Chapter 6
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.

Keywords Concentration of measure phenomenon · Poincaré inequalities

1991 Mathematics Subject Classification Primary 60E15, 60F10; Secondary 60B20

6.1 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 [2], 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). This approach was continued by Adamczak, Bednorz and Wolff for measures satisfying modified logarithmic Sobolev inequalities in [3]. While in [2], concentration around the mean is studied, the idea of sharpening concentration inequalities for Gaussian and related measures by requiring orthogonality

F. Götze (✉) · H. Sambale
Faculty of Mathematics, Bielefeld University, Bielefeld, Germany
e-mail: goetze@math.uni-bielefeld.de; hsambale@math.uni-bielefeld.de

N. Gozlan et al. (eds.), *High Dimensional Probability VIII*,
Progress in Probability 74, https://doi.org/10.1007/978-3-030-26391-1_6

to linear functions also appears in Wolff [16] as well as in Cordero-Erausquin et al. [9]. For a detailed overview of the concentration of measure phenomenon, see [8, 14].

Our research started with second order results for functions on the n-sphere orthogonal to linear functions [6], with an approach which has been extended in [10] 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 [7].

While in [7], 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,

$$\mathrm{Var}_\mu(f) \le \sigma^2 \int |\nabla f|^2 \, d\mu. \tag{6.1}$$

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

$$f^{(d)}_{i_1 \ldots i_d}(x) = \partial_{i_1 \ldots i_d} f(x), \qquad d = 1, 2, \ldots \tag{6.2}$$

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

$$|f^{(d)}(x)|_{\mathrm{Op}} = \sup\left\{ f^{(d)}(x)[v_1, \ldots, v_d] \colon |v_1| = \ldots |v_d| = 1 \right\}. \tag{6.3}$$

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

$$\|f^{(d)}\|_{\mathrm{Op},p} = \left(\int_G |f^{(d)}|^p_{\mathrm{Op}} \, d\mu \right)^{1/p}, \qquad p \in (0, \infty]. \tag{6.4}$$

For $p = \infty$, the right-hand side has to be read as the L^∞-norm of $|f^{(d)}|_{\mathrm{Op}}$.

We now have the following:

Theorem 6.1.1 *Let μ be a probability measure on an open set $G \subset \mathbb{R}^n$ 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*

$$\|f^{(k)}\|_{\mathrm{Op},2} \le \sigma^{d-k} \qquad \forall k = 1, \dots, d-1, \tag{6.5}$$

$$\|f^{(d)}\|_{\mathrm{Op},\infty} \le 1, \tag{6.6}$$

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 6.1.1 to its analogue in presence of a logarithmic Sobolev inequality, i.e. Theorem 1.6 in [7], we see that under the same assumptions (6.5) and (6.6), 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 (6.5) 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 (6.4), $\|f^{(d)}\|_{\mathrm{HS},2}$ then denotes the L^2-norm of $|f^{(d)}|_{\mathrm{HS}}$. In detail:

Theorem 6.1.2 *Let* μ *be a probability measure on an open set* $G \subset \mathbb{R}^n$ *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*

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

for all $k = 1, \dots, d-1$ *and* $1 \le i_1, \dots, i_k \le n$*. 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 6.1.1 immediately yields

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

for any $t \ge 0$. For small values of t, it is possible to obtain refined tail estimates in the spirit of Adamczak [1], Theorem 7, or Adamczak and Wolff [2], Theorem 3.3 (with $\gamma = 1$ using their notation), by analyzing the proof of Theorem 6.1.1:

Corollary 6.1.3 *Let μ be a probability measure on an open set $G \subset \mathbb{R}^n$ satisfying a Poincaré-type inequality with constant $\sigma^2 > 0$, and let $f : G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function with $\int_G f \, d\mu = 0$. For any $t \geq 0$, set*

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

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. Recall that 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,

$$\mathrm{Var}_\mu(f) \leq \int |\nabla f|^2 w^2 \, d\mu, \tag{6.7}$$

where $w : G \to [0, \infty)$ is some measurable function. Examples include Cauchy measures and Beta distributions. For a detailed discussion see Bobkov and Ledoux [5].

In these cases we cannot expect exponential integrability as in Theorem 6.1.1 any more, since distributions satisfying (6.7) 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 6.1.4 *Let μ be a probability measure on an open set $G \subset \mathbb{R}^n$ satisfying a weighted Poincaré-type inequality (6.7), and let $f : G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function with $\int_G f d\mu = 0$. Then, for any $p \geq 2$,*

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

Proposition 6.1.4 should be compared to (6.15) from the proof of Theorem 6.1.1 in Sect. 6.2. In particular, if the weight function w is bounded by some real number $\sigma > 0$, μ clearly satisfies a Poincaré-type inequality (6.1) with constant σ^2. In this case, Proposition 6.1.4 implies a slightly weaker version of (6.15), and it is possible to derive Theorem 6.1.1 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 [5] holds:

Corollary 6.1.5 *Let μ be a probability measure on an open set $G \subset \mathbb{R}^n$ satisfying a weighted Poincaré-type inequality (6.7), and let $f: G \to \mathbb{R}$ be a $\mathcal{C}^d$-smooth function with $\int_G f d\mu = 0$ and such that (6.5) (with $\sigma^2 = 1$) and (6.6) from Theorem 6.1.1 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. Note that if $t > (2^{\frac{d+5}{2}} C\mathrm{e}\, p)^d$, we may still give bounds on $\mu(|f| \ge t)$ by taking (6.23) for $q = p$ from the proof of Corollary 6.1.5. We omit details at this point. The assumption $C \ge 2^{-(d-1)/2}$ is needed for technical reasons. In fact, it guarantees that the quantities $(2^{\frac{k-1}{2}} C)^k$, $k \le d-1$, are bounded by $(2^{\frac{d-1}{2}} C)^d$. For $d = 1$ it can be removed. It is possible to adapt the proof for $0 < C < 2^{-(d-1)/2}$ and obtain similar bounds.

For $d = 1$, Corollary 6.1.5 gives back a version of Corollary 4.2 from [5] up to constants, though with a boundedness condition on $\|w\|_{2p}$ rather than $\|w\|_p$. This may be adjusted by working with the first inequality from Proposition 6.1.4, in which case we directly get back the [5] result. In the same way, it is possible to derive a result similar to Corollary 6.1.5 which requires bounds on $\|w\|_{2^{d-1}p}$. We have chosen to work with the second inequality from Proposition 6.1.4 instead (and thus need bounds on $\|w\|_{2^d p}$) since this is technically slightly more convenient.

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 [5].

Outline In Sect. 6.2, we give the proofs of the results stated above. In Sect. 6.3, we provide some applications, including homogeneous multilinear polynomials of order d and linear eigenvalue statistics in random matrix theory.

6.2 Proofs

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

$$|\nabla f(x)| = \limsup_{x \to y} \frac{|f(x) - f(y)|}{|x - y|}, \qquad x \in G, \tag{6.8}$$

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"

$$|\nabla T(f)| \le |T'(f)||\nabla f|, \tag{6.9}$$

where $|T'|$ is understood according to (6.8) again.

As shown in [7], Lemma 4.1, using the generalized modulus of the gradient, the operator norms of the derivatives of consecutive orders are related as follows:

Lemma 6.2.1 *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,

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

while, by the Taylor expansion,

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

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

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

□

Following the scheme of proof developed in [7], 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. Here, similarly to (6.4), we write

$$\|\nabla g\|_{\mathrm{Op},p} = \left(\int_G |\nabla g|^p \, d\mu\right)^{1/p}, \qquad p \in (0, \infty],$$

for any locally Lipschitz function g on G with generalized modulus of gradient $|\nabla g|$. In detail:

Lemma 6.2.2 *Let μ be a probability measure on an open set $G \subset \mathbb{R}^n$ 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 $p \geq 2$,*

$$\int_G |g|^p d\mu \leq \Big(\frac{\sigma p}{\sqrt{2}}\Big)^p \int_G |\nabla g|^p d\mu. \tag{6.10}$$

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

$$\|g\|_p \;\leq\; \|g\|_2 + \frac{\sigma p}{\sqrt{2}}\, \|\nabla g\|_p. \tag{6.11}$$

Note that in (6.11), 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 and bounded. 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.

$$\mathrm{Var}_{\mu^2}(u) \leq \sigma^2 \iint |\nabla u(x,y)|^2 d\mu(x) d\mu(y) \tag{6.12}$$

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

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

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 (6.12) 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,

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

or, equivalently,

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

In particular, the latter inequality shows that any locally Lipschitz function g such that the right-hand side is finite is integrable (if g is unbounded, we may perform a simple truncation argument). 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 (6.10). To see (6.11), it remains to note that by the triangle inequality,

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

□

Combining Lemma 6.2.1 and (6.11), we are able to prove Theorem 6.1.1. Recall that if a relation of the form

$$\|f\|_k \le \gamma k \qquad (k \in \mathbb{N}) \tag{6.13}$$

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

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

Proof of Theorem 6.1.1 Using (6.11) with f replaced by $|f^{(k-1)}|_{\mathrm{Op}}$, $2 \le k \le d$, we get

$$\begin{aligned} \|f^{(k-1)}\|_{\mathrm{Op},p} &\le \|f^{(k-1)}\|_{\mathrm{Op},2} + \frac{\sigma p}{\sqrt{2}} \|\nabla |f^{(k-1)}|_{\mathrm{Op}}\|_p \\ &\le \|f^{(k-1)}\|_{\mathrm{Op},2} + \frac{\sigma p}{\sqrt{2}} \|f^{(k)}\|_{\mathrm{Op},p}, \end{aligned} \tag{6.14}$$

where Lemma 6.2.1 was applied on the last step. Consequently, using (6.10) and then (6.14) iteratively,

$$\|f\|_p \ \leq \ \sum_{k=1}^{d-1} \Big(\frac{\sigma p}{\sqrt{2}}\Big)^k \ \|f^{(k)}\|_{\mathrm{Op},2} + \Big(\frac{\sigma p}{\sqrt{2}}\Big)^d \ \|f^{(d)}\|_{\mathrm{Op},p}. \tag{6.15}$$

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

$$\|f\|_p \ \leq \ \sigma^d \sum_{k=1}^{d} (p/\sqrt{2})^k \ \leq \ \frac{1}{1-(p/\sqrt{2})^{-1}} \, (\sigma p/\sqrt{2})^d \ \leq \ 4\,(\sigma p/\sqrt{2})^d \tag{6.16}$$

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. (6.13) holds with $\gamma = 6\sigma$ (and $|f|^{1/d}$ in place of f). This yields the assertion of the theorem. □

Proof of Theorem 6.1.2 Starting as in the proof of Theorem 6.1.1, we arrive at

$$\|f\|_p \leq \sum_{k=1}^{d-1} (\sigma p/\sqrt{2})^k \ \|f^{(k)}\|_{\mathrm{HS},2} + (\sigma p/\sqrt{2})^d \ \|f^{(d)}\|_{\mathrm{Op},p}, \tag{6.17}$$

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,

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

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

$$\|f^{(k)}\|_{\mathrm{HS},2}^2 = \int_G |f^{(k)}|_{\mathrm{HS}}^2 \, d\mu \leq \sigma^2 \int_G |f^{(k+1)}|_{\mathrm{HS}}^2 \, d\mu = \sigma^2 \, \|f^{(k+1)}\|_{\mathrm{HS},2}^2. \tag{6.18}$$

Using (6.18) in (6.17) 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 (6.16), from where we may proceed as in the proof of Theorem 6.1.1. □

Proof of Corollary 6.1.3 First note that by Chebyshev's inequality, for any $p \ge 1$

$$\mu(|f| \ge \mathrm{e}\,\|f\|_p) \le \mathrm{e}^{-p}. \tag{6.19}$$

Moreover, if $p \ge 2$, it follows from (6.15) 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 (6.19) 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 $p = \eta_f(t) < 2$, we obtain

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

for all $t \ge 0$. 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 6.1.4, we have to adapt the first steps of the proof of Theorem 6.1.1. First, we have the following generalization of Lemma 6.2.2 (in fact, this is a version of Theorem 4.1 in [5]):

Lemma 6.2.3 *Let μ be a probability measure on an open set $G \subset \mathbb{R}^n$ satisfying a weighted Poincaré-type inequality* (6.7)*, and let $g\colon G \to \mathbb{R}$ be locally Lipschitz with $\int_G g d\mu = 0$. Then, for any $p \ge 2$,*

$$\int_G |g|^p d\mu \le \Big(\frac{p}{\sqrt{2}}\Big)^p \int_G |\nabla g|^p w^p \, d\mu. \tag{6.20}$$

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

$$\|g\|_p \;\le\; \|g\|_2 + \frac{p}{\sqrt{2}}\,\|w|\nabla g|\|_p. \tag{6.21}$$

The proof of Lemma 6.2.3 uses similar arguments as the proof of Lemma 6.2.2, and we therefore omit it. In particular, by Hölder's inequality, (6.21) implies

$$\|g\|_p \;\le\; \|g\|_2 + \frac{p}{\sqrt{2}}\,\|w\|_{2p}\|\nabla g\|_{2p}. \tag{6.22}$$

Starting with (6.20)–(6.22) and iterating as in (6.14) and (6.15), 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 6.1.4. Again, we omit the details.

Finally, the proof of Corollary 6.1.5 is similar to the proof of Corollary 4.2 in [5].

Proof of Corollary 6.1.5 First let $2 \le q \le p$. Using the assumptions and Proposition 6.1.4, 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 (6.16), 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 $0 < q \le p$,

$$\mu(|f| \ge t) \le \frac{\mathbb{E}|f|^q}{t^q} \le \mathrm{e}^{d/\mathrm{e}}\left(\frac{(2^{\frac{d+5}{2}}Cq)^d}{t}\right)^q. \tag{6.23}$$

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 concave 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. □

6.3 Applications

Let $X_1, \ldots, X_n$ be independent random variables with distributions satisfying a Poincaré-type inequality (6.1) with common constant $\sigma^2 > 0$. For real numbers $a_{i_1 \ldots i_d}$, $i_1 < \ldots < i_d$, consider the function

$$f(X_1, \ldots, X_n) := \sum_{i_1 < \ldots < i_d} a_{i_1 \ldots i_d} X_{i_1} \cdots X_{i_d}, \tag{6.24}$$

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 (6.24). Here it is convenient to assume for the random variables X_i to have mean zero:

Proposition 6.3.1 *Let $X_1, \ldots, X_n$ be independent random variables with distributions satisfying a Poincaré-type inequality (6.1) 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 (6.24). 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 6.1.2. 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 6.3.1 follows immediately from Theorem 6.1.2 and Corollary 6.1.3. Note that for non-centered random variables $X_1, \dots, X_n$, applying Proposition 6.3.1 to the random variables $X_i - \mathbb{E}X_i$ means removing certain "lower order" terms in (6.24), 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 [10], 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 (6.1) 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

$$\sigma_N^2 = \frac{2\sigma^2}{N} \tag{6.25}$$

(see for instance Bobkov and Götze [4]). Note that similar observations also hold for Hermitian 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

$$S_N = \sum_{j=1}^{N} (f(\lambda_j) - \mathbb{E}f(\lambda_j)) \Rightarrow \mathcal{N}(0, \sigma_f^2) \tag{6.26}$$

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 by Johansson [12] for the case of β-ensembles and, for general Wigner matrices, by Khorunzhy et al. [13] as well as Sinai and Soshnikov [15]. Concentration of measure results have been studied by Guionnet and Zeitouni [11], in particular proving fluctuations of order $\mathcal{O}_{\mathbb{P}}(1)$. Our results yield a second order concentration bound:

Proposition 6.3.2 *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 (6.26). 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 6.1.2.

Since $\tilde{S}_N$ is "centered" in the sense of Theorem 6.1.2, Proposition 6.3.2 immediately follows from elementary calculus, using (6.25). 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 6.3.2 to [10], 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 6.1.3, 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 $d \ge 3$.

Acknowledgement This research was supported by CRC 1283.

References

1. R. Adamczak, Moment inequalities for U-statistics. Ann. Probab. **34**(6), 2288–2314 (2006)
2. R. Adamczak, P. Wolff, Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order. Probab. Theory Relat. Fields **162**(3), 531–586 (2015)
3. R. Adamczak, W. Bednorz, P. Wolff, Moment estimates implied by modified log-Sobolev inequalities. ESAIM Probab. Stat. **21**, 467–494 (2017)
4. S.G. Bobkov, F. Götze, Concentration of empirical distribution functions with applications to non-i.i.d. models. Bernoulli **16**(4), 1385–1414 (2010)
5. S.G. Bobkov, M. Ledoux, Weighted Poincaré-type inequalities for Cauchy and other convex measures. Ann. Probab. **37**(2), 403–427 (2009)
6. S.G. Bobkov, G.P. Chistyakov, F. Götze, Second order concentration on the sphere. Commun. Contemp. Math. **19**(5), 1650058, 20 pp. (2017)
7. S.G. Bobkov, F. Götze, H. Sambale, Higher order concentration of measure. Commun. Contemp. Math. **21**(3), 1850043, 36 (2019)
8. S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities. A Nonasymptotic Theory of Independence* (Oxford University Press, Oxford, 2013)
9. D. Cordero-Erausquin, M. Fradelizi, B. Maurey, The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems. J. Funct. Anal. **214**(2), 410–427 (2004)
10. F. Götze, H. Sambale, Second order concentration via logarithmic Sobolev inequalities. Bernoulli (to appear). arXiv:1605.08635 (preprint)
11. A. Guionnet, O. Zeitouni, Concentration of the spectral measure for large matrices. Electron. Commun. Prob. **5**, 119–136 (2000)

12. K. Johansson, On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. **91**(1), 151–204 (1998)
13. A.M. Khorunzhy, B.A. Khoruzhenko, L.A. Pastur, Asymptotic properties of large random matrices with independent entries. J. Math. Phys. **37**, 5033–5060 (1996)
14. M. Ledoux, *The Concentration of Measure Phenomenon* (American Mathematical Society, Providence, 2001)
15. Ya. Sinai, A. Soshnikov, A central limit theorem for traces of large random matrices with independent matrix elements. Bol. Soc. Brasil. Mat. **29**, 1–24 (1998)
16. P. Wolff, On some Gaussian concentration inequality for non-Lipschitz functions. High Dimens. Probab. VI **66**, 103–110 (2013)