HIGHER ORDER CONCENTRATION IN PRESENCE OF POINCARÉ-TYPE INEQUALITIES

FRIEDRICH GÖTZE AND HOLGER SAMBALE

Abstract. We show sharpened forms of the concentration of measure phenomenon typically centered at stochastic expansions of order $d-1$ for any $d \in \mathbb{N}$. Here we focus on differentiable functions on the Euclidean space in presence of a Poincaré-type inequality. The bounds are based on d-th order derivatives.

1. Introduction

In this note, we study higher order versions of the concentration of measure phenomenon. Instead of the classical problem of deviations of f around the mean $\mathbb{E}f$, we study potentially smaller fluctuations of $\tilde{f}_d := f - \mathbb{E}f - f_1 - \ldots - f_d$, where $f_1, \ldots, f_d$ are "lower order terms" of f with respect to a suitable decomposition, such as a Taylor-type decomposition of f. In order to study the concentration of $\tilde{f}_d$ around 0, which we call higher order concentration of measure, we use derivatives up to order d.

Previous work includes Adamczak and Wolff [A-W], who exploited certain Sobolev-type inequalities or subgaussian tail conditions to derive exponential tail inequalities for functions with bounded higher-order derivatives (evaluated in terms of some tensor-product matrix norms). While in [A-W], concentration around the mean is studied, the idea of sharpening concentration inequalities for Gaussian and related measures by requiring orthogonality to linear functions also appears in Wolff [W] as well as in Cordero-Erausquin, Fradelizi and Maurey [CE-F-M].

Our research started with second order results for functions on the n-sphere orthogonal to linear functions [B-C-G], with an approach which has been extended in [G-S] for measures satisfying logarithmic Sobolev inequalities. This includes discrete models as well as differentiable functions on open subsets of $\mathbb{R}^n$. These results were extended to arbitrary higher orders in [B-G-S].

While in [B-G-S], measures satisfying a logarithmic Sobolev inequality were considered, the aim of this note is to prove similar results for measures satisfying a Poincaré-type inequality, i. e. a weaker assumption. To this end, let us recall that a Borel probability measure μ on an open set $G \subset \mathbb{R}^n$ is said to satisfy a *Poincaré-type inequality* with constant $\sigma^2 > 0$ if for any bounded smooth function f on G with gradient ∇f,

$$\mathrm{Var}_\mu(f) \le \sigma^2 \int |\nabla f|^2 \, d\mu. \tag{1.1}$$

Date: March 28, 2018.

1991 *Mathematics Subject Classification.* Primary 60E15, 60F10; secondary 60B20.

Key words and phrases. Concentration of measure phenomenon, Poincare inequalities.

This research was supported by CRC 1283.

Here, $\mathrm{Var}_\mu(f) = \int f^2 \, d\mu - (\int f \, d\mu)^2$ denotes the variance. When considering σ instead of σ^2 itself, we will always assume it to be positive.

Given a function $f \in \mathcal{C}^d(G)$, we define $f^{(d)}$ to be the (hyper-) matrix whose entries

$$f^{(d)}_{i_1 \ldots i_d}(x) = \partial_{i_1 \ldots i_d} f(x), \qquad d = 1, 2, \ldots \tag{1.2}$$

represent the d-fold (continuous) partial derivatives of f at $x \in G$. By considering $f^{(d)}(x)$ as a symmetric multilinear d-form, we define operator-type norms by

$$|f^{(d)}(x)|_{\mathrm{Op}} = \sup \left\{ f^{(d)}(x)[v_1, \ldots, v_d] \colon |v_1| = \ldots |v_d| = 1 \right\}. \tag{1.3}$$

For instance, $|f^{(1)}(x)|_{\mathrm{Op}}$ is the Euclidean norm of the gradient $\nabla f(x)$, and $|f^{(2)}(x)|_{\mathrm{Op}}$ is the operator norm of the Hessian $f''(x)$. Furthermore, we will use the short-hand notation

$$\|f^{(d)}\|_{\mathrm{Op},p} = \left(\int_G |f^{(d)}|^p_{\mathrm{Op}} \, d\mu \right)^{1/p}, \qquad p \in (0, \infty]. \tag{1.4}$$

We now have the following:

Theorem 1.1. *Let μ be a probability measure on G satisfying a Poincaré-type inequality with constant $\sigma^2 > 0$, and let $f \colon G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function with $\int_G f \, d\mu = 0$. Assuming the conditions*

$$\|f^{(k)}\|_{\mathrm{Op},2} \le \sigma^{d-k} \qquad \forall k = 1, \ldots, d-1, \tag{1.5}$$

$$\|f^{(d)}\|_{\mathrm{Op},\infty} \le 1, \tag{1.6}$$

there exists some universal constant $c > 0$ such that

$$\int_G \exp\Big(\frac{c}{\sigma} |f|^{1/d}\Big) \, d\mu \ \le \ 2.$$

Here, a possible choice is $c = 1/(12\mathrm{e})$. Comparing Theorem 1.1 to its analogue in presence of a logarithmic Sobolev inequality, i. e. Theorem 1.6 in [B-G-S], we see that under the same assumptions (1.5) and (1.6), logarithmic Sobolev inequalities yield exponential moment bounds for $|f|^{2/d}$, whereas Poincaré-type inequalities provide exponential moments for $|f|^{1/d}$ only. This corresponds to the well-known behaviour in case of $d = 1$.

If f has centered partial derivatives of order up to $d - 1$, it is possible to replace (1.5) by a somewhat simpler condition. To this end, we need to involve Hilbert–Schmidt-type norms $|f^{(d)}(x)|_{\mathrm{HS}}$ defined as the Euclidean norm of $f^{(d)}(x) \in \mathbb{R}^{n^d}$. Similarly to (1.4), $\|f^{(d)}\|_{\mathrm{HS},2}$ then denotes the L^2-norm of $|f^{(d)}|_{\mathrm{HS}}$. In detail:

Theorem 1.2. *Let μ be a probability measure on G satisfying a Poincaré-type inequality with constant σ^2, and let $f \colon G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function such that*

$$\int_G f \, d\mu = 0 \qquad \text{and} \qquad \int_G \partial_{i_1 \ldots i_k} f \, d\mu = 0$$

for all $k = 1, \ldots, d-1$ and $1 \le i_1, \ldots, i_k \le n$. Assuming that

$$\|f^{(d)}\|_{\mathrm{HS},2} \le 1 \qquad \text{and} \qquad \|f^{(d)}\|_{\mathrm{Op},\infty} \le 1,$$

there exists some universal constant $c > 0$ such that

$$\int_G \exp\Big(\frac{c}{\sigma} |f|^{1/d}\Big) \, d\mu \le 2.$$

Here again, a possible choice is $c = 1/(12\mathrm{e})$.

By Chebyshev's inequality, Theorem 1.1 immediately yields

$$\mu(|f| \geq t) \leq 2\mathrm{e}^{-ct^{1/d}/\sigma}$$

for any $t \geq 0$. For small values of t, it is possible to obtain refined tail estimates in the spirit of R. Adamczak [A], Theorem 7, or R. Adamczak and P. Wolff [A-W], Theorem 3.3, by analyzing the proof of Theorem 1.1:

Corollary 1.3. *Let μ be a probability measure on G satisfying a Poincaré-type inequality with constant $\sigma^2 > 0$, and let $f\colon G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function with $\int_G f\, d\mu = 0$. For any $t \geq 0$, set*

$$\eta_f(t) := \min\Big(\frac{\sqrt{2}t^{1/d}}{\sigma\|f^{(d)}\|^{1/d}_{\mathrm{Op},\infty}}, \min_{k=1,\ldots,d-1} \frac{\sqrt{2}t^{1/k}}{\sigma\|f^{(k)}\|^{1/k}_{\mathrm{Op},2}}\Big).$$

Then,

$$\mu(|f| \geq t) \leq \mathrm{e}^2 \exp(-\eta_f(t)/(d\mathrm{e})).$$

As a generalization of these bounds, we may consider measures satisfying weighted Poincaré-type inequalities. Indeed, a Borel probability measure μ on an open set $G \subset \mathbb{R}^n$ is said to satisfy a *weighted Poincaré-type inequality* if for any bounded smooth function f on G with gradient ∇f,

$$\mathrm{Var}_\mu(f) \leq \int |\nabla f|^2 w^2\, d\mu, \tag{1.7}$$

where $w\colon G \to [0, \infty)$ is some measurable function. Examples include Cauchy measures and Beta distributions. For a detailed discussion see S. G. Bobkov and M. Ledoux [B-L2].

In these cases we cannot expect exponential integrability as in Theorem 1.1 any more, since distributions satisfying (1.7) may have a slow, say, polynomial, decay at infinity. Nevertheless, it is still possible to obtain higher order concentration results by controlling the L^p-norms of f and its derivatives. In detail:

Proposition 1.4. *Let μ be a probability measure on G satisfying a weighted Poincaré-type inequality* (1.7), *and let $f\colon G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function with $\int_G f d\mu = 0$. Then,*

$$\begin{aligned}\|f\|_p \;&\leq\; \sum_{k=1}^{d-1}(2^{\frac{k-2}{2}}p\|w\|_{2^kp})^k\,\|f^{(k)}\|_{\mathrm{Op},2} + (2^{\frac{d-2}{2}}p)^d\|w\|^{d-1}_{2^{d-1}p}\,\|w|f^{(d)}|_{\mathrm{Op}}\|_{2^{d-1}p}\\ &\leq\; \sum_{k=1}^{d-1}(2^{\frac{k-2}{2}}p\|w\|_{2^kp})^k\,\|f^{(k)}\|_{\mathrm{Op},2} + (2^{\frac{d-2}{2}}p\|w\|_{2^dp})^d\,\|f^{(d)}\|_{\mathrm{Op},2^dp}.\end{aligned}$$

Proposition 1.4 should be compared to (2.8) from the proof of Theorem 1.1 in Section 2. In particular, if the weight function w is bounded by some real number $\sigma > 0$, μ clearly satisfies a Poincaré-type inequality (1.1) with constant σ^2. In this case, Proposition 1.4 implies a slightly weaker version of (2.8), and it is possible to derive Theorem 1.1 again though with a somewhat weaker constant $c = c_d > 0$.

Suitable conditions on the weight function w may still yield exponential-type tails at least in certain intervals. For instance, the following higher order analogue of Corollary 4.2 in [B-L2] holds:

Corollary 1.5. *Let* μ *be a probability measure on* G *satisfying a weighted Poincaré-type inequality* (1.7)*, and let* $f\colon G \to \mathbb{R}$ *be a* $\mathcal{C}^d$*-smooth function with* $\int_G f d\mu = 0$ *and such that* (1.5) *and* (1.6) *from Theorem 1.1 hold. Assume* $\|w\|_{2^d p} \le C$ *for some* $p \ge 2$ *and some* $C \ge 2^{-(d-1)/2}$*. Then, for any* $0 \le t \le (2^{\frac{d+5}{2}} C \mathrm{e} p)^d$*,*

$$\mu(|f| \ge t) \le \mathrm{e}^{d/\mathrm{e}} \exp(-dt^{1/d}/(2^{\frac{d+5}{2}} C\mathrm{e})).$$

Hence, we obtain exponential-type tail bounds on an interval of length proportional to p^d. The assumption $C \ge 2^{-(d-1)/2}$ is needed for technical reasons. If $0 < C < 2^{-(d-1)/2}$, an inspection of the proof of Corollary 1.5 yields similar bounds e. g. by replacing C by $C^{1/d}$. Under stronger moment conditions on the weight function w, e. g. $\int \mathrm{e}^{w^2/\alpha} d\mu \le 2$ for some $\alpha > 0$, it is possible to obtain exponential-type tail bounds even on the whole positive half-line, cf. Corollary 4.3 in [B-L2].

Outline. In Section 2, we give the proofs of the results stated above. In Section 3, we provide some applications, including homogeneous multilinear polynomials of order d and linear eigenvalue statistics in random matrix theory.

2. Proofs

Given a continuous function on an open subset $G \subset \mathbb{R}^n$, the equality

$$|\nabla f(x)| = \limsup_{x \to y} \frac{|f(x) - f(y)|}{|x - y|}, \qquad x \in G, \tag{2.1}$$

may be used as definition of the generalized modulus of the gradient of f. The function $|\nabla f|$ is Borel measurable, and if f is differentiable at x, the generalized modulus of the gradient agrees with the Euclidean norm of the usual gradient. This operator preserves many identities from calculus in form of inequalities, such as a "chain rule inequality"

$$|\nabla T(f)| \le |T'(f)||\nabla f|, \tag{2.2}$$

where $|T'|$ is understood according to (2.1) again.

As shown in [B-G-S], Lemma 4.1, using the generalized modulus of the gradient, the derivatives of consecutive orders are related as follows:

Lemma 2.1. *Given a* $\mathcal{C}^d$*-smooth function* $f\colon G \to \mathbb{R}$*,* $d \in \mathbb{N}$*, at all points* $x \in G$*,*

$$|\nabla |f^{(d-1)}(x)|_{\mathrm{Op}}| \le |f^{(d)}(x)|_{\mathrm{Op}}.$$

Proof. Indeed, for any $h \in \mathbb{R}^n$, by the triangle inequality,

$$\begin{aligned} \big| |f^{(d-1)}(x+h)|_{\mathrm{Op}} - |f^{(d-1)}(x)|_{\mathrm{Op}} \big| \;&\le\; |f^{(d-1)}(x+h) - f^{(d-1)}(x)|_{\mathrm{Op}} \\ = \sup\{(f^{(d-1)}(x+h) - f^{(d-1)}(x))&[v_1, \ldots, v_{d-1}] \colon v_1, \ldots, v_{d-1} \in S^{n-1}\}, \end{aligned}$$

while, by the Taylor expansion,

$$(f^{(d-1)}(x+h) - f^{(d-1)}(x))[v_1, \ldots, v_{d-1}] = f^{(d)}(x)[v_1, \ldots, v_{d-1}, h] + o(|h|)$$

as $h \to 0$. Here, the o-term can be bounded by a quantity which is independent of $v_1, \ldots, v_{d-1} \in S^{n-1}$. As a consequence,

$$\begin{aligned} &\limsup_{h \to 0} \frac{\big| |f^{(d-1)}(x+h)|_{\mathrm{Op}} - |f^{(d-1)}(x)|_{\mathrm{Op}} \big|}{|h|} \\ &\le \sup\{f^{(d)}(x)[v_1, \ldots, v_{d-1}, v_d] \colon v_1, \ldots, v_d \in S^{n-1}\} = |f^{(d)}(x)|_{\mathrm{Op}}. \end{aligned}$$

□

Following the scheme of proof developed in [B-G-S], we moreover need to establish a recursion for the L^p-norms of the derivatives of f of consecutive orders. To this end, we recall a classical result on the moments of Lipschitz functions in the presence of Poincaré-type inequalities. In detail:

Lemma 2.2. *Let μ be a probability measure on G satisfying a Poincaré-type inequality with constant $\sigma^2 > 0$, and let $g \colon G \to \mathbb{R}$ be locally Lipschitz with $\int_G g d\mu = 0$. Then, for any $p \geq 2$,*

$$\int_G |g|^p d\mu \leq \Big(\frac{\sigma p}{\sqrt{2}}\Big)^p \int_G |\nabla g|^p d\mu. \tag{2.3}$$

In particular, for any $g \colon G \to \mathbb{R}$ locally Lipschitz,

$$\|g\|_p \ \leq \ \|g\|_2 + \frac{\sigma p}{\sqrt{2}} \, \|\nabla g\|_p. \tag{2.4}$$

Note that in (2.4), g is not required to have mean 0. For the reader's convenience, let us briefly recall the proof.

Proof. By standard arguments, we may assume g to be $\mathcal{C}^1$-smooth. Moreover, by the subadditivity property of the variance functional, the Poincaré-type inequality for the probability measure μ on G is extended to the same relation on $G \times G$, i. e.

$$\mathrm{Var}_{\mu^2}(u) \leq \sigma^2 \iint |\nabla u(x,y)|^2 d\mu(x) d\mu(y) \tag{2.5}$$

for the product measure $\mu^2 = \mu \otimes \mu$. Here, for any $\mathcal{C}^1$-smooth function $u = u(x,y)$, the modulus of the gradient is given by

$$|\nabla u(x,y)|^2 = |\nabla_x u(x,y)|^2 + |\nabla_y u(x,y)|^2.$$

Now consider the function

$$u(x,y) = |g(x) - g(y)|^{\frac{p}{2}} \mathrm{sign}(g(x) - g(y)),$$

which is $\mathcal{C}^1$-smooth for $p > 2$ with modulus of gradient

$$|\nabla u(x,y)| = \frac{p}{2} |g(x) - g(y)|^{\frac{p}{2}-1} \sqrt{|\nabla g(x)|^2 + |\nabla g(y)|^2}.$$

Since u has a symmetric distribution under μ^2, applying (2.5) together with Hölder's inequality yields

$$\begin{aligned}
&\frac{1}{\sigma^2} \iint |g(x) - g(y)|^p d\mu^2(x,y) \\
&\leq \frac{p^2}{4} \iint |g(x) - g(y)|^{p-2} \big(|\nabla g(x)|^2 + |\nabla g(y)|^2\big) d\mu^2(x,y) \\
&\leq \frac{p^2}{4} \Big(\iint |g(x) - g(y)|^p d\mu^2(x,y) \Big)^{\frac{p-2}{p}} \Big(\iint \big(|\nabla g(x)|^2 + |\nabla g(y)|^2\big)^{\frac{p}{2}} d\mu^2(x,y) \Big)^{\frac{2}{p}}.
\end{aligned}$$

By Jensen's inequality, the last integral may be bounded by

$$2^{\frac{p}{2}-1} \iint (|\nabla g(x)|^p + |\nabla g(y)|^p) d\mu^2(x,y) = 2^{\frac{p}{2}} \int |\nabla g|^p d\mu.$$

Consequently,

$$\Big(\iint |g(x) - g(y)|^p d\mu^2(x,y) \Big)^{\frac{2}{p}} \leq \frac{\sigma^2 p^2}{2} \Big(\int |\nabla g|^p d\mu \Big)^{\frac{2}{p}},$$

or, equivalently,

$$\iint |g(x)-g(y)|^p d\mu^2(x,y) \leq \Big(\frac{\sigma p}{\sqrt{2}}\Big)^p \int |\nabla g|^p d\mu.$$

If the right integral is finite, then so is the left one, which implies g is integrable. Moreover, if $\int g d\mu = 0$, it follows from Jensen's inequality that the left integral can be bounded below by $\int |g|^p d\mu$, which proves (2.3). To see (2.4), it remains to note that by the triangle inequality,

$$\Big\| g - \int g d\mu \Big\|_p \ \geq\ \|g\|_p - \Big|\int g d\mu\Big| \ \geq\ \|g\|_p - \|g\|_2.$$

$\square$

Combining Lemma 2.1 and (2.4), we are able to prove Theorem 1.1. Recall that if a relation of the form

$$\|f\|_k \leq \gamma k \qquad (k \in \mathbb{N}) \tag{2.6}$$

holds true with some constant $\gamma > 0$, then f has sub-exponential tails, i. e. $\int \mathrm{e}^{c|f|} d\mu \leq 2$ for some constant $c = c(\gamma) > 0$, e. g. $c = \frac{1}{2\gamma \mathrm{e}}$. Indeed, using $k! \geq (\frac{k}{\mathrm{e}})^k$, we have

$$\int \exp(c|f|) d\mu = 1 + \sum_{k=1}^\infty c^k \frac{\int |f|^k d\mu}{k!} \leq 1 + \sum_{k=1}^\infty (c\gamma)^k \frac{k^k}{k!} \leq 1 + \sum_{k=1}^\infty (c\gamma \mathrm{e})^k = 2.$$

Proof of Theorem 1.1. Using (2.4) with f replaced by $|f^{(k-1)}|_{\mathrm{Op}}$, $2 \leq k \leq d$, we get

$$\begin{aligned} \|f^{(k-1)}\|_{\mathrm{Op},p} &\leq \|f^{(k-1)}\|_{\mathrm{Op},2} + \frac{\sigma p}{\sqrt{2}}\, \|\nabla |f^{(k-1)}|_{\mathrm{Op}}\|_p \\ &\leq \|f^{(k-1)}\|_{\mathrm{Op},2} + \frac{\sigma p}{\sqrt{2}}\, \|f^{(k)}\|_{\mathrm{Op},p}, \end{aligned} \tag{2.7}$$

where Lemma 2.1 was applied on the last step. Consequently, using (2.3) and then (2.7) iteratively,

$$\|f\|_p \ \leq\ \sum_{k=1}^{d-1} \Big(\frac{\sigma p}{\sqrt{2}}\Big)^k \|f^{(k)}\|_{\mathrm{Op},2} + \Big(\frac{\sigma p}{\sqrt{2}}\Big)^d \|f^{(d)}\|_{\mathrm{Op},p}. \tag{2.8}$$

Since $\|f^{(k)}\|_{\mathrm{Op},2} \leq \sigma^{d-k}$ for all $k = 1, \ldots, d-1$ and $\|f^{(d)}\|_{\mathrm{Op},\infty} \leq 1$ by assumption, we obtain

$$\|f\|_p \ \leq\ \sigma^d \sum_{k=1}^d (p/\sqrt{2})^k \ \leq\ \frac{1}{1-(p/\sqrt{2})^{-1}}\, (\sigma p/\sqrt{2})^d \ \leq\ 4\, (\sigma p/\sqrt{2})^d \tag{2.9}$$

and therefore $\|f\|_p \leq (3\sigma p)^d$ for all $p \geq 2$. Moreover, $\|f\|_p \leq \|f\|_2 \leq (6\sigma)^d$ for $p < 2$. It follows that

$$\||f|^{1/d}\|_k = \|f\|_{k/d}^{1/d} \leq \gamma k$$

for all $k \in \mathbb{N}$, i. e. (2.6) with $\gamma = 6\sigma$. $\square$

Proof of Theorem 1.2. Starting as in the proof of Theorem 1.1, we arrive at

$$\|f\|_p \leq \sum_{k=1}^{d-1} (\sigma p/\sqrt{2})^k\, \|f^{(k)}\|_{\mathrm{HS},2} + (\sigma p/\sqrt{2})^d\, \|f^{(d)}\|_{\mathrm{Op},p}, \tag{2.10}$$

where we used that operator norms are dominated by Hilbert–Schmidt norms. Moreover, since $\int_G \partial_{i_1\dots i_k} f\, d\mu = 0$, by the Poincaré-type inequality,

$$\int_G (\partial_{i_1\dots i_k} f)^2\, d\mu \le \sigma^2 \sum_{j=1}^n \int_G (\partial_{i_1\dots i_k j} f)^2\, d\mu$$

whenever $1 \le i_1, \dots, i_k \le n$, $k \le d-1$. Summing over all $1 \le i_1, \dots, i_k \le n$, we get

$$\|f^{(k)}\|_{\mathrm{HS},2}^2 = \int_G |f^{(k)}|_{\mathrm{HS}}^2\, d\mu \le \sigma^2 \int_G |f^{(k+1)}|_{\mathrm{HS}}^2\, d\mu = \sigma^2\, \|f^{(k+1)}\|_{\mathrm{HS},2}^2. \tag{2.11}$$

Using (2.11) in (2.10) and iterating, we thus obtain

$$\|f\|_p \le \sum_{k=1}^{d-1} \sigma^d (p/\sqrt{2})^k\, \|f^{(d)}\|_{\mathrm{HS},2} + (\sigma p/\sqrt{2})^d\, \|f^{(d)}\|_{\mathrm{Op},p}.$$

Noting that $\|f^{(d)}\|_{\mathrm{HS},2} \le 1$ and $\|f^{(d)}\|_{\mathrm{Op},\infty} \le 1$, we arrive at (2.9), from where we may proceed as in the proof of Theorem 1.1. □

Proof of Corollary 1.3. First note that by Chebychev's inequality, for any $p \ge 1$

$$\mu(|f| \ge \mathrm{e}\, \|f\|_p) \le \mathrm{e}^{-p}. \tag{2.12}$$

Moreover, if $p \ge 2$, it follows from (2.8) that

$$\mathrm{e}\, \|f\|_p \le \ \mathrm{e}\Big(\sum_{k=1}^{d-1} (\sigma p/\sqrt{2})^k\, \|f^{(k)}\|_{\mathrm{Op},2} + (\sigma p/\sqrt{2})^d\, \|f^{(d)}\|_{\mathrm{Op},\infty}\Big).$$

Assuming $\eta_f(t) \ge 2$, we therefore arrive at

$$\mathrm{e}\, \|f\|_{\eta_f(t)} \le \mathrm{e}\big(\sum_{k=1}^{d-1} t + t\big) = (d\mathrm{e})t.$$

Hence, applying (2.12) to $p = \eta_f(t)$ (if $p \ge 2$) yields

$$\mu(|f| \ge (d\mathrm{e})t) \le \mu(|f| \ge \mathrm{e}\, \|f\|_{\eta_f(t)}) \le \exp(-\eta_f(t)).$$

Using a trivial estimate provided that $p < 2$, we obtain

$$\mu(|f| \ge (d\mathrm{e})t) \le \mathrm{e}^2 \exp(-\eta_f(t)).$$

The proof now easily follows by rescaling f by $d\mathrm{e}$ and using that $\eta_{d\mathrm{e}f}(t) \ge \eta_f(t)/(d\mathrm{e})$. □

In order to prove Proposition 1.4, we have to adapt the first steps of the proof of Theorem 1.1. First, we have the following generalization of Lemma 2.2 (in fact, this is a version of Theorem 4.1 in [B-L2]):

Lemma 2.3. *Let μ be a probability measure on G satisfying a weighted Poincaré-type inequality (1.7), and let $g\colon G \to \mathbb{R}$ be locally Lipschitz with $\int_G g d\mu = 0$. Then, for any $p \ge 2$,*

$$\int_G |g|^p d\mu \le \Big(\frac{p}{\sqrt{2}}\Big)^p \int_G |\nabla g|^p w^p\, d\mu. \tag{2.13}$$

In particular, for any $g\colon G \to \mathbb{R}$ locally Lipschitz,

$$\|g\|_p \ \le\ \|g\|_2 + \frac{p}{\sqrt{2}}\, \|w|\nabla g|\|_p. \tag{2.14}$$

The proof of Lemma 2.3 uses similar arguments as the proof of Lemma 2.2, and we therefore omit it. In particular, by Hölder's inequality, (2.14) implies

$$\|g\|_p \;\le\; \|g\|_2 + \frac{p}{\sqrt{2}}\,\|w\|_{2p}\|\nabla g\|_{2p}. \tag{2.15}$$

Starting with (2.13)–(2.15) and iterating as in (2.7) and (2.8), we obtain

$$\|f\|_p \;\le\; \sum_{k=1}^{d-1} 2^{\binom{k}{2}}\Big(\frac{p\|w\|_{2^k p}}{\sqrt{2}}\Big)^k\,\|f^{(k)}\|_{\mathrm{Op},2} + 2^{\binom{d}{2}}\Big(\frac{p\|w\|_{2^{d-1}p}}{\sqrt{2}}\Big)^d\,\|w|f^{(d)}|_{\mathrm{Op}}\|_{2^{d-1}p},$$

hence we easily arrive at the conclusions of Proposition 1.4. Again, we omit the details.

Finally, the proof of Corollary 1.5 is similar to the proof of Corollary 4.2 in [B-L2].

Proof of Corollary 1.5. First let $2 \le q \le p$. Using the assumptions and Proposition 1.4, we arrive at

$$\|f\|_q \;\le\; \sum_{k=1}^{d-1}(2^{\frac{k-2}{2}}qC)^k + (2^{\frac{d-2}{2}}qC)^d$$

and hence

$$\|f\|_q \le 4\,(2^{\frac{d-1}{2}}Cq)^d \le (2^{\frac{d+3}{2}}Cq)^d$$

(this follows as in (2.9), substituting σ by $2^{\frac{d-1}{2}}C \ge 1$). Moreover, if $0 < q \le 2$, we have

$$\|f\|_q \le \|f\|_2 \le (2^{\frac{d+5}{2}}C)^d.$$

Since the function $q \mapsto \mathrm{e}^{d/\mathrm{e}}q^{dq}$, $q > 0$, is minimized at $q = 1/\mathrm{e}$ with minimum value 1, it follows that $\mathbb{E}|f|^q \le \mathrm{e}^{d/\mathrm{e}}\,(2^{\frac{d+5}{2}}Cq)^{dq}$ for all $0 < q \le p$. Therefore, for any $t > 0$ and any $0 < q \le p$,

$$\mu(|f| \ge t) \le \frac{\mathbb{E}|f|^q}{t^q} \le \mathrm{e}^{d/\mathrm{e}}\left(\frac{(2^{\frac{d+5}{2}}Cq)^d}{t}\right)^q.$$

Now set $s = t^{1/d}/(2^{\frac{d+5}{2}}C)$ and write $\mu(|f| \ge t) \le \mathrm{e}^{d/\mathrm{e}}\mathrm{e}^{-\varphi(q)}$ with $\varphi(q) = dq(\log(s) - \log(q))$. It is easy to check that φ is a convex function on $(0, \infty)$ which attains its maximum at $q_0 = s/\mathrm{e}$ with $\varphi(q_0) = ds/\mathrm{e} = dt^{1/d}/(2^{\frac{d+5}{2}}C\mathrm{e})$. Noting that $q_0 \le p$ is equivalent with $t \le (2^{\frac{d+5}{2}}C\mathrm{e}p)^d$ completes the proof. □

3. APPLICATIONS

Let $X_1, \ldots, X_n$ be independent random variables with distributions satisfying a Poincaré-type inequality (1.1) with common constant $\sigma^2 > 0$. For real numbers $a_{i_1\ldots i_d}$, $i_1 < \ldots < i_d$, consider the function

$$f(X_1, \ldots, X_n) := \sum_{i_1<\ldots<i_d} a_{i_1\ldots i_d}X_{i_1}\cdots X_{i_d}, \tag{3.1}$$

which is a homogeneous multilinear polynomial of order d. For any $i_1 < \ldots < i_d$ and any permutation $\sigma \in S^d$, set $a_{\sigma(i_1)\ldots\sigma(i_d)} \equiv a_{i_1\ldots i_d}$. Moreover, set $a_{i_1\ldots i_d} = 0$ whenever the indexes $i_1, \ldots, i_d$ are not pairwise different. This gives rise to a hypermatrix $A = (a_{i_1\ldots i_d}) \in \mathbb{R}^{n^d}$, whose Euclidean norm we denote by $\|A\|_{\mathrm{HS}}$. Moreover, set $\|A\|_\infty := \max_{i_1<\ldots<i_d}|a_{i_1\ldots i_d}|$.

As a first example, we may apply our results to functions of type (3.1). Here it is convenient to assume for the random variables X_i to have mean zero:

Proposition 3.1. *Let $X_1, \ldots, X_n$ be independent random variables with distributions satisfying a Poincaré-type inequality* (1.1) *with common constant $\sigma^2 > 0$. Assume $\mathbb{E}X_i = 0$ for all $i = 1, \ldots, n$. Let $d \in \mathbb{N}$, and consider a function f of type* (3.1)*. Then,*

$$\mathbb{E}\exp\Big(\frac{c}{\sigma \|A\|_{\mathrm{HS}}^{1/d}} |f|^{1/d}\Big) \le 2.$$

Here, $\mathbb{E}$ denotes the expectation with respect to the random variables $X_1, \ldots, X_n$, and c is the absolute constant appearing in Theorem 1.2. In particular,

$$\mathbb{E}\exp\Big(\frac{c}{\sigma n^{1/2} \|A\|_{\infty}^{1/d}} |f|^{1/d}\Big) \le 2.$$

Moreover, if $\mathbb{E}X_i^2 = 1$ for all $i = 1, \ldots, n$,

$$\begin{aligned}\mathbb{P}(|f - \mathbb{E}f| \ge t) &\le \mathrm{e}^2 \exp\Big(-\frac{\sqrt{2}}{\sigma d \mathrm{e}} \min\Big(\frac{t}{\|A\|_{\mathrm{HS}}}, \frac{t^{1/d}}{\|A\|_{\mathrm{HS}}^{1/d}}\Big)\Big) \\ &\le \mathrm{e}^2 \exp\Big(-\frac{\sqrt{2}}{\sigma d \mathrm{e}} \min\Big(\frac{t}{n^{d/2}\|A\|_{\infty}}, \frac{t^{1/d}}{n^{1/2}\|A\|_{\infty}^{1/d}}\Big)\Big).\end{aligned}$$

Proposition 3.1 follows immediately from Theorem 1.2 and Corollary 1.3. Note that for non-centered random variables $X_1, \ldots, X_n$, applying Proposition 3.1 to the random variables $X_i - \mathbb{E}X_i$ means removing certain “lower order” terms in (3.1), which is in accordance with the ideas sketched in the introduction.

We may furthermore apply our results in the context of random matrix theory. Here we extend an example on second order concentration bounds for linear eigenvalue statistics in presence of a logarithmic Sobolev inequality [G-S], Proposition 1.10, to the situation where only a Poincaré-type inequality is available.

Indeed, let $\{\xi_{jk}, 1 \le j \le k \le N\}$ be a family of independent random variables on some probability space. Assume that the distributions of the ξ_{jk}'s all satisfy a (one-dimensional) Poincaré-type inequality (1.1) with common constant σ^2. Put $\xi_{jk} = \xi_{kj}$ for $1 \le k < j \le N$ and consider a symmetric $N \times N$ random matrix $\Xi = (\xi_{jk}/\sqrt{N})_{1 \le j,k \le N}$ and denote by $\mu^{(N)}$ the joint distribution of its ordered eigenvalues $\lambda_1 \le \ldots \le \lambda_N$ on $\mathbb{R}^N$ (in fact, $\lambda_1 < \ldots < \lambda_N$ a.s.). Recall that by a simple argument using the Hoffman–Wielandt theorem, $\mu^{(N)}$ satisfies a Poincaré-type inequality with constant

$$\sigma_N^2 = \frac{2\sigma^2}{N} \tag{3.2}$$

(see for instance S. G. Bobkov and F. Götze [B-G3]). Note that similar observations also hold for Hermitean random matrices.

Considering the probability space $(\mathbb{R}^N, \mathbb{B}^N, \mu^{(N)})$, if $f \colon \mathbb{R} \to \mathbb{R}$ is a $\mathcal{C}^1$-smooth function, it is well-known that asymptotic normality

$$S_N = \sum_{j=1}^{N} (f(\lambda_j) - \mathbb{E}f(\lambda_j)) \Rightarrow \mathcal{N}(0, \sigma_f^2) \tag{3.3}$$

holds for the self-normalized linear eigenvalue statistics S_N. Here, “$\Rightarrow$” denotes weak convergence, $\mathbb{E}$ means taking the expectation with respect to $\mu^{(N)}$ and $\mathcal{N}(0, \sigma_f^2)$ denotes a normal distribution with mean zero and variance σ_f^2 depending on f. This result was established K. Johansson [J] for the case of β-ensembles and, for general

Wigner matrices, A. M. Khorunzhy, B. A. Khoruzhenko and L. A. Pastur [K-K-P] as well as Ya. Sinai and A. Soshnikov [S-S]. Concentration of measure results have been studied by A. Guionnet and O. Zeitouni [G-Z], in particular proving fluctuations of order $\mathcal{O}_{\mathbb{P}}(1)$. Our results yield a second order concentration bound:

Proposition 3.2. *Let $\mu^{(N)}$ be the joint distribution of the ordered eigenvalues of Ξ. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^2$-smooth function with $f'(\lambda_j) \in L^1(\mu^{(N)})$ and bounded second derivatives, and let*

$$\tilde{S}_N := S_N - \sum_{j=1}^N (\lambda_j - \mathbb{E}(\lambda_j))\mathbb{E}f'(\lambda_j)$$

with S_N as in (3.3)*. Then, we have*

$$\mathbb{E}\exp\Big(\frac{cN^{1/4}}{\sqrt{2}\sigma\|f''\|_\infty^{1/2}}|\tilde{S}_N|^{1/2}\Big) \le 2,$$

where $c > 0$ is the absolute constant from Theorem 1.2.

Since $\tilde{S}_N$ is "centered" in the sense of Theorem 1.2, Proposition 3.2 immediately follows from elementary calculus, using (3.2). Note that in view of the self-normalizing property of S_N, the fluctuation result for $\tilde{S}_N$ is of the next order, although the scaling is of order $\sqrt{N}$ only. Comparing Proposition 3.2 to [G-S], Proposition 1.10, we see that we essentially arrive at the same result though for $|\tilde{S}_N|^{1/2}$ instead of $|\tilde{S}_N|$ due to the assumption of a Poincaré-type inequality.

Using Corollary 1.3, we can in fact slightly sharpen the results on the tail behavior of S_N. Indeed, an easy calculation yields

$$\mu_N(|S_N| \ge t) \le \ \mathrm{e}^2 \exp\Big(-\frac{1}{\sigma d\mathrm{e}}\min\Big(\frac{tN^{1/2}}{(\int\sum_i (f'(\lambda_i))^2 d\mu_N)^{1/2}}, \frac{t^{1/2}N^{1/4}}{\|f''\|_\infty^{1/2}}\Big)\Big)$$

for any $t \ge 0$. Similar results may be obtained for higher orders $d \ge 3$.

References

[A] Adamczak, R.: *Moment inequalities for U-statistics.* The Annals of Probability 34(6) (2006), 2288–2314.

[A-W] Adamczak, R., Wolff, P.: *Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order.* Probability Theory Related Fields 162(3) (2015), 531–586.

[A-S] Aida, S., Stroock, D.: *Moment estimates derived from Poincaré and logarithmic Sobolev inequalities.* Math. Res. Lett. 1(1) (1994), 75–86.

[B] Bobkov, S. G.: *The growth of L^p-norms in presence of logarithmic Sobolev inequalities.* Vestnik Syktyvkar Univ. 11(1) (2010), 92–111.

[B-C-G] Bobkov, S. G., Chistyakov, G. P., Götze, F.: *Second Order Concentration on the Sphere.* Commun. Contemp. Math. 19(5) (2017), 1650058, 20 pp.

[B-G1] Bobkov, S. G., Götze, F.: *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.* J. Func. Anal. 163(1) (1999), 1–28.

[B-G2] Bobkov, S. G., Götze, F.: *Concentration inequalities and limit theorems for randomized sums.* Probability Theory Related Fields 137 (2007), 49–81.

[B-G3] Bobkov, S. G., Götze, F.: *Concentration of empirical distribution functions with applications to non-i.i.d. models.* Bernoulli 16(4) (2010), 1385–1414.

[B-G-S] Bobkov, S. G., Götze, F., Sambale, H.: *Higher order concentration of measure.* Preprint, arXiv:1709.06838.

[B-L1] Bobkov, S. G., Ledoux, M.: *From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities.* Geom. Funct. Anal. 10(5) (2000), 1028–1052.

[B-L2] Bobkov, S. G., Ledoux, M.: *Weighted Poincaré-type inequalities for Cauchy and other convex measures.* The Annals of Probability 37(2) (2009), 403–427.

[B-B-L-M] Boucheron, S., Bousquet, O., Lugosi, G., Massart, P.: *Moment inequalities for functions of independent random variables.* The Annals of Probability 33(2) (2005), 514–560.

[B-L-M] Boucheron, S., Lugosi, G., Massart, P.: *Concentration Inequalities. A Nonasymptotic Theory of Independence.* Oxford University Press, 2013.

[CE-F-M] Cordero-Erausquin, D., Fradelizi, M., Maurey, B.: *The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems.* J. Funct. Anal. 214(2) (2004), 410–427.

[G-S] Götze, F., Sambale, H.: *Second order concentration via logarithmic Sobolev inequalities.* Preprint, arXiv:1605.08635.

[G-Z] Guionnet, A., Zeitouni, O.: *Concentration of the spectral measure for large matrices.* Electron. Commun. Prob. 5 (2000), 119–136.

[Gr] Gross, L.: *Logarithmic Sobolev inequalities.* Amer. J. Math 97(4) (1975), 1061–1083.

[J] Johansson, K.: *On fluctuations of eigenvalues of random Hermitean matrices.* Duke Math. J. 91(1) (1998), 151–204.

[K-K-P] Khorunzhy, A. M., Khoruzhenko, B. A., Pastur, L. A.: *Asymptotic properties of large random matrices with independent entries.* J. Math. Phys. 37 (1996), 5033-5060.

[L-O] Latała, R., Oleszkiewicz, K.: *Between Sobolev and Poincaré. Geometric aspects of functional analysis.* Lecture Notes in Math. 1745, 147–168. Springer, 2000.

[L1] Ledoux, M.: *On Talagrand's deviation inequalities for product measures.* ESAIM Prob. & Stat. 1 (1996), 63–87.

[L2] Ledoux, M.: *Concentration of measure and logarithmic Sobolev inequalities. Séminaire de Probabilités XXXIII.* Lecture Notes in Math. 1709, 120–216. Springer, 1999.

[L3] Ledoux, M.: *The Concentration of Measure Phenomenon.* American Mathematical Society, 2001.

[S-S] Sinai, Ya., Soshnikov, A.: *A central limit theorem for traces of large random matrices with independent matrix elements.* Bol. Soc. Brasil. Mat (N.S.) 29 (1998), 1–24.

[W] Wolff, P.: *On some Gaussian concentration inequality for non-Lipschitz functions.* High Dimensional Probability VI, Progress in Probability 66, 103–110.

Friedrich Götze, Faculty of Mathematics, Bielefeld University, Bielefeld, Germany
E-mail address: `goetze@math.uni-bielefeld.de`

Holger Sambale, Faculty of Mathematics, Bielefeld University, Bielefeld, Germany
E-mail address: `hsambale@math.uni-bielefeld.de`