Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Functional Analysis 262 (2012) 4103–4119
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

Linear n-widths of diagonal matrices in the average and probabilistic settings ☆

Feng Dai [a], Heping Wang [b,*]

[a] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada*
[b] *School of Mathematical Sciences, Capital Normal University, Beijing 100048, China*

Received 25 September 2011; accepted 13 February 2012
Available online 18 February 2012
Communicated by G. Schechtman

Abstract

Exact values of the average linear n-widths with respect to the standard Gaussian measure on $\mathbb{R}^m$ are determined for diagonal matrices, and are applied to deduce several new results on linear n-widths in the average and probabilistic settings, including the sharp upper and lower estimates of the linear (n, δ)-widths of diagonal matrices.

Keywords: Diagonal matrices; p-Average linear n-widths; Probabilistic linear (n, δ)-widths; Gaussian measure

1. Introduction

Let ℓ_q^m denote the m-dimensional normed linear space of vectors $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ with norm

$$\|x\|_q \equiv \|x\|_{\ell_q^m} := \begin{cases} (\sum_{i=1}^m |x_i|^q)^{\frac{1}{q}}, & 1 \leqslant q < \infty, \\ \max_{1 \leqslant i \leqslant m} |x_i|, & q = \infty, \end{cases}$$

☆ The first author was partially supported by the NSERC Canada under grant RGPIN 311678-2010. The second author was supported by the National Natural Science Foundation of China (Project No. 10871132), the Beijing Natural Science Foundation (1102011), and Specialized Research Fund for the Doctoral Program of Higher Education (20091108110004).
* Corresponding author.
E-mail addresses: dfeng@math.ualberta.ca (F. Dai), wanghp@yahoo.cn (H. Wang).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2012.02.010

and let $B_q^m := \{x \in \mathbb{R}^m\colon \|x\|_q \leqslant 1\}$ denote the unit ball of ℓ_q^m. As usual, we identify $\mathbb{R}^m$ with the space ℓ_2^m, and use the notation $\langle x, y\rangle$ to denote the Euclidean inner product of $x, y \in \mathbb{R}^m$. Let $\mathbb{S}^{m-1} := \{x \in \mathbb{R}^m\colon \|x\|_2 = 1\}$ denote the unit sphere of $\mathbb{R}^m$ equipped with the usual rotation invariant measure $d\sigma_m(x)$ normalized by $\int_{\mathbb{S}^{m-1}} d\sigma_m(x) = 1$. Given $1 \leqslant q \leqslant \infty$, the classic linear n-width of a linear operator $T : \ell_2^m \to \ell_q^m$ is defined by

$$\lambda_n\big(T : \ell_2^m \to \ell_q^m\big) := \inf_{T_n} \sup_{x\in\mathbb{S}^{m-1}} \|Tx - T_n x\|_q, \tag{1.1}$$

with the infimum being taken over all linear operators $T_n : \ell_2^m \to \ell_q^m$ with rank $\leqslant n$. In this paper, we shall discuss the average linear n-widths, and the probabilistic linear (n, δ)-widths, whose definitions will be given below.

Let γ_m denote the standard Gaussian measure on $\mathbb{R}^m$ given by

$$\gamma_m(G) = (2\pi)^{-m/2} \int_G \exp\big(-\|x\|_2^2/2\big)\, dx, \quad \text{for each Borel subset } G \subset \mathbb{R}^m.$$

For $0 \leqslant n \leqslant m$, $0 < p < \infty$ and $1 \leqslant q \leqslant \infty$, the p-average linear n-width of a linear operator $T : \mathbb{R}^m \to \ell_q^m$ is defined by

$$\lambda_n^{(a)}\big(T : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p := \inf_{T_n}\bigg(\int_{\mathbb{R}^m} \|Tx - T_n x\|_q^p\, d\gamma_m(x)\bigg)^{1/p},$$

where the infimum is taken over all linear mappings $T_n : \mathbb{R}^m \to \ell_q^m$ with rank $\leqslant n$. Equivalently, the p-average linear n-width $\lambda_n^{(a)}(T : \mathbb{R}^m \to \ell_q^m, \gamma_m)_p$ can be expressed as

$$\lambda_n^{(a)}\big(T : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p := \inf_{T_n}\big(\mathbb{E}\|TX - T_n X\|_{\ell_q^m}^p\big)^{\frac{1}{p}},$$

where $X \sim N_m(0, I_m)_{\mathbb{R}^m}$ is an $\mathbb{R}^m$-valued Gaussian random vector with mean 0 and covariance matrix I_m, the m by m identity matrix.

A connection between the p-average linear n-width $\lambda_n^{(a)}(T : \mathbb{R}^m \to \ell_q^m, \gamma_m)_p$ and the classic linear n-width $\lambda_n(T : \ell_2^m \to \ell_q^m)$ can be seen as follows:

$$\lambda_n^{(a)}\big(T : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p = c(m, p) \inf_{T_n}\bigg(\int_{\mathbb{S}^{m-1}} \|Tx - T_n x\|_{\ell_q^m}^p\, d\sigma_m(x)\bigg)^{\frac{1}{p}} \tag{1.2}$$

$$\leqslant c(m, p)\lambda_n\big(T : \ell_2^m \to \ell_q^m\big), \tag{1.3}$$

where $c(m, p) = \sqrt{2}\big(\frac{\Gamma(\frac{m+p}{2})}{\Gamma(\frac{m}{2})}\big)^{\frac{1}{p}} \asymp \sqrt{m}$, and the constant of equivalence is independent of m. Here $A \asymp B$ means that there exists a constant $C > 0$, which is called the constant of equivalence, such that $C^{-1} \leqslant A/B \leqslant C$. Indeed, (1.2) can be verified by straightforward calculations, whereas (1.3) follows directly from (1.1). It turns out that in many cases, the quantity

$$\inf_{T_n}\Bigg(\int\limits_{\mathbb{S}^{m-1}} \|Tx - T_n x\|_{\ell_q^m}^p \, d\sigma_m(x)\Bigg)^{\frac{1}{p}}$$

is significantly smaller than the linear n-width

$$\lambda_n\big(T : \ell_2^m \to \ell_q^m\big) := \inf_{T_n} \sup_{x\in\mathbb{S}^{m-1}} \|Tx - T_n x\|_{\ell_q^m}.$$

(See, for instance, [4,3,8].) Results on classic linear widths can be found in the references [5,9] on widths, whereas known results on average linear widths can be found in [4,8,3,11].

For $\delta \in (0,1)$ and $1 \leqslant q \leqslant \infty$, the probabilistic linear (n,δ)-width of a linear mapping $T : \mathbb{R}^m \to \ell_q^m$ is defined by

$$\lambda_{n,\delta}\big(T : \mathbb{R}^m \to \ell_q^m, \gamma_m\big) = \inf_{G_\delta} \inf_{T_n} \sup_{x\in\mathbb{R}^m\setminus G_\delta} \|Tx - T_n x\|_q, \quad 1 \leqslant n \leqslant m,$$

where the first infimum is taken over all Borel subsets of $\mathbb{R}^m$ with $\gamma_m(G_\delta) \leqslant \delta$, and the second infimum is taken over all linear operators $T_n : \mathbb{R}^m \to \ell_q^m$ with rank $\leqslant n$. Thus,

$$\lambda_{n,\delta}\big(T : \mathbb{R}^m \to \ell_q^m, \gamma_m\big) < \lambda$$

if and only if there exists a linear operator $T_n : \mathbb{R}^m \to \ell_q^m$ with rank $\leqslant n$ such that with probability $\geqslant 1 - \delta$, one has

$$\|TX - T_n X\|_q < \lambda,$$

where $X \sim N_m(0, I_m)_{\mathbb{R}^m}$. For more information on the probabilistic linear widths and their connections to the average linear widths, we refer to [3,4,6–8] and the references therein.

The main purpose in this paper is to study the average linear n-widths and the probabilistic linear (n,δ)-widths for diagonal operators. Throughout the paper, we use the letter D to denote an $m \times m$ real diagonal matrix $\operatorname{diag}(d_1, \dots, d_m)$ with $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_m > 0$, and the letter D_n to denote the diagonal matrix $\operatorname{diag}(d_1, \dots, d_n, 0, \dots, 0)$ for $1 \leqslant n \leqslant m$. Moreover, $\{e_1, \dots, e_m\}$ denotes the standard orthonormal basis in $\mathbb{R}^m$:

$$e_1 = (1, 0, \dots, 0), \quad \dots, \quad e_m = (0, \dots, 0, 1).$$

In Section 2, we determine the exact values of the average linear n-widths $\lambda_n^{(a)}(D : \mathbb{R}^m \to \ell_q^m, \gamma_m)_q$ for all $1 \leqslant q < \infty$. Our result is new even in the case of the identity operator, where we prove that for $1 \leqslant q < \infty$, and $1 \leqslant n \leqslant m$,

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q = C(q) m^{1/q} (1 - n/m)^{\max(1/q, 1/2)}, \tag{1.4}$$

with

$$C(q) = \bigg(\pi^{-\frac{1}{2}} 2^{\frac{q}{2}} \Gamma\bigg(\frac{q+1}{2}\bigg)\bigg)^{1/q}. \tag{1.5}$$

As corollaries, we also deduce sharp upper and lower estimates of the average linear n-widths $\lambda_n^{(a)}(D:\mathbb{R}^m \to \ell_q^m, \gamma_m)_p$ for all $1 \leqslant q < \infty$ and $0 < p < \infty$ in Section 2. Section 3 is devoted to the sharp estimates of the average linear n-width $\lambda_n^{(a)}(I_m:\mathbb{R}^m \to \ell_\infty^m, \gamma_m)_p$ for $0 < p < \infty$. Finally, in Section 4, using our result on average linear n-widths, we obtain the sharp upper and lower estimates of the probabilistic linear (n,δ)-widths $\lambda_{n,\delta}(D:\mathbb{R}^m \to \ell_q^m, \gamma_m)$ for $1 \leqslant q \leqslant 2$, which improves previously known results even in the case of the identity matrix.

2. Exact values of the average linear n-widths of diagonal matrices

2.1. Main results and corollaries

For the classic linear n-widths, the following exact values are known in the case of $0 \leqslant n < m$ and $1 \leqslant q \leqslant 2$:

$$\lambda_n\big(D:\ell_2^m \to \ell_q^m\big) = \sup_{x\in\mathbb{S}^{m-1}} \|Dx - D_n x\|_q = \left(\sum_{k=n+1}^m d_k^r\right)^{1/r}, \tag{2.1}$$

where $1/r = 1/q - 1/2$ (see [9, Theorem 2.2]). For the case of $2 < q < \infty$, the exact value of $\lambda_n(D:\ell_2^m \to \ell_q^m)$ remains unknown.

In this section, we shall determine exact values of the average linear n-widths $\lambda_n^{(a)}(D:\mathbb{R}^m \to \ell_q^m, \gamma_m)_q$ for all $1 \leqslant q < \infty$.

Theorem 2.1. *Let $C(q)$ be given in* (1.5).

(i) *If $1 \leqslant q \leqslant 2$ and $1 \leqslant n \leqslant m$, then*

$$\lambda_n^{(a)}\big(D:\mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q = C(q)\left(\sum_{k=n+1}^m d_k^q\right)^{\frac{1}{q}}. \tag{2.2}$$

Moreover, D_n is the optimal linear operator for the average linear n-widths $\lambda_n^{(a)}(I_m:\mathbb{R}^m \to \ell_q^m, \gamma_m)_q$ for all $1 \leqslant q \leqslant 2$ in the sense that

$$\lambda_n^{(a)}\big(D:\mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q = \left(\int_{\mathbb{R}^m} \|Dx - D_n x\|_{\ell_q^m}^q \, d\gamma_m(x)\right)^{1/q}.$$

(ii) *If $2 < q < \infty$ and $1 \leqslant n \leqslant m$, then*

$$\lambda_n^{(a)}\big(D:\mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q = C(q)\left[(\kappa - n)^{\frac{q}{2}}\left(\sum_{j=1}^{\kappa} d_j^r\right)^{1-\frac{q}{2}} + \sum_{j=\kappa+1}^m d_j^q\right]^{\frac{1}{q}}, \tag{2.3}$$

where $\frac{1}{r} = \frac{1}{q} - \frac{1}{2}$, *and* $\kappa = \kappa(m, n, q, D)$ *is the biggest integer in* $[n+1, m]$ *such that*

$$(\kappa - n)\, d_\kappa^r \leqslant \sum_{j=1}^{\kappa} d_j^r.$$

According to [12, Corollary 1],

$$\lambda_n^{(a)}\big(D : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p \asymp \lambda_n^{(a)}\big(D : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q$$

for all $1 \leqslant q \leqslant \infty$ and $0 < p < \infty$, where the constants of equivalence depend only on p and q. Thus, Theorem 2.1 also gives the sharp asymptotic orders of $\lambda_n^{(a)}(D : \mathbb{R}^m \to \ell_q^m, \gamma_m)_p$ for all $1 \leqslant q < \infty$ and $0 < p < \infty$:

Corollary 2.2. *Let* $J_{m,n}(D)_q$ *denote the expression on the right-hand side of* (2.2) *and* (2.3) *for* $1 \leqslant q \leqslant 2$ *and* $2 < q < \infty$ *respectively. Then for all* $0 < p < \infty$ *and* $1 \leqslant q < \infty$,

$$\lambda_n^{(a)}\big(D : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p \asymp J_{m,n}(D)_q, \quad 1 \leqslant n \leqslant m. \tag{2.4}$$

We point out that the sharp estimates (2.4) for $2 \leqslant q < \infty$ were previously obtained in [3] under the additional assumptions that

$$m \geqslant 2n \quad \text{and} \quad m^{-1} \sum_{j=n+2}^{m} d_j^{\frac{2q}{2-q}} \leqslant d_{n+1}^{\frac{2q}{2-q}}. \tag{2.5}$$

It is worthwhile to mention that Theorem 2.1 and Corollary 2.2 are new even in the case of the identity matrix I_m, where our results can be stated more explicitly as follows:

Corollary 2.3. *If* $1 \leqslant q < \infty$, *and* $1 \leqslant n \leqslant m$, *then*

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q = C(q) m^{1/q} (1 - n/m)^{\max(1/q, 1/2)}, \tag{2.6}$$

and

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p \asymp m^{1/q} (1 - n/m)^{\max(1/q, 1/2)}, \quad 0 < p < \infty, \tag{2.7}$$

where the constant of equivalence depends only on p *and* q.

The following sharp estimate

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_p \asymp m^{1/q}, \quad 1 \leqslant q < \infty,\ 0 < p < \infty$$

was previously obtained in [4] under the additional assumption that $m \geqslant 2n$. Our result (2.7) here applies to the full range of $1 \leqslant n \leqslant m$.

2.2. *Proof of Theorem 2.1*

For the proof of Theorem 2.1, we need a series of lemmas.

Lemma 2.4. *(See [9, Lemma 2.8, pp. 207–208].) If n is a nonnegative integer $\leqslant m$, and $(\xi_1, \dots, \xi_m) \in [0,1]^m$ satisfies $\sum_{k=1}^m \xi_k = n$, then there exists an n-dimensional linear subspace V of $\mathbb{R}^m$ such that $\xi_j = \|P_V e_j\|_2^2$ for all $1 \leqslant j \leqslant m$, where $P_V : \mathbb{R}^m \to V$ denotes the orthogonal projection onto V.*

Lemma 2.5. *If $T : \mathbb{R}^m \to \ell_q^m$ is an arbitrary linear mapping with $1 \leqslant q < \infty$, then*

$$\left(\int_{\mathbb{R}^m} \|Dx - Tx\|_q^q \, d\gamma_m(x) \right)^{\frac{1}{q}} = C(q) \left(\sum_{i=1}^m \|d_i e_i - T^* e_i\|_2^q \right)^{1/q}, \tag{2.8}$$

where T^* denotes the conjugate operator of $T : \mathbb{R}^m \to \ell_q^m$, and $C(q)$ is given in (1.5)*. In particular, if $T = D_n$ with $1 \leqslant n < m$, then*

$$\left(\int_{\mathbb{R}^m} \|Dx - D_n x\|_q^q \, d\gamma_m(x) \right)^{1/q} = C(q) \left(\sum_{k=n+1}^m d_k^q \right)^{1/q}. \tag{2.9}$$

Proof. Letting $X \sim N_m(0, I_m)_{\mathbb{R}^m}$, we have

$$\begin{aligned} \int_{\mathbb{R}^m} \|Dx - Tx\|_q^q \, d\gamma_m(x) &= \mathbb{E}\|DX - TX\|_q^q = \mathbb{E}\left(\sum_{j=1}^m \left|\langle DX - TX, e_j\rangle\right|^q \right) \\ &= \sum_{j=1}^m \mathbb{E}\left|\langle X, De_j - T^* e_j\rangle\right|^q = \sum_{j=1}^m \mathbb{E}\left|\langle X, d_j e_j - T^* e_j\rangle\right|^q. \end{aligned} \tag{2.10}$$

Since $\langle X, d_j e_j - T^* e_j\rangle = \|d_j e_j - T^* e_j\|_2 \eta_j$ with $\eta_j \sim N(0,1)_{\mathbb{R}^1}$, it follows that

$$\begin{aligned} \sum_{j=1}^m \mathbb{E}\left|\langle X, d_j e_j - T^* e_j\rangle\right|^q &= \sum_{j=1}^m \|d_j e_j - T^* e_j\|_2^q \mathbb{E}|\eta_j|^q \\ &= C(q)^q \sum_{j=1}^m \|d_j e_j - T^* e_j\|_2^q, \end{aligned} \tag{2.11}$$

where

$$C(q)^q = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} |t|^q e^{-t^2/2} \, dt = \left(\pi^{-\frac{1}{2}} 2^{\frac{q}{2}} \Gamma\left(\frac{q+1}{2} \right) \right)^{1/q}.$$

A combination of (2.10) and (2.11) yields (2.8).

Finally, to complete the proof, we note that (2.9) follows directly from (2.8) since $D_n^* e_j = D_n e_j = d_j e_j$ for $1 \leqslant j \leqslant n$, and $D_n^* e_j = D_n e_j = 0$ for $n < j \leqslant m$. □

Lemma 2.6. *If $1 \leqslant q < \infty$, then*

$$\lambda_n^{(a)}\big(D : \mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q = C(q) I_q(m,n), \tag{2.12}$$

where $I_q(m,n)$ is defined by

$$I_q(m,n) := \min\left\{ \left(\sum_{i=1}^m d_i^q x_i^{q/2} \right)^{\frac{1}{q}} :\ (x_1, \dots, x_m) \in [0,1]^m \text{ and } \sum_{i=1}^m x_i = m-n \right\}. \tag{2.13}$$

Proof. We start with the proof of the lower estimate $\lambda_n^{(a)}(T : \mathbb{R}^m \to \ell_q^m, \gamma_m)_q \geqslant C(q) I_q(m,n)$. To this end, let $T_n : \mathbb{R}^m \to \mathbb{R}^m$ be an arbitrary linear operator with rank $\leqslant n$, and let $V = T_n^*(\mathbb{R}^m)$ be the range of the conjugate operator T_n^*. Then $\dim V \leqslant n$, and

$$\sum_{j=1}^m \big\| d_j e_j - T_n^* e_j \big\|_2^q \geqslant \sum_{j=1}^m d_j^q \min_{y \in V} \|e_j - y\|_2^q = \sum_{j=1}^m |d_j|^q \|P_{V^\perp} e_j\|_2^q, \tag{2.14}$$

where $P_{V^\perp}$ denotes the orthogonal projection onto the orthogonal complement $V^\perp$ of V in $\mathbb{R}^m$. Let

$$x_j = (m-n) \|P_{V^\perp} e_j\|_2^2 \left(\sum_{i=1}^m \|P_{V^\perp} e_i\|_2^2 \right)^{-1}, \quad \text{for } 1 \leqslant j \leqslant m.$$

Clearly, $\sum_{j=1}^m x_j = m-n$. Since

$$\sum_{i=1}^m \|P_{V^\perp} e_i\|_2^2 = \sum_{i=1}^m \langle P_{V^\perp} e_i, e_i \rangle = \operatorname{trace}(P_{V^\perp}) \geqslant m-n,$$

it follows that $0 \leqslant x_i \leqslant \|P_{V^\perp} e_i\|_2^2 \leqslant 1$ for each $1 \leqslant i \leqslant m$. Thus, using (2.14), we deduce

$$\sum_{j=1}^m \big\| d_j e_j - T_n^* e_j \big\|_2^q \geqslant \sum_{j=1}^m |d_j|^q x_j^{q/2} \geqslant I_q(m,n)^q,$$

which, together with (2.8), implies that

$$\left(\int_{\mathbb{R}^m} |Dx - T_n x|^q \, d\gamma_m(x) \right)^{\frac{1}{q}} \geqslant C(q) I_q(m,n).$$

Since T_n is an arbitrary linear operator with rank $\leqslant n$, the desired lower estimate follows.

It remains to show the upper estimate $\lambda_n^{(a)}(D:\mathbb{R}^m \to \ell_q^m, \gamma_m)_q \leqslant C(q) I_q(m,n)$. Let $(y_1,\dots,y_m) \in [0,1]^m$ be the minimizer of $\sum_{j=1}^m d_j^q x_j^{\frac{q}{2}}$ subject to the conditions $(x_1,\dots,x_m) \in [0,1]^m$ and $\sum_{j=1}^m x_j = m-n$. Setting $\xi_j = 1 - y_j$ for $1 \leqslant j \leqslant m$, we have $\xi_j \in [0,1]$ and $\sum_{j=1}^m \xi_j = n$. Hence, by Lemma 2.4, there exists a linear subspace V of $\mathbb{R}^m$ with $\dim V = n$ and such that $\xi_j = \|P_V e_j\|_2^2$ for all $1 \leqslant j \leqslant m$. Clearly, $\|e_j - P_V e_j\|_2^2 = \|e_j\|_2^2 - \|P_V e_j\|_2^2 = 1 - \xi_j = y_j$. Thus, using Lemma 2.5, we have that

$$\begin{aligned}\lambda_n^{(a)}\big(D:\mathbb{R}^m \to \ell_q^m, \gamma_m\big)_q^q &\leqslant \int\limits_{\mathbb{R}^m} \|Dx - DP_V x\|_q^q \, d\gamma_m(x) \\ &= C(q)^q \sum_{j=1}^m \big\|D^* e_j - P_V^* D^* e_j\big\|_2^q = C(q)^q \sum_{j=1}^m d_j^q \|e_j - P_V e_j\|_2^q \\ &= C(q)^q \sum_{j=1}^m d_j^q y_j^{\frac{q}{2}} = C(q)^q I_q(m,n)^q,\end{aligned}$$

which gives the desired upper estimate, and hence completes the proof of Lemma 2.6. □

Now we are in a position to prove Theorem 2.1.

Proof of Theorem 2.1. By Lemma 2.6, it suffices to prove that

$$I_q(m,n)^q = \begin{cases} \sum_{k=n+1}^m d_k^q, & \text{if } 1 \leqslant q \leqslant 2, \\ (\kappa - n)^{\frac{q}{2}} (\sum_{j=1}^\kappa d_j^r)^{1-\frac{q}{2}} + \sum_{j=\kappa+1}^m d_j^q, & \text{if } 2 < q < \infty, \end{cases} \tag{2.15}$$

where $\frac{1}{r} = \frac{1}{q} - \frac{1}{2}$, and $\kappa = \kappa(m,n,q,D)$ is as defined in Theorem 2.1.

If $1 \leqslant q \leqslant 2$, $x_1,\dots,x_m \in [0,1]$ and $\sum_{j=1}^m x_j = m-n$, then using the monotonicity of the d_j,

$$\sum_{i=1}^m d_i^q x_i^{q/2} \geqslant \sum_{i=1}^m d_i^q x_i \geqslant d_n^q \sum_{i=1}^n x_i + \sum_{i=n+1}^m d_i^q - d_n^q \sum_{i=n+1}^m (1 - x_i) = \sum_{i=n+1}^m d_i^q,$$

where the equalities can be achieved whenever $x_1 = \cdots = x_n = 0$ and $x_{n+1} = \cdots = x_m = 1$. (2.15) for $1 \leqslant q \leqslant 2$ then follows.

To prove (2.15) for $2 < q < \infty$, we claim that

$$I_q(m,n)^q := \min\left\{ \sum_{j=1}^\kappa d_j^q x_j^{\frac{q}{2}} + \sum_{j=\kappa+1}^m d_j^q \colon x_1,\dots,x_\kappa \in [0,1] \text{ and } \sum_{j=1}^\kappa x_j = \kappa - n \right\}. \tag{2.16}$$

For the moment, we take (2.16) for granted and proceed with the proof of (2.15). Using the following reverse Hölder inequality

$$\sum_j |a_j b_j| \geqslant \left(\sum_j |a_j|^p\right)^{\frac{1}{p}} \left(\sum_j |b_j|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}, \quad 0 < p < 1,$$

and recalling $\frac{2}{q} \in (0,1)$, we deduce that

$$\sum_{j=1}^{\kappa} d_j^q x_j^{\frac{q}{2}} \geqslant \left(\sum_{j=1}^{\kappa} x_j\right)^{\frac{q}{2}} \left(\sum_{j=1}^{\kappa} d_j^{\frac{2q}{2-q}}\right)^{1-\frac{q}{2}} = (\kappa-n)^{\frac{q}{2}} \left(\sum_{j=1}^{\kappa} d_j^r\right)^{1-\frac{q}{2}}$$

whenever $x_1,\dots,x_\kappa \in [0,1]$ and $\sum_{j=1}^{\kappa} x_j = \kappa - n$. This combined with the claim (2.16) yields the desired lower estimate of (2.15) for $2<q<\infty$. To prove the desired upper estimate, let

$$z_i = (\kappa-n) d_i^r \left(\sum_{j=1}^{\kappa} d_j^r\right)^{-1}, \quad 1 \leqslant i \leqslant \kappa.$$

Since $r = \frac{2q}{2-q} < 0$ and $d_1 \geqslant \dots \geqslant d_m > 0$, we deduce from the definition of κ that $0 \leqslant z_i \leqslant z_\kappa \leqslant 1$ for all $1 \leqslant i \leqslant \kappa$, and $\sum_{i=1}^{\kappa} z_i = \kappa - n$. Thus, using (2.16), we obtain

$$I_q(m,n)^q \leqslant \sum_{j=1}^{\kappa} d_j^q z_j^{\frac{q}{2}} + \sum_{j=\kappa+1}^{m} d_j^q = (\kappa-n)^{\frac{q}{2}} \left(\sum_{j=1}^{\kappa} d_j^r\right)^{1-\frac{q}{2}} + \sum_{j=\kappa+1}^{m} d_j^q.$$

The desired upper estimate of (2.15) for $2<q<\infty$ then follows.

Now it remains to show the claim (2.16). Let $(y_1,\dots,y_m) \in [0,1]^m$ be the minimizer of $\sum_{j=1}^{m} d_j^q x_j^{\frac{q}{2}}$ subject to the conditions $(x_1,\dots,x_m) \in [0,1]^m$ and $\sum_{j=1}^{m} x_j = m-n$. Note that $\sum_{j=1}^{m} d_j^q x_j^{\frac{q}{2}} \geqslant \sum_{j=1}^{m} d_j^q (x_j^*)^{\frac{q}{2}}$, where $0 \leqslant x_1^* \leqslant \dots \leqslant x_m^*$ is a nondecreasing rearrangement of $x_1,\dots,x_m$. Thus, without loss of generality, we may assume that $0 \leqslant y_1 \leqslant \dots \leqslant y_m \leqslant 1$. For the proof of the claim (2.16), it suffices to show that $y_{\kappa+1} = 1$. To this end, let's first observe that $y_1 > 0$. Indeed, if $y_1 = 0$, then the function $\varphi(t) = d_1^q t^{\frac{q}{2}} + d_m^q (y_m - t)^{\frac{q}{2}}$, $0 \leqslant t \leqslant \delta$ achieves its minimum at $t = 0$, whenever $\delta > 0$ is sufficiently small. However, this is impossible since $q > 2$ and $\varphi'(t) = \frac{q}{2} d_1^q t^{\frac{q}{2}-1} - \frac{q}{2} d_m^q (y_m - t)^{\frac{q}{2}-1} < 0$ for $t > 0$ being sufficiently small. Next, we show that $y_{\kappa+1} = 1$. Assuming that $y_{\kappa+1} < 1$. We shall obtain a contradiction as follows. Let k_0 be the largest integer j such that $\kappa + 1 \leqslant j \leqslant m$ and $y_j < 1$. Then, clearly, $(y_1,\dots,y_{k_0})$ is a minimizer of $\sum_{j=1}^{k_0} d_j^q x_j^{\frac{q}{2}}$ subject to the conditions $(x_1,\dots,x_{k_0}) \in (0,1)^{k_0}$ and $\sum_{j=1}^{k_0} x_j = k_0 - n$. Thus, applying the method of the Lagrange multipliers, we deduce

$$d_1^q y_1^{\frac{q}{2}-1} = \dots = d_{k_0}^q y_{k_0}^{\frac{q}{2}-1}, \quad \text{and} \quad \sum_{j=1}^{k_0} y_j = k_0 - n,$$

which in turn implies

$$y_{k_0} = (k_0 - n) d_{k_0}^r \left(\sum_{j=1}^{k_0} d_j^r\right)^{-1}.$$

On the one hand, since $y_{k_0} \in [0,1]$, it follows that

$$\sum_{j=1}^{k_0} d_j^r \geqslant (k_0 - n) d_{k_0}^r. \tag{2.17}$$

On the other hand, however, using the definition of κ, and the facts that $r < 0$ and $d_1 \geqslant \cdots \geqslant d_m > 0$, we deduce

$$\begin{aligned} \sum_{j=1}^{k_0} d_j^r &= \sum_{1 \leqslant j \leqslant \kappa+1} d_j^r + \sum_{\kappa+2 \leqslant j \leqslant k_0} d_j^r < (\kappa + 1 - n) d_{\kappa+1}^r + (k_0 - \kappa - 1) d_{k_0}^r \\ &\leqslant (k_0 - n) d_{k_0}^r, \end{aligned}$$

which contradicts (2.17), and hence completes the proof of Theorem 2.1. $\square$

3. The average linear n-widths $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m)_p$

Corollary 2.3 in the last section gives the sharp lower and upper estimates of $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m)_p$ for all $1 \leqslant q < \infty$ and $0 < p < \infty$. The remaining case $q = \infty$ is more difficult to deal with. It was shown by Maiorov and Wasilkowski [8] that

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m\big)_1 \asymp \sqrt{\ln m}, \quad m \geqslant 2n. \tag{3.1}$$

Our first result in this section gives a slightly better upper estimate of $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m)_p$ for the full range of $1 \leqslant n \leqslant m$:

Theorem 3.1. *If* $0 < p < \infty$*, and* $1 \leqslant n \leqslant m$ *then*

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m\big)_p \leqslant C_p \min\big((m-n)^{1/2}, (\ln m)^{1/2}\big)(1 - n/m)^{1/2}. \tag{3.2}$$

Proof. It suffices to prove (3.2) with $p = 1$ by [12, Corollary 1]. By Lemma 2.4, there exists an n-dimensional linear subspace V of $\mathbb{R}^m$ such that $\|P_V e_k\|_2^2 = \frac{n}{m}$ for all $1 \leqslant k \leqslant m$, where $P_V e_k$ denotes the orthogonal projection of the kth unit vector $e_k \in \mathbb{R}^m$ onto V. By Stirling's formula (see [1, p. 18]),

$$\lim_{x \to +\infty} \frac{\Gamma(x+1)e^x}{\sqrt{2\pi} x^{x+\frac{1}{2}}} = 1,$$

we obtain

$$\lim_{q \to +\infty} \frac{C(q)}{q^{1/2}} = \lim_{q \to +\infty} \left(\frac{q-1}{q}\right)^{1/2} 2^{\frac{1}{2q}} \exp\left(-\frac{q-1}{2q}\right) = e^{-1/2}.$$

From the proofs of Lemma 2.6 and Theorem 2.1, it follows that

$$\left(\int_{\mathbb{R}^m} \|x - P_V x\|_q^q \, d\gamma_m(x)\right)^{1/q} = C(q) m^{1/q} (1 - n/m)^{1/2}$$
$$\leqslant C q^{1/2} m^{1/q} (1 - n/m)^{1/2}, \quad 2 \leqslant q < \infty.$$

Thus, taking $q = \ln(e^2 m)$, we get

$$\begin{aligned}\lambda_n^{(a)}\left(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m\right)_1 &\leqslant \int_{\mathbb{R}^m} \|x - P_V x\|_\infty \, d\gamma_m(x) \\ &\leqslant \left(\int_{\mathbb{R}^m} \|x - P_V x\|_q^q \, d\gamma_m(x)\right)^{1/q} \\ &\leqslant C(\ln m)^{1/2}(1 - n/m)^{1/2}. \end{aligned} \tag{3.3}$$

On the other hand, however, a direct computation shows

$$\begin{aligned}\lambda_n^{(a)}\left(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m\right)_1 &\leqslant \int_{\mathbb{R}^m} \|x - P_V x\|_\infty \, d\gamma_m(x) \\ &= \int_{\mathbb{R}^m} \max_{1 \leqslant i \leqslant m} \left|\langle P_{V^\perp} x, e_i\rangle\right| d\gamma_m(x) \\ &\leqslant \max_{1 \leqslant i \leqslant m} \|P_{V^\perp} e_i\|_2 \int_{\mathbb{R}^m} \|P_{V^\perp} x\|_2 \, d\gamma_m(x) \\ &= (1 - n/m)^{1/2} \int_{\mathbb{R}^{m-n}} \|x\|_2 \, d\gamma_{m-n}(x) \\ &\leqslant C(m - n)^{1/2}(1 - n/m)^{1/2}, \end{aligned} \tag{3.4}$$

where $P_{V^\perp}$ denotes the orthogonal projection onto the orthogonal complement $V^\perp$ of V in $\mathbb{R}^m$, and in the fourth step, we have used the rotation invariance of the Gaussian measure γ_m. Thus, a combination of (3.3) and (3.4) yields the desired upper estimate of $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m)_1$. $\square$

It is still unclear whether or not the upper estimate (3.2) is sharp for the full range of $1 \leqslant m \leqslant n$. On the other hand, however, from (3.1) and our proofs of Theorem 2.1 and Corollary 3.1, there exists an n-dimensional linear subspace V of $\mathbb{R}^m$ independent of q such that

$$\lambda_n^{(a)}\left(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m\right)_q = \left(\int_{\mathbb{R}^m} \|x - P_V x\|_q^q \, d\gamma_m(x)\right)^{1/q}, \quad 2 \leqslant q < \infty$$

and

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m\big)_1 \asymp \int_{\mathbb{R}^m} \|x - P_V x\|_\infty \, d\gamma_m(x), \qquad m \geqslant 2n, \tag{3.5}$$

where $P_V : \mathbb{R}^m \to V$ denotes the orthogonal projection onto V. This means that P_V is the simultaneous optimal linear operator and asymptotically optimal linear operator for the average linear n-widths $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m)_q$ and $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_q^m, \gamma_m)_1$ with $2 \leqslant q < \infty$, and is asymptotically optimal for $\lambda_n^{(a)}(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m)_1$ when $m \geqslant 2n$. It remains unclear whether the restriction $m \geqslant 2n$ in (3.5) can be dropped. Nevertheless, as our next theorem shows, (3.5) is true for some n-dimensional linear space V.

Theorem 3.2. *For $1 \leqslant n \leqslant m$ and $0 < p < \infty$, we have*

$$\lambda_n^{(a)}\big(I_m : \mathbb{R}^m \to \ell_\infty^m, \gamma_m\big)_p^p = \inf_{\substack{V \subset \mathbb{R}^m \\ \dim V = n}} \int_{\mathbb{R}^m} \|x - P_V x\|_\infty^p \, d\gamma_m(x). \tag{3.6}$$

The proof of Theorem 3.2 relies on the following lemma, which is essentially contained in [2, p. 29]:

Lemma 3.3. *Let X and Y be two $\mathbb{R}^m$-valued Gaussian random vectors with mean zero. If*

$$\mathbb{E}\big|\langle x, X\rangle\big|^2 \leqslant \mathbb{E}\big|\langle x, Y\rangle\big|^2, \quad \forall x \in \mathbb{R}^m, \tag{3.7}$$

then for any $0 < p < \infty$, and any semi-norm $\varphi : \mathbb{R}^m \to [0, \infty)$ on $\mathbb{R}^m$,

$$\mathbb{E}\varphi(X)^p \leqslant \mathbb{E}\varphi(Y)^p.$$

Proof. By Theorem 1.8.9 of [2, p. 29], (3.7) is equivalent to the following condition: for every convex symmetric subset $A \subset \mathbb{R}^m$,

$$\mathbb{P}(X \in A) \geqslant \mathbb{P}(Y \in A). \tag{3.8}$$

Given any semi-norm φ on $\mathbb{R}^m$, and any $t > 0$, the set $A_t := \{x \in \mathbb{R}^m\colon \varphi(x) \leqslant t\}$ is symmetric and convex. Thus, using (3.8),

$$\mathbb{P}\big(\varphi(X) > t\big) = 1 - \mathbb{P}(X \in A_t) \leqslant 1 - \mathbb{P}(Y \in A_t) = \mathbb{P}\big(\varphi(Y) > t\big).$$

It then follows that

$$\begin{aligned} \mathbb{E}\varphi(X)^p &= p \int_0^\infty t^{p-1} \mathbb{P}\big(\varphi(X) > t\big)\, dt \\ &\leqslant p \int_0^\infty t^{p-1} \mathbb{P}\big(\varphi(Y) > t\big)\, dt = \mathbb{E}\varphi(Y)^p. \qquad \square \end{aligned}$$

Proof of Theorem 3.2. Let $T:\mathbb{R}^m\to\ell_\infty^m$ be a linear mapping with rank n, and let $V=T^*(\mathbb{R}^m)$ denote the range of T^*, the conjugate of T. Clearly, for the proof of Theorem 3.2, it is enough to show that for $Z\sim N_m(0,I_m)_{\mathbb{R}^m}$,

$$\mathbb{E}\|Z-TZ\|_\infty^p\geqslant\mathbb{E}\|Z-P_VZ\|_\infty^p.$$

Since $Z-TZ$ and $Z-P_VZ$ are both centered Gaussian random vectors on $\mathbb{R}^m$, using Lemma 3.3, we reduce to showing that

$$\mathbb{E}\bigl|\langle x,Z-TZ\rangle\bigr|^2\geqslant\mathbb{E}\bigl|\langle x,Z-P_VZ\rangle\bigr|^2,\quad\forall x\in\mathbb{R}^m.\tag{3.9}$$

To see this, we use the fact that $\langle\xi,Z\rangle\sim N(0,1)$ whenever $\xi\in\mathbb{S}^{m-1}$ and $Z\sim N_m(0,I_m)_{\mathbb{R}^m}$. We then get, for any $x\in\mathbb{R}^m$,

$$\begin{aligned}\mathbb{E}\bigl|\langle x,Z-TZ\rangle\bigr|^2&=\mathbb{E}\bigl|\bigl\langle x-T^*x,Z\bigr\rangle\bigr|^2=\bigl\|x-T^*x\bigr\|_2^2\geqslant\|x-P_Vx\|_2^2\\&=\mathbb{E}\bigl|\langle x,Z-P_VZ\rangle\bigr|^2.\qquad\square\end{aligned}$$

4. The probabilistic linear (n,δ)-widths

In this section, we shall consider the probabilistic linear (n,δ)-widths. Our main result is the following:

Theorem 4.1. *If $1\leqslant q\leqslant 2$, then there exists a constant $\delta_0\in(0,1/2]$ such that for all $\delta\in(0,\delta_0]$,*

$$\lambda_{n,\delta}\bigl(D:\mathbb{R}^m\to\ell_q^m,\gamma_m\bigr)\asymp\left(\sum_{k=n+1}^m d_k^q\right)^{1/q}+\sqrt{\ln(1/\delta)}\left(\sum_{k=n+1}^m d_k^r\right)^{1/r},\tag{4.1}$$

where $\frac1r=\frac1q-\frac12$ and the constants of equivalence depend only on q.

Theorem 4.1 gives a negative answer to a conjecture in [3, p. 339], where the authors proved sharp upper and lower estimates of $\lambda_{n,\delta}(D:\mathbb{R}^m\to\ell_q^m,\gamma_m)$ for $2\leqslant q<\infty$ under the additional assumptions (2.5). In the case of the identity matrix I_m, sharp estimates of $\lambda_{n,\delta}(I_m:\mathbb{R}^m\to\ell_q^m,\gamma_m)$ were previously obtained in [7,8] for $2\leqslant q\leqslant\infty$, and in [4] for $1\leqslant q\leqslant 2$, under the additional assumption $m\geqslant 2n$.

For the proof of Theorem 4.1, we need the following two known lemmas.

Lemma 4.2. *(See [2, p. 2], [7].) If $\eta\sim N(0,1)_{\mathbb{R}}$ then for all $t>0$*

$$\sqrt2\pi^{-1/2}\bigl(t^{-1}-t^{-3}\bigr)e^{-t^2/2}\leqslant\mathbb{P}\bigl(|\eta|\geqslant t\bigr)\leqslant\sqrt2\pi^{-1/2}t^{-1}e^{-t^2/2}.$$

In particular, if $\delta\in(0,e^{-1}]$ then $\mathbb{P}(|\eta|\geqslant c_0\sqrt{\ln\delta^{-1}})>\delta$ for some absolute constant $c_0>0$.

Lemma 4.3. *(See [2, (1.7.7)], [10, p. 47].) Let $F:\mathbb{R}^m\to\mathbb{R}$ be a function satisfying the following Lipschitz condition*

$$\left|F(x)-F(y)\right|\leqslant\sigma\|x-y\|_2,\quad x,y\in\mathbb{R}^m,$$

for some $\sigma>0$ independent of x and y. If $X\sim N_m(0,I_m)_{\mathbb{R}^m}$ is an $\mathbb{R}^m$-valued Gaussian random vector with mean 0 and covariance matrix I_m, then for all $t>0$

$$\mathbb{P}\left(\left|F(X)-\mathbb{E}F(X)\right|>t\right)\leqslant 2\exp\left(-\frac{t^2}{K^2\sigma^2}\right),\tag{4.2}$$

with $K>0$ being an absolute constant.

(4.2) is called the Gaussian concentration inequality.

Now we are in a position to prove Theorem 4.1.

Proof of Theorem 4.1. We start with the proof of the upper estimates, which is simpler. Let $F(x)=\|Dx-D_nx\|_q$ for $x\in\mathbb{R}^m$ and $1\leqslant q\leqslant 2$. Recalling $\frac{1}{r}+\frac{1}{2}=\frac{1}{q}$, we use Hölder's inequality to obtain

$$\left|F(x)-F(y)\right|\leqslant\left\|(D-D_n)(x-y)\right\|_q=\left(\sum_{k=n+1}^m d_k^q|x_k-y_k|^q\right)^{\frac{1}{q}}\leqslant\sigma_1\|x-y\|_2,$$

where $\sigma_1:=(\sum_{k=n+1}^m d_k^r)^{1/r}$. Thus, the Gaussian concentration inequality (4.2) yields

$$\mathbb{P}\left(\left|F(X)-\mathbb{E}F(X)\right|>t\right)\leqslant 2\exp\left(-\frac{t^2}{K^2\sigma_1^2}\right),\quad\forall t>0,$$

where, here and in what follows, $X\sim N_m(0,I_m)_{\mathbb{R}^m}$. In particular, this implies that for $Q_\delta=\{x\in\mathbb{R}^m:\ F(x)>\mathbb{E}F(X)+K\sigma_1\sqrt{\ln(2/\delta)}\}$ with $\delta\in(0,1)$,

$$\gamma_m(Q_\delta)\leqslant\mathbb{P}\left(\left|F(X)-\mathbb{E}F(X)\right|>K\sigma_1\sqrt{\ln(2/\delta)}\right)\leqslant\delta.$$

By the definition of the linear (n,δ)-widths, this last equation further implies that

$$\lambda_{n,\delta}\left(D:\mathbb{R}^m\to\ell_q^m,\gamma_m\right)\leqslant\sup_{\mathbb{R}^m\setminus Q_\delta}\|Dx-D_nx\|_q\leqslant\mathbb{E}F(X)+K\sigma_1\sqrt{\ln(2/\delta)}.\tag{4.3}$$

On the other hand, however, using (2.9), we have

$$\mathbb{E}F(X)=\mathbb{E}\|DX-D_nX\|_q\asymp\left(\mathbb{E}\|DX-D_nX\|_q^q\right)^{1/q}=C(q)\sigma_2,\tag{4.4}$$

where $\sigma_2:=(\sum_{k=n+1}^m d_k^q)^{1/q}$. Thus, combining (4.3) with (4.4), we deduce the desired upper estimate

$$\lambda_{n,\delta}\left(D:\mathbb{R}^m\to\ell_q^m,\gamma_m\right)\leqslant C\sigma_2+K\sigma_1\sqrt{\ln(2/\delta)}.$$

To show the desired lower estimates, we need only to prove both

$$\lambda_{n,\delta}\big(D:\mathbb{R}^m \to \ell_q^m, \gamma_m\big) \geqslant c\sigma_1\sqrt{\ln(2/\delta)}, \tag{4.5}$$

and

$$\lambda_{n,\delta}\big(D:\mathbb{R}^m \to \ell_q^m, \gamma_m\big) \geqslant c\sigma_2, \tag{4.6}$$

for some absolute constant $c > 0$.

For the proof of (4.5), it suffices to show that for any linear mapping $T_n : \mathbb{R}^m \to \ell_q^m$ with rank $\leqslant n$,

$$\mathbb{P}\big(\|DX - T_n X\|_q \geqslant c\sigma_1\sqrt{\ln(2/\delta)}\,\big) > \delta, \tag{4.7}$$

since (4.7) implies that the set $\{x \in \mathbb{R}^m\colon\ \|D_n x - T_n x\|_q \geqslant c\sigma_1\sqrt{\ln(2/\delta)}\}$ cannot be entirely contained in any Borel subset $G \subset \mathbb{R}^m$ with $\gamma_m(G) \leqslant \delta$. Letting $\eta \sim N(0,1)_{\mathbb{R}}$, we observe that

$$\begin{aligned}
\mathbb{P}\big(\|DX - T_n X\|_q \geqslant c\sigma_1\sqrt{\ln(2/\delta)}\,\big) &= \mathbb{P}\Big(\sup_{\|y\|_{q'}\leqslant 1} \big|\big\langle X, \big(D - T_n^*\big)y\big\rangle\big| \geqslant c\sigma_1\sqrt{\ln(2/\delta)}\,\Big) \\
&\geqslant \sup_{\|y\|_{q'}\leqslant 1} \mathbb{P}\big(\big|\big\langle X, \big(D - T_n^*\big)y\big\rangle\big| \geqslant c\sigma_1\sqrt{\ln(2/\delta)}\,\big) \\
&= \sup_{\|y\|_{q'}\leqslant 1} \mathbb{P}\bigg(|\eta| \geqslant c\frac{\sigma_1\sqrt{\ln(2/\delta)}}{\|(D - T_n^*)y\|_2}\bigg) \\
&\geqslant \mathbb{P}\bigg(|\eta| \geqslant c\frac{\sigma_1\sqrt{\ln(2/\delta)}}{B}\bigg),
\end{aligned}$$

where

$$B := \sup_{\|y\|_{q'}\leqslant 1} \big\|\big(D - T_n^*\big)y\big\|_2 = \sup_{\|y\|_2\leqslant 1} \big\|(D - T_n)y\big\|_q \geqslant \lambda_n\big(D:\ell_2^m \to \ell_q^m\big) = \sigma_1$$

with the last step using (2.1). It then follows by Lemma 4.2 that

$$\mathbb{P}\big(\|DX - T_n X\|_q \geqslant c\sigma_1\sqrt{\ln(2/\delta)}\,\big) \geqslant \mathbb{P}\big(|\eta| \geqslant c\sqrt{\ln(2/\delta)}\,\big) > \delta$$

with $c = c_0$ being the same as in Lemma 4.2. This proves (4.7), and hence the lower estimate (4.5).

Now it remains to prove the lower estimate (4.6). Again, it suffices to show that

$$\mathbb{P}\big(\|DX - T_n X\|_q \geqslant c\sigma_2\big) > \delta, \tag{4.8}$$

where $T_n : \mathbb{R}^m \to \mathbb{R}^m$ is an arbitrary linear mapping with rank $\leqslant n$. Letting

$$e(D, T_n, \gamma_m, t) := \inf\big\{\rho > 0\colon\ \mathbb{P}\big(\|DX - T_n X\|_q > \rho\big) \leqslant t\big\}, \quad 0 \leqslant t \leqslant 1, \tag{4.9}$$

we reduce the proof of (4.8) to showing

$$e(D, T_n, \gamma_m, \delta) > c\sigma_2. \tag{4.10}$$

Since the function $t \to e(D, T_n, \gamma_m, t)$ defined by (4.9) is a nonincreasing rearrangement of the function $G(x) := \|Dx - T_n x\|_q$ with respect to the standard Gaussian measure γ_m on $\mathbb{R}^m$, it follows that

$$\mathbb{E}G(X) = \int_{\mathbb{R}^m} G(x)\,d\gamma_m(x) = \int_0^1 e(D, T_n, \gamma_m, t)\,dt. \tag{4.11}$$

Next, we note that for $x, y \in \mathbb{R}^m$,

$$\bigl|G(x) - G(y)\bigr| \leqslant \bigl\|(D - T_n)(x - y)\bigr\|_q \leqslant \tau\|x - y\|_2,$$

where

$$\begin{aligned}
\tau &:= \sup_{\|z\|_2=1} \bigl\|(D - T_n)z\bigr\|_q = \sup_{\|z\|_{q'}=1} \bigl\|\bigl(D - T_n^*\bigr)z\bigr\|_2 = \sup_{\|z\|_{q'}=1} \left\| \sum_{j=1}^m z_j\bigl(D - T_n^*\bigr)e_j \right\|_2 \\
&\leqslant \sup_{\|z\|_{q'}=1} \sum_{j=1}^m |z_j| \bigl\|\bigl(D - T_n^*\bigr)e_j\bigr\|_2 = \left(\sum_{j=1}^m \bigl\|\bigl(D - T_n^*\bigr)e_j\bigr\|_2^q \right)^{1/q} \\
&= C(q)^{-1}\bigl(\mathbb{E}\bigl\|(D - T_n)X\bigr\|_q^q\bigr)^{1/q} \leqslant c\mathbb{E}G(X) =: \sigma_3.
\end{aligned}$$

Thus, it follows from the Gaussian concentration inequality (4.2) that for any $s > 0$

$$\mathbb{P}\bigl(\bigl|G(X) - \mathbb{E}G(X)\bigr| > s\bigr) \leqslant 2\exp\left(-\frac{s^2}{K^2\sigma_3^2}\right). \tag{4.12}$$

Given $t \in (0, 1]$, we choose $s = K\sigma_3\sqrt{\ln(2/t)} = C\mathbb{E}G(X)\sqrt{\ln(2/t)}$ so that $2\exp(-s^2/K^2\sigma_3^2) = t$, and then apply (4.12) to deduce

$$\mathbb{P}\bigl(\bigl|G(X)\bigr| > \bigl(1 + C\sqrt{\ln(2/t)}\,\bigr)\mathbb{E}G(X)\bigr) \leqslant t.$$

Using (4.9), we conclude that

$$e(D, T, \gamma_m, t) \leqslant \bigl(1 + C\sqrt{\ln(2/t)}\,\bigr)\mathbb{E}G(X). \tag{4.13}$$

Now letting $\delta_0 \in (0, 1)$ be such that $\int_0^{\delta_0}(1 + \sqrt{\ln(2/t)}\,)\,dt = \frac{1}{2}$, we deduce from (4.11) and (4.13) that for any $\delta \in (0, \delta_0)$,

$$\begin{aligned}\mathbb{E}G(X) &= \int_0^1 e(D, T_n, \gamma_m, t)\, dt \\ &\leqslant \left(\int_0^\delta \bigl(1 + C\sqrt{\ln(2/t)}\bigr)\, dt\right)\mathbb{E}G(X) + \int_\delta^1 e(D, T, \gamma_m, t)\, dt \\ &\leqslant \frac{1}{2}\mathbb{E}G(X) + e(D, T, \gamma_m, \delta).\end{aligned}$$

It follows that

$$\begin{aligned}e(D, T_n, \gamma_m, \delta) &\geqslant \frac{1}{2}\mathbb{E}G(X) \geqslant C\bigl(\mathbb{E}\|DX - T_n X\|_q^q\bigr)^{\frac{1}{q}} \\ &\geqslant C \inf_{T_n}\bigl(\mathbb{E}\|DX - T_n X\|_q^q\bigr)^{\frac{1}{q}} \\ &= C\lambda_n^{(a)}\bigl(D : \mathbb{R}^m \to \ell_q^m, \gamma_m\bigr)_q = CC(q)\sigma_2,\end{aligned}$$

which proves (4.10), and hence the desired lower estimate (4.6). This completes the proof of Theorem 4.1. □

References

[1] G.E. Andrews, R. Askey, R. Roy, Special Functions, Cambridge Univ. Press, Cambridge, 1999.

[2] V.I. Bogachev, Gaussian Measures, Math. Surveys Monogr., vol. 62, American Mathematical Society, Providence, RI, 1998.

[3] Guanggui Chen, Pengjuan Nie, Xinjian Luo, The approximation characteristic of diagonal matrix in probabilistic setting, J. Complexity 26 (2010) 336–343.

[4] Gensun Fang, Peixin Ye, Probabilistic and average linear widths of Sobolev space with Gaussian measure, J. Complexity 19 (2003) 73–84.

[5] G.G. Lorentz, M.V. Golitschek, Y. Makovoz, Constructive Approximation, Advanced Problems, Springer-Verlag, New York, 1996.

[6] V.E. Maiorov, Kolmogorov's (n, δ)-widths of spaces of the smooth functions, Russian Acad. Sci. Sb. Math. 79 (2) (1994) 265–279.

[7] V.E. Maiorov, Linear widths of function spaces equipped with the Gaussian measure, J. Approx. Theory 77 (1) (1994) 74–88.

[8] V.E. Maiorov, G.W. Wasilkowski, Probabilistic and average linear widths in L_∞-norm with respect to r-fold Wiener measure, J. Approx. Theory 84 (1) (1996) 31–40.

[9] A. Pinkus, n-Widths in Approximation Theory, Springer, Berlin, 1985.

[10] G. Pisier, The Volume of Convex Bodies and Banach Space Geometry, Cambridge Univ. Press, Cambridge, UK, 1989.

[11] K. Ritter, Average-Case Analysis of Numerical Problems, Lecture Notes in Math., vol. 1733, Springer-Verlag, 2000.

[12] Heping Wang, Xuebo Zhai, Best approximation of functions on the ball on the weighted Sobolev space equipped with a Gaussian measure, J. Approx. Theory 162 (2010) 1160–1177.