and $\theta \in \mathbb{R}^S$. Equation (2.5) in [1] can be verified as for Example 2.5, using that G is Lipschitz. Checking (2.6) in [1] is also similar, except for the term $\log(a_{|S|}/a_{|S'|})$. Set $H(\theta) = \sum_{i=1}^p h(\theta_i)$. The assumptions on G imply that $c \leq e^{G(t)} \leq Ce^{dt}$ for positive c, C, d, so, since $\|\theta\| \leq \|\theta\|_1$,

$$\int e^{H(\theta) - d\|\theta\|_1} \, d\theta \lesssim a_{|S|}^{-1} = \int e^{H(\theta) - G(\|\theta\|)} \, d\theta \lesssim \int e^{H(\theta)} d\theta.$$

The integrals on the far left and right sides factorize, where e^h is integrable by assumption. Hence $|S| \lesssim \log a_{|S|}^{-1} \lesssim |S|$, which implies (2.6) in [1].

Next we verify (2.7) in [1] in the special case that $h(\theta) = -\|\theta\|_1$ and G that is Lipschitz with constant $a < 1$. For $\theta = (\theta_1, \theta_2)$, we have that $\big|G(\|\theta\|) - G(\|\theta_1\|)\big| \leq a\big|\|\theta\| - \|\theta_1\|\big| \leq a\|\theta_2\| \leq a\|\theta_2\|_1$. Thus

$$\frac{g_{S_1,S_2}(\theta_1,\theta_2)}{g_{S_1}(\theta_1)} = \frac{a_{|S_1|+|S_2|}}{a_{|S_1|}} e^{-\|\theta_2\|_1 - G(\|\theta\|) + G(\|\theta_1\|)} \leq \frac{a_{|S_1|+|S_2|}}{a_{|S_1|}} e^{-(1-a)\|\theta_2\|_1}.$$

When G is a-Lipschitz, the previous bounds on $a_{|S|}$ can be written precisely as $C + |S|\log(2/(1+a)) \leq \log a_{|S|}^{-1} \leq C' + |S|\log 2$. Deduce that (2.7) in [1] is satisfied with γ_{S_2} a product of $|S_2|$ univariate Laplace densities with scale parameter $1-a$ and $m_1 = (1+a)/(1-a)$.

6. Results on the posterior coordinate-wise median. First, we show that under the conditions of Theorem 2.1 or 2.4 in [1], the posterior coordinate-wise median has at most a constant times p_n non-zero coefficients, with high probability. We call this dimension reduction property of the posterior coordinate-wise median.

Second, under a slightly faster decrease for the prior on dimension, namely for complexity priors satisfying (2.8) in [1], we show that the posterior coordinate-wise is rate-minimax over $\ell_0[p_n]$ for any d_q-risk with $0 < q \leq 2$ (as can be seen from the proof of Lemma 6.1 below, the exact behavior (2.8) in [1] can be relaxed to an approximation of (2.8) in [1].)

For any $1 \leq i \leq n$, let $m_i(X)$ denote the marginal posterior median on the coordinate i. The vector $m(X) = (m_1(X), \ldots, m_n(X))$ is the posterior coordinate-wise median at stake. Let $\Sigma_n(X)$ denote the support of this vector, that is

$$\Sigma_n(X) := \{i \in \{1,\ldots,n\},\ m_i(X) \neq 0\}.$$

Let $\mathbb{P}_{P_{n,\theta_0}}$ denote the probability under P_{n,θ_0},

LEMMA 6.1 (Dimension reduction for the median). *Under the conditions of Theorem 2.1 or 2.4 in [1], for M large enough, as $n \to +\infty$,*

$$\mathbb{P}_{P_{n,\theta_0}}(|\Sigma_n(X)| > 4Mp_n) = o(1).$$

For the complexity prior defined by (2.8) in [1] with a large enough and $b > e$, we also have the more precise estimate

$$\mathbb{P}_{P_{n,\theta_0}}(|\Sigma_n(X)| > 4Mp_n) = o(p_n/n). \tag{6.1}$$

PROOF. The posterior probability that the posterior selects (in the sense that it picks non-zero coefficients from) a subset S of size larger than Mp_n tends to 0 as $n \to +\infty$. This follows from the proof of Theorem 2.1 in [1]. For M large enough constant, as $n \to +\infty$,

$$\sup_{\theta_0 \in \ell_0[p_n]} P_{n,\theta_0} \Pi_n(\theta : \ |S_\theta| > Mp_n | X) = o(1). \tag{6.2}$$

Below we see that the $o(1)$ can be refined to some rate of convergence to 0.

Let us denote by $|\mathcal{M}|$ the cardinality of the set of indexes $\mathcal{M}$ and $E^\Pi[\cdot|\, X]$ the expectation under the posterior distribution. Also, S in the sequel denotes the (random) set of non-zero coordinates selected by the posterior. Given X, for any index j in $\Sigma_n(X)$, we have $\Pi[j \in S|\, X] := \Pi[\theta_j \neq 0|\, X] \geq 1/2$, by definition of the median. Then

$$E^\Pi[\ |S \cap \Sigma_n(X)|\ |\, X] = E^\Pi[\sum_{j \in \Sigma_n(X)} 1_{j \in S}\ |\, X] \geq \frac{|\Sigma_n(X)|}{2}.$$

We also have that

$$\begin{aligned}
&E^\Pi[\ |S \cap \Sigma_n(X)|\ |\, X] \\
&= E^\Pi\left[\ |S \cap \Sigma_n(X)| \{1_{|S \cap \Sigma_n(X)| \leq |\Sigma_n(X)|/4} + 1_{|S \cap \Sigma_n(X)| > |\Sigma_n(X)|/4}\}\ |\, X\right] \\
&\leq |\Sigma_n(X)|/4 + |\Sigma_n(X)| \Pi\left[|S \cap \Sigma_n(X)| > |\Sigma_n(X)|/4\ |\, X\right]
\end{aligned}$$

Putting together the two previous bounds, one obtains, if $|\Sigma_n(X)| > 0$,

$$\Pi\Big[|S \cap \Sigma_n(X)| > |\Sigma_n(X)|/4\ |\, X\Big] \geq 1/4.$$

Using this and (6.2) we obtain,

$$\begin{aligned}
&\mathbb{P}_{P_{n,\theta_0}}(|\Sigma_n(X)| > 4Mp_n) \\
&\leq \mathbb{P}_{P_{n,\theta_0}}\Big(\Pi\left[|S \cap \Sigma_n(X)| > Mp_n\ |\, X\right] \geq 1/4\Big) \\
&\leq 4P_{n,\theta_0}\Pi\left[|S| > Mp_n\ |\, X\right] = o(1).
\end{aligned}$$

Let us see how to refine this into a $o(p_n/n)$. To obtain (6.2), one uses Proposition 4.1 in [1] with $A = Cp_n$ and C large enough. It holds $\Pi_n(\theta : \ |S_\theta| > 2Cp_n|X) \leq \Pi_n(\theta : \ |S_\theta \cap S_{\theta_0}^c| > Cp_n|X)$. Proposition 4.1 thus gives

$$P_{n,\theta_0}\Pi_n(\theta : \ |S_\theta| > 2Cp_n|X) \lesssim \sum_{p=Cp_n}^{n-p_n} m_1^p \max_{0 \leq k \leq p_n} \frac{\pi_{n,k}(p)}{\pi_{n,k}(0)}.$$

For any prior on dimension with strict exponential decrease, it follows from the proof of Lemma 4.1 in [1] that for C large enough, the previous display is bounded by e^{-cp_n} for some $c > 0$. This is a $o(p_n/n)$ if we assume the mild condition $p_n \geq d \log n$ for some constant $d > 0$. This condition is in fact not needed under (2.8) in [1], as we see below.

In the case of the complexity prior (2.8) in [1], we proceed as in the proof of Theorem 2.7 in Section 3 above. This enables, if both C and a are larger than some universal constants, to bound from above the last display by

$$e^{cp_n} \sum_{p=Cp_n}^{n-p_n} \exp\Big(- ap \log \frac{ne}{m_1^{1/a}(p_n+p)}\Big) \leq e^{cp_n} \sum_{p=Cp_n}^{n-p_n} e^{-ap\log(n\sqrt{e}/p)}.$$

This last bound is a $o(e^{-\log(n/p_n)})$ as $n \to +\infty$ (for instance split $a = (a-1)+1$ and use the convergence of the geometric series. In fact one can get a more precise bound but the previous one is enough for our needs). So, working with the complexity prior (6.1) always holds (for this prior we therefore do not need to assume that $p_n \gtrsim \log n$). $\square$

Now we can show that the posterior coordinatewise median is rate-minimax for the d_q-distance over $\ell_0[p_n]$ *for any* $0 < q \leq 2$.

THEOREM 6.1 (Minimaxity of the coordinate-wise median). *Suppose that the prior on dimension is given by (2.8) in [1] with parameters $b > e$ and $a > 1$ large enough, and that the prior densities g_S are of the product form $\otimes_S g$ for g square-integrable and satisfying (2.3) in [1]. Then for any $0 < q \leq 2$,*

$$\sup_{\theta_0 \in \ell_0[p_n]} P_{n,\theta_0} d_q(m(X), \theta_0) \lesssim r_{n,q}^*.$$

Remark. One can relax the condition $g_S = \otimes_S g$ by assuming only conditions (2.5) up to (2.7) in [1].

PROOF. Since the conditions of Theorem 2.5 in [1] are satisfied, by the consequence written below Theorem 2.5 in [1], we have

$$P_{n,\theta_0} \int \|\theta - \theta_0\|_2^2 d\Pi(\theta|\, X) \lesssim r_{n,2}^*.$$

The case $q = 2$. For any real a and any $i = 1, \ldots, n$, it holds

$$(m_i(X) - a)^2 \leq 2 \int (\theta_i - a)^2 d\Pi(\theta|\, X).$$

For $m_i(X) \leq a$, this is obtained by bounding the integral from below by restricting it to the set $\{\theta_i \leq m_i(X)\}$ and using $\Pi(\theta_i \leq m_i(X) | X) \geq 1/2$ by definition of the median. The case $m_i(X) \leq a$ is similar using the indicator $\{\theta_i \geq m_i(X)\}$. Applying the previous inequality with the choice $a = \theta_{0,i}$, summing over i and finally taking expectations leads to,

$$P_{n,\theta_0} \|m(X) - \theta_0\|_2^2 \leq 2P_{n,\theta_0} \int \|\theta - \theta_0\|_2^2 d\Pi(\theta | X) \lesssim r_{n,2}^*.$$

uniformly over $\ell_0[p_n]$.

The case $0 < q < 2$. From the case $q = 2$ we know that

(6.3)
$$P_{n,\theta_0} \|m(X) - \theta_0\|_2^2 = P_{n,\theta_0} \sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 + P_{n,\theta_0} \sum_{i \notin \Sigma_n(X)} \theta_{0,i}^2 \lesssim r_{n,2}^*$$

Similarly, for the d_q-distance,

$$\begin{aligned} P_{n,\theta_0} d_q(m(X), \theta_0) =& P_{n,\theta_0} \sum_{i \in \Sigma_n(X)} |m_i(X) - \theta_{0,i}|^q + P_{n,\theta_0} \sum_{i \notin \Sigma_n(X)} |\theta_{0,i}|^q \\ =& (I) + (II). \end{aligned}$$

To bound the term (I), note that, using Hölder inequality,

$$\begin{aligned} & P_{n,\theta_0} \left[\sum_{i \in \Sigma_n(X)} |m_i(X) - \theta_{0,i}|^q 1_{|\Sigma_n(X)| \leq 4Mp_n} \right] \\ & \leq P_{n,\theta_0} \left[\Big(\sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 \Big)^{q/2} (4Mp_n)^{1-q/2} \right] \\ & \leq (4Mp_n)^{1-q/2} \left[P_{n,\theta_0} \sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 \right]^{q/2}, \end{aligned}$$

where the last line follows from Jensen's inequality and the concavity of the map $u \to u^{q/2}$ on the positive real line when $q < 2$. Using (6.3) it follows that this expression is bounded from above by $(4Mp_n)^{1-q/2} {r_{n,2}^*}^{q/2}$ which is nothing but the minimax rate $r_{n,q}^*$ for the d_q-distance over $\ell_0[p_n]$, up to some multiplicative constant.

The part of (I) involving large cardinalities of $\Sigma_n(X)$ is treated first similarly, by simply bounding the number of terms in the sum by n,

$$\begin{aligned}
&P_{n,\theta_0} \sum_{i \in \Sigma_n(X)} |m_i(X) - \theta_{0,i}|^q 1_{|\Sigma_n(X)|>4Mp_n} \\
&\leq P_{n,\theta_0} \left[\Big(\sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 \Big)^{q/2} n^{1-q/2} 1_{|\Sigma_n(X)|>4Mp_n} \right].
\end{aligned}$$

By Hölder inequality again, this time with respect to P_{n,θ_0}, one obtains the bound

$$n^{1-q/2} \left[P_{n,\theta_0} \sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 \right]^{q/2} \left[\mathbb{P}_{P_{n,\theta_0}}(|\Sigma_n(X)| > 4Mp_n) \right]^{1-q/2}.$$

Combine this with (6.1)-(6.3) to obtain that this term is $n^{1-q/2} {r^*_{n,2}}^{q/2} o((p_n/n)^{1-q/2})$. Thus this term is a $o(r^*_{n,q})$ as $n \to +\infty$.

It remains to bound (II). This is done by noticing that the number of non-zero terms in the sum over $i \notin \Sigma_n(X)$ is certainly at most p_n, since θ_0 belongs to $\ell_0[p_n]$. Thus using Hölder inequality,

$$\begin{aligned}
(II) &= P_{n,\theta_0} \sum_{i \notin \Sigma_n(X)} |\theta_{0,i}|^q \\
&\leq P_{n,\theta_0} \left[\Big(\sum_{i \notin \Sigma_n(X)} \theta_{0,i}^2 \Big)^{q/2} p_n^{1-q/2} \right] \\
&\leq p_n^{1-q/2} \left[P_{n,\theta_0} \sum_{i \notin \Sigma_n(X)} \theta_{0,i}^2 \right]^{q/2} \lesssim p_n^{1-q/2} (r^*_{n,2})^{q/2} \lesssim r^*_{n,q}.
\end{aligned}$$

using Jensen's inequality as above and (6.3). This concludes the proof for the case $0 < q < 2$ and the above claim is proved. □

REFERENCES

[1] Castillo, I., and van der Vaart, A. W. Needles and Straw in a Haystack: Posterior concentration for possibly sparse sequences. *Ann. Statist.*.

[2] Ghosal, S., Ghosh, J. K., and van der Vaart, A. W. Convergence rates of posterior distributions. *Ann. Statist. 28*, 2 (2000), 500–531.

[3] Ghosal, S., and van der Vaart, A. Convergence rates of posterior distributions for non-i.i.d. observations. *Ann. Statist. 35*, 1 (2007), 192–223.

[4] Ghosal, S., and van der Vaart, A. Posterior convergence rates of Dirichlet mixtures at smooth densities. *Ann. Statist. 35*, 2 (2007), 697–723.

[5] Impens, C. Stirling's series made easy. *Amer. Math. Monthly 110*, 8 (2003), 730–735.

[6] POLLARD, D. *Empirical processes: theory and applications.* NSF-CBMS Regional Conference Series in Probability and Statistics, 2. Institute of Mathematical Statistics, Hayward, CA, 1990.

CNRS & LABORATOIRE DE PROBABILITÉS ET MODÈLES ALÉATOIRES, UNIVERSITÉS PARIS VI & VII,
E-MAIL: ismael.castillo@upmc.fr
DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCES, VU UNIVERSITY AMSTERDAM,
E-MAIL: aad@cs.vu.nl