IMPROVED BOUNDS IN THE METRIC COTYPE INEQUALITY FOR BANACH SPACES

OHAD GILADI, MANOR MENDEL, AND ASSAF NAOR

Abstract. It is shown that if $(X, \|\cdot\|_X)$ is a Banach space with Rademacher cotype q then for every integer n there exists an even integer $m \lesssim n^{1+\frac{1}{q}}$ such that for every $f : \mathbb{Z}_m^n \to X$ we have

$$\sum_{j=1}^{n} \mathbb{E}_x\left[\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q\right] \lesssim m^q \mathbb{E}_{\varepsilon,x}\Big[\|f(x+\varepsilon)-f(x)\|_X^q\Big], \qquad (1)$$

where the expectations are with respect to uniformly chosen $x \in \mathbb{Z}_m^n$ and $\varepsilon \in \{-1,0,1\}^n$, and all the implied constants may depend only on q and the Rademacher cotype q constant of X. This improves the bound of $m \lesssim n^{2+\frac{1}{q}}$ from [13]. The proof of (1) is based on a "smoothing and approximation" procedure which simplifies the proof of the metric characterization of Rademacher cotype of [13]. We also show that any such "smoothing and approximation" approach to metric cotype inequalities must require $m \gtrsim n^{\frac{1}{2}+\frac{1}{q}}$.

Contents

2010 *Mathematics Subject Classification.* 46B80,46B85,51F99,05C12.

1. Introduction

A metric space $(\mathscr{M}, d_{\mathscr{M}})$ is said [13] to have metric cotype $q > 0$ with constant $\Gamma > 0$ if for every integer n there exists an even integer m such that for every $f : \mathbb{Z}_m^n \to X$ we have

$$\sum_{j=1}^{n} \mathbb{E}_x \left[d_{\mathscr{M}} \left(f\left(x + \frac{m}{2}e_j\right), f(x) \right)^q \right] \leqslant \Gamma^q m^q \mathbb{E}_{\varepsilon,x} \Big[d_{\mathscr{M}} \left(f(x+\varepsilon), f(x) \right)^q \Big]. \tag{2}$$

In (2) the expectations are taken with respect to x chosen uniformly at random from the discrete torus $\mathbb{Z}_m^n$, and ε chosen uniformly at random from $\{-1,0,1\}^n$ (the ℓ_∞ generators of $\mathbb{Z}_m^n$). Also, in (2) and in what follows, $\{e_j\}_{j=1}^n$ denotes the standard basis of $\mathbb{Z}_m^n$.

A Banach space $(X, \|\cdot\|_X)$ is said to have Rademacher cotype $q > 0$ if there exists a constant $C < \infty$ such that for every $n \in \mathbb{N}$ and for every $x_1, x_2, \ldots., x_n \in X$,

$$\sum_{j=1}^{n} \|x_j\|_X^q \leqslant C^q \mathbb{E}_\varepsilon \left[\left\| \sum_{j=1}^{n} \varepsilon_j x_j \right\|_X^q \right]. \tag{3}$$

X is said to have Rademacher type $p > 0$ if there exists a constant $T < \infty$ such that for every $n \in \mathbb{N}$ and for every $x_1, x_2, \ldots., x_n \in X$,

$$\mathbb{E}_\varepsilon \left[\left\| \sum_{j=1}^{n} \varepsilon_j x_j \right\|_X^p \right] \leqslant T^p \sum_{j=1}^{n} \|x_j\|_X^p. \tag{4}$$

The smallest possible constants C, T in (3), (4) are denoted $C_p(X), T_p(X)$, respectively. We refer to [16, 9] for more information on the notions of type and cotype, though the present paper requires minimal background of this theory. We shall use throughout standard Banach space notation and terminology, as appearing in, say, [21].

The following theorem was proved in [13]:

Theorem 1.1 ([13]). *A Banach space $(X, \|\cdot\|_X)$ has Rademacher cotype q if and only if it has metric cotype q.*

Thus, for Banach spaces the *linear* notion of Rademacher cotype q is equivalent to the notion of metric cotype q, which ignores all the structure of the Banach space except for its metric properties. Theorem 1.1 belongs to a comprehensive program, first formulated by Bourgain in [2], which is known as the Ribe program, whose goal is to recast the local theory of Banach spaces as a purely metric theory. A byproduct of this program is that linear properties such as Rademacher cotype can be made to make sense in general metric spaces, with applications to metric geometry in situations which lack any linear structure. We refer to [13] and the references therein for more information on the Ribe program and its applications.

Definition (2) and Theorem 1.1 suppress the value of m, since it is irrelevant for the purpose of a metric characterization of Rademacher cotype. Nevertheless, good bounds on m are important for applications of metric cotype to embedding theory, some of which will be recalled in Section 1.3. It was observed in [13] that if the metric space $\mathscr{M}$ contains at least two points then the value of m in (2) must satisfy $m \gtrsim n^{1/q}$ (where the implied constant depends only on Γ). If X is a Banach space with Rademacher type $p > 1$ and Rademacher cotype q, then it was shown in [13] that X satisfies the metric cotype q inequality (2) for every $m \geqslant n^{1/q}$ (in which case Γ depends only on $p, q, T_p(X), C_q(X)$). Such a sharp bound on m is crucial for certain applications [13, 14] of metric cotype, and perhaps the most important open problem in [13] is whether this sharp bound on m holds true even when the condition that X has type $p > 1$ is dropped. The bound on m from [13] in Theorem 1.1 is $m \gtrsim n^{2+\frac{1}{q}}$. Our main result improves this bound to $m \gtrsim n^{1+\frac{1}{q}}$:

Theorem 1.2. *Let X be a Banach space with Rademacher cotype $q \geqslant 2$. Then for every $n \in \mathbb{N}$, every integer $m \geqslant 6n^{1+\frac{1}{q}}$ which is divisible by 4, and every $f : \mathbb{Z}_m^n \to X$, we have*

$$\sum_{j=1}^{n} \mathbb{E}_x \left[\left\| f\left(x + \frac{m}{2} e_j\right) - f(x) \right\|_X^q \right] \lesssim_X m^q \mathbb{E}_{\varepsilon,x} \Big[\| f(x+\varepsilon) - f(x) \|_X^q \Big]. \tag{5}$$

In (5), and in what follows, $\lesssim_X, \gtrsim_X$ indicate the corresponding inequalities up to constants which may depend only on q and $C_q(X)$. Similarly, we will use the notation $\lesssim_q, \gtrsim_q$ to indicate the corresponding inequalities up to constants which may depend only on q.

Though a seemingly modest improvement over the result of [13], the strengthened metric cotype inequality (5) does yield some new results in embedding theory, as well as a new proof of a result of Bourgain [3]; these issues are discussed in Section 1.3. More importantly, our proof of Theorem 1.2 is based on a better understanding and sharpening of the underlying principles behind the proof of Theorem 1.1 in [13]. As a result, we isolate here the key approach to the metric characterization of Rademacher cotype in [13], yielding a simpler and clearer proof of Theorem 1.1, in addition to the improved bound on m. This is explained in detail in Section 1.1. While the bound $m \gtrsim n^{1+\frac{1}{q}}$ is far from the conjectured optimal bound $m \geqslant n^{1/q}$, our second main result is that (a significant generalization of) the scheme for proving Theorem 1.1 and Theorem 1.2 (which is implicit in [13] and formulated explicitly here) cannot yield a bound better than $m \gtrsim n^{\frac{1}{2}+\frac{1}{q}}$. Our method for proving this lower bound is presented in Section 4, and might be of independent interest.

We remark in passing that in [13] a one parameter family of variants of the notion of metric cotype is studied, corresponding to raising the distances to powers other than q, and modifying the right-hand side of (2), (5) accordingly (we refer to [13] for more details). The argument presented here can be modified to yield simplifications and improvements of all the corresponding variants of Theorem 1.1. While these variants are crucial for certain applications of metric cotype [13, 11], we chose to present Theorem 1.2 only for the simplest "vanilla" version of metric cotype (2), for the sake of simplicity of exposition.

Notation for measures. Since our argument uses a variety of averaging procedures over several spaces, it will be convenient to depart from the expectation notation that we used thus far. In particular, throughout this paper μ will denote the uniform probability measure on $\mathbb{Z}_m^n$ (m, n will always be clear from the context), σ will denote the uniform probability measure on $\{-1,0,1\}^n$, and τ will denote the uniform probability measure on $\{-1,1\}^n$.

1.1. **The smoothing and approximation scheme.** We start with a description of an abstraction of the approach of [13] to proving the metric characterization of Rademacher cotype of Theorem 1.1.

For a Banach space X, a function $f : \mathbb{Z}_m^n \to X$ and a probability measure ν on $\mathbb{Z}_m^n$, we use the standard notation for the convolution $f * \nu : \mathbb{Z}_m^n \to X$:

$$f * \nu(x) = \int_{\mathbb{Z}_m^n} f(x-y) d\nu(y).$$

Assume that we are given n probability measures $\nu_1,\ldots,\nu_n$ on $\mathbb{Z}_m^n$, and two additional probability measures β_1,β_2 on the pairs in $\mathbb{Z}_m^n\times\mathbb{Z}_m^n$ of ℓ_∞ distance 1, i.e., on the set

$$E_\infty(\mathbb{Z}_m^n)\stackrel{\mathrm{def}}{=}\Big\{(x,y)\in\mathbb{Z}_m^n\times\mathbb{Z}_m^n:\ x-y\in\{-1,0,1\}^n\Big\}. \tag{6}$$

For $A,S,q\geqslant 1$, we shall say that the measures $\nu_1,\ldots,\nu_n,\beta_1,\beta_2$ are a (q,A,S)-*smoothing and approximation scheme* on $\mathbb{Z}_m^n$ if for every Banach space $(X,\|\cdot\|_X)$ and every $f:\mathbb{Z}_m^n\to X$ we have the following two inequalities:

(A) Approximation property:

$$\frac{1}{n}\sum_{j=1}^n\int_{\mathbb{Z}_m^n}\|f*\nu_j(x)-f(x)\|_X^q\,d\mu(x)\leqslant A^q\int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\beta_1(x,y). \tag{7}$$

(S) Smoothing property:

$$\int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{j=1}^n\varepsilon_j\Big(f*\nu_j(x+e_j)-f*\nu_j(x-e_j)\Big)\right\|_X^q d\tau(\varepsilon)d\mu(x)$$
$$\leqslant S^q\int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\beta_2(x,y). \tag{8}$$

Often, when the underlying space $\mathbb{Z}_m^n$ is obvious from the context, we will not mention it explicitly, and simply call $\nu_1,\ldots,\nu_n,\beta_1,\beta_2$ a (q,A,S)-smoothing and approximation scheme. In some cases, however, it will be convenient to mention the underlying space $\mathbb{Z}_m^n$ so as to indicate certain restrictions on m.

We introduce these properties for the following simple reason. We wish to deduce the metric cotype inequality (5) from the Rademacher cotype inequality (3). In essence, the Rademacher cotype condition (3) is the same as the metric cotype inequality (5) when restricted to *linear* mappings. This statement is not quite accurate, but it suffices for the purpose of understanding the intuition behind the ensuing argument; we refer to Section 5.1 in [13] for the precise argument. In any case, it stands to reason that in order to prove (5) from (3), we should first smooth out f, so that it will be locally well approximated (on average) by a linear function. As we shall see momentarily, it turns out that the appropriate way to measure the quality of such a smoothing procedure is our smoothing property (8). Of course, while the averaging operators corresponding to convolution with the measures $\nu_1,\ldots,\nu_n$ yield a better behaved function, we still need the resulting averaged function to be close enough to the original function f, so as to deduce a meaningful inequality such as (5) for f itself. Our approximation property (7) is what's needed for carrying out such an approach.

The above general scheme is implicit in [13]. Once we have isolated the crucial approximation and smoothing properties, it is simple to see how they relate to metric cotype. For this purpose, assume that the Banach space X has Rademacher cotype q, and for each $x\in\mathbb{Z}_m^n$ apply the Rademacher cotype q inequality to the vectors $\{f*\nu_j(x+e_j)-f*\nu_j(x-e_j)\}_{j=1}^n$ (where the averaging in (3) is with respect to $\varepsilon\in\{-1,1\}^n$, rather than $\varepsilon\in\{-1,0,1\}^n$; it is an easy standard fact that these two variants of Rademacher cotype q coincide):

$$\int_{\{-1,1\}^n} \left\| \sum_{j=1}^n \varepsilon_j \Big(f*\nu_j(x+e_j) - f*\nu_j(x-e_j) \Big) \right\|_X^q d\tau(\varepsilon)$$
$$\gtrsim_X \sum_{j=1}^n \|f*\nu_j(x+e_j) - f*\nu_j(x-e_j)\|_X^q. \quad (9)$$

The triangle inequality, combined with the convexity of the function $t \mapsto t^q$, implies that for every $x \in \mathbb{Z}_m^n$ and $j \in \{1,\ldots,n\}$ we have

$$\left\| f\left(x+\frac{m}{2}e_j\right) - f(x) \right\|_X^q \leqslant 3^{q-1} \left\| f*\nu_j\left(x+\frac{m}{2}e_j\right) - f*\nu_j(x) \right\|_X^q$$
$$+ 3^{q-1} \left\| f*\nu_j\left(x+\frac{m}{2}e_j\right) - f\left(x+\frac{m}{2}e_j\right) \right\|_X^q + 3^{q-1}\|f*\nu_j(x) - f(x)\|_X^q. \quad (10)$$

At the same time (recalling that m is divisible by 4), a combination of the triangle inequality and Hölder's inequality bounds the first term in the right hand side of (10) as follows:

$$\left\| f*\nu_j\left(x+\frac{m}{2}e_j\right) - f*\nu_j(x) \right\|_X^q \leqslant \left(\sum_{t=1}^{m/4} \|f*\nu_j(x+2te_j) - f*\nu_j(x+2(t-1)e_j)\|_X \right)^q$$
$$\leqslant \left(\frac{m}{4}\right)^{q-1} \sum_{t=1}^{m/4} \|f*\nu_j(x+2te_j) - f*\nu_j(x+2(t-1)e_j)\|_X^q. \quad (11)$$

Substituting (11) into (10), summing up over $j \in \{1,\ldots,n\}$, and integrating with respect to $x \in \mathbb{Z}_m^n$ while using the translation invariance of the measure μ, we deduce the inequality

$$\sum_{j=1}^n \int_{\mathbb{Z}_m^n} \left\| f\left(x+\frac{m}{2}e_j\right) - f(x) \right\|_X^q d\mu(x) \lesssim 3^q \sum_{j=1}^n \int_{\mathbb{Z}_m^n} \|f*\nu_j(x) - f(x)\|_X^q \, d\mu(x)$$
$$+ m^q \sum_{j=1}^n \int_{\mathbb{Z}_m^n} \|f*\nu_j(x+e_j) - f*\nu_j(x-e_j)\|_X^q \, d\mu(x). \quad (12)$$

We can now bound the first term in the right hand side of (12) using the approximation property (7), and the second term in the right hand side of (12) using (9) and the smoothing property (8). The inequality thus obtained is

$$\sum_{j=1}^n \int_{\mathbb{Z}_m^n} \left\| f\left(x+\frac{m}{2}e_j\right) - f(x) \right\|_X^q d\mu(x)$$
$$\lesssim_X (nA^q + m^q S^q) \int_{E_\infty(\mathbb{Z}_m^n)} \|f(x)-f(y)\|_X^q d\beta_3(x,y), \quad (13)$$

where $\beta_3 = (\beta_1+\beta_2)/2$. Note in passing that when $m \lesssim A$, an inequality such as (13), with perhaps a different measure β_3 on $E_\infty(\mathbb{Z}_m^n)$, is a consequence of the triangle inequality, and therefore holds trivially on any Banach space X. Thus, for our purposes, we may assume throughout that a (q,A,S)-smoothing and approximation scheme on $\mathbb{Z}_m^n$ satisfies $m \gtrsim A$.

Assuming that

$$m \gtrsim \frac{A}{S} \cdot n^{1/q}, \tag{14}$$

inequality (13) becomes

$$\sum_{j=1}^{n} \int_{\mathbb{Z}_m^n} \left\| f\left(x + \frac{m}{2} e_j\right) - f(x) \right\|_X^q d\mu(x) \lesssim_X S^q m^q \int_{E_\infty(\mathbb{Z}_m^n)} \|f(x) - f(y)\|_X^q d\beta_3(x,y). \tag{15}$$

If we could come up with a smoothing and approximation scheme for which $S \lesssim 1$, and m satisfied (14), then inequality (15) would not quite be the desired metric cotype inequality (5), but it would be rather close to it. The difference is that the probability measure β_3 is not uniformly distributed on all ℓ_∞ edges $E_\infty(\mathbb{Z}_m^n)$, as required in (5). Nevertheless, for many measures β_3, elementary triangle inequality and symmetry arguments can be used to "massage" inequality (15) into the desired inequality (5). This last point is a technical issue, but it is not the heart of our argument: we wish to design a smoothing and approximation scheme satisfying $S \lesssim 1$ with A as small as possible. In [13] such a scheme was designed with $A \lesssim n^2$. Here we carefully optimize the approach of [13] to yield a smoothing and approximation scheme with $A \lesssim n$, in which case (14) becomes the desired bound $m \gtrsim n^{1+\frac{1}{q}}$.

The bounds that we need in order to establish this improved estimate on m are based on the analysis of some quite delicate cancelations; indeed the bounds that we obtain are sharp for our smoothing and approximation scheme, as discussed in Section 1.2. In proving such sharp bounds, a certain bivariate extension of the Bernoulli numbers arises naturally; these numbers, together with some basic asymptotic estimates for them, are presented in Section 3.1. The cancelations in the Rademacher sums corresponding to our convolution kernels are analyzed via certain combinatorial identities in Section 3.2.

1.2. **A lower bound on smoothing and approximation with general kernels.** One might wonder whether our failure to prove the bound $m \gtrsim n^{1/q}$ without the non-trivial Rademacher type assumption is due to the fact we chose the wrong smoothing and approximation scheme. This is not the case. In Section 4 we show that any approach based on smoothing and approximation is doomed to yield a sub-optimal dependence of m on n (assuming that the conjectured $n^{1/q}$ bound is indeed true). Specifically, we show that for *any* (q, A, S)-smoothing and approximation scheme on $\mathbb{Z}_m^n$, with $m \gtrsim A$, we must have $AS \gtrsim_q \sqrt{n}$. Thus the bound $S \lesssim 1$ forces the bound $A \gtrsim_q \sqrt{n}$, and correspondingly (14) becomes $m \gtrsim_q n^{\frac{1}{2}+\frac{1}{q}}$. Additionally, we show in Section 4 that for the specific smoothing and approximation scheme used here, the bound $m \gtrsim n^{1+\frac{1}{q}}$ is sharp.

It remains open what is the best bound on m that is achievable via a smoothing and approximation scheme. While this question is interesting from an analytic perspective, our current lower bound shows that we need to use more than averaging with respect to positive measures in order to prove the desired bound $m \gtrsim n^{1/q}$.

Note that the lower bound $m \gtrsim_q n^{\frac{1}{2}+\frac{1}{q}}$ for smoothing and approximation schemes rules out the applicability of this method to some of the most striking potential applications of metric cotype to embedding theory in the coarse, uniform, or quasisymmetric categories, as explained in Section 1.3; these applications rely crucially on the use of a metric cotype inequality with $m \asymp n^{1/q}$.

The cancelation that was exploited in [13] in order to prove the sharp bound on m in the presence of non-trivial Rademacher type was also related to smoothing properties of convolution kernels, but with respect to signed measures: the smoothed Rademacher sums in the left hand side of (8) are controlled in [13] via the Rademacher projection, and the corresponding smoothing inequality (for signed measures) is proved via an appeal to Pisier's K-convexity theorem [15]. It would be of great interest to understand combinatorially/geometrically the cancelations that underly the estimate $m \geqslant n^{1/q}$ from [13], though there seems to be a lack of methods to handle smoothing properties of signed convolution kernels in spaces with trivial Rademacher type and finite Rademacher cotype.

1.3. **The relation to nonembeddability results and some open problems.** We recall some standard terminology. Let (X, d_X) and (Y, d_Y) be metric spaces. X is said to embed with distortion D into Y if there exists a mapping $f : X \to Y$ and (scaling factor) $\lambda > 0$, such that for all $x, y \in X$ we have $\lambda d_X(x, y) \leqslant d_Y(f(x), f(y)) \leqslant D\lambda d_X(x, y)$. X is said to embed uniformly into Y if there exists an into homeomorphism $f : X \to Y$ such that both f and f^{-1} are uniformly continuous. X is said to embed coarsely into Y if there exists a mapping $f : X \to Y$ and two non-decreasing functions $\alpha, \beta : [0, \infty) \to [0, \infty)$ such that $\lim_{t\to\infty} \alpha(t) = \infty$, and for all $x, y \in X$ we have $\alpha(d_X(x, y)) \leqslant d_Y(f(x), f(y)) \leqslant \beta(d_X(x, y))$. X is said to admit a quasisymmetric embedding into Y if there exists a mapping $f : X \to Y$ and an increasing (modulus) $\eta : (0, \infty) \to (0, \infty)$ such that for all distinct $x, y, z \in X$ we have $\frac{d_Y(f(x),f(y))}{d_Y(f(x),f(z))} \leqslant \eta\left(\frac{d_X(x,y)}{d_X(x,z)}\right)$.

For a Banach space X, let q_X denote the infimum over those $q \geqslant 2$ such that X has Rademacher (equiv. metric) cotype q. It was shown in [13, 14] that if X, Y are Banach spaces, Y has Rademacher type $p > 1$, and X embeds uniformly, coarsely, or quasisymmetrically into Y, then $q_X \leqslant q_Y$. Thus, under the Rademacher type > 1 assumption on the target space, Rademacher cotype q is an invariant that is stable under embeddings of Banach spaces, provided that the embedding preserves distances in a variety of (seemingly quite weak) senses. The role of the assumption that Y has non-trivial Rademacher type is via the metric cotype inequality with optimal m: the proofs of these results only use that Y satisfies the metric cotype q inequality (2) for some $m \asymp n^{1/q}$ (under this assumption, Y can be a general metric space and not necessarily a Banach space). This fact motivates our conjecture that for any Banach space Y with Rademacher cotype q, the metric cotype inequality (5) holds for every $m \gtrsim_Y n^{1/q}$. The same assertion for general metric spaces of metric cotype q is too much to hope for; see [19].

Perhaps the simplest Banach spaces for which we do not know how to prove a sharp metric cotype inequality are L_1 and the Schatten-von Neumann trace class S_1 (see, e.g., [21]). Both of these spaces have Rademacher cotype 2 (for S_1 see [18]), yet the currently best known bound on m in the metric cotype inequality (5) (with $q = 2$) for both of these spaces is the bound $m \gtrsim n^{3/2}$ obtained here. The above embedding results in the uniform, coarse or quasisymmetric categories do hold true for embeddings into L_1 (i.e., a Banach space X that embeds in one of these senses into L_1 satisfies $q_X = 2$). This fact is due to an ad-hoc argument, which fails for S_1 (see Section 8 in [13] for an explanation). We can thus ask the following natural questions (many of which were already raised in [13]):

Question 1. Can L_r admit a uniform, coarse, or quasisymmetric embedding into S_1 when $r > 2$? More ambitiously, can a Banach space X with $q_X > 2$ embed in one of these senses into S_1? In greatest generality: can a Banach space X embed in one of these senses into a Banach space Y with $q_Y < q_X$?

If a Banach space X admits a uniform or coarse embedding into S_1, then X must have finite cotype. This fact, which could be viewed as a (non-quantitative) step towards Question 1, was communicated to us by Nigel Kalton. To prove it, note that it follows from [17, Lem. 3.2] that for any ultrapower $(S_1)^{\mathscr{U}}$ of S_1, the unit ball of $(S_1)^{\mathscr{U}}$ is uniformly homeomorphic to a subset of Hilbert space. Thus $(S_1)^{\mathscr{U}}$ has Kalton's property $\mathcal{Q}$ (see [8] for a detailed discussion of this property). If the unit ball of X is uniformly homeomorphic to a subset of S_1 (resp. X admits a coarse embedding into S_1), then the unit ball in any ultrapower of X is uniformly homeomorphic to a subset of $(S_1)^{\mathscr{U}}$ (resp. any ultrapower of X admits a coarse embedding into $(S_1)^{\mathscr{U}}$). By the proof of [8, Thm. 4.2], it follows that any ultrapower of X has property $\mathcal{Q}$, and hence it cannot contain c_0. Thus X cannot have infinite cotype by the Maurey-Pisier theorem [10] and standard Banach space ultrapower theory (see [4, Thm. 8.12]).

Question 2. Does S_1 admit a uniform, coarse, or quasisymmetric embedding into a Banach space Y with Rademacher type $p > 1$? More ambitiously, does S_1 embed in one of these senses into Banach space Y with Rademacher type $p > 1$ and $q_Y = 2$? In greatest generality: does every Banach space X embed in one of these senses into a Banach space Y with Rademacher type $p > 1$? Perhaps we can even ensure in addition that $q_Y = q_X$?

Question 2 relates to Question 1 since embeddings into spaces with type > 1 would allow us to use the nonembeddability results of [13].

While the improved bound on m in Theorem 1.2 does not solve any of these fundamental questions, it does yield new restrictions on the possible moduli of embeddings in the uniform, coarse, or quasisymmetric categories. Instead of stating our nonembedding corollaries in greatest generality, let us illustrate our (modest) improved nonembeddability results for snowflake embeddings of L_4 into S_1 (this is just an illustrative example; the method of [13] yields similar results for embeddings of any Banach space X with $q_X > 2$ into S_1, and S_1 itself can be replaced by general Banach spaces of finite cotype). Take $\theta \in (0, 1)$ and assume the metric space $(L_4, \|x - y\|_4^\theta)$ admits a bi-Lipschitz embedding into S_1. Our strong conjectures imply that this cannot happen, but at present the best we can do is give bounds on θ. An application of Theorem 1.2 shows that $\theta \leqslant 4/5$, i.e., we have a definite quantitative estimate asserting that a uniform embedding of L_4 into S_1 must be far from bi-Lipschitz. The previous bound from [13] for S_1 was $m = n^{5/2}$, yielding $\theta \leqslant 8/9$. Our lower bound shows that by using a smoothing and approximation scheme we cannot hope to get a bound of $\theta < 2/3$.

Turning to bi-Lipschitz embeddings, consider the grid $\{0, 1, \ldots, m\}^n \subseteq \mathbb{R}^n$, equipped with the ℓ_∞^n metric. We denote this metric space by $[m]_\infty^n$. Bourgain [3] proved that if Y is a Banach space with Rademacher cotype q, then any embedding of $\left[n^{1+\frac{1}{q}}\right]_\infty^n$ into Y incurs distortion $\gtrsim_Y n^{1/q}$. The same result follows from Theorem 1.2, while the previous estimate on m from [13] only yields the weaker distortion lower bound of $\gtrsim_Y n^{\frac{q+1}{q(2q+1)}}$ for embeddings of $\left[n^{1+\frac{1}{q}}\right]_\infty^n$ into Y. The sharp bound on m from [13] when Y has Rademacher type > 1 implies that in this case, any embedding of $\left[n^{1/q}\right]_\infty^n$ into Y incurs distortion $\gtrsim n^{1/q}$ (where the implied constant is now allowed to depend also on the Rademacher type parameters of

Y). Our main conjecture implies the same improvement of Bourgain's result without the assumption that Y has non-trivial Rademacher type.

Bourgain's theorem [3] is part of his more general investigation of embeddings of ε-nets in unit balls of finite dimensional normed spaces. Bourgain's approach in [3] is based on ideas similar to ours, that are carried out in the continuous domain. Specifically, given a mapping $f : [m]_\infty^n \to Y$, he finds a mapping $g : \mathbb{R}^n \to Y$ which is L-Lipschitz and close in an appropriate sense (depending on L, m, n) to f on points of the grid $[m]_\infty^n$. Once this is achieved, it is possible to differentiate g to obtain the desired distortion lower bound. Bourgain's approximate Lipschitz extension theorem (an alternative proof of which was found in [1]) is a continuous version of a smoothing and approximation scheme; it seems plausible that our method in Section 4 for proving impossibility results for such schemes can be used to prove similar restrictions on Bourgain's approach to approximate Lipschitz extension. When Y has non-trivial Rademacher type, the improvement in [13] over Bourgain's nonembeddability result for grids is thus based on a more delicate cancelation than was used in [3, 1].

Question 3. Is it true that for any Banach space Y of Rademacher cotype q, any embedding of $\left[n^{1/q}\right]_\infty^n$ into Y incurs distortion $\gtrsim_Y n^{1/q}$ (if true, this is a sharp bound). Specializing to the Schatten-von Neumann trace class S_1, we do not even know whether the distortion of $[\sqrt{n}]_\infty^n$ in S_1 is $\gtrsim \sqrt{n}$. Theorem 1.2 implies a distortion lower bound of $\gtrsim n^{1/6}$, while the bound on m from [13] only yields a distortion lower bound of $\gtrsim n^{1/10}$. Our results in Section 4 show that one cannot get a distortion lower bound asymptotically better than $n^{1/4}$ by using smoothing and approximation schemes.

We did not discuss here metric characterizations of Rademacher type. We refer to [12] for more information on this topic. It turns out that our approach to Theorem 1.2 yields improved bounds in [12] as well; see [5].

Acknowledgements. O. G. was partially supported by NSF grant CCF-0635078. M. M. was partially supported by ISF grant no. 221/07, BSF grant no. 2006009, and a gift from Cisco research center. A. N. was supported in part by NSF grants CCF-0635078 and CCF-0832795, BSF grant 2006009, and the Packard Foundation.

2. Proof of Theorem 1.2

For $n \in \mathbb{N}$ denote $[n] = \{1, \ldots, n\}$. When $B \subseteq [n]$, and $x \in \mathbb{Z}_m^B$, we will sometimes slightly abuse notation by treating x as an element of $\mathbb{Z}_m^n$, with the understanding that for $i \in [n] \setminus B$ we have $x_i = 0$. For $y \in \mathbb{Z}_m^n$, we denote by y_B the restriction of y to the coordinates in B.

As in [13], for $j \in [n]$ and an odd integer $k < m/2$, we define $S(j,k) \subseteq \mathbb{Z}_m^n$ by

$$S(j,k) \stackrel{\mathrm{def}}{=} \left\{ y \in [-k,k]^n \subseteq \mathbb{Z}_m^n :\ y_j \text{ is even } \wedge\ \forall \ell \in [n] \setminus \{j\}\ y_\ell \text{ is odd} \right\}. \tag{16}$$

The parameter k will be fixed throughout the ensuing argument, and will be specified later. For every $j \in [n]$ let ν_j be the uniform probability measure on $S(j,k)$. Following the notation of [13], for a Banach space $(X, \|\cdot\|_X)$ and $f : \mathbb{Z}_m^n \to X$, we write $f * \nu_j = \mathcal{E}_j f$, that is,

$$\mathcal{E}_j f(x) \stackrel{\mathrm{def}}{=} \frac{1}{\mu(S(j,k))} \int_{S(j,k)} f(x+y)\, d\mu(y). \tag{17}$$

Recall that $E_\infty(\mathbb{Z}_m^n)$, defined in (6), is the set of all ℓ_∞ edges of $\mathbb{Z}_m^n$. Similarly, we denote the ℓ_1 edges of $\mathbb{Z}_m^n$ by $E_1(\mathbb{Z}_m^n)$, i.e.,

$$E_1(\mathbb{Z}_m^n) \stackrel{\mathrm{def}}{=} \Big\{(x,y)\in \mathbb{Z}_m^n\times\mathbb{Z}_m^n:\ x-y\in\{\pm e_1,\ldots,\pm e_n\}\Big\}. \tag{18}$$

Clearly $E_1(\mathbb{Z}_m^n)\subseteq E_\infty(\mathbb{Z}_m^n)$.

Let β_1° denote the uniform probability distribution on the pairs $(x,y)\in E_\infty(\mathbb{Z}_m^n)$ with $x-y\in\{-1,1\}^n$, and let $\beta_1^{\circ\circ}$ denote the uniform probability distribution on $E_1(\mathbb{Z}_m^n)$. We shall consider the probability measure on $E_\infty(\mathbb{Z}_m^n)$ given by $\beta_1=(\beta_1^\circ+\beta_1^{\circ\circ})/2$.

Lemma 5.1 in [13] implies that for all $q\geqslant 1$ and $f:\mathbb{Z}_m^n\to X$ we have:

$$\frac{1}{n}\sum_{j=1}^n\int_{\mathbb{Z}_m^n}\|\mathcal{E}_jf-f\|_X^q\,d\mu\lesssim(2k)^q\int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\beta_1(x,y). \tag{19}$$

Inequality (19) corresponds to the approximation property (7), with $A\lesssim k$. The relevant smoothing inequality is the main new ingredient in our proof of Theorem 1.2, and it requires a more delicate choice of probability measure β_2 on $E_\infty(\mathbb{Z}_m^n)$. If $(x,y)\in E_\infty(\mathbb{Z}_m^n)$ then $x-y\in\{-1,0,1\}^n$. Let $S=\{i\in[n]:\ x_i=y_i\}$, and define

$$\beta_2(x,y)\stackrel{\mathrm{def}}{=}\frac{1}{Z}\cdot\frac{(n/k)^{q|S|}}{2^{n-|S|}m^n\binom{n}{|S|}}, \tag{20}$$

where Z is a normalization factor ensuring that β_2 is a probability measure, i.e.,

$$Z=\sum_{\ell=0}^n\left(\frac{n}{k}\right)^{q\ell}\asymp 1, \tag{21}$$

provided that, say,

$$k\geqslant 2n. \tag{22}$$

Our final choice of k will satisfy (22), so we may assume throughout that Z satisfies (21).

The key smoothing property of the averaging operators $\{\mathcal{E}_j\}_{j=1}^n$ is contained in the following lemma:

Lemma 2.1. *Let X be a Banach space, $q\geqslant 1$, $n,m\in\mathbb{N}$, where $m>4n$ is divisible by 4, and $f:\mathbb{Z}_m^n\to X$. Suppose that k is an odd integer satisfying $2n\leqslant k<\frac{m}{2}$. Then,*

$$\int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{j=1}^n\varepsilon_j\left[\mathcal{E}_jf(x+e_j)-\mathcal{E}_jf(x-e_j)\right]\right\|_X^q d\tau(\varepsilon)d\mu(x)$$
$$\leqslant S^q\int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\beta_2(x,y), \tag{23}$$

where $S\lesssim_q 1$.

We shall postpone the proof of Lemma 2.1 to Section 3, and proceed now to deduce Theorem 1.2 assuming its validity. Before doing so, we recall for future use the following simple lemma from [13]:

Lemma 2.2 (Lemma 2.6 from [13]). *For every $q \geqslant 1$ and for every $f : \mathbb{Z}_m^n \to X$,*

$$\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\|f(x+e_j)-f(x)\|_X^q d\mu(x) \lesssim 2^q \int_{\{-1,0,1\}^n}\int_{\mathbb{Z}_m^n}\|f(x+\delta)-f(x)\|_X^q d\mu(x)d\sigma(\delta). \quad (24)$$

Proof of Theorem 1.2. The argument in the introduction leading to (15), when specialized to our smoothing and approximation scheme using (19) and (23), shows that if $k \geqslant 2n$ and $m \geqslant 2kn^{1/q}$, then

$$\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x) \lesssim_X m^q \int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\beta_3(x,y), \qquad (25)$$

where

$$\beta_3 = \frac{\beta_1+\beta_2}{2} \leqslant \beta_1^\circ + \beta_1^{\circ\circ} + \beta_2.$$

Note that $\beta_1^\circ \lesssim \beta_2$ due to the contribution of $S=\emptyset$ in (20). Thus, (25) implies the following bound:

$$\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x) \lesssim_X \frac{m^q}{n}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\|f(x+e_j)-f(x)\|_X^q d\mu(x)$$
$$+ m^q \sum_{S\subseteq[n]}\frac{(n/k)^{q|S|}}{\binom{n}{|S|}}\int_{\{-1,1\}^{[n]\setminus S}}\int_{\mathbb{Z}_m^n}\|f(x+\varepsilon)-f(x)\|_X^q d\mu(x)d\tau(\varepsilon), \quad (26)$$

where the first term in the right hand side of (26) corresponds to $\beta_1^{\circ\circ}$.

In order to deduce the desired metric cotype inequality (5) from (26), we shall apply (26) to lower dimensional sub-tori of $\mathbb{Z}_m^n$. Note that we are allowed to do so since our requirements on k, namely $k \geqslant 2n$ and $m \geqslant 2kn^{1/q}$, remain valid for smaller n.

Fix $\emptyset \neq B \subseteq [n]$ and $x_{[n]\setminus B} \in \mathbb{Z}_m^{[n]\setminus B}$. We can then consider the mapping $g : \mathbb{Z}_m^B \to X$ given by $g(x_B) = f(x_{[n]\setminus B}, x_B)$. Applying (26) to g, and averaging the resulting inequality over $x_{[n]\setminus B} \in \mathbb{Z}_m^{[n]\setminus B}$, we obtain

$$\sum_{j\in B}\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x) \lesssim_X \frac{m^q}{|B|}\sum_{j\in B}\int_{\mathbb{Z}_m^n}\|f(x+e_j)-f(x)\|_X^q d\mu(x)$$
$$+ m^q \sum_{S\subseteq B}\frac{(|B|/k)^{q|S|}}{\binom{|B|}{|S|}}\int_{\{-1,1\}^{B\setminus S}}\int_{\mathbb{Z}_m^n}\|f(x+\varepsilon)-f(x)\|_X^q d\mu(x)d\tau(\varepsilon). \quad (27)$$

For $B \subseteq [n]$ define the weight $W_{|B|} \stackrel{\mathrm{def}}{=} \frac{2^{|B|-1}}{3^{n-1}}$. Multiplying (27) by $W_{|B|}$ and summing over $\emptyset \neq B \subseteq [n]$, we obtain the bound

$$\frac{1}{m^q}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x) \lesssim_X \sum_{\substack{B\subseteq[n]\\ B\neq\emptyset}}\frac{W_{|B|}}{|B|}\sum_{j\in B}\int_{\mathbb{Z}_m^n}\|f(x+e_j)-f(x)\|_X^q d\mu(x)$$
$$+ \sum_{\substack{B\subseteq[n]\\ B\neq\emptyset}} W_{|B|}\sum_{S\subseteq B}\frac{(|B|/k)^{q|S|}}{\binom{|B|}{|S|}}\int_{\{-1,1\}^{B\setminus S}}\int_{\mathbb{Z}_m^n}\|f(x+\varepsilon)-f(x)\|_X^q d\mu(x)d\tau(\varepsilon), \quad (28)$$

where we used the identity

$$\sum_{B\subseteq[n]} W_{|B|}\sum_{j\in B}\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x)=\sum_{j=1}^n\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x).$$

The first term in the right hand side of (28) is easy to bound, using Lemma 2.2, as follows:

$$\begin{aligned}\sum_{\substack{B\subseteq[n]\\ B\neq\emptyset}}\frac{W_{|B|}}{|B|}\sum_{j\in B}\int_{\mathbb{Z}_m^n}\|f(x+e_j)-f(x)\|_X^q d\mu(x)&\lesssim\frac{1}{n}\sum_{j=1}^n\int_{\mathbb{Z}_m^n}\|f(x+e_j)-f(x)\|_X^q d\mu(x)\\ &\overset{(24)}{\lesssim} 2^q\int_{\{-1,0,1\}^n}\int_{\mathbb{Z}_m^n}\|f(x+\delta)-f(x)\|_X^q d\mu(x)d\sigma(\delta),\end{aligned}\tag{29}$$

where in the first inequality of (29) we used the fact that $\sum_{\ell=1}^n\binom{n-1}{\ell-1}\frac{2^{\ell-1}}{3^{n-1}}\lesssim\frac{1}{n}$. To bound the second term in the right hand side of (28), note that it equals

$$\begin{aligned}C&\stackrel{\mathrm{def}}{=}\sum_{S\subseteq[n]}\sum_{\substack{S\subseteq B\subseteq[n]\\ B\neq\emptyset}}\sum_{\varepsilon\in\{-1,1\}^{B\setminus S}}\frac{2^{|B|-1}}{3^{n-1}}\cdot\frac{(|B|/k)^{q|S|}}{2^{|B|-|S|}\binom{|B|}{|S|}}\int_{\mathbb{Z}_m^n}\|f(x+\varepsilon)-f(x)\|_X^q d\mu(x)\\ &\lesssim\frac{1}{3^n}\sum_{T\subseteq[n]}\sum_{\varepsilon\in\{-1,1\}^T}a_T\int_{\mathbb{Z}_m^n}\|f(x+\varepsilon)-f(x)\|_X^q d\mu(x),\end{aligned}\tag{30}$$

where we used the change of variable $T=B\setminus S$, and for every $T\subseteq[n]$ we write,

$$a_T\stackrel{\mathrm{def}}{=}\sum_{B\supseteq T}\frac{2^{|B|-|T|}(|B|/k)^{q(|B|-|T|)}}{\binom{|B|}{|B|-|T|}}=\sum_{\ell=|T|}^n\binom{n-|T|}{\ell-|T|}\cdot\frac{2^{\ell-|T|}(\ell/k)^{q(\ell-|T|)}}{\binom{\ell}{\ell-|T|}}.$$

Fix $T\subseteq[n]$. Using the standard bounds $\left(\frac{u}{v}\right)^v\leqslant\binom{u}{v}\leqslant\left(\frac{eu}{v}\right)^v$, which hold for all integers $0\leqslant v\leqslant u$, we can bound a_T as follows:

$$\begin{aligned}a_T&\leqslant\sum_{\ell=|T|}^n\left(\frac{e(n-|T|)}{\ell-|T|}\right)^{\ell-|T|}\left(\frac{\ell-|T|}{\ell}\right)^{\ell-|T|}\left(\frac{\ell}{k}\right)^{q(\ell-|T|)}2^{\ell-|T|}\\ &=\sum_{\ell=|T|}^n\left(\frac{2e(n-|T|)\ell^{q-1}}{k^q}\right)^{\ell-|T|}.\end{aligned}$$

Thus, assuming that $k\geqslant 3n$, and recalling that $q\geqslant 2$, we get the bound

$$a_T\leqslant\sum_{\ell=|T|}^n\left(\frac{2e(n-|T|)n^{q-1}}{(3n)^q}\right)^{\ell-|T|}\leqslant\sum_{\ell=|T|}^n\left(\frac{2e}{9}\right)^{(\ell-|T|)}\lesssim 1.\tag{31}$$

Combining (31) with (30), we see that the second term in the right hand side of (28) is

$$\begin{aligned}C&\lesssim\frac{1}{3^n}\sum_{T\subseteq[n]}\sum_{\varepsilon\in\{-1,1\}^T}\int_{\mathbb{Z}_m^n}\|f(x+\varepsilon)-f(x)\|_X^q d\mu(x)\\ &=\int_{\{-1,0,1\}^n}\int_{\mathbb{Z}_m^n}\|f(x+\delta)-f(x)\|_X^q d\mu(x)d\sigma(\delta).\end{aligned}$$

In combination with (29), inequality (28) implies that

$$\frac{1}{m^q}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\left\|f\left(x+\frac{m}{2}e_j\right)-f(x)\right\|_X^q d\mu(x)$$
$$\lesssim_X \int_{\{-1,0,1\}^n}\int_{\mathbb{Z}_m^n}\|f(x+\delta)-f(x)\|_X^q d\mu(x)d\sigma(\delta),$$

which is precisely the desired inequality (5). Recall that in the above argument, our requirement on k was $k \geqslant 3n$, and our requirement on m was $m \geqslant 2kn^{1/q}$ (and that it is divisible by 4). This implies the requirement $m \geqslant 6n^{1+\frac{1}{q}}$ of Theorem 1.2. □

3. Proof of Lemma 2.1

Lemma 2.1 is the main new ingredient of the proof of Theorem 1.2. Its proof is based on combinatorial identities which relate the "smoothed out Rademacher sum"

$$\sum_{j=1}^{n}\varepsilon_j\left[\mathcal{E}_j f(x+e_j)-\mathcal{E}_j f(x-e_j)\right] \tag{32}$$

to a certain bivariate extension of the Bernoulli numbers. We shall therefore first, in Section 3.1, do some preparatory work which introduces these numbers and establishes estimates that we will need in the ensuing argument. We shall then derive, in Section 3.2, certain combinatorial identities that relate (32) to the bivariate Bernoulli numbers. In Section 3.3 we shall combine the results of Section 3.1 and Section 3.2 to complete the proof of Lemma 2.1.

3.1. **Estimates for the bivariate Bernoulli numbers.** There are two commonly used definitions of the Bernoulli numbers $\{B_r\}_{r=0}^{\infty}$. For more information on these two conventions, we refer to `http://en.wikipedia.org/wiki/Bernoulli_number`. Here we shall refer to the variant of the Bernoulli numbers that was originally defined by J. Bernoulli, for which $B_1=\frac{1}{2}$, and which is defined via the recursion

$$r=\sum_{a=0}^{r-1}B_a\binom{r}{a}. \tag{33}$$

Observe that (33) contains the base case $B_0=1$ when substituting $r=1$. The recursion (33) extends naturally to a bivariate sequence $\{B_{r,s}\}_{r,s=0}^{n}$, given by

$$r-s=\sum_{a=0}^{r-1}B_{a,s}\binom{r}{a}-\sum_{b=0}^{s-1}B_{r,b}\binom{s}{b}. \tag{34}$$

It is well-known (cf. [20, Sec. 2.5]) that the exponential generating function for $\{B_r\}_{r=0}^{\infty}$ is

$$F(x)\stackrel{\mathrm{def}}{=}\frac{xe^x}{e^x-1}=\sum_{r=0}^{\infty}B_r\frac{x^r}{r!}.$$

We shall require the following analogous computation of the bivariate exponential generating function of $\{B_{r,s}\}_{r,s=0}^{n}$:

Lemma 3.1. *For all $x, y \in \mathbb{C}$ with $|x|, |y| < \pi$ we have*

$$F(x,y) \stackrel{\text{def}}{=} \frac{(x-y)e^{x+y}}{e^x - e^y} = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} B_{r,s}\frac{x^r y^s}{r! \cdot s!}, \tag{35}$$

where the series in (35) *is absolutely convergent on* $\{(x,y) \in \mathbb{C}\times\mathbb{C} :\ |x|, |y| \leqslant r\}$ *for all* $r < \pi$.

Proof. The function $F(x,y)$ is analytic on $D_\pi = \{(x,y) \in \mathbb{C}\times\mathbb{C} :\ |x|, |y| < \pi\}$, since its only non-removable singularities are when $x - y \in 2\pi i(\mathbb{Z}\setminus\{0\})$. It follows that we can write $F(x,y) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} z_{r,s}x^r y^s$, for some $\{z_{r,s}\}_{r,s=0}^{\infty} \subseteq \mathbb{C}$, where the series converges absolutely on any compact subset of D_π (see, e.g., [6, Thm. 2.2.6]). Note that

$$\begin{aligned}(e^x - e^y)F(x,y) &= \left(\sum_{n=1}^{\infty}\frac{x^n - y^n}{n!}\right)\left(\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} z_{r,s}x^r y^s\right)\\ &= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\left(\sum_{a=0}^{r-1}\frac{z_{a,s}}{(r-a)!} - \sum_{b=0}^{s-1}\frac{z_{r,b}}{(s-b)!}\right)x^r y^s. \end{aligned}\tag{36}$$

At the same time,

$$(e^x - e^y)F(x,y) = (x-y)e^x e^y = (x-y)\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{x^r y^s}{r!s!} = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}(r-s)\,\frac{x^r y^s}{r!s!}. \tag{37}$$

By equating coefficients in (36) and (37), we see that for all $r, s \in \mathbb{N}\cup\{0\}$,

$$r - s = r!s!\left(\sum_{a=0}^{r-1}\frac{z_{a,s}}{(r-a)!} - \sum_{b=0}^{s-1}\frac{z_{r,b}}{(s-b)!}\right) = \sum_{a=0}^{r-1}\binom{r}{a}a!s!z_{a,s} - \sum_{b=0}^{s-1}\binom{s}{b}r!b!z_{r,b}.$$

Since $z_{0,0} = 1$, the recursive definition (34) implies that $z_{r,s} = \frac{B_{r,s}}{r!s!}$, as required. $\square$

An immediate corollary of Lemma 3.1 is that since $F(x,y) = F(y,x)$,

$$\forall r, s \in \mathbb{N}\cup\{0\}, \quad B_{r,s} = B_{s,r}. \tag{38}$$

Another (crude) corollary of Lemma 3.1 is that since the power series in (35) converges absolutely on $\{(x,y) \in \mathbb{C}\times\mathbb{C} :\ |x|, |y| \leqslant 2\}$, for all but at most finitely many $r, s \in \mathbb{N}\cup\{0\}$ we have $|B_{r,s}/(r!s!)|^{1/(r+s)} \leqslant 1/2$. Thus,

$$\forall r, s \in \mathbb{N}\cup\{0\}, \quad |B_{r,s}| \lesssim \frac{r!s!}{2^{r+s}}. \tag{39}$$

Remark 3.1. Since $B_{2m} = \frac{(-1)^{m-1}2\zeta(2m)(2m)!}{(2\pi)^{2m}}$, where $\zeta(s)$ is the Riemann zeta function (and $B_{2m+1} = 0$ for $m \geqslant 1$), one has the sharp asymptotics $|B_{2m}| \sim \frac{2(2m)!}{(2\pi)^{2m}}$ for the classical Bernoulli numbers. We did not investigate the question whether similar sharp asymptotics can be obtained for the bivariate Bernoulli numbers.

3.2. **Some combinatorial identities.** We start by introducing some notation. For $y \in \mathbb{Z}_m^n$ write:

$$\uparrow y \uparrow \stackrel{\mathrm{def}}{=} \left|\{l : y_l = k \pmod m\}\right|,$$

$$\downarrow y \downarrow \stackrel{\mathrm{def}}{=} \left|\{l : y_l = -k \pmod m\}\right|,$$

$$\updownarrow y \updownarrow \stackrel{\mathrm{def}}{=} \uparrow y \uparrow + \downarrow y \downarrow \quad \text{and} \quad \uparrow y \downarrow \stackrel{\mathrm{def}}{=} \uparrow y \uparrow - \downarrow y \downarrow .$$

We also define

$$\mathbb{S} \stackrel{\mathrm{def}}{=} \Big\{ y \in [-k,k]^n \subseteq \mathbb{Z}_m^n : \; y_t \text{ is odd } \forall t \in [n] \Big\}.$$

For $x \in \mathbb{Z}_m^n$ and $\varepsilon \in \{-1,1\}^n$, let $x \odot \varepsilon \in \mathbb{Z}_m^n$ be the coordinate wise multiplication, i.e., $(x \odot \varepsilon)_j = x_j \varepsilon_j$. Also for $\varepsilon, \varepsilon' \in \{-1,1\}^n$ let $\langle \varepsilon, \varepsilon' \rangle = \sum_{j=1}^n \varepsilon_j \varepsilon_j'$.

We need to define additional auxiliary averaging operators.

Definition 3.2. For $f : \mathbb{Z}_m^n \to X$, $k < \frac{m}{2}$ odd, and $B \subseteq [n]$, let

$$\Delta_B f(x) \stackrel{\mathrm{def}}{=} \frac{1}{\mu(L_B)} \int_{L_B} f(x+y) d\mu(y),$$

where

$$L_B \stackrel{\mathrm{def}}{=} \Big\{ y \in (-k,k)^n \subseteq \mathbb{Z}_m^n : \; \forall i \notin B, \; y_i = 0 \; \wedge \; \forall i \in [n] \; y_i \text{ is even} \Big\}.$$

Definition 3.3. Define for $z \in \mathbb{Z}_m^n$, $\varepsilon \in \{-1,1\}^n$, $i \in [n]$, and $0 \leqslant j \leqslant i$,

$$b_{i,j}(z,\varepsilon) \stackrel{\mathrm{def}}{=} \sum_{\substack{S \subseteq [n] \\ |S|=i}} \sum_{\substack{\delta \in \{-1,1\}^S \\ \langle \delta, \varepsilon_S \rangle = i-2j}} \Big(\mathbf{1}_{\delta k + \varepsilon_{[n]\setminus S} + L_{[n]\setminus S}}(z) - \mathbf{1}_{\delta k - \varepsilon_{[n]\setminus S} + L_{[n]\setminus S}}(z) \Big), \tag{40}$$

$$a(z,\varepsilon) \stackrel{\mathrm{def}}{=} \sum_{j=1}^n \varepsilon_j \left(\mathbf{1}_{e_j + S(j,k)}(z) - \mathbf{1}_{-e_j + S(j,k)}(z) \right), \tag{41}$$

where we recall that $S(j,k)$ was defined in (16).

The next lemma follows immediately from an inspection of our definitions.

Lemma 3.2. *The following identities hold true:*

$$\sum_{y \in \mathbb{Z}_m^n} b_{i,j}(y-x,\varepsilon) f(y) = k^{n-i} \sum_{\substack{S \subseteq [n] \\ |S|=i}} \sum_{\substack{\delta \in \{-1,1\}^S \\ \langle \delta, \varepsilon_S \rangle = i-2j}} \Big(\Delta_{[n]\setminus S} f(x + \delta k + \varepsilon_{[n]\setminus S}) - \Delta_{[n]\setminus S} f(x + \delta k - \varepsilon_{[n]\setminus S}) \Big), \tag{42}$$

$$\sum_{y \in \mathbb{Z}_m^n} a(y-x,\varepsilon) f(y) = k(k+1)^{n-1} \sum_{j=1}^n \varepsilon_j \big(\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j) \big). \tag{43}$$

Claim 3.3. *If there exits $t \in [n]$ such that z_t is even, then $a(z,\varepsilon) = b_{i,j}(z,\varepsilon) = 0$ for all $\varepsilon \in \{-1,1\}^n$.*

Proof. This follows directly from the definitions of the sets $S(j,k)$, and L_B, since all values of the coordinates are odd in all the points of the sets

$$\delta k+\varepsilon_{[n]\setminus S}+L_{[n]\setminus S},\ \delta k-\varepsilon_{[n]\setminus S}+L_{[n]\setminus S},\ e_j+S(j,k),\ -e_j+S(j,k),$$

for every $S\subseteq[n]$, $\delta\in\{-1,1\}^S$ and $\varepsilon\in\{-1,1\}^n$. $\square$

Claim 3.4. *If z_t is odd and either $z_t\neq\pm k \pmod m$ for all $t\in[n]$, or $|z_{t_0}|>k$ for some $t_0\in[n]$, then $a(z,\varepsilon)=b_{i,j}(z,\varepsilon)=0$.*

Proof. We may assume that $z\in[-m/2,m/2]^n$. If there is $t_0\in[n]$ for which $|z_{t_0}|>k$ then all the terms in the right hand side of (40) and (41) are 0. If $|z_t|<k$ for all $t\in[n]$, then all the terms in the right hand side of (40) and (41) cancel out. $\square$

It follows that for $z\notin\mathbb{S}$ we have $a(z,\varepsilon)=b_{i,j}(z,\varepsilon)=0$ for every $\varepsilon\in\{-1,1\}^n$ and every $0\leqslant j\leqslant i\leqslant n$. Thus, in particular, identity (42) can be rewritten as:

$$\sum_{y\in x+\mathbb{S}} b_{i,j}(y-x,\varepsilon)f(y)$$
$$=k^{n-i}\sum_{\substack{S\subseteq[n]\\|S|=i}}\ \sum_{\substack{\delta\in\{-1,1\}^S\\ \langle\delta,\varepsilon_S\rangle=i-2j}}\Big(\Delta_{[n]\setminus S}f(x+\delta k+\varepsilon_{[n]\setminus S})-\Delta_{[n]\setminus S}f(x+\delta k-\varepsilon_{[n]\setminus S})\Big). \quad (44)$$

Note that the definition (41) shows that for $z\in\mathbb{S}$ we have

$$a(z,\varepsilon)=\sum_{\substack{t\in[n]\\ z_t=k}}\varepsilon_t-\sum_{\substack{t\in[n]\\ z_t=-k}}\varepsilon_t=\uparrow z\odot\varepsilon\downarrow. \quad (45)$$

Using Claim 3.3 and Claim 3.4, in conjunction with (43) and (45), we conclude that:

Lemma 3.5. *The following identity holds for all $x\in\mathbb{Z}_m^n$ and $\varepsilon\in\{-1,1\}^n$:*

$$\sum_{j=1}^n\varepsilon_j\big(\mathcal{E}_jf(x+e_j)-\mathcal{E}_jf(x-e_j)\big)=\frac{1}{k(k+1)^{n-1}}\sum_{y\in x+\mathbb{S}}\Big\uparrow(y-x)\odot\varepsilon\Big\downarrow f(y). \quad (46)$$

Lemma 3.6. *If $z\in\mathbb{S}$ and $i\geqslant\updownarrow z\updownarrow$ then $\forall j\in\{0,\dots,i\}$ and $\forall\varepsilon\in\{-1,1\}^n$, we have $b_{i,j}(z,\varepsilon)=0$.*

Proof. If $i>\updownarrow z\updownarrow$ then

$$z\notin\Big(\delta k+\varepsilon_{[n]\setminus S}+L_{[n]\setminus S}\Big)\bigcup\Big(\delta k-\varepsilon_{[n]\setminus S}+L_{[n]\setminus S}\Big),$$

for every $S\subseteq[n]$ with $|S|=i$, and $\delta\in\{-1,1\}^S$. If $i=\updownarrow z\updownarrow$ then there exists exactly one subset $S\subseteq[n]$ in (40) where z can appear, namely $S=\{\ell\in[n]:\ z_\ell\in\{-k,k\}\}$. If

$$z\in\Big(\delta k+\varepsilon_{[n]\setminus S}+L_{[n]\setminus S}\Big)\bigcup\Big(\delta k-\varepsilon_{[n]\setminus S}+L_{[n]\setminus S}\Big),$$

for some $\delta\in\{-1,1\}^S$, then

$$z\in\Big(\delta k+\varepsilon_{[n]\setminus S}+L_{[n]\setminus S}\Big)\bigcap\Big(\delta k-\varepsilon_{[n]\setminus S}+L_{[n]\setminus S}\Big),$$

since for all coordinates $i\in[n]\setminus S$ we have $|z_i|<k$. Hence in this case the terms in the sum in the right hand side of (40) cancel out. $\square$

Lemma 3.7. *For every $z \in \mathbb{S}$, $\varepsilon \in \{-1,1\}^n$, and $0 \leqslant j \leqslant i < \updownarrow z \updownarrow$,*

$$b_{i,j}(z,\varepsilon) = \begin{cases} \binom{\updownarrow z \updownarrow - j}{i-j} & \downarrow z \odot \varepsilon \downarrow = j, \\ -\binom{\updownarrow z \updownarrow - (i-j)}{j} & \uparrow z \odot \varepsilon \uparrow = i-j, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. By looking at the elements of

$$\Big(\delta k + \varepsilon_{[n]\setminus S} + L_{[n]\setminus S}\Big) \bigcup \Big(\delta k - \varepsilon_{[n]\setminus S} + L_{[n]\setminus S}\Big),$$

it is clear that we must have $S \subseteq \{h : \ z_h \in \{-k,k\}\}$ in order to get a nonzero contribution to the right hand side of (40). For such an S there is at most one $\delta \in \{-1,1\}^S$ which can contribute to the sum, namely $\delta_h = \operatorname{sgn}(z_h)$ for every $h \in S$. But since this δ should also satisfy $\langle \delta, \varepsilon_S \rangle = i - 2j$, we conclude that a non-zero contribution can occur only when $\downarrow z_S \odot \varepsilon_S \downarrow = j$. In those cases, there is an actual contribution only if either $\operatorname{sgn}(z_h \varepsilon_h) = 1$ for every $h \in \{\ell : \ z_\ell \in \{-k,k\}\} \setminus S$, or $\operatorname{sgn}(z_h\varepsilon_h) = -1$ for every $h \in \{\ell : \ z_\ell \in \{-k,k\}\} \setminus S$, and those contributions have different signs. The claim now follows. $\square$

The following lemma relates, via Lemma 3.7, what we have done so far to the bivariate Bernoulli numbers.

Lemma 3.8. *There exists a sequence $\{h_{\alpha,\beta}\}_{\substack{0\leqslant\alpha\leqslant n \\ 0 \leqslant \beta \leqslant \alpha}} \subseteq \mathbb{R}$ such that for all $y \in \mathbb{Z}_m^n$ and all $\varepsilon \in \{-1,1\}^n$,*

$$\uparrow y \odot \varepsilon \downarrow = \sum_{\alpha=0}^{n} \sum_{\beta=0}^{\alpha} h_{\alpha,\beta} b_{\alpha,\beta}(y,\varepsilon), \tag{47}$$

$$|h_{\alpha,\beta}| \lesssim \frac{(\alpha-\beta)!\beta!}{2^\alpha}, \ \textit{for all } 0 \leqslant \beta \leqslant \alpha \tag{48}$$

$$h_{\alpha,\beta} = h_{\alpha,\alpha-\beta}. \tag{49}$$

Proof. Write $r = \uparrow z \odot \varepsilon \uparrow$ and $s = \downarrow z \odot \varepsilon \downarrow$. Thus $r + s = \updownarrow z \updownarrow$ and $r - s = \uparrow z \odot \varepsilon \downarrow$. With this notation, if we substitute the values of $b_{\alpha,\beta}(y,\varepsilon)$ from Lemma 3.7, the desired identity (47) becomes:

$$\begin{aligned} r - s = \sum_{\alpha=s}^{n} h_{\alpha,s} \binom{r}{\alpha - s} - \sum_{\beta=0}^{n} h_{\beta+r,\beta} \binom{s}{\beta} &\overset{(\spadesuit)}{=} \sum_{a=0}^{n-s} h_{a+s,s} \binom{r}{a} - \sum_{b=0}^{n} h_{b+r,b} \binom{s}{b} \\ &\overset{(\clubsuit)}{=} \sum_{a=0}^{r-1} h_{a+s,s} \binom{r}{a} - \sum_{b=0}^{s-1} h_{b+r,b} \binom{s}{b}, \end{aligned} \tag{50}$$

where in $(\spadesuit)$ we used the change of variable $\beta = b$, $\alpha = a + s$, and in $(\clubsuit)$ we noted that $r + s = \updownarrow z \updownarrow \leqslant n$ and that the terms corresponding to $a > r$ or $b > s$ vanish, while the terms corresponding to $a = r$ and $b = s$ cancel out. Thus, the desired identity (50) shows that we must take $h_{a+b,b} = B_{a,b}$, or $h_{\alpha,\beta} = B_{\alpha-\beta,\beta}$. The bound (48) is now the same as (39), and the identity (49) is the same as (38). $\square$

3.3. **Putting things together.** We are now ready to complete the proof of Lemma 2.1 using the tools developed in the previous two sections.

Lemma 3.9. *Let $\{h_{\alpha,\beta}\}_{\substack{0\leqslant\alpha\leqslant n\\ 0\leqslant\beta\leqslant\alpha}}$ be the sequence from Lemma 3.8. Then for all $f:\mathbb{Z}_m^n\to X$ and all $\varepsilon\in\{-1,1\}^n$ we have,*

$$\int_{\mathbb{Z}_m^n}\left\|\frac{1}{k(k+1)^{n-1}}\left(\sum_{i=0}^n\sum_{j=0}^i h_{i,j}\sum_{y\in x+\mathbb{S}} b_{i,j}(y-x,\varepsilon)f(y)\right)\right\|_X^q d\mu(x)$$
$$\lesssim_q \sum_{\ell=0}^n \frac{(n/k)^{\ell q}}{\binom{n}{\ell}}\sum_{\substack{S\subseteq[n]\\ |S|=\ell}}\int_{\mathbb{Z}_m^n}\left\|f(x+\varepsilon_{[n]\setminus S})-f(x)\right\|_X^q d\mu(x). \quad (51)$$

Proof. For every $x\in\mathbb{Z}_m^n$ and $0\leqslant j\leqslant i\leqslant n$ write

$$D_{i,j}(x)\stackrel{\mathrm{def}}{=}\left\|\frac{1}{k(k+1)^{n-1}}\left(\sum_{y\in x+\mathbb{S}} b_{i,j}(y-x,\varepsilon)f(y)\right)\right\|_X.$$

Note that,

$$\left(\sum_{i=0}^n\sum_{j=0}^i |h_{i,j}|D_{i,j}(x)\right)^q=\left(\sum_{i=0}^n 2^{-(i+1)}\sum_{j=0}^i 2^{i+1}|h_{i,j}|D_{i,j}(x)\right)^q$$
$$\stackrel{(*)}{\leqslant}\sum_{i=0}^n 2^{-(i+1)}\left(\sum_{j=0}^i 2^{i+1}|h_{i,j}|D_{i,j}(x)\right)^q\stackrel{(**)}{\leqslant}\sum_{i=0}^n\sum_{j=0}^i 2^{(i+1)(q-1)}(i+1)^{q-1}|h_{i,j}|^q D_{i,j}(x)^q, \quad (52)$$

where in $(*)$ we used the convexity of the function $t\mapsto t^q$ and that $\sum_{i=0}^\infty 2^{-(i+1)}=1$, and in $(**)$ we used Hölder's inequality. It follows from (52), combined with the bound (48) on $h_{i,j}$, that,

$$\left\|\frac{1}{k(k+1)^{n-1}}\left(\sum_{i=0}^n\sum_{j=0}^i h_{i,j}\sum_{y\in\mathbb{Z}_m^n} b_{i,j}(y-x,\varepsilon)f(y)\right)\right\|_X^q\leqslant\left(\sum_{i=0}^n\sum_{j=0}^i |h_{i,j}|D_{i,j}(x)\right)^q$$
$$\lesssim\sum_{i=0}^n\sum_{j=0}^i 2^{(i+1)(q-1)}(i+1)^{q-1}\left(\frac{(i-j)!j!}{2^i}\right)^q D_{i,j}(x)^q. \quad (53)$$

Now, $D_{i,j}(x)$ can be estimated using the identity (44) as follows:

$$k^iD_{i,j}(x)\leqslant\frac{(k+1)^{n-1}}{k^{n-i-1}}D_{i,j}(x)$$
$$\leqslant\sum_{\substack{S\subseteq[n]\\ |S|=i}}\sum_{\substack{\delta\in\{-1,1\}^S\\ \langle\delta,\varepsilon_S\rangle=i-2j}}\left\|\Delta_{[n]\setminus S}f(x+\delta k+\varepsilon_{[n]\setminus S})-\Delta_{[n]\setminus S}f(x+\delta k-\varepsilon_{[n]\setminus S})\right\|_X. \quad (54)$$

Note that the number of terms in the sum in the right hand side of (54) is $\binom{n}{i}\binom{i}{j}$. Thus

$$D_{i,j}(x)^q \leqslant \frac{1}{k^{iq}}\binom{n}{i}^{q-1}\binom{i}{j}^{q-1} \cdot \sum_{\substack{S\subseteq[n]\\|S|=i}}\sum_{\substack{\delta\in\{-1,1\}^S\\ \langle\delta,\varepsilon_S\rangle=i-2j}} \left\|\Delta_{[n]\setminus S}f(x+\delta k+\varepsilon_{[n]\setminus S})-\Delta_{[n]\setminus S}f(x+\delta k-\varepsilon_{[n]\setminus S})\right\|_X^q. \quad (55)$$

If we integrate inequality (55) with respect to x, use the translation invariance of μ to eliminate the additive term δk in the argument of the integrands, and use the fact that Δ_B is an averaging operator for all $B\subseteq[n]$, we obtain the bound

$$\begin{aligned}\int_{\mathbb{Z}_m^n} D_{i,j}(x)^q d\mu(x) &\leqslant \frac{1}{k^{iq}}\binom{n}{i}^{q-1}\binom{i}{j}^{q}\sum_{\substack{S\subseteq[n]\\|S|=i}}\int_{\mathbb{Z}_m^n}\left\|f(x+\varepsilon_{[n]\setminus S})-f(x-\varepsilon_{[n]\setminus S})\right\|_X^q d\mu(x)\\ &\leqslant \frac{2^q}{k^{iq}}\binom{n}{i}^{q-1}\binom{i}{j}^{q}\sum_{\substack{S\subseteq[n]\\|S|=i}}\int_{\mathbb{Z}_m^n}\left\|f(x+\varepsilon_{[n]\setminus S})-f(x)\right\|_X^q d\mu(x), \quad (56)\end{aligned}$$

where in the last step of (56) we used the triangle inequality as follows:

$$\left\|f(x+\varepsilon_{[n]\setminus S})-f(x-\varepsilon_{[n]\setminus S})\right\|_X^q \leqslant 2^{q-1}\left\|f(x+\varepsilon_{[n]\setminus S})-f(x)\right\|_X^q+2^{q-1}\left\|f(x)-f(x-\varepsilon_{[n]\setminus S})\right\|_X^q,$$

while noticing that upon integration with respect to x, by translation invariance, both terms become equal.

Integrating (53) with respect to x, and using (56), we see that that the left hand side of (51) is at most

$$\begin{aligned}&\sum_{i=0}^n\sum_{j=0}^i \frac{2^{(i+1)(q-1)+q}(i+1)^{q-1}\left(\frac{(i-j)!j!}{2^ik^i}\binom{n}{i}\binom{i}{j}\right)^q}{\binom{n}{i}}\sum_{\substack{S\subseteq[n]\\|S|=i}}\int_{\mathbb{Z}_m^n}\left\|f(x+\varepsilon_{[n]\setminus S})-f(x)\right\|_X^q d\mu(x)\\ &\qquad = 2^{2q-1}\sum_{i=0}^n\frac{(i+1)^q}{2^i\binom{n}{i}}\left(\frac{n!}{k^i(n-i)!}\right)^q\sum_{\substack{S\subseteq[n]\\|S|=i}}\int_{\mathbb{Z}_m^n}\left\|f(x+\varepsilon_{[n]\setminus S})-f(x)\right\|_X^q d\mu(x). \quad (57)\end{aligned}$$

Inequality (57) implies the desired bound (51), since $(i+1)^q2^{-i}\lesssim_q 1$ and $n!/(n-i)!\leqslant n^i$. $\square$

Proof of Lemma 2.1. It follows from (43) and (46) that

$$\begin{aligned}&\int_{\{-1,1\}^n}\int_{\mathbb{Z}_m^n}\left\|\sum_{l=1}^n\varepsilon_j\left[\mathcal{E}_jf(x+e_j)-\mathcal{E}_jf(x+e_j)\right]\right\|_X^q d\mu(x)d\tau(\varepsilon)\\ &\qquad = \int_{\{-1,1\}^n}\int_{\mathbb{Z}_m^n}\left\|\frac{1}{k(k+1)^{n-1}}\sum_{y\in x+\mathbb{S}}\uparrow(y-x)\odot\varepsilon\downarrow f(y)\right\|_X^q d\mu(x)d\tau(\varepsilon). \quad (58)\end{aligned}$$

An application of identity (47) now shows that

$$\int_{\{-1,1\}^n}\int_{\mathbb{Z}_m^n}\left\|\frac{1}{k(k+1)^{n-1}}\sum_{y\in x+\mathbb{S}}\uparrow(y-x)\odot\varepsilon\downarrow f(y)\right\|_X^q d\mu(x)d\tau(\varepsilon)$$
$$=\int_{\{-1,1\}^n}\int_{\mathbb{Z}_m^n}\left\|\frac{1}{k(k+1)^{n-1}}\left(\sum_{i=0}^{n}\sum_{j=0}^{i}h_{i,j}\sum_{y\in x+\mathbb{S}}b_{i,j}(y-x,\varepsilon)f(y)\right)\right\|_X^q d\mu(x)d\tau(\varepsilon). \quad (59)$$

Lemma 2.1 now follows from Lemma 3.9. □

4. Lower bounds

In this section we establish lower bounds for the best possible value of m in Theorem 1.2 that is achievable via a smoothing and approximation scheme. Our first result deals with general convolution kernels:

Proposition 4.1. *Assume that the probability measures $\nu_1,\ldots,\nu_n,\beta_1,\beta_2$ are a (q,A,S)-smoothing and approximation scheme on $\mathbb{Z}_m^n$, i.e., conditions* (7) *and* (8) *are satisfied for every Banach space X and every $f:\mathbb{Z}_m^n\to X$. Assume also that $m>cA$ for a large enough universal constant $c>0$. Then*

$$S\gtrsim_q\frac{\sqrt{n}}{A}. \quad (60)$$

Recall, as explained in Section 1.1, that in order for a smoothing and approximation scheme to yield the metric cotype inequality (5), we require $S\lesssim 1$, in which case the bound on m becomes $m\gtrsim An^{1/q}$. Proposition 4.1 shows that $S\lesssim 1$ forces the bound $A\gtrsim_q\sqrt{n}$, and correspondingly $m\gtrsim_q n^{\frac{1}{2}+\frac{1}{q}}$.

For the particular smoothing and approximation scheme used in our proof of Theorem 1.2, the following proposition establishes asymptotically sharp bounds.

Proposition 4.2. *Fix an odd integer $k\leqslant m/2$ and consider the averaging operators $\{\mathcal{E}_j\}_{j=1}^n$ used in our proof of Theorem 1.2, i.e., they are defined as in* (17)*. If there exist probability measures β_1,β_2 on $E_\infty(\mathbb{Z}_m^n)$ for which the associated approximation and smoothing inequalities* (7) *and* (8) *are satisfied for every Banach space X and every $f:\mathbb{Z}_m^n\to X$, then*

$$A\gtrsim k\quad\text{and}\quad S\gtrsim\min\left\{\sqrt{\frac{n}{k}},\frac{n}{k}\right\}. \quad (61)$$

Proposition 4.2 shows that in order to have $S\lesssim 1$ we need to require $k\gtrsim n$, in which case $A\gtrsim n$, and correspondingly $m\gtrsim n^{1+\frac{1}{q}}$, matching the bound obtained in Theorem 1.2.

4.1. A lower bound for general convolution kernels: Proof of Proposition 4.1.

Assume that the probability measures $\nu_1,\ldots,\nu_n,\beta_1,\beta_2$ are a (q,A,S)-smoothing and approximation scheme, i.e., they satisfy (7) and (8). It will be convenient to think of these measures as functions defined on the appropriate (finite) spaces, i.e., $\nu_1,\ldots,\nu_n:\mathbb{Z}_m^n\to[0,1]$ and $\beta_1,\beta_2:E_\infty(\mathbb{Z}_m^n)\to[0,1]$.

For a probability measure ν on $\mathbb{Z}_m^n$, let $P_j(\nu)$ be the probability measure on $\mathbb{Z}_m$ which is the marginal of ν on the jth coordinate, i.e.,

$$P_j(\nu)(r) \stackrel{\mathrm{def}}{=} \sum_{\substack{x\in\mathbb{Z}_m^n\\ x_j=r}} \nu(x).$$

Define the absolute value of $x \in \mathbb{Z}_m$ to be $|x| = \min\{x, m-x\}$.

Lemma 4.3. *Assume that $\nu_1, \ldots, \nu_n, \beta_1$ satisfy* (7)*. Then for every $s \in \mathbb{N}$ we have:*

$$\frac{1}{n}\sum_{j=1}^{n}\sum_{\substack{x\in\mathbb{Z}_m^n\\ |x_j|>s}} \nu_j(x) \lesssim \frac{A}{s}. \tag{62}$$

Proof. We shall apply (7) with $X = \ell_\infty^n$. Let $g_s : \mathbb{R} \to \mathbb{R}$ be the truncated jigsaw function with period $12s$, depicted in Figure 1.

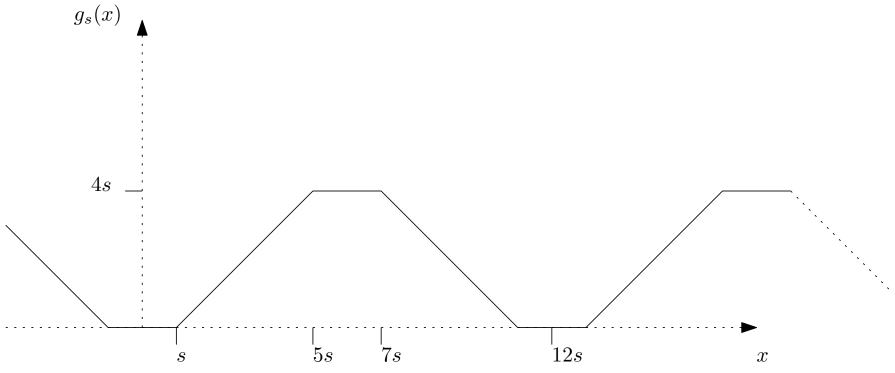

FIGURE 1. g_s is truncated jigsaw function.

Define $f_s : \mathbb{Z}_m^n \to X$ by

$$f_s(x) \stackrel{\mathrm{def}}{=} (g_s(x_1), g_s(x_2), \ldots, g_s(x_n)).$$

The Lipschitz constant of f_s with respect to the ℓ_∞ metric on $\mathbb{Z}_m^n$ is 1, and therefore it follows from (7) that

$$\left(\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n} \|f_s * \nu_j - f_s\|_{\ell_\infty^n}\, d\mu\right)^q \leqslant \frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n} \|f_s * \nu_j - f_s\|_{\ell_\infty^n}^q\, d\mu \lesssim A^q. \tag{63}$$

For every $x \in \mathbb{Z}_m^n$ and $j \in [n]$,

$$\begin{aligned}(f_s * \nu_j - f_s)(x) &= \sum_{y\in\mathbb{Z}_m^n} \nu_j(y)(f_s(x-y) - f_s(x))\\ &= \sum_{y\in\mathbb{Z}_m^n} \nu_j(y)\Big(g_s(x_1-y_1) - g_s(x_1), \ldots, g_s(x_n-y_n) - g_s(x_n)\Big).\end{aligned}$$

Assume that

$$(x_j \bmod 12s) \in [0, s] \cup [12s - s, 12s - 1]. \tag{64}$$

When $3s \leqslant |y_j| \leqslant 4s$, we have $g_s(x_j - y_j) - g_s(x_j) \geqslant s$, and for every $y_j \in \mathbb{Z}_m$, we have $g_s(x_j - y_j) - g_s(x_j) \geqslant 0$. Hence,

$$\begin{aligned}\|(f_s * \nu_j - f_s)(x)\|_{\ell_\infty^n} &\geqslant \sum_{y \in \mathbb{Z}_m^n} \nu_j(y)\big(g_s(x_j - y_j) - g_s(x_j)\big) \\ &\geqslant sP_j(\nu_j)\Big(\{z \in \mathbb{Z}_m :\ 3s \leqslant |z| \leqslant 4s\}\Big). \quad (65)\end{aligned}$$

Note that (64) holds for a constant fraction of $x \in \mathbb{Z}_m^n$, and hence by integrating (65) over $\mathbb{Z}_m^n$ we obtain:

$$\sum_{\substack{y \in \mathbb{Z}_m^n \\ 3s \leqslant |y_j| \leqslant 4s}} \nu_j(y) \lesssim \frac{1}{s} \int_{\mathbb{Z}_m^n} \|(f_s * \nu_j - f_s)(x)\|_{\ell_\infty^n} d\mu(x). \quad (66)$$

Averaging (66) over $j \in \{1, \ldots, n\}$ and using (63) we get,

$$\frac{1}{n}\sum_{j=1}^{n} \sum_{\substack{y \in \mathbb{Z}_m^n \\ 3s \leqslant |y_j| \leqslant 4s}} \nu_j(y) \overset{(66)\wedge(63)}{\lesssim} \frac{A}{s}. \quad (67)$$

Therefore

$$\frac{1}{n}\sum_{j=1}^{n} \sum_{\substack{y \in \mathbb{Z}_m^n \\ |y_j| \geqslant 3s}} \nu_j(y) = \sum_{\ell=0}^{\infty} \frac{1}{n}\sum_{j=1}^{n} \sum_{3\left\lfloor\left(\frac{4}{3}\right)^\ell s\right\rfloor \leqslant |y_j| \leqslant 4\left\lfloor\left(\frac{4}{3}\right)^\ell s\right\rfloor} \nu_j(y) \overset{(67)}{\lesssim} \sum_{\ell=0}^{\infty} \frac{A}{s \cdot (4/3)^\ell} \lesssim \frac{A}{s}. \qquad \square$$

Corollary 4.4. *Assume that $m > cA$ for a large enough universal constant $c \in \mathbb{N}$. Then:*

$$\frac{1}{n}\sum_{j=1}^{n} \sum_{z \in \mathbb{Z}_m} |P_j(\nu_j)(z+1) - P_j(\nu_j)(z-1)| \gtrsim \frac{1}{A}. \quad (68)$$

Proof. We may assume that A is an integer. By Lemma 4.3, for c large enough we have

$$\frac{1}{n}\sum_{j=1}^{n} \sum_{|z| \leqslant cA} P_j(\nu_j)(z) \geqslant \frac{3}{4} \quad \text{and} \quad \frac{1}{n}\sum_{j=1}^{n} \sum_{z=cA+2}^{3cA+2} P_j(\nu_j)(z) \leqslant \frac{1}{4}.$$

Therefore,

$$\begin{aligned}\frac{1}{2} &\leqslant \frac{1}{n}\sum_{j=1}^{n} \sum_{|z| \leqslant cA} P_j(\nu_j)(z) - \frac{1}{n}\sum_{j=1}^{n} \sum_{|z-2cA-2| \leqslant cA} P_j(\nu_j)(z) \\ &= \frac{1}{n}\sum_{j=1}^{n} \sum_{|z| \leqslant cA} [P_j(\nu_j)(z) - P_j(\nu_j)(z+2cA+2)] \\ &= \frac{1}{n}\sum_{j=1}^{n} \sum_{|z| \leqslant cA} \sum_{t=1}^{cA+1} [P_j(\nu_j)(z+2(t-1)) - P_j(\nu_j)(z+2t)] \\ &\lesssim \frac{A}{n}\sum_{j=1}^{n} \sum_{z \in \mathbb{Z}_m} |P_j(\nu_j)(z+1) - P_j(\nu_j)(z-1)|,\end{aligned}$$

as required. $\square$

Proof of Proposition 4.1. We shall apply the smoothing inequality (8) when $X = L_1(\mathbb{Z}_m^n, \mu)$ and $f : \mathbb{Z}_m^n \to X$ is defined as $f(x) = m^n \cdot \delta_{\{x\}}$, i.e., for $x \in \mathbb{Z}_m^n$ the function $f(x) : \mathbb{Z}_m^n \to \mathbb{R}$ is

$$f(x)(y) \stackrel{\mathrm{def}}{=} \begin{cases} m^n & x = y, \\ 0 & \text{otherwise}. \end{cases} \tag{69}$$

For every $\varepsilon \in \{-1,1\}^n$ and $x \in \mathbb{Z}_m^n$ we have:

$$\sum_{j=1}^n \varepsilon_j \left(f * \nu_j(x+e_j) - f * \nu_j(x-e_j) \right)$$
$$= \sum_{j=1}^n \varepsilon_j \left(\sum_{y \in \mathbb{Z}_m^n} \big(\nu_j(y+e_j) - \nu_j(y-e_j)\big) f(x-y) \right) \tag{70}$$

By Kahane's inequality [7, 21] and the fact that $L_1(\mathbb{Z}_m^n, \mu)$ has cotype 2 (see [21]),

$$\int_{\{-1,1\}^n} \left\| \sum_{j=1}^n \varepsilon_j \left(\sum_{y \in \mathbb{Z}_m^n} \big(\nu_j(y-e_j) - \nu_j(y+e_j)\big) f(x-y) \right) \right\|_{L_1(\mathbb{Z}_m^n,\mu)}^q d\tau(\varepsilon)$$
$$\gtrsim_q \left(\sum_{j=1}^n \left\| \sum_{y \in \mathbb{Z}_m^n} \big(\nu_j(y-e_j) - \nu_j(y+e_j)\big) f(x-y) \right\|_{L_1(\mathbb{Z}_m^n,\mu)}^2 \right)^{q/2} \tag{71}$$

Note that by the definition of f, for every $x \in \mathbb{Z}_m^n$ and $j \in [n]$ we have,

$$\left\| \sum_{y \in \mathbb{Z}_m^n} \big(\nu_j(y-e_j) - \nu_j(y+e_j)\big) f(x-y) \right\|_{L_1(\mathbb{Z}_m^n,\mu)} = \sum_{z \in \mathbb{Z}_m^n} |\nu_j(z-e_j) - \nu_j(z+e_j)|$$
$$\geqslant \sum_{w \in \mathbb{Z}_m} \left| \sum_{\substack{z \in \mathbb{Z}_m^n \\ z_j = w}} \big(\nu_j(z-e_j) - \nu_j(z+e_j)\big) \right| = \sum_{w \in \mathbb{Z}_m} |P_j(\nu_j)(w-1) - P_j(\nu_j)(w+1)|. \tag{72}$$

Hence,

$$\frac{1}{n} \sum_{j=1}^n \left\| \sum_{y \in \mathbb{Z}_m^n} \big(\nu_j(y-e_j) - \nu_j(y+e_j)\big) f(x-y) \right\|_{L_1(\mathbb{Z}_m^n,\mu)}^2$$
$$\stackrel{(72)}{\geqslant} \left(\frac{1}{n} \sum_{j=1}^n \sum_{w \in \mathbb{Z}_m} |P_j(\nu_j)(w-1) - P_j(\nu_j)(w+1)| \right)^2 \stackrel{(68)}{\gtrsim} \frac{1}{A^2}. \tag{73}$$

Finally, since $\|f(x)-f(y)\|_{L_1(\mathbb{Z}_m^n,\mu)} \leqslant \|f(x)\|_{L_1(\mathbb{Z}_m^n,\mu)} + \|f(y)\|_{L_1(\mathbb{Z}_m^n,\mu)} \leqslant 2$ for all $x,y\in\mathbb{Z}_m^n$, we can use the smoothing inequality (8) to deduce that

$$
\begin{aligned}
S^q &\gtrsim S^q \int_{E_\infty(\mathbb{Z}_m^n)} \|f(x)-f(y)\|^q_{L_1(\mathbb{Z}_m^n,\mu)} d\beta_2(x,y)\\
&\overset{(8)}{\geqslant} \int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n} \left\|\sum_{j=1}^n \varepsilon_j\left(f*\nu_j(x+e_j)-f*\nu_j(x-e_j)\right)\right\|^q_{L_1(\mathbb{Z}_m^n,\mu)} d\tau(\varepsilon)d\mu(x) \gtrsim_q \frac{n^{q/2}}{A^q},
\end{aligned}
$$

where in the last step we used the identity (70), combined with the inequalities (71) and (73). The proof of Proposition 4.1 is complete. □

4.2. **A sharp lower bound for $\mathcal{E}_j$ averages: Proof of Proposition 4.2.** Recall that $S(j,k)$ is defined in (16), and in the setting of Proposition 4.2 we have:

$$
\nu_j(x) = \frac{\mathbf{1}_{S(j,k)}(x)}{k(k+1)^{n-1}}.
$$

Let $s\in\{(k+1)/2,(k+3)/2\}$ be an odd integer. By the definition of $S(j,k)$ we have

$$
\sum_{\substack{x\in\mathbb{Z}_m^n\\ |x_j|>s}} \nu_j(x) = \frac{(k-s)(k+1)^{n-1}}{k(k+1)^{n-1}} \gtrsim 1.
$$

Plugging this estimate into (62) we see that $A/k\gtrsim 1$, proving the first assertion in (61).

To prove the second assertion of Proposition 4.2, we shall apply the smoothing inequality (8), as in Section 4.1, to the Banach space $X=L_1(\mathbb{Z}_m^n,\mu)$ and the function f from (69), i.e., $f(x)=m^n\delta_{\{x\}}\in L_1(\mathbb{Z}_m^n,\mu)$. We shall use here notation from Section 3.2.

In our setting, the value of

$$
\left\|\sum_{j=1}^n \varepsilon_j\left[\mathcal{E}_jf(x+e_j)-\mathcal{E}_jf(x-e_j)\right]\right\|_{L_1(\mathbb{Z}_m^n,\mu)}
$$

does not depend on $x\in\mathbb{Z}_m^n$ and $\varepsilon\in\{-1,1\}^n$. Thus the left hand side of (8) equals (by Lemma 3.5),

$$
\left\|\sum_j\left(\mathcal{E}_jf(e_j)-\mathcal{E}_jf(-e_j)\right)\right\|^q_{L_1(\mathbb{Z}_m^n,\mu)} = \left(\frac{1}{k(k+1)^{n-1}}\sum_{y\in\mathbb{S}}\Big|\!\uparrow\! y\!\downarrow\!\Big|\right)^q.
$$

At the same time, as noted in Section 4.1, the right hand side of (8) is $\lesssim S^q$. It follows that

$$
S \gtrsim \frac{1}{k(k+1)^{n-1}}\sum_{y\in\mathbb{S}}\Big|\!\uparrow\! y\!\downarrow\!\Big| = \mathbb{E}\left[Z\right], \tag{74}
$$

where $Z=|\sum_{i=1}^n\xi_i|$, and $\{\xi_i\}_{i=1}^n$ are i.i.d. random variables taking the 0 with probability $\frac{k-1}{k+1}$, and each of the values $\{-1,1\}$ with probability $\frac{1}{k+1}$. The last equality in (74) is an immediate consequence of the definitions of $\mathbb{S}$ and $\Big|\!\uparrow\! y\!\downarrow\!\Big|$. Writing $p=\frac{2}{k+1}$, we have $\mathbb{E}\left[Z^2\right]=np$ and $\mathbb{E}\left[Z^4\right]=np+n(n-1)p^2$. By Hölder's inequality it then follows that we have $S\gtrsim\mathbb{E}[Z]\geqslant\|Z\|_2^3/\|Z\|_4^2\asymp\min\left\{\sqrt{np},np\right\}$, completing the proof of Proposition 4.2. □

4.3. **Symmetrization.** We do not know what is the smallest m for which the metric cotype inequality (5) can be shown to hold true via a smoothing and approximation scheme: all we know is that it is between $n^{1+\frac{1}{q}}$ and $n^{\frac{1}{2}+\frac{1}{q}}$. In this short section, we note that the special symmetric structure of the smoothing and approximation scheme that we used in the proof of Theorem 1.2 can be always assumed to hold true without loss of generality. This explains why our choice of convolution kernels is natural. Additionally, this fact might be useful in improving the lower bound on m of Proposition 4.1, though we do not know how to use it in our current proof of Proposition 4.1.

For $\pi \in S_n$, i.e., a permutation of $[n]$, and $x \in \mathbb{Z}_m^n$, write

$$x^\pi \stackrel{\mathrm{def}}{=} \left(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}\right).$$

For $f : \mathbb{Z}_m^n \to X$ we define $f^\pi : \mathbb{Z}_m^n \to X$ by $f^\pi(x) = f(x^\pi)$. Note that if ν is a probability measure on $\mathbb{Z}_m^n$ then for all $x \in \mathbb{Z}_m^n$ we have

$$f * \nu^\pi = \left(f^{\pi^{-1}} * \nu\right)^\pi. \tag{75}$$

Indeed,

$$\begin{aligned} f * \nu^\pi(x) &= \int_{\mathbb{Z}_m^n} f(x-y)\nu(y^\pi)d\mu(y) = \int_{\mathbb{Z}_m^n} f\left(x - z^{\pi^{-1}}\right)\nu(z)d\mu(z) \\ &= \int_{\mathbb{Z}_m^n} f^{\pi^{-1}}\left(x^\pi - z\right)\nu(z)d\mu(z) = f^{\pi^{-1}} * \nu(x^\pi) = \left(f^{\pi^{-1}} * \nu\right)^\pi(x). \end{aligned}$$

It follows from (75) that

$$\|f * \nu^\pi - f\|_{L_q(\mathbb{Z}_m^n, X)} = \left\|f^{\pi^{-1}} * \nu - f^{\pi^{-1}}\right\|_{L_q(\mathbb{Z}_m^n, X)}. \tag{76}$$

Lemma 4.5. *Assume that the probability measures $\nu_1, \ldots, \nu_n, \beta_1, \beta_2$ are a (q, A, S)-smoothing and approximation scheme. Then there exist probability measures $\bar{\nu}_1, \ldots, \bar{\nu}_n$ on $\mathbb{Z}_m^n$ and two probability measures $\bar{\beta}_1, \bar{\beta}_2$ on $E_\infty(\mathbb{Z}_m^n)$, such that*

1. *The sequence $\bar{\nu}_1, \ldots, \bar{\nu}_n, \bar{\beta}_1, \bar{\beta}_2$ is also a (q, A, S)-smoothing and approximation scheme,*
2. *For any $j, h \in [n]$ we have $\bar{\nu}_j = \bar{\nu}_h^{(j,h)}$, where $(j, h) \in S_n$ is the transposition of j and h.*
3. *For every $j, h \in [n] \setminus \{i\}$ we have $P_j(\bar{\nu}_i) = P_h(\bar{\nu}_i)$.*

Proof. Define for $j \in [n]$,

$$\bar{\nu}_j \stackrel{\mathrm{def}}{=} \frac{1}{n!} \sum_{\pi \in S_n} \nu_{\pi(j)}^{\pi^{-1}}. \tag{77}$$

We also define for $(x, y) \in E_\infty(\mathbb{Z}_m^n)$,

$$\bar{\beta}_1(x, y) \stackrel{\mathrm{def}}{=} \frac{1}{n!} \sum_{\pi \in S_n} \beta_1\left(x^\pi, y^\pi\right) \quad \text{and} \quad \bar{\beta}_2(x, y) \stackrel{\mathrm{def}}{=} \frac{1}{n!} \sum_{\pi \in S_n} \beta_2\left(x^\pi, y^\pi\right). \tag{78}$$

Fix $f : \mathbb{Z}_m^n \to X$ and assume the validity of the approximation and smoothing inequalities (7), (8). Then, by the convexity of $\|\cdot\|_X^q$,

$$\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{Z}_m^n}\|f*\bar{\nu}_j - f\|_X^q d\mu \overset{(76)\wedge(77)}{\leqslant} \frac{1}{n!}\sum_{\pi\in S_n}\frac{1}{n}\sum_{j=1}^{n}\left\|f^\pi * \nu_{\pi(j)} - f^\pi\right\|_{L_q(\mathbb{Z}_m^n,X)}^q$$
$$\overset{(7)\wedge(78)}{\leqslant} A^q \int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\bar{\beta}_1(x,y). \quad (79)$$

This is precisely the approximation property for $\bar{\nu}_1,\ldots,\bar{\nu}_n,\bar{\beta}_1,\bar{\beta}_2$.

Similarly,

$$\int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{j=1}^{n}\varepsilon_j\Big(f*\bar{\nu}_j(x+e_j) - f*\bar{\nu}_j(x-e_j)\Big)\right\|_X^q d\tau(\varepsilon)d\mu(x)$$
$$\overset{(78)}{\leqslant} \frac{1}{n!}\sum_{\pi\in S_n}\int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{j=1}^{n}\varepsilon_j\Big(f*\nu_{\pi(j)}^{\pi^{-1}}(x+e_j) - f*\nu_{\pi(j)}^{\pi^{-1}}(x-e_j)\Big)\right\|_X^q d\tau(\varepsilon)d\mu(x). \quad (80)$$

Note that

$$\sum_{j=1}^{n}\varepsilon_j\Big(f*\nu_{\pi(j)}^{\pi^{-1}}(x+e_j) - f*\nu_{\pi(j)}^{\pi^{-1}}(x-e_j)\Big)$$
$$\overset{(75)}{=} \sum_{i=1}^{n}\varepsilon_{\pi^{-1}(i)}\left(f^\pi*\nu_i\left(x^{\pi^{-1}}+e_i\right) - f^\pi*\nu_i\left(x^{\pi^{-1}}-e_i\right)\right), \quad (81)$$

where we made the change of variable $j=\pi^{-1}(i)$ and used the fact that $e_r^{\pi^{-1}} = e_{\pi(r)}$ for all $r\in[n]$ and $\pi\in S_n$. Hence,

$$\int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{j=1}^{n}\varepsilon_j\Big(f*\nu_{\pi(j)}^{\pi^{-1}}(x+e_j) - f*\nu_{\pi(j)}^{\pi^{-1}}(x-e_j)\Big)\right\|_X^q d\tau(\varepsilon)d\mu(x)$$
$$\overset{(81)}{=} \int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{r=1}^{n}\varepsilon_r\Big(f^\pi*\nu_r(x+e_r) - f^\pi*\nu_r(x-e_r)\Big)\right\|_X^q d\tau(\varepsilon)d\mu(x). \quad (82)$$

The smoothing inequality for $\bar{\nu}_1,\ldots,\bar{\nu}_n,\bar{\beta}_1,\bar{\beta}_2$ now follows:

$$\int_{\mathbb{Z}_m^n}\int_{\{-1,1\}^n}\left\|\sum_{j=1}^{n}\varepsilon_j\Big(f*\bar{\nu}_j(x+e_j) - f*\bar{\nu}_j(x-e_j)\Big)\right\|_X^q d\tau(\varepsilon)d\mu(x)$$
$$\overset{(80)\wedge(82)\wedge(8)}{\leqslant} S^q\int_{E_\infty(\mathbb{Z}_m^n)}\|f(x)-f(y)\|_X^q d\bar{\beta}_2(x,y).$$

Assertions 2. and 3. of Lemma 4.5 follow directly from the definition (77). $\square$

References

[1] B. Begun. A remark on almost extensions of Lipschitz functions. *Israel J. Math.*, 109:151–155, 1999.

[2] J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56(2):222–230, 1986.

[3] J. Bourgain. Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 157–167. Springer, Berlin, 1987.

[4] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely summing operators*, volume 43 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[5] O. Giladi and A. Naor. Improved bounds in the scaled Enflo type inequality for Banach spaces. Available at `http://arxiv.org/abs/1004.4221`, 2010.

[6] L. Hörmander. *An introduction to complex analysis in several variables*, volume 7 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, third edition, 1990.

[7] J.-P. Kahane. Sur les sommes vectorielles $\sum \pm u_n$. *C. R. Acad. Sci. Paris*, 259:2577–2580, 1964.

[8] N. J. Kalton. Coarse and uniform embeddings into reflexive spaces. *Q. J. Math.*, 58(3):393–414, 2007.

[9] B. Maurey. Type, cotype and K-convexity. In *Handbook of the geometry of Banach spaces, Vol. 2*, pages 1299–1332. North-Holland, Amsterdam, 2003.

[10] B. Maurey and G. Pisier. Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach. *Studia Math.*, 58(1):45–90, 1976.

[11] M. Mendel. Metric dichotomies. In *Limits of graphs in group theory and computer science*, pages 59–76. EPFL Press, Lausanne, 2009.

[12] M. Mendel and A. Naor. Scaled Enflo type is equivalent to Rademacher type. *Bull. Lond. Math. Soc.*, 39(3):493–498, 2007.

[13] M. Mendel and A. Naor. Metric cotype. *Ann. of Math. (2)*, 168(1):247–298, 2008.

[14] A. Naor. An application of metric cotype to quasisymmetric embeddings. Preprint, 2006. Available at `http://arxiv.org/abs/math/0607644`.

[15] G. Pisier. Holomorphic semigroups and the geometry of Banach spaces. *Ann. of Math. (2)*, 115(2):375–392, 1982.

[16] G. Pisier. Probabilistic methods in the geometry of Banach spaces. In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 167–241. Springer, Berlin, 1986.

[17] Y. Raynaud. On ultrapowers of non commutative L_p spaces. *J. Operator Theory*, 48(1):41–68, 2002.

[18] N. Tomczak-Jaegermann. The moduli of smoothness and convexity and the Rademacher averages of trace classes $S_p (1 \leqslant p < \infty)$. *Studia Math.*, 50:163–182, 1974.

[19] A. Veomett and K. Wildrick. Spaces of small metric cotype. Preprint, 2010. Available at `http://arxiv.org/abs/1001.3326`.

[20] H. S. Wilf. *generatingfunctionology*. A K Peters Ltd., Wellesley, MA, third edition, 2006.

[21] P. Wojtaszczyk. *Banach spaces for analysts*, volume 25 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1991.

Courant Institute, New York University
E-mail address: `giladi@cims.nyu.edu`

Computer Science Division, The Open University of Israel
E-mail address: `mendelma@gmail.com`

Courant Institute, New York University
E-mail address: `naor@cims.nyu.edu`