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 \( b_p := \binom{n-p}{p}/\binom{n}{p+k} \) satisfy

\[
\frac{b_p}{b_{p-1}} = \frac{n-p-n+p+k}{n-p-k+1} = \left(1 - \frac{p_n-k}{n-p-k+1}\right) \left(1 + \frac{k}{p}\right).
\]

Hence

\[
b_p \leq b_{p-1}(1+C_1^{-1}) \quad \text{for} \quad p \geq C_1p_n \quad \text{and any} \quad k \leq p_n,
\]

whence the numbers \( a_p = \pi_n(p+k)b_p \) satisfy

\[
a_p \leq a_{p-1}D(1+C_1^{-1}) \leq a_{C_1p_n}F^{p-C_1p_n},
\]

for \( p \geq C_1p_n \) 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

\[
\sum_{p=C_1p_n}^{n-p_n} \frac{p\alpha_p}{\sum_{p=0}^{n-p_n} \alpha_p} \leq C_1p_n + \frac{\sum_{p=C_1p_n}^{n-p_n} (p-C_1p_n)a_{C_1p_n}F^{p-C_1p_n}}{a_{C_1p_n}} \lesssim C_1p_n + 1.
\]

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 \( r_k \).

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

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 $\|\theta - \theta_0\| \leq \|\theta_1\|/2$, then $\|\theta\| \geq \|\theta_1\|/2$ and hence $\langle \theta, \theta_1 \rangle = (\|\theta\|^2 + \|\theta_1\|^2 - \|\theta - \theta_1\|^2)/2 \geq \|\theta_1\|^2/2$. Therefore, the test $\phi = 1_{\theta^T X > D\|\theta_1\|}$ satisfies

$$P_{n,\theta_0}\phi = 1 - \Phi(D),$$

$$P_{n,\theta}(0 - \phi) = \Phi((D\|\theta_1\| - \langle \theta, \theta_1 \rangle)/\|\theta_1\|) \leq \Phi(D - \rho),$$

for $\rho = \|\theta_1\|/2$. The infimum over $D$ of $\alpha(1 - \Phi(D)) + \beta\Phi(D - \rho)$ 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 $\alpha(1 - \Phi(D))$ 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 $\alpha$.

Proof of Lemma 5.2. To prove the first assertion, we apply Jensen’s inequality to see that

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

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

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 $\sqrt{2\pi} < \Gamma(x + 1)/(x/e)^x \sqrt{x} < \sqrt{2\pi} e^{1/(12x)}$, which holds for any $x > 0$, see e.g. [5]. \qed

Proof of Lemma 5.7. By Lemma 5.2 in [1], there exist events $A_n$ with $\Pr(A_n^c) \leq e^{-r_n^2/8} \to 0$ such that $\int p_{n,\theta}/p_{n,\theta_0} d\Pi_n(\theta) \geq e^{-r_n^2/2} \Pi_n(\theta : \|\theta - \theta_0\| < r_n)$ By the definition of the posterior as a quotient, it follows that

$$P_{n,\theta_0}\Pi_n(\theta : \|\theta - \theta_0\| < s_n | X) \leq \frac{P_{n,\theta_0}\int_{\|\theta - \theta_0\| < s_n} p_{n,\theta}/p_{n,\theta_0} d\Pi_n(\theta)}{e^{-r_n^2/2} \Pi_n(\theta : \|\theta - \theta_0\| < r_n)} \leq \frac{\Pi_n(\theta : \|\theta - \theta_0\| < s_n)}{e^{-r_n^2/2} \Pi_n(\theta : \|\theta - \theta_0\| < r_n)}.$$
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

\[
\frac{v_p}{v_{p_n}} = \frac{n^{p-p_n}}{\Gamma(p/2 + 1)} \leq c_1 \sqrt{\frac{2\pi (p_n/2) (p_n/2)^{p_n/2}}{2\pi (p/2) (p/2)^{p/2}}}.
\]

Thus for any \( 1 \leq p \leq p_n \),

\[
\frac{v_p}{v_{p_n}} \leq c_1 \exp \left( \frac{p_n - p}{2} \log \frac{p_n}{2} - \frac{p}{2} \log \frac{p}{2} \right).
\]

Since the function \( p \mapsto (p/2) \log(p_n/p) \) is inferior to \( p_n/(2e) \) for any \( 1 \leq p \leq p_n \), we conclude, using the assumption on \( \alpha_n \), 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 \( \pi_n \) 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 \( \phi_p \) based on \( X \sim N_n(\theta, I) \) with, for any \( r > 1 \) and every integer \( j \geq 1 \),

(2.1) \[
P_{n,\theta_0} \phi_p \leq 66\sqrt{\frac{\beta_p}{\alpha_p}} \left( \frac{n}{p} \right) 48^p \exp(-r^2/32),
\]

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

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

(2.3) \[
C_{j,p}(r) = \{ \theta \in \Theta_p : jr < \|\theta - \theta_0\| \leq (j + 1)r \}.
\]

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 \( \phi_p \) 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 \( \theta \) in such a ball with center \( \theta_1 \) satisfies \( \| \theta - \theta_1 \| \leq jr/2 \leq \| \theta_0 - \theta_1 \|/2 \), since \( \theta_1 \in C_{j,p}(r) \). Hence each test satisfies the inequalities of Lemma 5.1 in [1].

If \( S_0 \subset S \), 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

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

(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

\[
P_{n,\theta_0} \phi_p \leq 2 \sqrt{\frac{\beta_p}{\alpha_p}} \left( \frac{n}{p} \right) 48^p \frac{e^{-r^2/32}}{1 - e^{-r^2/32}}.
\]

\[
\leq \frac{2}{1 - e^{-1/32}} \sqrt{\frac{\beta_p}{\alpha_p}} \left( \frac{n}{p} \right) 48^p e^{-r^2/32}.
\]

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

\[
e^{-2c_1|S|-c_1|S_0^c|} \leq \frac{g_S(\theta)}{g_{S_0}(\pi_{S_0} \theta_0)} \leq e^{2c_1|S|+c_1|S_0^c|+r^2/64}.
\]

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

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

Next we use condition (2.6) in [1] twice, first with \( S' = S_0 \cap S \subset S \) and next with \( S'' = S_0 \cap S \subset S_0 \), to see that

\[
| \log \frac{g_S(\pi_S \theta_0)}{g_{S_0 \cap S}(\pi_{S_0 \cap S} \theta_0)} | \leq c_1|S|,
\]

\[
| \log \frac{g_{S_0 \cap S}(\pi_{S_0 \cap S} \theta_0)}{g_{S_0}(\pi_{S_0} \theta_0)} | \leq c_1|S_0| + \| \pi_{S_0} \theta_0 \|^2 / 64.
\]

The lemma follows upon combining these three inequalities with the observation that \( \| \pi_{S_0} \theta_0 \|^2 + \| \pi_{S_0^c} \theta_0 \|^2 = \| (\theta, 0_{S^c}) - \theta_0 \|^2 \).
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 \( p_n = |S_{\theta_0}| \),

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

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

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

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

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

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 \( p = |S| \))

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

by Lemma 2.2, since \( ||(\theta, 0_{S'c}) - \theta_0|| = ||\pi_{S \theta_0}|| \). 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 \((n-p_n)\) subsets of size \( |S| = p \) that contain \( S_{\theta_0} \).

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 \( I_C(q) \leq I_C([q]) / C^{[q]-q} \).

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

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

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 \( M \geq 1 \),
\[
\sum_{j \geq M} (jD)^p e^{-j^2D^2} \leq 2(2p + 1)e^{-\left(\frac{\pi}{4} \sqrt{M^2D^2}\right)} (\frac{p}{2} \sqrt{M^2D^2})^{p/2}.
\]

PROOF. The function \( \psi_p : u \to (uD)^p e^{-u^2D^2} \) 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
\[
\sum_{j \geq M} (jD)^p e^{-j^2D^2} \leq (MD)^p e^{-M^2D^2} + \int_M^{\infty} (xD)^p e^{-x^2D^2} \, dx.
\]

It suffices to control the integral, with \( C = M^2D^2 \)
\[
\int_M^{\infty} (xD)^p e^{-x^2D^2} \, dx = \frac{1}{2D} I_C\left(\frac{p - 1}{2}\right).
\]

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

If \( M \leq u^* \), we split the sum in two parts corresponding to \( M \leq j \leq \lfloor u^* \rfloor \) and \( j \geq \lceil u^* \rceil + 1 \). 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
\[
\sum_{j \geq M} (jD)^p e^{-j^2D^2} \leq \lfloor u^* \rfloor - M + 2 \bar{\psi}_p + \int_{u^*}^{\infty} (xD)^p e^{-x^2D^2} \, dx.
\]

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 \( 3\sqrt{p/2} \bar{\psi}_p \leq 3p\bar{\psi}_p \). \( \square \)

Given any prior \( \pi_n \) on dimension, \( p_n \) an integer between 0 and \( n \), setting \( h = e^{5+2c_1} \) and \( d = \sqrt{2e\pi} \), define
\[
C_n(r; \pi_n, p_n) = 20e^{c_1p_n} \sum_{p=0}^{n} \sqrt{\pi_n(p)} (hn/p)^{p/2} (1 \vee r^2/p)^{p/4} \left( \sum_{p=p_n}^{n} \frac{(n-p_n)}{(p)} \pi_n(p) d^p(r^2/p)^{p/2} \right)^{1/2}.
\]

PROPOSITION 2.1. If the densities \( f_S \) have a finite second moment and satisfy (2.5)-(2.6) in [1], then for any \( 1 \leq p_n \leq n \) and any \( r \geq 1 \),
\[
\sup_{\theta_0 \in \ell_0[p_n]} P_{n, \theta_0} \Pi_n(\theta : ||\theta - \theta_0|| > 10r | X) \leq e^{-r^2/9} (C_n(r, \pi_n, p_n) + 1).
\]
Proof. For any \( p \in \mathbb{N} \) fix a test \( \phi_p \) as given in Lemma 2.1, with \( \alpha_p, \beta_p \) to be chosen later. Then, for \( \Theta_p = \{ \theta \in \mathbb{R}^n : |S_\theta| = p \} \) and any events \( \mathcal{A}_n \),
\[
P_{n,0_0} \Pi_n(\theta : \| \theta - \theta_0 \| > 10r \mid X)
= P_{n,0_0} \sum_{p=1}^{n} \Pi_n(\theta \in \Theta_p : \| \theta - \theta_0 \| > 10r \mid X)
= \sum_{p=0}^{n} P_{n,0_0} \phi_p + \sum_{p=1}^{n} P_{n,0_0} \Pi_n(\theta \in \Theta_p : \| \theta - \theta_0 \| > 10r \mid X) (1 - \phi_p) 1_{A_n}
= A + B + C.
\]

Using (2.1) and that \( (\begin{array}{c} n \\ p \end{array}) \leq \sum_{j=0}^{p} \left( \begin{array}{c} n \\ j \end{array} \right) \leq (ne/p)^p \), we obtain, for \( 10r > 1 \),
\[
A \leq a \sum_{p=1}^{n} \sqrt{\frac{\beta_p}{\alpha_p}} \left( \frac{n}{p} \right)^{48p} e^{-r^2/32} \leq a \sum_{p=0}^{n} \sqrt{\frac{\beta_p}{\alpha_p}} e^{5p} p \log n/p - r^2/32,
\]
with \( a = 66 \). In view of the second assertion of Lemma 5.2 in [1] applied with \( r \) equal to \( \eta r \) for a small number \( \eta \), to be determined later, there exist events \( \mathcal{A}_n \) such that
\[
C = P_{n,0_0} (A^c_n) \leq e^{-\eta r^2/8},
\]
while on the event \( \mathcal{A}_n \),
\[
\int \frac{P_{n,\theta}}{P_{n,0_0}} d\Pi_n(\theta) \geq e^{-\eta r^2} \Pi_n(\theta : \| \theta - \theta_0 \| < \eta r).
\]
Therefore, for \( C_{j,p}(r) \) as defined in (2.3), in view of Fubini’s theorem,
\[
B = \sum_{p=0}^{n} \sum_{j \geq 10} P_{n,0_0} \left[ (1 - \phi_p) 1_{A_n} \frac{\int C_{j,p}(r) P_{n,\theta} d\Pi_n(\theta)}{\int P_{n,\theta} d\Pi_n(\theta)} \right]
\leq e^{r^2} \sum_{p=0}^{n} \sum_{j \geq 10} \frac{\int C_{j,p}(r) P_{n,\theta} (1 - \phi_p) d\Pi_n(\theta)}{\Pi_n(\theta : \| \theta - \theta_0 \| < \eta r)}
\leq 2e^{r^2} \sum_{p=0}^{n} \sqrt{\frac{\alpha_p}{\beta_p}} \sum_{j \geq 10} \Pi_n(C_{j,p}(r)) e^{-j r^2/32},
\]
where the last inequality follows from inequality (2.2). Here, with \( v_p \) the
volume of the $p$-dimensional unit ball, 
\[
\Pi_n(C_{j,p}(r)) = \frac{\pi_n(p)}{p^p} \sum_{S: |S|=p} G_S(\theta \in \mathbb{R}^S : jr < \|\theta, 0_{S^c}\| - \theta_0 \leq (j+1)r) 
\leq \frac{\pi_n(p)}{p^p} \sum_{S: |S|=p} v_p(j + 1)^{p-r} \max_{\theta \in \mathbb{R}^S: \|\theta, 0_{S^c}\| - \theta_0 \leq (j+1)r} g_S(\theta) 
\leq \pi_n(p) v_p (2jr)^p g_{S_{\theta_0}}(\pi_{S_{\theta_0}} \theta_0) e^{2c_1p + c_1n + (j+1)^2r^2/64},
\]
by Lemma 2.2. Because $j^2/32 - (j+1)^2/64 \geq j^2/100$ for $j \geq 10$, we obtain
\[
\sum_{j \geq 10} \Pi_n(C_{j,p}(r)) e^{-j^2r^2/32} 
\leq \sum_{j \geq 10} \pi_n(p) v_p (jr/10)^p 20^p g_{S_{\theta_0}}(\pi_{S_{\theta_0}} \theta_0) e^{2c_1p + c_1n - j^2r^2/100} 
\leq 4d_2 \pi_n(p) p^{-p/2} e^{c_1n + 2c_1p} g_{S_{\theta_0}}(\pi_{S_{\theta_0}} \theta_0) F^p e^{-r^2\left(\frac{p}{2} \lor r\right)^{p/2}},
\]
where $F = 20\sqrt{2e\pi} \leq e^5$, 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
\[
B \leq \frac{\sum_{p=0}^n \sqrt{\alpha_p/\beta_p} \pi_n(p) e^{2c_1(p_n+p)} F^p (1 \lor (r^2/p))^{p/2}}{\sum_{p=p_n}^n \left(\frac{n-p_n}{n}\right) \pi_n(p) d^{p/2}(r^2/p)^{p/2}} e^{-r^2/9} b, 
\]
where $b = 4d_2/d_1 \leq 5$. We now balance $A$ and $B$ by choosing
\[
\alpha_p = e^{p \log n/p} a, \\
\beta_p = \pi_n(p) e^{2c_1(p_n+p)} F^p (1 \lor (r^2/p))^{p/2} \left(\sum_{p=p_n}^n \left(\frac{n-p_n}{n}\right) \pi_n(p) (d^{p/2}/p)^{p/2}\right) b.
\]
With this choice, for $h = e^{5+2c_1}$,
\[
A + B \leq e^{c_1p_n} \sum_{p=0}^n \sqrt{\pi_n(p) (hn/p)^{p/2} (1 \lor (r^2/p))^{p/4}} \left(\sum_{p=p_n}^n \left(\frac{n-p_n}{n}\right) \pi_n(p) (d^{p/2}/p)^{p/2}\right)^{1/2} e^{-r^2/9} \sqrt{ab}.
\]
Together with the bound (2.4) this concludes the proof of the first assertion. 
\[\square\]
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 \( b \geq e^{2h} \), 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 \( p_n \)-th 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 \( r \geq r_n \),
\[
e^{-r^2/9}(C_n(r; \pi_n, p_n) + 1) \lesssim e^{-r^2/10 + 2r_n^2 + c_1 p_n}.
\]
For \( r \) replaced by \( r + 4.5r_n \) 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 \( \theta_0 \) into \( \ell_0[p_n^*] \) by \( \theta_1 \), we have, for \( |S_0| \leq (A - 1)p_n^* \),
\[
d_q(\theta, \theta_0) \lesssim d_q(\theta, \theta_1) + d_q(\theta_1, \theta_0) \lesssim (Ap_n^*)^{1-q/2} \| \theta - \theta_1 \|^q + d_q(\theta_1, \theta_0)
\lesssim (Ap_n^*)^{1-q/2} \| \theta - \theta_0 \|^q + (Ap_n^*)^{1-q/2} \| \theta_0 - \theta_1 \|^q + d_q(\theta_1, \theta_0).
\]
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 \( \theta_0 \) in absolute value. We take \( B = \{ \theta : |S_0 \cap S_0^c| \geq Rp_n^* \} \). The denominator in (4.2) in [1] is bounded from below as below (4.2) in [1],
\[
(3.1) \quad \Pi_n(B \mid X, \theta_{S_0} = \tilde{\theta}_1) \leq \int_B \frac{p_{\bar{n}_2, \tilde{\theta}_2}(X_{S_0^c})}{p_{n_2, \bar{\theta}_{S_0}^c}} \frac{d\Pi_n(\theta_2 \mid \tilde{\theta}_1)}{\Pi_n(\theta_2 = 0 \mid \tilde{\theta}_1)},
\]
where now \( \bar{n}_2 = n - p_n^* \). If \( S_2 \) are the indices of the nonzero coordinates of \( \tilde{\theta}_2 \in \mathbb{R}^{S_0^c} \), \( \theta_2 \) the vector of their values and \( n_2 = |S_2| \), then
\[
\frac{p_{n_2, \bar{\theta}_2}(X_{S_0^c})}{p_{n_2, \bar{\theta}_2, \theta_{S_0^c}}(X_{S_0^c})} = \frac{p_{n_2, \bar{\theta}_2}(X_{S_0^c})}{p_{n_2, \bar{\theta}_2, \theta_{S_0}^c, \theta_{S_0^c}}(X_{S_0^c})},
\]

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 $E_n^2(S_2) = 3|S_2| \log((n - p_n^*)e/|S_2|)$,

$$A = \bigcap_{p = R_{p_n^*}^*} \bigcap_{|S_2| = p} \{(X_{S_2} - \theta_{0,S_2})^T \theta_{0,S_2} \leq \|\theta_{0,S_2}\| E_n(S_2)\}.$$ 

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

$$\|\theta_{0,S_2}\|^2 \leq \|\theta_{0,S_0^*}\|^2 \lesssim \mu_{n,s,2}^*.$$ 

Using the inequality $ab \leq b^2 + a^2/4$, deduce that on the event $A$,

$$\frac{p_{n_2,\theta_0,S_2}}{p_{n_2,\theta_0}} (X_{S_2}) \leq e^{C\mu_{n,s,2}^* + \frac{E_n(S_2)}{4}}.$$

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

$$\sum_{p = R_{p_n^*}^*} \sum_{|S_2| = p} \pi_{n,k}(p) \frac{m_1^{p+p_n}}{\pi_{n,k}(0)} e^{C\mu_{n,s,2}^* + \frac{E_n(S_2)}{4}} \int \frac{p_{n_2,\theta_1}}{p_{n_2,\theta_0,S_2}} (X_{S_2}) \gamma_{S_2}(\theta_2) \, d\theta_2.$$

Taking the maximum over $k$ inside and next the expectation under $P_{n,\theta_0}$,

$$P_{n,\theta_0}(B|X)1_A \leq e^{C\mu_{n,s,2}^*} \sum_{p = R_{p_n^*}^*} \max_{0 \leq k \leq p_n^*} \frac{\pi_{n,k}(p)}{\pi_{n,k}(0)} C^{p+p_n} e^{\frac{3}{4}p \log \left(\frac{n-p_n^*}{p}\right)}.$$

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 $p_n^*$,

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

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 $c$,

$$P_{n,\theta_0}(B|X)1_A \leq e^{C\mu_{n,s,2}^*} \sum_{p = R_{p_n^*}^*} e^{-\frac{ap \log\left(\frac{en}{p+p_n^*}\right)}{p+p_n^*}} \lesssim e^{-c\mu_{n,s,2}^*}.$$

The expectation over the complement of $A$ is bounded above by

$$P_{n,\theta_0}(A^c) \leq \sum_{p = R_{p_n^*}^*} \left(\frac{n - p_n^*}{p}\right) e^{-3p \log((n - p_n^*)e/p)/2} \leq \sum_{p = R_{p_n^*}^*} e^{-[3/2 - 1]p \log((n - p_n^*)e/p)} \lesssim ((n - p_n^*)e/p_n^*)^{R_{p_n^*}^*}. \qed$$
4. Proof of Theorem 2.9 in [1]. Let us define \( \alpha_n = c \varepsilon_n \) and \( \gamma_n = d \varepsilon_n \) 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 \( \theta_0 \) 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,
\[
\theta_{0,k} = \begin{cases} 
M n \varepsilon_n^2 & \text{if } k \in \{1, \ldots, d_{1,n}\} = S_1 \\
4n \alpha_n^2/(d_{2,n} - d_{1,n}) & \text{if } k \in \{d_{1,n} + 1, \ldots, d_{3,n}\}
\end{cases}
\]
Note that if the support of \( \theta \) is \( S \) and if \( \| \theta - \theta_0 \|^2 \leq n \alpha_n^2 \), then \( S \) belongs to the set \( Q \) of supports defined by
\[
Q = \{ S : S \supset S_1, |S \cap \{d_{1,n} + 1, \ldots, d_{3,n}\}| \geq d_{2,n} - d_{1,n} \}.
\]
In particular, for any \( S \in Q \), 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
\[
Q_n \leq \sum_{k=d_{2,n}}^{n} \sum_{S \in Q, |S|=k} \frac{\Pi_n(|\theta - \theta_0|^2 \leq n \alpha_n^2 |S| \pi_n(k)}{\Pi_n(|\theta - \theta_0|^2 \leq n \gamma_n^2 \pi_n(d_{1,n})}} \sum_{k=d_{2,n}}^{n} \sum_{S \in Q, |S|=k} \frac{\Pi_n(|\theta - \theta_0|^2 \leq n \alpha_n^2 |S| \pi_n(k)}{\Pi_n(|\theta - \pi_{S_1} \theta_0|^2 \leq n \gamma_n^2/2 |S_1| \pi_n(d_{1,n})}} \pi_n(d_{1,n})^2/n}
\]
Now note that if \( \gamma_n^2 \geq 2 \alpha_n^2 \) (which amounts to impose \( d \geq \sqrt{2c} \)), the first ratio of probabilities in the last display is bounded above by 1. Using the monotonicity of the prior \( \pi_n \), one obtains that \( \pi_n(k) \leq \pi_n(d_{2,n}) \) for any \( k \geq d_{2,n} \). Thus \( Q_n \leq \exp(-n \varepsilon_n^2) \leq \exp(-2n \gamma_n^2) \), 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 \( S' \subset S \subset \{1, \ldots, n\} \) 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 \( H(\theta) = \sum_{i=1}^{n} h(\theta_i) \). The assumptions on \( G \) imply that \( c \leq e^{G(\theta)} \leq Ce^{d t} \) for positive \( c, C, d \), so, since \( \| \theta \| \leq \| \theta \|_1 \),
\[
\int e^{H(\theta)-d\|\theta\|_1} d\theta \leq a_{|S|}^{-1} \int e^{H(\theta)-G(\|\theta\|)} d\theta \leq \int e^{H(\theta)} d\theta.
\]
The integrals on the far left and right sides factorize, where $e^h$ is integrable by assumption. Hence $|S| \lesssim \log a^{-1}|S| \lesssim |S|$, 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 $|G(||\theta||) - G(||\theta_1||)| \leq a\|\theta\| - ||\theta_1|| \leq a\|\theta_2\| \leq a\|\theta_2\|_1$. Thus

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

When $G$ is $a$-Lipschitz, the previous bounds on $a|S|$ can be written precisely as $C + |S| \log(2/(1 + a)) \leq \log a^{-1}|S| \leq C' + |S| \log 2$. Deduce that (2.7) in [1] is satisfied with $\gamma_{S_2}$ 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 $P_{P_n, \theta_0}$ denote the probability under $P_{n, \theta_0}$.

**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 $n \to +\infty$,

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

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

$$\mathbb{P}_{P_{n, \theta_0}}(|\Sigma_n(X)| > 4M p_n) = o(p_n/n).$$
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 $n \to +\infty$,

\begin{equation}
\sup_{\theta_0 \in \ell_{0[p_n]}} P_{n, \theta_0} \Pi_n(\theta : |S_\theta| > Mp_n | X) = o(1).
\end{equation}

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

Let us denote by $|M|$ the cardinality of the set of indexes $M$ and $E\Pi[|X]$ 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

\begin{equation}
E\Pi[|S \cap \Sigma_n(X)| | X] = E\Pi[\sum_{j \in S} 1_j | X] \geq \frac{|\Sigma_n(X)|}{2}.
\end{equation}

We also have that

\begin{align*}
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}\right] | X] \\
&\leq |\Sigma_n(X)|/4 + |\Sigma_n(X)|\Pi[|S \cap \Sigma_n(X)| > |\Sigma_n(X)|/4 | X]
\end{align*}

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

\begin{equation}
\Pi\left[|S \cap \Sigma_n(X)| > |\Sigma_n(X)|/4 | X\right] \geq 1/4.
\end{equation}

Using this and (6.2) we obtain,

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

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

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

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

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 $p_n \gtrsim \log n$).

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
$0 < q \leq 2$,

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

**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

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

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,

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

uniformly over $\ell_0[p_n]$.

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

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

Similarly, for the $d_q$-distance,

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

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

$$P_{n,\theta_0} \left[ \sum_{i \in \Sigma_n(X)} |m_i(X) - \theta_{0,i}|^{q/2} |\Sigma_n(X)|^{1/2} \right] \leq P_{n,\theta_0} \left[ \left( \sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 \right)^{q/2} (4M_{p_n})^{1-q/2} \right] \leq (4M_{p_n})^{1-q/2} \left[ P_{n,\theta_0} \sum_{i \in \Sigma_n(X)} (m_i(X) - \theta_{0,i})^2 \right]^{q/2},$$

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 $(4M_{p_n})^{1-q/2} r^*_n, q/2$ 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 sim-
ilarly, by simply bounding the number of terms in the sum by \( n \),

\[
P_{n, \theta_0} \sum_{i \in \Sigma_n(X)} |m_i(X) - \theta_{0,i}|^{q/1} |\Sigma_n(X)| > 4M_p
\]

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

By Hölder inequality again, this time with respect to \( P_{n, \theta_0} \), one obtains the bound

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

Combine this with (6.1)-(6.3) to obtain that this term is \( n^{1-q/2} r_{n}^{*} \). 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 \( \theta_0 \) belongs to \( \ell_0[p_n] \). Thus using Hölder inequality,

\[
(II) = P_{n, \theta_0} \sum_{i \notin \Sigma_n(X)} |\theta_{0,i}|^q
\]

\[
\leq P_{n, \theta_0} \left[ \sum_{i \notin \Sigma_n(X)} \theta_{0,i}^2 \right]^{q/2} n^{1-q/2}
\]

\[
\leq n^{1-q/2} \left[ P_{n, \theta_0} \sum_{i \notin \Sigma_n(X)} \theta_{0,i}^2 \right]^{q/2} \leq n^{1-q/2} (r_{n,2}^{*})^{q/2} \leq r_{n,q}^{*},
\]

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


