The Integrality Gap of the Goemans–Linial SDP Relaxation for Sparsest Cut Is at least a Constant Multiple of $\sqrt{\log n}$*

Assaf Naor†
Princeton University
Princeton, NJ, USA 08544-1000
naor@math.princeton.edu

Robert Young‡
New York University
New York, NY, USA 10012-1185
ryoung@cims.nyu.edu

ABSTRACT

We prove that the integrality gap of the Goemans–Linial semidefinite programming relaxation for the Sparsest Cut Problem is $\Omega(\sqrt{\log n})$ on inputs with n vertices, thus matching the previously best known upper bound $(\log n)^{\frac{1}{2}+o(1)}$ up to lower-order factors. This statement is a consequence of the following new isoperimetric-type inequality. Consider the 8-regular graph whose vertex set is the 5-dimensional integer grid $\mathbb{Z}^5$ and where each vertex $(\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e}) \in \mathbb{Z}^5$ is connected to the 8 vertices $(\mathsf{a}\pm 1,\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e})$, $(\mathsf{a},\mathsf{b}\pm 1,\mathsf{c},\mathsf{d},\mathsf{e})$, $(\mathsf{a},\mathsf{b},\mathsf{c}\pm 1,\mathsf{d},\mathsf{e}\pm\mathsf{a})$, $(\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d}\pm 1,\mathsf{e}\pm\mathsf{b})$. This graph is known as the Cayley graph of the 5-dimensional discrete Heisenberg group. Given $\Omega \subseteq \mathbb{Z}^5$, denote the size of its edge boundary in this graph (a.k.a. the *horizontal perimeter* of Ω) by $|\partial_{\mathsf{h}}\Omega|$. For $t \in \mathbb{N}$, denote by $|\partial^t_{\mathsf{v}}\Omega|$ the number of $(\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e}) \in \mathbb{Z}^5$ such that exactly one of the two vectors $(\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e}), (\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e}+t)$ is in Ω. The *vertical perimeter* of Ω is defined to be $|\partial_{\mathsf{v}}\Omega| = \sqrt{\sum_{t=1}^{\infty} |\partial^t_{\mathsf{v}}\Omega|^2/t^2}$. We show that every subset $\Omega \subseteq \mathbb{Z}^5$ satisfies $|\partial_{\mathsf{v}}\Omega| = O(|\partial_{\mathsf{h}}\Omega|)$. This *vertical-versus-horizontal isoperimetric inequality* yields the above-stated integrality gap for Sparsest Cut and answers several geometric and analytic questions of independent interest.

The theorem stated above is the culmination of a program whose aim is to understand the performance of the Goemans–Linial semidefinite program through the embeddability properties of Heisenberg groups. These investigations have mathematical significance even beyond their established relevance to approximation algorithms and combinatorial optimization. In particular they contribute to a range of mathematical disciplines including functional analysis, geometric group theory, harmonic analysis, sub-Riemannian geometry, geometric measure theory, ergodic theory, group representations, and metric differentiation. This article builds on the above cited works, with the "twist" that while those works were equally valid for any finite dimensional Heisenberg group, our result holds for the Heisenberg group of dimension 5 (or higher) but *fails* for the 3-dimensional Heisenberg group. This insight leads to our core contribution, which is a deduction of an endpoint L_1-boundedness of a certain singular integral on $\mathbb{R}^5$ from the (local) L_2-boundedness of the corresponding singular integral on $\mathbb{R}^3$. To do this, we devise a corona-type decomposition of subsets of a Heisenberg group, in the spirit of the construction that David and Semmes performed in $\mathbb{R}^n$, but with two main conceptual differences (in addition to more technical differences that arise from the peculiarities of the geometry of Heisenberg group). Firstly, the"atoms" of our decomposition are perturbations of *intrinsic Lipschitz graphs* in the sense of Franchi, Serapioni, and Serra Cassano (plus the requisite "wild" regions that satisfy a Carleson packing condition). Secondly, we control the local overlap of our corona decomposition by using quantitative monotonicity rather than Jones-type β-numbers.

CCS CONCEPTS

• **Theory of computation** → **Approximation algorithms analysis**; **Mathematical optimization**;

KEYWORDS

Sparsest Cut Problem, approximation algorithms, semidefinite programming, metric embeddings.

ACM Reference format:
Assaf Naor and Robert Young. 2017. The Integrality Gap of the Goemans–Linial SDP Relaxation for Sparsest Cut Is at least a Constant Multiple of $\sqrt{\log n}$. In *Proceedings of 49th Annual ACM SIGACT Symposium on the Theory of Computing, Montreal, Canada, June 2017 (STOC'17),* 12 pages.
DOI: 10.1145/3055399.3055413

1 INTRODUCTION

Fix $n \in \mathbb{N}$. The input of the *Sparsest Cut Problem* consists of two n by n symmetric matrices with nonnegative entries $C = (C_{ij})$, $D = (D_{ij}) \in M_n([0,\infty))$, which are commonly called capacities and demands, respectively. The goal is to design a polynomial-time algorithm to evaluate the quantity

$$\mathrm{OPT}(C,D) \stackrel{\mathrm{def}}{=} \min_{\varnothing \subsetneq A \subsetneq \{1,\ldots,n\}} \frac{\sum_{(i,j)\in A\times(\{1,\ldots,n\}\smallsetminus A)} C_{ij}}{\sum_{(i,j)\in A\times(\{1,\ldots,n\}\smallsetminus A)} D_{ij}}. \quad (1)$$

In view of the extensive literature on the Sparsest Cut Problem, it would be needlessly repetitive to recount here the rich and multifaceted impact of this optimization problem on computer science and mathematics; see instead the articles [1, 59], the surveys [16, 62, 73, 85], Chapter 10 of the monograph [32], Chapter 15

*The research presented here was conducted under the auspices of the Simons Algorithms and Geometry (A&G) Think Tank. A full version of this extended abstract, titled "Vertical perimeter versus horizontal perimeter," that contains complete proofs and additional results is available at https://arxiv.org/abs/1701.00620. Nevertheless, this extended abstract contains material that is not included in the full version.

†Supported by BSF grant 2010021, the Packard Foundation and the Simons Foundation.
‡Supported by the NSF and the Sloan Foundation. This material is based upon work supported by the National Science Foundation under Grant No. 1612061.

DOI: 10.1145/3055399.3055413

of the monograph [68], Chapter 1 of the monograph [80], and the references therein. It suffices to say that by tuning the choice of matrices C, D to the problem at hand, the minimization in (1) finds a partition of the "universe" $\{1, \ldots, n\}$ into two parts, namely the sets A and $\{1, \ldots, n\} \setminus A$, whose appropriately weighted interface is as small as possible, thus allowing for inductive solutions of various algorithmic tasks, a procedure known as *divide and conquer.* (Not all of the uses of the Sparsest Cut Problem fit into this framework. A recent algorithmic application of a different nature can be found in [66].)

It is NP-hard to compute $\mathsf{OPT}(C, D)$ in polynomial time [84]. By [25] there exists $\varepsilon_0 > 0$ such that it is even NP-hard to compute $\mathsf{OPT}(C, D)$ within a multiplicative factor of less than $1 + \varepsilon_0$. If one assumes Khot's Unique Games Conjecture [46, 47, 88] then by [18, 49] there does not exist a polynomial-time algorithm that can compute $\mathsf{OPT}(C, D)$ within any universal constant factor.

By the above hardness results, a much more realistic goal would be to design a polynomial-time algorithm that takes as input the capacity and demand matrices $C, D \in M_n([0, \infty))$ and outputs a number $\mathsf{ALG}(C, D)$ that is guaranteed to satisfy

$$\mathsf{ALG}(C, D) \leqslant \mathsf{OPT}(C, D) \leqslant \rho(n)\mathsf{ALG}(C, D),$$

with (hopefully) the quantity $\rho(n)$ growing to ∞ slowly as $n \to \infty$. Determining the best possible asymptotic behaviour of $\rho(n)$ (assuming $P \neq NP$) is an open problem of major importance.

In [6, 64] an algorithm was designed, based on linear programming (through the connection to multicommodity flows) and Bourgain's embedding theorem [12], which yields $\rho(n) = O(\log n)$. An algorithm based on semidefinite programming (to be described precisely below) was proposed by Goemans and Linial in the mid-1990s. To the best of our knowledge this idea first appeared in the literature in [38, page 158], where it was speculated that it might yield a constant factor approximation for Sparsest Cut (see also [62, 63]). In what follows, we denote the approximation ratio of the Goemans–Linial algorithm on inputs of size at most n by $\rho_{\mathsf{GL}}(n)$. The hope that $\rho_{\mathsf{GL}}(n) = O(1)$ was dashed in the remarkable work [49], where the lower bound $\rho_{\mathsf{GL}}(n) \gtrsim \sqrt[6]{\log \log n}$ was proven.[1] An improved analysis of the ideas of [49] was conducted in [52], yielding the estimate $\rho_{\mathsf{GL}}(n) \gtrsim \log \log n$. An entirely different approach based on the geometry of the Heisenberg group was introduced in [56]. In combination with the important works [19, 20] it gives a different proof that $\lim_{n\to\infty} \rho_{\mathsf{GL}}(n) = \infty$. In [21, 22] the previously best-known bound $\rho_{\mathsf{GL}}(n) \gtrsim (\log n)^{\delta}$ was obtained for an effective (but small) positive universal constant δ.

Despite these lower bounds, the Goemans–Linial algorithm yields an approximation ratio of $o(\log n)$, so it is asymptotically more accurate than the linear program of [6, 64]. Specifically, in [17] it was shown that $\rho_{\mathsf{GL}}(n) \lesssim (\log n)^{\frac{3}{4}}$. This was improved in [3] to $\rho_{\mathsf{GL}}(n) \lesssim (\log n)^{\frac{1}{2}+o(1)}$. See Section 1.7 below for additional background on the results quoted above. No other polynomial-time algorithm for the Sparsest Cut problem is known (or conjectured) to have an approximation ratio that is asymptotically better than that of the Goemans–Linial algorithm. However, despite major scrutiny by researchers in approximation algorithms, the asymptotic behavior of $\rho_{\mathsf{GL}}(n)$ as $n \to \infty$ remained unknown. Theorem 1.1 below resolves this question up to lower-order factors.

Theorem 1.1. *The approximation ratio of the Goemans–Linial algorithm satisfies* $\rho_{\mathsf{GL}}(n) \gtrsim \sqrt{\log n}$.

1.1 The SDP Relaxation

The Goemans–Linial algorithm is simple to describe. It takes as input the symmetric matrices $C, D \in M_n([0, \infty))$ and proceeds to compute the following quantity.

$$\mathsf{SDP}(C, D) \stackrel{\mathrm{def}}{=} \inf_{(v_1,\ldots,v_n)\in\mathsf{NEG}_n} \frac{\sum_{i=1}^n \sum_{j=1}^n C_{ij}\|v_i - v_j\|_2^2}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}\|v_i - v_j\|_2^2},$$

where

$$\mathsf{NEG}_n \stackrel{\mathrm{def}}{=} \Big\{(v_1, \ldots v_n) \in (\mathbb{R}^n)^n : \\ \|v_i - v_j\|_2^2 \leqslant \|v_i - v_k\|_2^2 + \|v_k - v_j\|_2^2 \\ \text{for all } i, j, k \in \{1, \ldots, n\}\Big\}.$$

Thus NEG_n is the set of n-tuples $(v_1, \ldots v_n)$ of vectors in $\mathbb{R}^n$ such that $(\{v_1, \ldots, v_n\}, \nu_n)$ is a semi-metric space, where $\nu_n : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ is defined by $\nu_n(x, y) = \sum_{j=1}^n (x_j - y_j)^2 = \|x - y\|_2^2$ for every $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. A semi-metric space (X, d_X) is said [32] to be of *negative type* if $(X, \sqrt{d_X})$ embeds isometrically into a Hilbert space. So, NEG_n can be described as the set of all (ordered) negative type semi-metrics of size n. It is simple to check that the evaluation of the quantity $\mathsf{SDP}(C, D)$ can be cast as a semidefinite program (SDP), so it can be achieved (up to $o(1)$ precision) in polynomial time [39]. One has $\mathsf{SDP}(C, D) \leqslant \mathsf{OPT}(C, D)$ for all symmetric matrices $C, D \in M_n([0, \infty))$. See e.g. [69, Section 15.9] or [73, Section 4.3] for an explanation of the above assertions about $\mathsf{SDP}(C, D)$, as well as additional background and motivation. The pertinent question is therefore to evaluate the asymptotic behavior as $n \to \infty$ of the sequence

$$\rho_{\mathsf{GL}}(n) \stackrel{\mathrm{def}}{=} \sup_{\substack{C, D\in M_n([0,\infty)) \\ C, D \text{ symmetric}}} \frac{\mathsf{OPT}(C, D)}{\mathsf{SDP}(C, D)}.$$

This is the quantity $\rho_{\mathsf{GL}}(n)$ appearing in Theorem 1.1, also known as the *integrality gap* of the Goemans–Linial semidefinite programming relaxation for the Sparsest Cut Problem.

1.2 Bi-Lipschitz Embeddings

A duality argument of Rabinovich (see [73, Lemma 4.5] or [21, Section 1]) establishes that $\rho_{\mathsf{GL}}(n)$ is equal to the largest possible L_1-distortion of an n-point semi-metric of negative type. If $d : \{1, \ldots, n\}^2 \to [0, \infty)$ is a semi-metric, its L_1 distortion, denoted $c_1(\{1, \ldots, n\}, d)$, is the smallest $D \in [1, \infty)$ for which there are integrable functions[2] $f_1, \ldots, f_n : [0, 1] \to \mathbb{R}$ such that $\int_0^1 |f_i(t) - f_j(t)|\, \mathrm{d}t \leqslant d(i, j) \leqslant D \int_0^1 |f_i(t) - f_j(t)|\, \mathrm{d}t$ for every $i, j \in \{1, \ldots, n\}$. Rabinovich's duality argument proves that $\rho_{\mathsf{GL}}(n)$ is equal to the

[1] Here, and in what follows, we use the following (standard) asymptotic notation. Given $a, b > 0$, the notations $a \lesssim b$ and $b \gtrsim a$ mean that $a \leqslant \mathsf{K}b$ for some universal constant $\mathsf{K} > 0$. The notation $a \asymp b$ stands for $(a \lesssim b) \wedge (b \lesssim a)$. Thus $a \lesssim b$ and $a \gtrsim b$ are the same as $a = O(b)$ and $a = \Omega(b)$, respectively, and $a \asymp b$ is the same as $a = \Theta(b)$.

[2] If one wishes to use finite-dimensional vectors rather than functions then by [89] there exist $v_1, \ldots, v_n \in \mathbb{R}^{n(n-1)/2}$ such that $\int_0^1 |f_i(t) - f_j(t)|\, \mathrm{d}t = \|v_i - v_j\|_1 = \sum_{k=1}^{n(n-1)/2} |v_{ik} - v_{jk}|$ for every $i, j \in \{1, \ldots, n\}$.

maximum of $c_1(\{1,\ldots,n\},d)$ over all possible semi-metrics d of negative type on $\{1,\ldots,n\}$. Hence, Theorem 1.1 is equivalent to the assertion that for every $n\in\mathbb{N}$ there exists a metric of negative type $d:\{1,\ldots,n\}^2\to[0,\infty)$ for which $c_1(\{1,\ldots,n\},d)\gtrsim\sqrt{\log n}$.

1.3 A Poorly-Embeddable Metric

The 5-dimensional discrete Heisenberg group, denoted $\mathbb{H}^5_{\mathbb{Z}}$, is the following group of 4 by 4 invertible matrices, equipped with the usual matrix multiplication.

$$\mathbb{H}^5_{\mathbb{Z}}\stackrel{\mathrm{def}}{=}\left\{\begin{pmatrix}1&\mathrm{a}&\mathrm{b}&\mathrm{e}\\0&1&0&\mathrm{c}\\0&0&1&\mathrm{d}\\0&0&0&1\end{pmatrix}:\mathrm{a,b,c,d,e}\in\mathbb{Z}\right\}\subseteq \mathrm{GL}_4(\mathbb{R}). \tag{2}$$

This group is generated by the symmetric set

$$S\stackrel{\mathrm{def}}{=}\{X_1,X_1^{-1},X_2,X_2^{-1},Y_1,Y_1^{-1},Y_2,Y_2^{-1}\},$$

where

$$\begin{aligned}X_1&\stackrel{\mathrm{def}}{=}\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}, & X_2&\stackrel{\mathrm{def}}{=}\begin{pmatrix}1&0&1&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\\ Y_1&\stackrel{\mathrm{def}}{=}\begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{pmatrix}, & Y_2&\stackrel{\mathrm{def}}{=}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&1\\0&0&0&1\end{pmatrix}.\end{aligned} \tag{3}$$

For notational convenience we shall identify the matrix in (2) with the vector $(\mathrm{a,b,c,d,e})\in\mathbb{Z}^5$. This yields an identification of $\mathbb{H}^5_{\mathbb{Z}}$ with the 5-dimensional integer grid $\mathbb{Z}^5$. We view $\mathbb{Z}^5$ as a (non-commutative) group equipped with the product that is inherited from matrix multiplication through the above identification, i.e., for every $(\mathrm{a,b,c,d,e}),(\alpha,\beta,\gamma,\delta,\epsilon)\in\mathbb{Z}^5$ we set

$$\begin{aligned}&(\mathrm{a,b,c,d,e})(\alpha,\beta,\gamma,\delta,\epsilon)\\&\qquad\stackrel{\mathrm{def}}{=}(\mathrm{a}+\alpha,\mathrm{b}+\beta,\mathrm{c}+\gamma,\mathrm{d}+\delta,\mathrm{e}+\epsilon+\mathrm{a}\gamma+\mathrm{b}\delta).\end{aligned} \tag{4}$$

Note that under the above identification the identity element of $\mathbb{H}^5_{\mathbb{Z}}$ is the zero vector $\mathbf{0}\in\mathbb{Z}^5$, the inverse of an element $h=(\mathrm{a,b,c,d,e})\in\mathbb{Z}^5$ is $h^{-1}=(-\mathrm{a},-\mathrm{b},-\mathrm{c},-\mathrm{d},-\mathrm{e}+\mathrm{ac}+\mathrm{bd})$, and the generators X_1,X_2,Y_1,Y_2 in (3) are the first four standard basis elements of $\mathbb{R}^5$. Let Z denote the fifth standard basis element of $\mathbb{R}^5$, i.e., $Z=(0,0,0,0,1)$. We then have the relations $[X_1,Y_1]=[X_2,Y_2]=Z$ and $[X_1,X_2]=[X_1,Y_2]=[X_1,Z]=[Y_1,X_2]=[Y_1,Y_2]=[Y_1,Z]=[X_2,Z]=[Y_2,Z]=\mathbf{0}$, where we recall the standard commutator notation $[g,h]=ghg^{-1}h^{-1}$ for every two group elements $g,h\in\mathbb{H}^5_{\mathbb{Z}}$. In other words, any two elements from $\{X_1,X_2,Y_1,Y_2,Z\}$ other than X_1,Y_1 or X_2,Y_2 commute, and the commutators of X_1,Y_1 and X_2,Y_2 are both equal to Z. In particular, Z commutes with all of the members of the generating set S, and therefore Z is in the *center* of $\mathbb{H}^5_{\mathbb{Z}}$. It is worthwhile to mention that these commutation relations could be used to define the group $\mathbb{H}^5_{\mathbb{Z}}$ abstractly using generators and relations, but this fact will not be needed in what follows.

This group structure induces a graph $\mathcal{X}_S(\mathbb{H}^5_{\mathbb{Z}})$ on $\mathbb{Z}^5$, called the *Cayley graph* of $\mathbb{H}^5_{\mathbb{Z}}$. The edges of this graph are defined to be the unordered pairs of the form $\{h,hs\}$, where $h\in\mathbb{Z}^5$ and $s\in S$. This is an 8-regular connected graph, and by the group law (4), the neighbors of each vertex $(\mathrm{a,b,c,d,e})\in\mathbb{Z}^5$ are $(\mathrm{a}\pm1,\mathrm{b,c,d,e})$, $(\mathrm{a},\mathrm{b}\pm1,\mathrm{c,d,e})$, $(\mathrm{a,b},\mathrm{c}\pm1,\mathrm{d},\mathrm{e}\pm\mathrm{a})$, $(\mathrm{a,b,c},\mathrm{d}\pm1,\mathrm{e}\pm\mathrm{b})$. The shortest-path metric on $\mathbb{Z}^5$ that is induced by this graph structure will be denoted below by $d_W:\mathbb{Z}^5\times\mathbb{Z}^5\to\mathbb{N}\cup\{0\}$. This metric is also known as the left-invariant *word metric* on the Heisenberg group $\mathbb{H}^5_{\mathbb{Z}}$. For every $R\in[0,\infty)$ denote the (closed) ball of radius R centered at the identity element by $\mathcal{B}_R=\{h\in\mathbb{Z}^5:\ d_W(h,\mathbf{0})\leqslant R\}$. It is well-known (see e.g. [10]) that $|\mathcal{B}_R|\asymp R^6$ and $d_W(\mathbf{0},Z^R)\asymp\sqrt{R}$ for every $R\in\mathbb{N}$. Our main result is the following theorem.

Theorem 1.2. *For all $R\geqslant 2$ we have $c_1(\mathcal{B}_R,d_W)\asymp\sqrt{\log R}$.*

The new content of Theorem 1.2 is the bound $c_1(\mathcal{B}_R,d_W)\gtrsim\sqrt{\log R}$. The matching upper bound $c_1(\mathcal{B}_R,d_W)\lesssim\sqrt{\log R}$ has several proofs in the literature; see e.g. the discussion immediately following Corollary 1.3 in [53] or Section 1.7.2 below. The previous best known estimate [22] was that there exists a universal constant $\delta>0$ such that $c_1(\mathcal{B}_R,d_W)\geqslant(\log R)^\delta$. By [56, Theorem 2.2] the metric d_W is bi-Lipschitz equivalent to a metric on $\mathbb{H}^5_{\mathbb{Z}}$ that is of negative type. We remark that [56] makes this assertion for a different metric on a larger continuous group that contains $\mathbb{H}^5_{\mathbb{Z}}$ as a discrete co-compact subgroup, but by a simple general result (e.g. [14, Theorem 8.3.19]) the word metric d_W is bi-Lipschitz equivalent to the metric considered in [56]. Since $|\mathcal{B}_R|\asymp R^6$, we have $\sqrt{\log|\mathcal{B}_R|}\asymp\sqrt{\log R}$, so Theorem 1.2 implies Theorem 1.1 through the duality result of Rabinovich that was recalled in Section 1.2.

The following precise theorem about L_1 embeddings that need not be bi-Lipschitz implies Theorem 1.2 by considering the special case of the modulus $\omega(t)=t/D$ for $D\geqslant1$ and $t\in[0,\infty)$.

Theorem 1.3. *There exists a universal constant $c\in(0,1)$ with the following property. Fix $R\geqslant2$ and a nondecreasing function $\omega:[1,\infty)\to[1,\infty)$. Then there exists $\phi:\mathcal{B}_R\to L_1$ for which every distinct $x,y\in\mathcal{B}_R$ satisfy*

$$\omega\big(d_W(x,y)\big)\lesssim\|\phi(x)-\phi(y)\|_1\leqslant d_W(x,y), \tag{5}$$

if and only if *$\omega(t)\lesssim t$ for all $t\in[1,\infty)$ and*

$$\int_1^{cR}\frac{\omega(s)^2}{s^3}\,\mathrm{d}s\lesssim1. \tag{6}$$

The fact that the integrability requirement (6) implies the existence of the desired embedding ϕ is due to [87, Corollary 5]. The new content of Theorem 1.3 is that the existence of the embedding ϕ implies (6). By letting $R\to\infty$ in Theorem 1.3 we see that there exists $\phi:\mathbb{Z}^5\to L_1$ that satisfies

$$\forall\,x,y\in\mathbb{Z}^5,\quad \omega\big(d_W(x,y)\big)\lesssim\|\phi(x)-\phi(y)\|_1\leqslant d_W(x,y), \tag{7}$$

if and only if

$$\int_1^{\infty}\frac{\omega(s)^2}{s^3}\,\mathrm{d}s\lesssim1. \tag{8}$$

In [22] it was shown that if $\phi:\mathbb{Z}^5\to L_1$ satisfies (7), then there must exist arbitrarily large $t\geqslant2$ for which $\omega(t)\lesssim t/(\log t)^\delta$, where $\delta>0$ is a universal constant. This follows from (8) with $\delta=\frac12$, which is the largest possible constant for which this conclusion holds true. This positively answers a question that was asked in [22, Remark 1.7]. In fact, it provides an even better conclusion, because (8) implies that, say, there must exist arbitrarily large $t \geqslant 4$ for which

$$\omega(t) \lesssim \frac{t}{\sqrt{(\log t)\log\log t}}.$$

(The precise criterion is the integrability condition (8).) Finally, by considering $\omega(t) = t^{1-\varepsilon}/D$ for $\varepsilon \in (0,1)$ and $D \geqslant 1$, we obtain the following notable corollary.

Corollary 1.4 (L_1 distortion of snowflakes). *For every $\varepsilon \in (0,1)$ we have $c_1\big(\mathbb{Z}^5, d_W^{1-\varepsilon}\big) \asymp \frac{1}{\sqrt{\varepsilon}}$.*

The fact that for every $O(1)$-doubling metric space (X,d) we have $c_1(X, d^{1-\varepsilon}) \lesssim 1/\sqrt{\varepsilon}$ follows from an argument of [55] (see also [78, Theorem 5.2]). Corollary 1.4 shows that this is sharp. More generally, it follows from Theorem 1.3 that for every $R \geqslant 2$ and $\varepsilon \in (0,1)$ we have

$$c_1\big(\mathcal{B}_R, d_W^{1-\varepsilon}\big) \asymp \min\left\{\frac{1}{\sqrt{\varepsilon}}, \sqrt{\log R}\right\}.$$

1.4 Vertical-versus-Horizontal Isoperimetry

Our new non-embeddability results are all consequences of an independently interesting isoperimetric-type inequality which we shall now describe. Roughly speaking, this inequality subtly quantifies the fact that for any $n \in \mathbb{Z}$ and any $h \in \mathbb{H}^5_{\mathbb{Z}}$, there are many paths in the Cayley graph $X_S(\mathbb{H}^5_{\mathbb{Z}})$ of length roughly $\sqrt{n}$ that connect h to hZ^n. Consequently, if a finite subset $\Omega \subseteq \mathbb{Z}^5$ has a small edge boundary in the Cayley graph, then the number of pairs $(x,y) \in \mathbb{Z}^5 \times \mathbb{Z}^5$ for which $|\{x,y\} \cap \Omega| = 1$ yet x and y differ only in their fifth (vertical) coordinate must also be small. It turns out that the proper interpretation of the term "small" is this context is not at all obvious, and it should be measured in a certain multi-scale fashion. Formally, we consider the following quantities.

Definition 1.5 (Discrete boundaries). *For $\Omega \subseteq \mathbb{Z}^5$, the* **horizontal boundary** *of Ω is defined by*

$$\partial_{\mathsf{h}}\Omega \stackrel{\mathrm{def}}{=} \big\{(x,y) \in \Omega \times \Big(\mathbb{Z}^5 \smallsetminus \Omega\Big) : x^{-1}y \in S\big\}. \tag{9}$$

*Given also $t \in \mathbb{N}$, the t-***vertical boundary** *of Ω is defined by*

$$\partial^t_{\mathsf{v}}\Omega \stackrel{\mathrm{def}}{=} \Big\{(x,y) \in \Omega \times \Big(\mathbb{Z}^5 \smallsetminus \Omega\Big) : x^{-1}y \in \big\{Z^t, Z^{-t}\big\}\Big\}. \tag{10}$$

The **horizontal perimeter** *of Ω is defined to be the cardinality $|\partial_{\mathsf{h}}\Omega|$ of its horizontal boundary. The* **vertical perimeter** *of Ω is defined to be the quantity*

$$|\partial_{\mathsf{v}}\Omega| \stackrel{\mathrm{def}}{=} \left(\sum_{t=1}^{\infty} \frac{|\partial^t_{\mathsf{v}}\Omega|^2}{t^2}\right)^{\frac12}. \tag{11}$$

The horizontal perimeter of Ω is nothing more than the size of its *edge boundary* in the Cayley graph $X_S(\mathbb{H}^5_{\mathbb{Z}})$. The vertical perimeter of Ω is a more subtle concept that does not have such a simple combinatorial description. The definition (11) was first published in [53, Section 4], where the isoperimetric-type conjecture that we resolve here as Theorem 1.6 below also appeared for the first time. These were formulated by the first named author and were circulating for several years before [53] appeared, intended as a possible route towards the algorithmic application that we indeed succeed to obtain here. That "vertical smallness" should be measured through the quantity $|\partial_{\mathsf{v}}\Omega|$, i.e., the ℓ_2 norm of the sequence $\{|\partial^t_{\mathsf{v}}\Omega|/t\}_{t=1}^{\infty}$, was arrived at through trial and error, inspired by functional inequalities that were obtained in [7, 53], as explained in [53, Section 4].

Theorem 1.6. *Every $\Omega \subseteq \mathbb{Z}^5$ satisfies $|\partial_{\mathsf{v}}\Omega| \lesssim |\partial_{\mathsf{h}}\Omega|$.*

The significance of Theorem 1.6 can only be fully appreciated through an examination of the geometric and analytic reasons for its validity. To facilitate this, we shall include in this extended abstract an extensive overview of the ideas of the proof of Theorem 1.6; see Section 1.6 below. Before doing so, we shall now demonstrate the utility of Theorem 1.6 by using it to deduce Theorem 1.3. As explained above, by doing so we shall conclude the proof of all of our new results (modulo Theorem 1.6), including the lower bound on the integrality gap for the Goemans–Linial SDP.

1.5 From Isoperimetry to Non-embeddability

An equivalent formulation of Theorem 1.6 is that every finitely supported function $\phi : \mathbb{Z}^5 \to L_1$ satisfies the following Poincaré-type inequality.

$$\left(\sum_{t=1}^{\infty} \frac{1}{t^2}\left(\sum_{h\in\mathbb{Z}^5} \big\|\phi\big(hZ^t\big) - \phi(h)\big\|_1\right)^2\right)^{\frac12} \lesssim \sum_{h\in\mathbb{Z}^5}\sum_{\sigma\in S} \big\|\phi(h\sigma) - \phi(h)\big\|_1. \tag{12}$$

Indeed, Theorem (1.6) is nothing more than the special case $\phi = \mathbf{1}_\Omega$ of (12). Conversely, the fact that (12) follows from Theorem (1.6) is a straightforward application of the cut-cone representation of L_1 metrics (see e.g. [32, Proposition 4.2.2] or [73, Corollary 3.2]), though our proof will yield the (seemingly) stronger statement (12) directly. Next, Section 3.2 of [53] shows that (12) formally implies its local counterpart, which asserts that there exists a universal constant $\alpha \geqslant 1$ such that for every $n \in \mathbb{N}$ and every $\phi : \mathbb{Z}^5 \to L_1$ we have

$$\left(\sum_{t=1}^{n^2} \frac{1}{t^2}\left(\sum_{h\in\mathcal{B}_n} \big\|\phi\big(hZ^t\big) - \phi(h)\big\|_1\right)^2\right)^{\frac12} \lesssim \sum_{h\in\mathcal{B}_{\alpha n}}\sum_{\sigma\in S} \big\|\phi(h\sigma) - \phi(h)\big\|_1. \tag{13}$$

To deduce Theorem 1.3, suppose that $R \geqslant 2$, that $\omega : [0,\infty) \to [0,\infty)$ is nondecreasing and that the mapping $\phi : \mathcal{B}_R \to L_1$ satisfies (5). For notational convenience, fix two universal constants $\beta \in (0,1)$ and $\gamma \in (1,\infty)$ such that $\beta\sqrt{t} \leqslant d_W(Z^t, \mathbf{0}) \leqslant \gamma\sqrt{t}$ for every $t \in \mathbb{N}$. Note that (5) implies in particular that $\omega(R) \lesssim R$, so for every $c \in (0,1)$ the left hand side of (6) is at most a universal constant multiple of R^2. Hence, it suffices to prove Theorem 1.3 when $R \geqslant 1 + \max\{\alpha, \gamma\}$, where α is the universal constant in (13). Denote $n = \lfloor \min\{R/(1+\gamma), (R-1)/\alpha\}\rfloor \in \mathbb{N}$. If $t \in \{1, \ldots, n^2\}$ and $h \in \mathcal{B}_n$ then $d_W(hZ^t, \mathbf{0}) \leqslant n + \gamma\sqrt{t} \leqslant (1+\gamma)n \leqslant R$, and therefore we may apply (5) with $x = hZ^t$ and $y = h$ to deduce that $\|\phi(hZ^t) - \phi(h)\|_1 \gtrsim \omega(d_W(Z^t, \mathbf{0})) \geqslant \omega(\beta\sqrt{t})$. Consequently,

$$\begin{aligned}\sum_{t=1}^{n^2} \frac{1}{t^2}\bigg(\sum_{h\in\mathcal{B}_n} \big\|\phi(hZ^t) - \phi(h)\big\|_1\bigg)^2 &\gtrsim \sum_{t=1}^{n^2} \frac{|\mathcal{B}_n|^2\omega\big(\beta\sqrt{t}\big)^2}{t^2}\\ \gtrsim n^{12}\sum_{t=1}^{n^2}\int_t^{t+1} \frac{\omega\big(\beta\sqrt{u/2}\big)^2}{u^2}\,\mathrm{d}u &= \beta^2 n^{12}\int_{\frac{\beta}{\sqrt{2}}}^{\frac{\beta\sqrt{n^2+1}}{\sqrt{2}}} \frac{\omega(s)^2}{s^3}\,\mathrm{d}s\\ \geqslant \frac{\beta^2(R/2)^{12}}{\max\{(1+\gamma)^{12},\alpha^{12}\}}\int_1^{\frac{\beta R}{2\max\{1+\gamma,\alpha\}}} \frac{\omega(s)^2}{s^3}\,\mathrm{d}s, \qquad (14)\end{aligned}$$

where the second inequality in (14) uses the fact that ω is non-decreasing, the penultimate step of (14) uses the change of variable $s = \beta\sqrt{u/2}$, and for the final step of (14) recall that $\beta < 1$ and the definition of n. At the same time, by our choice of n we have $h\sigma \in \mathcal{B}_{\alpha n+1} \subseteq \mathcal{B}_R$ for every $h \in \mathcal{B}_{\alpha n}$ and $\sigma \in S$, and so by (5) we have $\|\phi(h\sigma) - \phi(h)\|_1 \leqslant d_W(h\sigma, h) = 1$. The right hand side of (13) is therefore at most a universal constant multiple of $|\mathcal{B}_{\alpha n}| \cdot |S| \lesssim (\alpha n)^6 \lesssim R^6$. By contrasting (14) with (13) we obtain that the desired estimate (6) indeed holds true.

1.6 Overview of the Proof of Theorem 1.6

Our proof of (12), and hence also of Theorem 1.6, is carried out in a continuous setting that is equivalent to its discrete counterpart. Such a passage from continuous to discrete is commonplace, and in the present setting this was carried out in [7, 53]. The idea is to consider a continuous group that contains $\mathbb{H}^5_{\mathbb{Z}}$ and to deduce the discrete inequality (12) from its (appropriately formulated) continuous counterpart via a partition of unity argument. There is an obvious way to embed $\mathbb{H}^5_{\mathbb{Z}}$ in a continuous group, namely by considering the same group of matrices as in (2), but with the entries a, b, c, d, e now allowed to be arbitrary real numbers instead of integers. This is a indeed a viable route and the ensuing discussion could be carried out by considering the resulting continuous matrix group. Nevertheless, it is notationally advantageous to work with a different (standard) realization of $\mathbb{H}^5_{\mathbb{Z}}$ which is isomorphic to the one that we considered thus far. We shall now introduce the relevant notation.

Fix an orthonormal basis $\{X_1, X_2, Y_1, Y_2, Z\}$ of $\mathbb{R}^5$. If $h = \alpha_1 X_1 + \alpha_2 X_2 + \beta_1 Y_1 + \beta_2 Y_2 + \gamma Z \in \mathbb{R}^5$ then denote $x_i(h) = \alpha_i$, $y_i(h) = \beta_i$ for $i \in \{1, 2\}$ and $z(h) = \gamma$, i.e., $x_1, x_2, y_1, y_2, z : \mathbb{R}^5 \to \mathbb{R}$ are the coordinate functions corresponding to the above basis. The continuous Heisenberg group $\mathbb{H}^5$ is defined to be $\mathbb{R}^5$, equipped with the following group law.

$$\begin{aligned}uv &\stackrel{\mathrm{def}}{=} u + v\\ &\quad+ \frac{x_1(u)y_1(v) - y_1(u)x_1(v) + x_2(u)y_2(v) - y_2(u)x_2(v)}{2} Z. \qquad (15)\end{aligned}$$

The identity element of $\mathbb{H}^5$ is $\mathbf{0} \in \mathbb{R}^5$ and the inverse of $h \in \mathbb{R}^5$ under the group law (15) is equal to $-h$. By directly computing Jacobians, one checks that the Lebesgue measure on $\mathbb{R}^5$ is invariant under the group operation given in (15), i.e., it is a Haar measure of $\mathbb{H}^5$. In what follows, in order to avoid confusing multiplication by scalars with the group law of $\mathbb{H}^5$, for every $h \in \mathbb{H}^5$ and $t \in \mathbb{R}$ we shall use the exponential notation $h^t = (th_1, \ldots, th_5)$; this agrees with the group law when $t \in \mathbb{Z}$. (This convention is not strictly necessary, but without it the ensuing discussion could become somewhat notationally confusing.)

The subgroup of $\mathbb{H}^5$ that is generated by $\{X_1, X_2, Y_1, Y_2\}$ is the discrete Heisenberg group of dimension 5, denoted $\mathbb{H}^5_{\mathbb{Z}}$. The apparent inconsistency with (2) is not an actual issue because it is straightforward to check that the two groups in question are in fact isomorphic. The linear span of $\{X_1, Y_1, Z\}$ is a subgroup of $\mathbb{H}^5$ which is denoted $\mathbb{H}^3$ (the 3-dimensional Heisenberg group).

There is a canonical left-invariant metric on $\mathbb{H}^5$, commonly called the *Carnot–Carathéodory metric*, which we denote by d. We refer to [15] for a precise definition of this metric. For the purpose of the present discussion it suffices to know that d possesses the following properties. Firstly, for every $g, h \in \mathbb{H}^5$ and $\theta \in \mathbb{R}$ we have $d(\mathfrak{s}_\theta(g), \mathfrak{s}_\theta(h)) = |\theta| d(g, h)$. Here, $\mathfrak{s}_\theta$ denotes the *Heisenberg scaling* by θ, given by the formula

$$\mathfrak{s}_\theta(\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma) = (\theta\alpha_1, \theta\alpha_2, \theta\beta_1, \theta\beta_2, \theta^2\gamma)$$

Secondly, the restriction of d to the subgroup $\mathbb{H}^5_{\mathbb{Z}}$ is bi-Lipschitz to the word metric induced by its generating set $\{X_1^{\pm1}, X_2^{\pm1}, Y_1^{\pm1}, Y_2^{\pm1}\}$. Thirdly, there exists $C \in (1, \infty)$ such that every $h \in \mathbb{H}^5$ satisfies

$$d(h, \mathbf{0}) \leqslant |x_1(h)| + |x_2(h)| + |y_1(h)| + |y_2(h)| + 4\sqrt{|z(h)|} \leqslant \frac{C}{2} d(h, \mathbf{0}). \qquad (16)$$

Given $r \in (0, \infty)$ we shall denote by $B_r \subseteq \mathbb{H}^5$ the open ball in the metric d of radius r centered at the identity element, i.e., $B_r = \{h \in \mathbb{H}^5 :\ d(\mathbf{0}, h) < r\}$. For $\Omega \subseteq \mathbb{H}^5$ the Lipschitz constant of a mapping $f : \Omega \to \mathbb{R}$ relative to the metric d will be denoted by $\|f\|_{\mathrm{Lip}(\Omega)}$. For $s \in (0, \infty)$, the notation $\mathcal{H}^s$ will be used exclusively to denote the s-dimensional Hausdorff measure that is induced by the metric d (see e.g. [70]). One checks that $\mathcal{H}^6$ is proportional to the Lebesgue measure on $\mathbb{R}^5$ and that the restriction of $\mathcal{H}^4$ to the subgroup $\mathbb{H}^3$ is proportional to the Lebesgue measure on $\mathbb{H}^3$ (under the canonical identification of $\mathbb{H}^3$ with $\mathbb{R}^3$). For two measurable subsets $E, U \subseteq \mathbb{H}^5$ define the *normalized vertical perimeter* of E in U to be the function $\overline{\mathrm{v}}_U(E) : \mathbb{R} \to [0, \infty]$ given by setting for every $s \in \mathbb{R}$,

$$\begin{aligned}\overline{\mathrm{v}}_U(E)(s) &\stackrel{\mathrm{def}}{=} \frac{1}{2^s}\mathcal{H}^6\Big(\big(E \vartriangle \big(EZ^{2^{2s}}\big)\big) \cap U\Big)\\ &= \frac{1}{2^s}\int_U \Big|\mathbf{1}_E(u) - \mathbf{1}_E\big(uZ^{-2^{2s}}\big)\Big|\,\mathrm{d}\mathcal{H}^6(u). \qquad (17)\end{aligned}$$

where $A \vartriangle B \stackrel{\mathrm{def}}{=} (A \smallsetminus B) \cup (B \smallsetminus A)$ is the symmetric difference. We also denote $\overline{\mathrm{v}}(E) \stackrel{\mathrm{def}}{=} \overline{\mathrm{v}}_{\mathbb{H}^5}(E)$.

The isoperimetric-type inequality of Theorem 1.7 below implies Theorem 1.6. See [79] for an explanation of this (standard) deduction; the argument is a straightforward use of the co-area formula (see e.g. [2, 65]) to pass from sets to functions, followed by the partition of unity argument of [53, Section 3.3] to pass from the continuous setting to the desired discrete inequality (12).

Theorem 1.7. *$\big\|\overline{\mathrm{v}}(E)\big\|_{L_2(\mathbb{R})} \lesssim \mathcal{H}^5(\partial E)$ for all open $E \subseteq \mathbb{H}^5$.*

We shall now explain the overall strategy and main ideas of our proof of Theorem 1.7. Complete technical details are included in [79]. A key new ingredient appears in Section 1.6.1 below, which is the *only* place in our proof where we use the fact that we are dealing with $\mathbb{H}^5$ rather than $\mathbb{H}^3$. In fact, the analogue of Theorem 1.7 for

$\mathbb{H}^3$ (i.e., with $\mathcal{H}^5(\partial E)$ replaced by $\mathcal{H}^3(\partial E)$ and $\overline{\mathrm{v}}(E)(\cdot)$ defined in the same way as in (17) but with $\mathcal{H}^6$ replaced by the restriction of $\mathcal{H}^4$ to $\mathbb{H}^3$) is *false* (see Section 1.7.1 below). The crux of the matter is the special case of Theorem 1.7 where the boundary of E is (a piece of) an *intrinsic Lipschitz graph*. Such sets were introduced by Franchi, Serapioni, and Serra Cassano [36]. These sets can be quite complicated, and in particular they *are not the same* as graphs of functions (in the usual sense) that are Lipschitz with respect to the Carnot–Carathéodory metric. Our proof of this special case relies crucially on an L_2-variant of (12) for $\mathbb{H}^3$ that was proven in [7] using representation theory and in [53] using Littlewood–Paley theory. In essence, our argument "lifts" a certain L_2 inequality in lower dimensions to a formally stronger endpoint L_1 (or isoperimetric-type) inequality in higher dimensions. Once the special case is established, we prove Theorem 1.7 in its full generality by decomposing an open set E into parts whose boundaries are close to pieces of intrinsic Lipschitz graphs and applying the special case to each part of this decomposition. We deduce the desired estimate by summing up all the inequalities thus obtained. Such a "corona decomposition" is an important and widely-used tool in harmonic analysis on $\mathbb{R}^n$ that was formulated by David and Semmes in [29]. For the present purpose we need to devise an "intrinsic version" of a corona decomposition on the Heisenberg group. This step uses a different "coercive quantity" to control local overlaps, but for the most part it follows the lines of the well-understood methodology of David and Semmes, as described in the monographs [29, 30].

1.6.1 Intrinsic Lipschitz Graphs. Set $V \stackrel{\mathrm{def}}{=} \{h \in \mathbb{H}^5 :\ x_2(h) = 0\}$. For $f : V \to \mathbb{R}$ define

$$\Gamma_f \stackrel{\mathrm{def}}{=} \left\{vX_2^{f(v)} :\ v \in V\right\} \stackrel{(15)}{=} \left\{\left(\mathrm{a}, f(\mathrm{a},\mathrm{c},\mathrm{d},\mathrm{e}), \mathrm{c}, \mathrm{d}, \mathrm{e} - \frac{1}{2}\mathrm{d}f(\mathrm{a},\mathrm{c},\mathrm{d},\mathrm{e})\right) :\ \mathrm{a},\mathrm{c},\mathrm{d},\mathrm{e} \in \mathbb{R}\right\}, \tag{18}$$

where (18) uses the identification of $\mathrm{a}X_1 + \mathrm{b}X_2 + \mathrm{c}Y_1 + \mathrm{d}Y_2 + \mathrm{e}Z \in \mathbb{H}^5$ with $(\mathrm{a},\mathrm{b},\mathrm{c},\mathrm{d},\mathrm{e}) \in \mathbb{R}^5$ and the identification of $(\mathrm{a},0,\mathrm{c},\mathrm{d},\mathrm{e}) \in V$ with $(\mathrm{a},\mathrm{c},\mathrm{d},\mathrm{e}) \in \mathbb{R}^4$ (thus we think of the domain of f as equal to $\mathbb{R}^4$). The set Γ_f is a typical *intrinsic graph* in $\mathbb{H}^5$. See [79] for a discussion of the general case, which is equivalent to this case via a symmetry of $\mathbb{H}^5$ (so the ensuing discussion has no loss of generality). Suppose that $\lambda \in (0,\infty)$. We say that Γ_f is an *intrinsic λ-Lipschitz graph* over the vertical hyperplane V if

$$\forall\, w_1, w_2 \in \Gamma_f, \qquad |x_2(w_1) - x_2(w_2)| \leqslant \lambda d(w_1, w_2). \tag{19}$$

Due to (18) the condition (19) amounts to a point-wise inequality for f that is somewhat complicated, and in particular it does not imply that f must be Lipschitz with respect to the restriction of the Carnot–Carathéodory metric to the hyperplane V, as explained in [37, Remark 3.13].

Denote by $\Gamma_f^+ = \{vX_2^t :\ v \in V \ \wedge\ t > f(v)\}$ the half-space that is bounded by the intrinsic graph Γ_f. Suppose that Γ_f is an intrinsic λ-Lipschitz graph with $\lambda \in (0,1)$. We claim that

$$\forall\, r \in (0,\infty), \qquad \left\|\overline{\mathrm{v}}_{B_r}\left(\Gamma_f^+\right)\right\|_{L_2(\mathbb{R})} \lesssim \frac{r^5}{1-\lambda}. \tag{20}$$

When, say, $\lambda \in (0, \frac{1}{2})$, the estimate (20) is in essence the special case of Theorem 1.7 for pieces of Lipschitz graphs. This is so because, due to the isoperimetric inequality for the Heisenberg group [81], the right-hand side of (20) is at most a universal constant multiple of $\mathcal{H}^5(\partial(B_r \cap \Gamma_f^+))$ whenever $\mathcal{H}^6(B_r \cap \Gamma_f^+) \gtrsim r^6$, i.e., provided that Γ_f^+ occupies a constant fraction of the volume of the ball B_r. The estimate (20) will be used below only in such a non-degenerate situation.

The advantage of working in $\mathbb{H}^5$ rather than $\mathbb{H}^3$ is that $V \subseteq \mathbb{H}^5$ can be sliced into copies of $\mathbb{H}^3$. We will bound $\|\overline{\mathrm{v}}_{B_r}(\Gamma_f^+)\|_{L_2(\mathbb{R})}$ by decomposing Γ_f^+ into a corresponding family of slices. Write

$$\forall\, u \in \mathbb{H}^5, \quad h_u \stackrel{\mathrm{def}}{=} X_1^{x_1(u)} + Y_1^{y_1(u)} + Z^{z(u)+\frac{1}{2}x_2(u)y_2(u)} \in \mathbb{H}^3. \tag{21}$$

Recalling (15), one computes directly that $u = Y_2^{y_2(u)} h_u X_2^{x_2(u)}$. Let $C \in (1,\infty)$ be the universal constant in (16). A straightforward computation using (16) shows that $d(h_u, \mathbf{0}) \leqslant Cd(u, \mathbf{0})$. Also, (16) implies that $|y_2(u)| \leqslant Cd(u,\mathbf{0})$. These simple observations demonstrate that

$$\forall\, u \in \mathbb{H}^5, \qquad \mathbf{1}_{B_r}(u) \leqslant \mathbf{1}_{[-Cr,Cr]}\big(y_2(u)\big)\mathbf{1}_{\mathbb{H}^3 \cap B_{Cr}}(h_u). \tag{22}$$

For every $\chi \in \mathbb{R}$ define $f_\chi : \mathbb{H}^3 \to \mathbb{R}$ by $f_\chi(h) = f(Y_2^\chi h)$ (recall that $\mathbb{H}^3$ is the span of $\{X_1, Y_1, Z\}$, so $Y_2^\chi h \in V$ is in the domain of f). Under this notation $u \in \Gamma_f^+$ if and only if $x_2(u) > f_{y_2(u)}(h_u)$. Also, for every $\alpha \in \mathbb{R}$ we have $uZ^\alpha \in \Gamma_f^+$ if and only if $x_2(u) > f_{y_2(u)}(h_u Z^\alpha)$, since $h_{uZ^\alpha} = h_u Z^\alpha$ by (21). Due to (17) and (22), these observations imply that for every $s \in \mathbb{R}$ we have

$$\begin{aligned}\overline{\mathrm{v}}_{B_r}&\left(\Gamma_f^+\right)(s)\\ &\leqslant \frac{1}{2^s}\int_{\mathbb{H}^5}\left|\mathbf{1}_{\{x_2(u)>f_{y_2(u)}(h_u)\}} - \mathbf{1}_{\{x_2(u)>f_{y_2(u)}(h_u Z^{-2^{2s}})\}}\right|\\ &\qquad\qquad \times \mathbf{1}_{[-Cr,Cr]}\big(y_2(u)\big)\mathbf{1}_{\mathbb{H}^3\cap B_{Cr}}(h_u)\,\mathrm{d}\mathcal{H}^6(u).\end{aligned} \tag{23}$$

Recall that $\mathcal{H}^6$ is proportional to the Lebesgue measure on $\mathbb{H}^5$. Hence, if we continue to canonically identify $\mathrm{a}X_1 + \mathrm{b}X_2 + \mathrm{c}Y_1 + \mathrm{d}Y_2 + \mathrm{e}Z \in \mathbb{H}^5$ with $(\mathrm{a},\mathrm{b},\mathrm{c},\mathrm{d},\mathrm{e}) \in \mathbb{R}^5$ and $\mathrm{a}X_1 + \mathrm{c}Y_1 + \mathrm{e}Z \in \mathbb{H}^3$ with $(\mathrm{a},\mathrm{c},\mathrm{e}) \in \mathbb{R}^3$ then, recalling (21), the integral in the right hand side of (23) is proportional to

$$\begin{aligned}&\int_{\mathbb{R}^5}\left|\mathbf{1}_{\left\{\mathrm{b}>f_{\mathrm{d}}\left(\mathrm{a},\mathrm{c},\mathrm{e}+\frac{1}{2}\mathrm{bd}\right)\right\}} - \mathbf{1}_{\left\{\mathrm{b}>f_{\mathrm{d}}\left(\mathrm{a},\mathrm{c},\mathrm{e}+\frac{1}{2}\mathrm{bd}-2^{2s}\right)\right\}}\right|\\ &\qquad\qquad \times \mathbf{1}_{[-Cr,Cr]}(\mathrm{d})\mathbf{1}_{\mathbb{H}^3\cap B_{Cr}}\left(\mathrm{a},\mathrm{c},\mathrm{e}+\frac{1}{2}\mathrm{bd}\right)\mathrm{d}(\mathrm{a},\mathrm{b},\mathrm{c},\mathrm{d},\mathrm{e})\\ &= \int_{\mathbb{R}^5}\left|\mathbf{1}_{\{\mathrm{b}>f_{\mathrm{d}}(\alpha,\gamma,\epsilon)\}} - \mathbf{1}_{\{\mathrm{b}>f_{\mathrm{d}}(\alpha,\gamma,\epsilon-2^{2s})\}}\right|\\ &\qquad\qquad \times \mathbf{1}_{[-Cr,Cr]}(\mathrm{d})\mathbf{1}_{\mathbb{H}^3\cap B_{Cr}}(\alpha,\gamma,\epsilon)\,\mathrm{d}(\alpha,\mathrm{b},\gamma,\mathrm{d},\epsilon),\end{aligned}$$

where for each fixed $\mathrm{b},\mathrm{d} \in \mathbb{R}$ we made the change of variable $(\alpha,\gamma,\epsilon) = (\mathrm{a},\mathrm{c},\mathrm{e}+\mathrm{bd}/2)$. Since the restriction of the Hausdorff measure $\mathcal{H}^4$ to $\mathbb{H}^3$ is proportional to the Lebesgue measure on

$\mathbb{H}^3 \cong \mathbb{R}^3$, we conclude from the above considerations that for every $s \in \mathbb{R}$ we have

$$\begin{aligned}
&\overline{\mathrm{v}}_{B_r}\big(\Gamma_f^+\big)(s)\\
&\lesssim \frac{1}{2^s}\int_{-Cr}^{Cr}\int_{\mathbb{H}^3\cap B_{Cr}}\\
&\qquad\left(\int_{-\infty}^{\infty}\left|\mathbf{1}_{\{\xi>f_\chi(h)\}}-\mathbf{1}_{\{\xi>f_\chi(hZ^{-2^{2s}})\}}\right|\mathrm{d}\xi\right)\mathrm{d}\mathcal{H}^4(h)\,\mathrm{d}\chi\\
&=\frac{1}{2^s}\int_{-Cr}^{Cr}\int_{\mathbb{H}^3\cap B_{Cr}}\left|f_\chi(h)-f_\chi\big(hZ^{-2^{2s}}\big)\right|\mathrm{d}\mathcal{H}^4(h)\,\mathrm{d}\chi. \qquad (24)
\end{aligned}$$

Next, fix $h_1,h_2\in\mathbb{H}^3$ and $\chi\in\mathbb{R}$. Denote $w_1 \stackrel{\mathrm{def}}{=} Y_2^\chi h_1 X_2^{f_\chi(h_1)}$ and $w_2 \stackrel{\mathrm{def}}{=} Y_2^\chi h_2 X_2^{f_\chi(h_2)}$. By design we have $w_1,w_2\in\Gamma_f$ and therefore we may apply (20) to deduce that

$$\begin{aligned}
|f_\chi(h_1)-f_\chi(h_2)| &= |x_2(w_1)-x_2(w_2)| \leqslant \lambda d(w_1,w_2)\\
&=\lambda d\Big(Y_2^\chi h_1 X_2^{f_\chi(h_1)}, Y_2^\chi h_2 X_2^{f_\chi(h_2)}\Big)\\
&=\lambda d\Big(\mathbf{0}, h_1^{-1}h_2X_2^{f_\chi(h_2)-f_\chi(h_1)}\Big)\\
&\leqslant \lambda\Big(Cd(h_1,h_2)+|f_\chi(h_1)-f_\chi(h_2)|\Big), \qquad (25)
\end{aligned}$$

where the first inequality in (25) uses (20) , the penultimate step of (25) uses the left-invariance of the metric d and the fact that X_2 commutes with all of the elements of $\mathbb{H}^3$, and the final step of (25) uses (16) (twice). The estimate (25) simplifies to show that $|f_\chi(h_1)-f_\chi(h_2)|\lesssim d(h_1,h_2)/(1-\lambda)$, i.e., for every fixed $\chi\in\mathbb{R}$ the function f_χ is Lipschitz on $\mathbb{H}^3$ with $\|f_\chi\|_{\mathrm{Lip}(\mathbb{H}^3)}\lesssim 1/(1-\lambda)$.

In [7, Theorem 7.5] the following inequality was proved for a Lipschitz function $\psi:\mathbb{H}^3\to\mathbb{R}$ and $\rho\in(0,\infty)$ as a consequence of a continuous L_2-variant of (12). Due to its quadratic nature, this variant can be proved using a decomposition into irreducible representations (i.e., a spectral argument).

$$\int_0^{\rho^2}\int_{B_\rho\cap\mathbb{H}^3}\left|\psi(h)-\psi\big(hZ^{-t}\big)\right|^2\,\mathrm{d}\mathcal{H}^4(h)\frac{\mathrm{d}t}{t^2}\lesssim \rho^4\|\psi\|^2_{\mathrm{Lip}(\mathbb{H}^3)}. \qquad (26)$$

Consequently,

$$\begin{aligned}
&\frac{r^5}{(1-\lambda)^2} \qquad (27)\\
&\gtrsim \int_{-Cr}^{Cr}\int_0^{(Cr)^2}\int_{B_{Cr}\cap\mathbb{H}^3}\left|f_\chi(h)-f_\chi\big(hZ^{-t}\big)\right|^2\,\mathrm{d}\mathcal{H}^4(h)\frac{\mathrm{d}t}{t^2}\,\mathrm{d}\chi\\
&=\int_{-\infty}^{\log_2(Cr)}\frac{2\log 2}{2^{2s}}\\
&\quad\times\int_{-Cr}^{Cr}\int_{B_{Cr}\cap\mathbb{H}^3}\left|f_\chi(h)-f_\chi\big(hZ^{-2^{2s}}\big)\right|^2\,\mathrm{d}\mathcal{H}^4(h)\,\mathrm{d}\chi\,\mathrm{d}s \qquad (28)\\
&\gtrsim\int_{-\infty}^{\log_2(Cr)}\frac{1}{r^5} \qquad (29)\\
&\quad\times\left(\frac{1}{2^s}\int_{-Cr}^{Cr}\int_{B_{Cr}\cap\mathbb{H}^3}\left|f_\chi(h)-f_\chi\big(hZ^{-2^{2s}}\big)\right|\,\mathrm{d}\mathcal{H}^4(h)\,\mathrm{d}\chi\right)^2\mathrm{d}s\\
&\gtrsim\frac{1}{r^5}\int_{-\infty}^{\log_2(Cr)}\overline{\mathrm{v}}_{B_r}\big(\Gamma_f^+\big)(s)^2\,\mathrm{d}s, \qquad (30)
\end{aligned}$$

In (27) we applied (26) with $\psi=f_\chi$ for each $\chi\in[-Cr,Cr]$, while using $\|f_\chi\|_{\mathrm{Lip}(\mathbb{H}^3)}\lesssim 1/(1-\lambda)$. In (28) we made the change of variable $t=2^{2s}$. In (29) we used the Cauchy–Schwarz inequality while noting that $\mathcal{H}^4(B_{Cr}\cap\mathbb{H}^3)\asymp r^4$. Finally, (30) follows from an application of (24). Now,

$$\begin{aligned}
&\left\|\overline{\mathrm{v}}_{B_r}\big(\Gamma_f^+\big)\right\|^2_{L_2(\mathbb{R})}\\
&=\int_{-\infty}^{\log_2(Cr)}\overline{\mathrm{v}}_{B_r}\big(\Gamma_f^+\big)(s)^2\,\mathrm{d}s+\int_{\log_2(Cr)}^{\infty}\overline{\mathrm{v}}_{B_r}\big(\Gamma_f^+\big)(s)^2\,\mathrm{d}s\\
&\lesssim\frac{r^{10}}{(1-\lambda)^2}+\int_{\log_2(Cr)}^{\infty}\frac{\mathcal{H}^6(B_r)^2}{2^{2s}}\,\mathrm{d}s\\
&\asymp\frac{r^{10}}{(1-\lambda)^2}+\int_{\log_2(Cr)}^{\infty}\frac{r^{12}}{2^{2s}}\,\mathrm{d}s\asymp\frac{r^{10}}{(1-\lambda)^2}, \qquad (31)
\end{aligned}$$

where we estimated the second integral using the trivial bound $\overline{\mathrm{v}}_{B_r}(E)(s)\leqslant\mathcal{H}^6(B_r)/2^s\asymp r^6/2^s$. By taking square roots of both sides of (31) we obtain the desired estimate (20). It is important to stress that this proof does not work for functions on $\mathbb{H}^3$ because it relies on slicing $\mathbb{H}^5$ into copies of $\mathbb{H}^3$. There is no analogue of (26) for 1-dimensional vertical slices of $\mathbb{H}^3$.

1.6.2 An Intrinsic Corona Decomposition. In Section 1.6.1 we presented the complete details of the proof of a crucial new ingredient that underlies the validity of Theorem 1.7. This ingredient is the *only step* that relies on a property of $\mathbb{H}^5$ that is not shared by $\mathbb{H}^3$. We believe that it is important to fully explain this key ingredient within this extended abstract, but this means that we must defer the details of the formal derivation of Theorem 1.7 from its special case (20) to the full version [79]. The complete derivation requires additional terminology and notation, but the main idea is to produce "intrinsic corona decompositions" in the Heisenberg group. Corona decompositions are an established tool in analysis for reducing the study of certain singular integrals on $\mathbb{R}^n$ to the case of Lipschitz graphs, starting with seminal works of David [26, 27] and Jones [42, 43] on the Cauchy integral and culminating with the David–Semmes theory of quantitative rectifiability [29, 30]. Our adaptation of this technique is mostly technical, but it will also involve a conceptually new ingredient, namely the use of quantitative monotonicity for this purpose. We will now outline the remainder of the proof of Theorem 1.7.

Our arguments hold for Heisenberg groups of any dimension (including $\mathbb{H}^3$), but we avoid introducing new notation by continuing to work with $\mathbb{H}^5$ for now. The first step is to show that in order to establish Theorem 1.7 it suffices to prove that for every $r\in(0,\infty)$ and every $E\subseteq\mathbb{H}^5$, we have $\|\overline{\mathrm{v}}_{B_r}(E)\|_{L_2(\mathbb{R})}\lesssim r^5$ under the additional assumption that the sets E, $\mathbb{H}^5\smallsetminus E$, and ∂E are r-locally Ahlfors-regular, i.e., $\mathcal{H}^6(uB_\rho\cap E)\asymp\rho^6\asymp\mathcal{H}^6(vB_\rho\smallsetminus E)$ and $\mathcal{H}^5(wB_\rho\cap\partial E)\asymp\rho^5$ for all $\rho\in(0,r)$ and $(u,v,w)\in E\times(\mathbb{H}^5\smallsetminus E)\times\partial E$. We prove this by first applying a Heisenberg scaling and an approximation argument to reduce Theorem 1.7 to the case that E is a "cellular set," i.e., it is a union of parallelepipeds of the form $h[-\frac12,\frac12]^5$ as h ranges over a subset of the discrete Heisenberg group $\mathbb{H}^5_{\mathbb{Z}}\subseteq\mathbb{H}^5$. Any such set is Ahlfors-regular on sufficiently small balls. We next argue that E can be decomposed into sets that satisfy the desired local Ahlfors-regularity. The full construction of this decomposition appears in [79], but we remark briefly that it amounts to the following natural "greedy" iterative procedure. If one of the sets $E, \mathbb{H}^5 \smallsetminus E, \partial E$ were not locally Ahlfors-regular then there would be some smallest ball B such that the density of E, $\mathbb{H}^5 \smallsetminus E$ or ∂E is either too low or too high on B. By replacing E by either $E \cup B$ or $E \smallsetminus B$, we cut off a piece of ∂E and decrease $\mathcal{H}^5(\partial E)$. Since B was the smallest ball where Ahlfors-regularity fails, $E, \mathbb{H}^5 \smallsetminus E, \partial E$ are Ahlfors-regular on balls smaller than B. Repeating this process eventually reduces E to the empty set, and we arrive at the conclusion of Theorem 1.7 for the initial set E by proving the (local version of) the theorem for each piece of this decomposition, then summing the resulting inequalities. We will therefore suppose from now on that E, $\mathbb{H}^5 \smallsetminus E$ and ∂E are all locally Ahlfors-regular.

The next step is the heart of the matter: approximating ∂E by intrinsic Lipschitz graphs so that we can use the fact that Theorem 1.7 holds for (pieces of) such graphs. The natural way to do this is to construct (an appropriate Heisenberg version of) a corona decomposition in the sense of [29, 30]. Such a decomposition covers ∂E by two types of sets, called *stopping-time regions* and *bad cubes*. Stopping-time regions correspond to parts of ∂E that are close to intrinsic Lipschitz graphs, and bad cubes correspond to parts of ∂E, like sharp corners, that are not. The multiplicity of this cover depends on the shape of ∂E at different scales. For example, ∂E might look smooth on a large neighborhood of a point x, jagged at a medium scale, then smooth again at a small scale. If so, then x is contained in a large stopping-time region, a medium-sized bad cube, and a second small stopping-time region. A cover like this is a corona decomposition if it satisfies a *Carleson packing condition* (see [79]) that bounds its average multiplicity on any ball.

We construct our cover following the well-established methods of [29, 30]. We start by constructing a sequence of nested partitions of ∂E into pieces called *cubes*; this is a standard construction due to Christ [23] and David [28] and only uses the Ahlfors regularity of ∂E. These partitions are analogues of the standard tilings of $\mathbb{R}^n$ into dyadic cubes. Next, we classify the cubes into *good cubes*, which are close to a piece of a hyperplane, and *bad cubes*, which are not. In order to produce a corona decomposition, there cannot be too many bad cubes, i.e., they must satisfy a Carleson packing condition. In [29, 30], this condition follows from *quantitative rectifiability*; the surface in question is assumed to satisfy a condition that bounds the sum of its (appropriately normalized) local deviations from hyperplanes. These local deviations are higher-dimensional versions of Jones' β-numbers [42, 43], and the quantitative rectifiability assumption leads to the desired packing condition. In the present setting, the packing condition follows instead from *quantitative non-monotonicity*. The concept of the quantitative non-monotonicity of a set $E \subseteq \mathbb{H}^5$ (see [79]) was first defined in [21, 22], where the kinematic formula for the Heisenberg group was used to show that the total non-monotonicity of all of the cubes is at most a constant multiple of $\mathcal{H}^5(\partial E)$. This means that there cannot be many cubes that have large non-monotonicity. By a result of [21, 22], if a set has small non-monotonicity, then its boundary is close to a hyperplane. Consequently, most cubes are close to hyperplanes and are therefore good. (The result in [21, 22] is stronger than what we need for this proof; it provides power-type bounds on how closely a nearly-monotone surface approximates a hyperplane. For our purposes, it is enough to have *some* bound (not necessarily power-type) on the shape of nearly-monotone surfaces, and we can deduce the bound that we need by applying a quick compactness argument to a result from [20] that states that if a set is *precisely monotone* (i.e., every line intersects its boundary in at most one point), then it is a half-space.)

Next, we partition the good cubes into stopping-time regions by using an iterative construction that corrects overpartitioning that may have occurred when the Christ cubes were constructed. If Q is a largest good cube that hasn't been treated yet and if P is its approximating half-space, we find all of the descendants of Q with approximating half-spaces that are sufficiently close to P. If we glue these half-spaces together using a partition of unity, the result is an intrinsic Lipschitz half-space that approximates all of these descendants. By repeating this procedure for each untreated cube, we obtain a collection of stopping-time regions. These regions satisfy a Carleson packing condition because if a point $x \in \partial E$ is contained in many different stopping-time regions, then either x is contained in many different bad cubes, or x is contained in good cubes whose approximating hyperplanes point in many different directions. In either case, these cubes generate non-monotonicity, so there can only be a few points with large multiplicity.

The construction above leads to the proof of Theorem 1.7 as follows. The vertical perimeter of ∂E comes from three sources: the bad cubes, the approximating Lipschitz graphs, and the error incurred by approximating a stopping-time region by an intrinsic Lipschitz graph. By the Carleson packing condition, there are few bad cubes, and they contribute vertical perimeter on the order of $\mathcal{H}^5(\partial E)$. By the result of Section 1.6.1, the intrinsic Lipschitz graphs also contribute vertical perimeter on the order of $\mathcal{H}^5(\partial E)$. Finally, the vertical perimeter of the difference between a stopping-time region and an intrinsic Lipschitz graph is bounded by the size of the stopping-time region. The stopping-time regions also satisfy a Carleson packing condition, so these errors also contribute vertical perimeter on the order of $\mathcal{H}^5(\partial E)$. Summing these contributions, we obtain the desired bound.

1.7 Historical Overview and Directions for Further Research

Among the well-established deep and multifaceted connections between theoretical computer science and pure mathematics, the Sparsest Cut Problem stands out for its profound and often unexpected impact on a variety of areas. Indeed, previous research on this question came hand-in-hand with the development of remarkable mathematical and algorithmic ideas that spurred many further works of importance in their own right. Because the present work belongs to this tradition, we will try to put it into context by elaborating further on the history of these investigations and describing directions for further research and open problems. Some of these directions will appear in forthcoming work.

The first polynomial-time algorithm for Sparsest Cut with approximation ratio $O(\log n)$ was obtained in the important work [59], which studied the notable special case of *Sparsest Cut with Uniform Demands* (see Section 1.7.3 below). This work introduced a linear programming relaxation and developed influential techniques for its analysis, and it has led to a myriad of algorithmic applications. The seminal contributions [6, 64] obtained the upper bound

$\rho_{\mathrm{GL}}(n) \lesssim \log n$ in full generality by incorporating a classical embedding theorem of Bourgain [12], thus heralding the transformative use of metric embeddings in algorithm design. The matching lower bound on the integrality gap of this linear program was proven in [59, 64]. This showed for the first time that Bourgain's embedding theorem is asymptotically sharp and was the first demonstration of the power of expander graphs in the study of metric embeddings.

A $O(\sqrt{\log n})$ upper bound for the approximation ratio of the Goemans–Linial algorithm in the case of uniform demands was obtained in the important work [4]. This work relied on a clever use of the concentration of measure phenomenon and introduced influential techniques such as a "chaining argument" for metrics of negative type and the use of expander flows. [4] also had direct impact on results in pure mathematics, including combinatorics and metric geometry; see e.g. the "edge replacement theorem" and the estimates on the observable diameter of doubling metric measure spaces in [77]. The best-known upper bound $\rho_{\mathrm{GL}}(n) \lesssim (\log n)^{\frac{1}{2}+o(1)}$ of [3] built on the (then very recent) development of two techniques: The chaining argument of [4] (through its refined analysis in [54]) and the *measured descent* embedding method of [51] (through its statement as a gluing technique for Lipschitz maps in [54]). Another important input to [3] was a re-weighting argument of [17] that allowed for the construction of an appropriate "random zero set" from the argument of [4, 54] (see [73, 75] for more on this notion and its significance).

The impossibility result [49] that refuted the Goemans–Linial conjecture relied on a striking link to complexity theory through the Unique Games Conjecture (UGC), as well as an interesting use of discrete harmonic analysis (through [13]) in this context; see also [52] for an incorporation of a different tool from discrete harmonic analysis (namely [44], following [48]) for the same purpose, as well as [18, 24] for computational hardness. The best impossibility result currently known [45] for Sparsest Cut with Uniform Demands relies on the development of new pseudorandom generators.

The idea of using the geometry of the Heisenberg group to bound $\rho_{\mathrm{GL}}(n)$ from below originated in [56], where the relevant metric of negative type was constructed through a complex-analytic argument, and initial (qualitative) impossibility results were presented through the use of Pansu's differentiation theorem [82] and the Radon–Nikodým Property from functional analysis (see e.g. [11]). In [19], it was shown that the Heisenberg group indeed provides a proof that $\lim_{n\to\infty} \rho_{\mathrm{GL}}(n) = \infty$. This proof introduced a remarkable new notion of differentiation for L_1-valued mappings, which led to the use of tools from geometric measure theory [34, 35] to study the problem. A different proof that $\mathbb{H}^3$ fails to admit a bi-Lipschitz embedding into L_1 was found in [20], where a classical notion of metric differentiation [50] was used in conjunction with the novel idea to consider monotonicity of sets in this context, combined with a sub-Riemannian-geometric argument that yielded a classification of monotone subsets of $\mathbb{H}^3$. The main result of [22] finds a quantitative lower estimate for the scale at which this differentiation argument can be applied, leading to a lower bound of $(\log n)^{\Omega(1)}$ on $\rho_{\mathrm{GL}}(n)$. This result relies on a mixture of the methods of [19] and [20] and requires overcoming obstacles that are not present in the original qualitative investigation. In particular, [22] introduced the quantitative measures of non-monotonicity that we use in the present work to find crucial bounds in the construction of an intrinsic corona decomposition. The quantitative differentiation bound of [22] remains the best bound currently known, and it would be very interesting to discover the sharp behavior in this more subtle question.

The desire to avoid the (often difficult) need to obtain sharp bounds for quantitative differentiation motivated the investigations in [7, 53]. In particular, [7] devised a method to prove sharp (up to lower order factors) nonembeddability statements for the Heisenberg group based on a cohomological argument and a quantitative ergodic theorem. For Hilbert-space valued mappings, [7] used a cohomological argument in combination with representation theory to prove the following quadratic inequality for every finitely supported function $\phi : \mathbb{H}^5_Z \to L_2$.

$$\left(\sum_{t=1}^{\infty} \frac{1}{t^2} \sum_{h\in\mathbb{Z}^5} \left\|\phi(hZ^t) - \phi(h)\right\|_2^2\right)^{\frac{1}{2}} \lesssim \left(\sum_{h\in\mathbb{Z}^5}\sum_{\sigma\in S} \left\|\phi(h\sigma) - \phi(h)\right\|_2^2\right)^{\frac{1}{2}}. \quad (32)$$

In [53] a different approach based on Littlewood–Paley theory was devised, leading to the following generalization of (32) that holds true for every $p \in (1,2]$ and every finitely supported $\phi : \mathbb{H}^5 \to L_p$.

$$\left(\sum_{t=1}^{\infty} \frac{1}{t^2}\left(\sum_{h\in\mathbb{Z}^5} \left\|\phi(hZ^t) - \phi(h)\right\|_p^p\right)^{\frac{2}{p}}\right)^{\frac{1}{2}} \leqslant C(p)\left(\sum_{h\in\mathbb{Z}^5}\sum_{\sigma\in S} \left\|\phi(h\sigma) - \phi(h)\right\|_p^p\right)^{\frac{1}{p}}, \quad (33)$$

for some $C(p) \in (0,\infty)$. See [53] for a strengthening of (33) that holds for general uniformly convex targets (using the recently established [67] vector-valued Littlewood–Paley–Stein theory for the Poisson semigroup). These functional inequalities yield sharp non-embeddability estimates for balls in $\mathbb{H}^5_{\mathbb{Z}}$, but the method of [53] inherently yields a constant $C(p)$ in (33) that satisfies $\lim_{p\to 1} C(p) = \infty$. The estimate (13) that we prove here for L_1-valued mappings is an endpoint estimate corresponding to (33), showing that the best possible $C(p)$ actually remains bounded as $p \to 1$. This confirms a conjecture of [53] and is crucial for the results that we obtain here.

As explained in Section 1.6.1, our proof of (32) uses the $\mathbb{H}^3$-analogue of (32). It should be mentioned at this juncture that the proofs of (32) and (33) in [7, 53] were oblivious to the dimension of the underlying Heisenberg group.[3] An unexpected aspect of the present work is that the underlying dimension does play a role at the endpoint $p = 1$, with the analogue of (13) (or Theorem 1.6) for $\mathbb{H}^3$ being in fact *incorrect*; see Section 1.7.1 below. In the full version [79] of this paper we shall establish the $\mathbb{H}^{2k+1}$-analogue of Theorem 1.6 for every $k \in \{2, 3, \ldots\}$, in which case the implicit constant depends

[3] Thus far in this extended abstract we recalled the definitions of $\mathbb{H}^5$ and $\mathbb{H}^3$ but not of higher-dimensional Heisenberg groups (since they are not needed for any of the applications that are obtain here). Nevertheless, it is obvious how to generalize either the matrix group or the group modelled on $\mathbb{R}^5$ that we considered above to obtain the Heisenberg group $\mathbb{H}^{2k+1}$ for any $k \in \mathbb{N}$.

on k, and we shall also obtain the sharp asymptotic behavior as $k \to \infty$.

As we recalled above, past progress on the Sparsest Cut Problem came hand-in-hand with meaningful mathematical developments. The present work is a culmination of a long-term project that is rooted in mathematical phenomena that are interesting not just for their relevance to approximation algorithms but also for their connections to the broader mathematical world. In the ensuing subsections we shall describe some further results and questions related to this general direction.

1.7.1 The 3-dimensional Heisenberg Group. The investigation of the possible validity of an appropriate analogue of Theorem 1.6 with $\mathbb{H}^5_{\mathbb{Z}}$ replaced by $\mathbb{H}^3_{\mathbb{Z}}$ remains an intriguing mystery and a subject of ongoing research that will be published elsewhere. This ongoing work shows that Theorem 1.6 fails for $\mathbb{H}^3_{\mathbb{Z}}$, but that there exists $p \in (2,\infty)$ such that for every $\Omega \subseteq \mathbb{H}^3_{\mathbb{Z}}$ we have

$$\left(\sum_{t=1}^{\infty} \frac{|\partial^t_{\mathsf{v}}\Omega|^p}{t^{1+\frac{p}{2}}}\right)^{\frac{1}{p}} \lesssim |\partial_{\mathsf{h}}\Omega|. \tag{34}$$

A simple argument shows that $\sup_{s\in\mathbb{N}} |\partial^s_{\mathsf{v}}\Omega|/\sqrt{s} \leqslant \gamma|\partial_{\mathsf{h}}\Omega|$ for some universal constant $\gamma > 0$. Hence, for every $t \in \mathbb{N}$ we have

$$\frac{|\partial^t_{\mathsf{v}}\Omega|^p}{t^{1+\frac{p}{2}}} \leqslant \frac{|\partial^t_{\mathsf{v}}\Omega|^2}{t^2}\sup_{s\in\mathbb{N}}\left(\frac{|\partial^s_{\mathsf{v}}\Omega|}{\sqrt{s}}\right)^{p-2} \leqslant \frac{|\partial^t_{\mathsf{v}}\Omega|^2}{t^2}(\gamma|\partial_{\mathsf{h}}\Omega|)^{p-2}.$$

This implies that the left hand side of (34) is bounded from above by a universal constant multiple of $|\partial_{\mathsf{v}}\Omega|^{2/p}|\partial_{\mathsf{h}}\Omega|^{1-2/p}$. Therefore (34) is weaker than the estimate $|\partial_{\mathsf{v}}\Omega| \lesssim |\partial_{\mathsf{h}}\Omega|$ of Theorem 1.6. It would be interesting to determine the infimum over those p for which (34) holds true for every $\Omega \subseteq \mathbb{H}^3_{\mathbb{Z}}$, with our ongoing work showing that it is at least 4. In fact, this work shows that for every $R \geqslant 2$ the L_1 distortion of the ball of radius R in $\mathbb{H}^3_{\mathbb{Z}}$ is at most a constant multiple of $\sqrt[4]{\log R}$ — asymptotically *less* than the distortion of the ball of the same radius in $\mathbb{H}^5_{\mathbb{Z}}$. It would be interesting to determine the correct asymptotics of this distortion, with the best-known lower bound remaining that of [22], i.e., a constant multiple of $(\log R)^{\delta}$ for some universal constant $\delta > 0$. It should be stressed, however, that the algorithmic application of Theorem 1.6 that is obtained here uses Theorem (1.6) as stated for $\mathbb{H}^5_{\mathbb{Z}}$, and understanding the case of $\mathbb{H}^3_{\mathbb{Z}}$ would not yield any further improvement. So, while the above questions are geometrically and analytically interesting in their own right, they are not needed for applications that we currently have in mind.

1.7.2 Metric Embeddings. Theorem 1.2 also yields a sharp result for the general problem of finding the asymptotically largest-possible L_1 distortion of a finite doubling metric space with n points. A metric space (X, d_X) is said to be K-doubling for some $K \in \mathbb{N}$ if every ball in X (centered anywhere and of any radius) can be covered by K balls of half its radius. By [51],

$$c_1(X, d_X) \lesssim \sqrt{(\log K)\log|X|}. \tag{35}$$

As noted in [40], the dependence on $|X|$ in (35), but with a worse dependence on K, follows by combining results of [5] and [83] (the dependence on K that follows from [5, 83] was improved significantly in [40]). The metric space $(\mathbb{Z}^5, d_W)$ is $O(1)$-doubling because $|\mathcal{B}_R| \asymp R^6$ for every $R \geqslant 1$. Theorem 1.2 shows that (35) is sharp when $K = O(1)$, thus improving over the previously best-known construction [58] of arbitrarily large $O(1)$-doubling finite metric spaces $\{(X_i, d_i)\}_{i=1}^{\infty}$ for which $c_1(X_i, d_i) \gtrsim \sqrt{(\log|X_i|)/\log\log|X_i|}$. Probably (35) is sharp for every $K \leqslant |X|$; conceivably this could be proven by incorporating Theorem 1.2 into the argument of [41], but we shall not pursue this here. Theorem 1.2 establishes for the first time the existence of a metric space that simultaneously has several useful geometric properties and poor (indeed, worst possible) embeddability into L_1. By virtue of being $O(1)$-doubling, the metric space $(\mathbb{Z}^5, d_W)$ also has Markov type 2 due to [33] (which improves over [76], where the conclusion that it has Markov type p for every $p < 2$ was obtained). For more on the bi-Lipschitz invariant Markov type and its applications, see [8, 74]. The property of having Markov type 2 is shared by the construction of [58], which is also $O(1)$-doubling, but $(\mathbb{Z}^5, d_W)$ has additional features that the example of [58] fails to have. For one, it is a group; for another, by [60, 61] we know that $(\mathbb{Z}^5, d_W)$ has Markov convexity 4 (and no less). (See [57, 71] for background on the bi-Lipschitz invariant Markov convexity and its consequences.) By [71, Section 3] the example of [58] does not have Markov convexity p for any finite p. No examples of arbitrarily large finite metric spaces $\{(X_i, d_i)\}_{i=1}^{\infty}$ with bounded Markov convexity (and Markov convexity constants uniformly bounded) such that $c_1(X_i, d_i) \gtrsim \sqrt{\log|X_i|}$ were previously known to exist. Analogous statements are known to be impossible for Banach spaces [72], so it is natural in the context of the Ribe program (see the surveys [9, 74] for more on this research program) to ask whether there is a potential metric version of [72]; the above discussion shows that there is not.

1.7.3 The Sparsest Cut Problem with Uniform Demands. An important special case of the Sparsest Cut Problem is when the demand matrix D is the matrix $\mathbf{1}_{\{1,\ldots,n\}\times\{1,\ldots,n\}} \in M_n(\mathbb{R})$ all of whose entries equal 1 and the capacity matrix C lies in $M_n(\{0,1\})$, i.e., all its entries are either 0 or 1. This is known as the *Sparsest Cut Problem with Uniform Demands*. In this case C can also be described as the adjacency matrix of a graph G whose vertex set is $\{1,\ldots,n\}$ and whose edge set consists of those unordered pairs $\{i,j\} \subseteq \{1,\ldots,n\}$ for which $C_{ij} = 1$. With this interpretation, given $A \subseteq \{1,\ldots,n\}$ the numerator in (1) equals twice the number of edges that are incident to A in G. And, since $D = \mathbf{1}_{\{1,\ldots,n\}\times\{1,\ldots,n\}}$, the denominator in (1) is equal to $2|A|(n-|A|) \asymp n\min\{|A|, |\{1,\ldots,n\} \smallsetminus A|\}$. So, the Sparsest Cut Problem with Uniform Demands asks for an algorithm that takes as input a finite graph and outputs a quantity which is bounded above and below by universal constant multiples of its *conductance* [86] divided by n. The Goemans–Linial integrality gap corresponding to this special case is

$$\rho^{\mathrm{unif}}_{\mathrm{GL}}(n) \stackrel{\mathrm{def}}{=} \sup_{\substack{C\in M_n(\{0,1\})\\ C\text{ symmetric}}} \frac{\mathsf{OPT}(C, \mathbf{1}_{\{1,\ldots,n\}\times\{1,\ldots,n\}})}{\mathsf{SDP}(C, \mathbf{1}_{\{1,\ldots,n\}\times\{1,\ldots,n\}})}.$$

The Goemans–Linial algorithm furnishes the best-known approximation ratio also in the case of uniform demands. By the important work [4] we have $\rho^{\mathrm{unif}}_{\mathrm{GL}}(n) \lesssim \sqrt{\log n}$, improving over the previous bound $\rho^{\mathrm{unif}}_{\mathrm{GL}}(n) \lesssim \log n$ of [59]. As explained in [21], the present approach based on (fixed dimensional) Heisenberg groups cannot yield a lower bound on $\rho^{\mathrm{unif}}_{\mathrm{GL}}(n)$ that tends to ∞ with n. The best-known lower bound [45] is $\rho^{\mathrm{unif}}_{\mathrm{GL}}(n) \geqslant \exp(c\sqrt{\log\log n})$ for some universal constant $c > 0$, improving over the previous bound $\rho_{\mathrm{GL}}^{\mathrm{unif}}(n) \gtrsim \log\log n$ of [31]. Determining the asymptotic behavior of $\rho_{\mathrm{GL}}^{\mathrm{unif}}(n)$ remains an intriguing open problem.

REFERENCES

[1] A. Agrawal, P. Klein, R. Ravi, and S. Rao. 1990. Approximation through multicommodity flow. In *31st Annual Symposium on Foundations of Computer Science*. IEEE Computer Soc., Los Alamitos, CA, 726–737.

[2] Luigi Ambrosio. 2001. Some fine properties of sets of finite perimeter in Ahlfors regular metric measure spaces. *Adv. Math.* 159, 1 (2001), 51–67. DOI:http://dx.doi.org/10.1006/aima.2000.1963

[3] Sanjeev Arora, James R. Lee, and Assaf Naor. 2008. Euclidean distortion and the sparsest cut. *J. Amer. Math. Soc.* 21, 1 (2008), 1–21 (electronic). DOI:http://dx.doi.org/10.1090/S0894-0347-07-00573-5

[4] Sanjeev Arora, Satish Rao, and Umesh Vazirani. 2009. Expander flows, geometric embeddings and graph partitioning. *J. ACM* 56, 2 (2009), Art. 5, 37. DOI:http://dx.doi.org/10.1145/1502793.1502794

[5] Patrice Assouad. 1983. Plongements lipschitziens dans $\mathbf{R}^n$. *Bull. Soc. Math. France* 111, 4 (1983), 429–448. http://www.numdam.org/item?id=BSMF_1983__111__429_0

[6] Yonatan Aumann and Yuval Rabani. 1998. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.* 27, 1 (1998), 291–301 (electronic). DOI:http://dx.doi.org/10.1137/S0097539794285983

[7] Tim Austin, Assaf Naor, and Romain Tessera. 2013. Sharp quantitative nonembeddability of the Heisenberg group into superreflexive Banach spaces. *Groups Geom. Dyn.* 7, 3 (2013), 497–522. DOI:http://dx.doi.org/10.4171/GGD/193

[8] Keith Ball. 1992. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.* 2, 2 (1992), 137–172. DOI:http://dx.doi.org/10.1007/BF01896971

[9] Keith Ball. 2013. The Ribe programme. *Astérisque* 352 (2013), Exp. No. 1047, viii, 147–159. Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.

[10] Hyman Bass. 1972. The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc. (3)* 25 (1972), 603–614.

[11] Yoav Benyamini and Joram Lindenstrauss. 2000. *Geometric nonlinear functional analysis. Vol. 1*. American Mathematical Society Colloquium Publications, Vol. 48. American Mathematical Society, Providence, RI. xii+488 pages.

[12] Jean Bourgain. 1985. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.* 52, 1-2 (1985), 46–52. DOI:http://dx.doi.org/10.1007/BF02776078

[13] Jean Bourgain. 2002. On the distributions of the Fourier spectrum of Boolean functions. *Israel J. Math.* 131 (2002), 269–276. DOI:http://dx.doi.org/10.1007/BF02785861

[14] Dmitri Burago, Yuri Burago, and Sergei Ivanov. 2001. *A course in metric geometry*. Graduate Studies in Mathematics, Vol. 33. American Mathematical Society, Providence, RI. xiv+415 pages. DOI:http://dx.doi.org/10.1090/gsm/033

[15] Luca Capogna, Donatella Danielli, Scott D. Pauls, and Jeremy T. Tyson. 2007. *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*. Progress in Mathematics, Vol. 259. Birkhäuser Verlag, Basel. xvi+223 pages.

[16] Shuchi Chawla. 2008. Sparsest Cut. In *Encyclopedia of Algorithms*. Springer-Verlag US, 868–870.

[17] Shuchi Chawla, Anupam Gupta, and Harald Räcke. 2008. Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut. *ACM Trans. Algorithms* 4, 2 (2008), Art. 22, 18. DOI:http://dx.doi.org/10.1145/1361192.1361199

[18] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar. 2006. On the hardness of approximating multicut and sparsest-cut. *Comput. Complexity* 15, 2 (2006), 94–114. DOI:http://dx.doi.org/10.1007/s00037-006-0210-9

[19] Jeff Cheeger and Bruce Kleiner. 2010. Differentiating maps into L^1, and the geometry of BV functions. *Ann. of Math. (2)* 171, 2 (2010), 1347–1385. DOI:http://dx.doi.org/10.4007/annals.2010.171.1347

[20] Jeff Cheeger and Bruce Kleiner. 2010. Metric differentiation, monotonicity and maps to L^1. *Invent. Math.* 182, 2 (2010), 335–370. DOI:http://dx.doi.org/10.1007/s00222-010-0264-9

[21] Jeff Cheeger, Bruce Kleiner, and Assaf Naor. 2009. A $(\log n)^{\Omega(1)}$ integrality gap for the sparsest cut SDP. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)*. IEEE Computer Soc., Los Alamitos, CA, 555–564. DOI:http://dx.doi.org/10.1109/FOCS.2009.47

[22] Jeff Cheeger, Bruce Kleiner, and Assaf Naor. 2011. Compression bounds for Lipschitz maps from the Heisenberg group to L_1. *Acta Math.* 207, 2 (2011), 291–373. DOI:http://dx.doi.org/10.1007/s11511-012-0071-9

[23] Michael Christ. 1990. A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.* 60/61, 2 (1990), 601–628.

[24] Julia Chuzhoy and Sanjeev Khanna. 2007. Polynomial flow-cut gaps and hardness of directed cut problems [extended abstract]. In *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*. ACM, New York, 179–188.

[25] Julia Chuzhoy and Sanjeev Khanna. 2009. Polynomial flow-cut gaps and hardness of directed cut problems. *J. ACM* 56, 2 (2009), Art. 6, 28. DOI:http://dx.doi.org/10.1145/1502793.1502795

[26] Guy David. 1984. Opérateurs intégraux singuliers sur certaines courbes du plan complexe. *Ann. Sci. École Norm. Sup. (4)* 17, 1 (1984), 157–189. http://www.numdam.org/item?id=ASENS_1984_4_17_1_157_0

[27] Guy David. 1991. *Wavelets and singular integrals on curves and surfaces*. Lecture Notes in Mathematics, Vol. 1465. Springer-Verlag, Berlin. x+107 pages. DOI:http://dx.doi.org/10.1007/BFb0091544

[28] Guy David. 1991. *Wavelets and singular integrals on curves and surfaces*. Lecture Notes in Mathematics, Vol. 1465. Springer-Verlag, Berlin. x+107 pages. DOI:http://dx.doi.org/10.1007/BFb0091544

[29] Guy David and Stephen Semmes. 1991. Singular integrals and rectifiable sets in $\mathbf{R}^n$: Beyond Lipschitz graphs. *Astérisque* 193 (1991), 152.

[30] Guy David and Stephen Semmes. 1993. *Analysis of and on uniformly rectifiable sets*. Mathematical Surveys and Monographs, Vol. 38. American Mathematical Society, Providence, RI. xii+356 pages. DOI:http://dx.doi.org/10.1090/surv/038

[31] Nikhil R. Devanur, Subhash A. Khot, Rishi Saket, and Nisheeth K. Vishnoi. 2006. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*. ACM, New York, 537–546. DOI:http://dx.doi.org/10.1145/1132516.1132594

[32] Michel Marie Deza and Monique Laurent. 1997. *Geometry of cuts and metrics*. Algorithms and Combinatorics, Vol. 15. Springer-Verlag, Berlin. xii+587 pages. DOI:http://dx.doi.org/10.1007/978-3-642-04295-9

[33] Jian Ding, James R. Lee, and Yuval Peres. 2013. Markov type and threshold embeddings. *Geom. Funct. Anal.* 23, 4 (2013), 1207–1229. DOI:http://dx.doi.org/10.1007/s00039-013-0234-7

[34] Bruno Franchi, Raul Serapioni, and Francesco Serra Cassano. 2001. Rectifiability and perimeter in the Heisenberg group. *Math. Ann.* 321, 3 (2001), 479–531. DOI:http://dx.doi.org/10.1007/s002080100228

[35] Bruno Franchi, Raul Serapioni, and Francesco Serra Cassano. 2003. On the structure of finite perimeter sets in step 2 Carnot groups. *J. Geom. Anal.* 13, 3 (2003), 421–466. DOI:http://dx.doi.org/10.1007/BF02922053

[36] Bruno Franchi, Raul Serapioni, and Francesco Serra Cassano. 2006. Intrinsic Lipschitz graphs in Heisenberg groups. *J. Nonlinear Convex Anal.* 7, 3 (2006), 423–441.

[37] Bruno Franchi, Raul Serapioni, and Francesco Serra Cassano. 2011. Differentiability of intrinsic Lipschitz functions within Heisenberg groups. *J. Geom. Anal.* 21, 4 (2011), 1044–1084. DOI:http://dx.doi.org/10.1007/s12220-010-9178-4

[38] Michel X. Goemans. 1997. Semidefinite programming in combinatorial optimization. *Math. Programming* 79, 1-3, Ser. B (1997), 143–161. Lectures on mathematical programming (ismp97) (Lausanne, 1997).

[39] Martin Grötschel, László Lovász, and Alexander Schrijver. 1993. *Geometric algorithms and combinatorial optimization* (second ed.). Algorithms and Combinatorics, Vol. 2. Springer-Verlag, Berlin. xii+362 pages.

[40] Anupam Gupta, Robert Krauthgamer, and James R. Lee. 2003. Bounded Geometries, Fractals, and Low-Distortion Embeddings. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*. IEEE Computer Society, 534–543. DOI:http://dx.doi.org/10.1109/SFCS.2003.1238226

[41] Alexander Jaffe, James R. Lee, and Mohammad Moharrami. 2011. On the optimality of gluing over scales. *Discrete Comput. Geom.* 46, 2 (2011), 270–282. DOI:http://dx.doi.org/10.1007/s00454-011-9359-3

[42] Peter W. Jones. 1989. Square functions, Cauchy integrals, analytic capacity, and harmonic measure. In *Harmonic analysis and partial differential equations (El Escorial, 1987)*. Lecture Notes in Math., Vol. 1384. Springer, Berlin, 24–68. DOI:http://dx.doi.org/10.1007/BFb0086793

[43] Peter W. Jones. 1990. Rectifiable sets and the traveling salesman problem. *Invent. Math.* 102, 1 (1990), 1–15. DOI:http://dx.doi.org/10.1007/BF01233418

[44] Jeff Kahn, Gil Kalai, and Nathan Linial. 1988. The Influence of Variables on Boolean Functions (Extended Abstract). In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*. IEEE Computer Society, 68–80. DOI:http://dx.doi.org/10.1109/SFCS.1988.21923

[45] Daniel Kane and Raghu Meka. 2013. A PRG for Lipschitz functions of polynomials with applications to sparsest cut. In *STOC'13—Proceedings of the 2013 ACM Symposium on Theory of Computing*. ACM, New York, 1–10. DOI:http://dx.doi.org/10.1145/2488608.2488610

[46] Subhash Khot. 2002. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*. ACM, New York, 767–775. DOI:http://dx.doi.org/10.1145/509907.510017

[47] Subhash Khot. 2010. Inapproximability of NP-complete problems, discrete Fourier analysis, and geometry. In *Proceedings of the International Congress of Mathematicians. Volume IV*. Hindustan Book Agency, New Delhi, 2676–2697.

[48] Subhash Khot and Assaf Naor. 2006. Nonembeddability theorems via Fourier analysis. *Math. Ann.* 334, 4 (2006), 821–852. DOI:http://dx.doi.org/10.1007/s00208-005-0745-0

[49] Subhash A. Khot and Nisheeth K. Vishnoi. 2015. The unique games conjecture, integrability gap for cut problems and embeddability of negative-type metrics

into ℓ_1. *J. ACM* 62, 1 (2015), Art. 8, 39. DOI:http://dx.doi.org/10.1145/2629614

[50] Bernd Kirchheim. 1994. Rectifiable metric spaces: local structure and regularity of the Hausdorff measure. *Proc. Amer. Math. Soc.* 121, 1 (1994), 113–123. DOI:http://dx.doi.org/10.2307/2160371

[51] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. 2005. Measured descent: a new embedding method for finite metrics. *Geom. Funct. Anal.* 15, 4 (2005), 839–858. DOI:http://dx.doi.org/10.1007/s00039-005-0527-6

[52] Robert Krauthgamer and Yuval Rabani. 2009. Improved lower bounds for embeddings into L_1. *SIAM J. Comput.* 38, 6 (2009), 2487–2498. DOI:http://dx.doi.org/10.1137/060660126

[53] Vincent Lafforgue and Assaf Naor. 2014. Vertical versus horizontal Poincaré inequalities on the Heisenberg group. *Israel J. Math.* 203, 1 (2014), 309–339. DOI:http://dx.doi.org/10.1007/s11856-014-1088-x

[54] James R. Lee. 2005. On distance scales, embeddings, and efficient relaxations of the cut cone. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. ACM, New York, 92–101 (electronic).

[55] James R. Lee, Manor Mendel, and Assaf Naor. 2005. Metric structures in L_1: dimension, snowflakes, and average distortion. *European J. Combin.* 26, 8 (2005), 1180–1190. DOI:http://dx.doi.org/10.1016/j.ejc.2004.07.002

[56] James R. Lee and Assaf Naor. 2006. L_p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *Proceedings of 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006)*. 99–108. Available at https://web.math.princeton.edu/~naor/homepage%20files/L_pHGL.pdf.

[57] James R. Lee, Assaf Naor, and Yuval Peres. 2009. Trees and Markov convexity. *Geom. Funct. Anal.* 18, 5 (2009), 1609–1659. DOI:http://dx.doi.org/10.1007/s00039-008-0689-0

[58] James R. Lee and Anastasios Sidiropoulos. 2011. Near-optimal distortion bounds for embedding doubling spaces into L_1 [extended abstract]. In *STOC'11—Proceedings of the 43rd ACM Symposium on Theory of Computing*. ACM, New York, 765–772. DOI:http://dx.doi.org/10.1145/1993636.1993737

[59] Tom Leighton and Satish Rao. 1999. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM* 46, 6 (1999), 787–832. DOI:http://dx.doi.org/10.1145/331524.331526

[60] Sean Li. 2014. Coarse differentiation and quantitative nonembeddability for Carnot groups. *J. Funct. Anal.* 266, 7 (2014), 4616–4704. DOI:http://dx.doi.org/10.1016/j.jfa.2014.01.026

[61] Sean Li. 2016. Markov convexity and nonembeddability of the Heisenberg group. *Ann. Inst. Fourier (Grenoble)* 66, 4 (2016), 1615–1651.

[62] Nathan Linial. 2002. Finite metric-spaces—combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002)*. Higher Ed. Press, Beijing, 573–586.

[63] Nathan Linial. 2002. Squared ℓ_2 metrics into ℓ_1. In *Open problems on embeddings of finite metric spaces, edited by J. Matoušek*. 5.

[64] Nathan Linial, Eran London, and Yuri Rabinovich. 1995. The geometry of graphs and some of its algorithmic applications. *Combinatorica* 15, 2 (1995), 215–245.

[65] Valentino Magnani. 2011. Area implies coarea. *Indiana Univ. Math. J.* 60, 1 (2011), 77–100. DOI:http://dx.doi.org/10.1512/iumj.2011.60.4172

[66] Konstantin Makarychev, Yury Makarychev, and Aravindan Vijayaraghavan. 2014. Bilu-Linial stable instances of max cut and minimum multiway cut. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*. ACM, New York, 890–906. DOI:http://dx.doi.org/10.1137/1.9781611973402.67

[67] Teresa Martínez, José L. Torrea, and Quanhua Xu. 2006. Vector-valued Littlewood-Paley-Stein theory for semigroups. *Adv. Math.* 203, 2 (2006), 430–475. DOI:http://dx.doi.org/10.1016/j.aim.2005.04.010

[68] Jiří Matoušek. 2002. *Lectures on discrete geometry*. Graduate Texts in Mathematics, Vol. 212. Springer-Verlag, New York. xvi+481 pages. DOI:http://dx.doi.org/10.1007/978-1-4613-0039-7

[69] Jiří Matoušek. 2002. *Lectures on discrete geometry*. Graduate Texts in Mathematics, Vol. 212. Springer-Verlag, New York. xvi+481 pages. DOI:http://dx.doi.org/10.1007/978-1-4613-0039-7

[70] Pertti Mattila. 1995. *Geometry of sets and measures in Euclidean spaces*. Cambridge Studies in Advanced Mathematics, Vol. 44. Cambridge University Press, Cambridge. xii+343 pages. DOI:http://dx.doi.org/10.1017/CBO9780511623813 Fractals and rectifiability.

[71] Manor Mendel and Assaf Naor. 2013. Markov convexity and local rigidity of distorted metrics. *J. Eur. Math. Soc. (JEMS)* 15, 1 (2013), 287–337. DOI:http://dx.doi.org/10.4171/JEMS/362

[72] Vitali D. Milman and Haim Wolfson. 1978. Minkowski spaces with extremal distance from the Euclidean space. *Israel J. Math.* 29, 2-3 (1978), 113–131.

[73] Assaf Naor. 2010. L_1 embeddings of the Heisenberg group and fast estimation of graph isoperimetry. In *Proceedings of the International Congress of Mathematicians. Volume III*. Hindustan Book Agency, New Delhi, 1549–1575.

[74] Assaf Naor. 2012. An introduction to the Ribe program. *Jpn. J. Math.* 7, 2 (2012), 167–233. DOI:http://dx.doi.org/10.1007/s11537-012-1222-7

[75] Assaf Naor. 2014. Comparison of metric spectral gaps. *Anal. Geom. Metr. Spaces* 2 (2014), 1–52. DOI:http://dx.doi.org/10.2478/agms-2014-0001

[76] Assaf Naor, Yuval Peres, Oded Schramm, and Scott Sheffield. 2006. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.* 134, 1 (2006), 165–197. DOI:http://dx.doi.org/10.1215/S0012-7094-06-13415-4

[77] Assaf Naor, Yuval Rabani, and Alistair Sinclair. 2005. Quasisymmetric embeddings, the observable diameter, and expansion properties of graphs. *J. Funct. Anal.* 227, 2 (2005), 273–303. DOI:http://dx.doi.org/10.1016/j.jfa.2005.04.003

[78] Assaf Naor and Lior Silberman. 2011. Poincaré inequalities, embeddings, and wild groups. *Compos. Math.* 147, 5 (2011), 1546–1572. DOI:http://dx.doi.org/10.1112/S0010437X11005343

[79] Assaf Naor and Robert Young. 2017. Vertical perimeter versus horizontal perimeter. (2017). Preprint available at https://arxiv.org/abs/1701.00620.

[80] Mikhail I. Ostrovskii. 2013. *Metric embeddings*. De Gruyter Studies in Mathematics, Vol. 49. De Gruyter, Berlin. xii+372 pages. DOI:http://dx.doi.org/10.1515/9783110264012 Bilipschitz and coarse embeddings into Banach spaces.

[81] Pierre Pansu. 1982. Une inégalité isopérimétrique sur le groupe de Heisenberg. *C. R. Acad. Sci. Paris Sér. I Math.* 295, 2 (1982), 127–130.

[82] Pierre Pansu. 1989. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math. (2)* 129, 1 (1989), 1–60. DOI:http://dx.doi.org/10.2307/1971484

[83] Satish Rao. 1999. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry (Miami Beach, FL, 1999)*. ACM, New York, 300–306 (electronic). DOI:http://dx.doi.org/10.1145/304893.304983

[84] Farhad Shahrokhi and D. W. Matula. 1990. The maximum concurrent flow problem. *J. Assoc. Comput. Mach.* 37, 2 (1990), 318–334.

[85] David B. Shmoys. 1997. Cut problems and their application to divide-and-conquer. In *Approximation Algorithms for NP-hard Problems, (D.S. Hochbaum, ed.)*. PWS, 192–235.

[86] Alistair Sinclair and Mark Jerrum. 1989. Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. and Comput.* 82, 1 (1989), 93–133. DOI:http://dx.doi.org/10.1016/0890-5401(89)90067-9

[87] Romain Tessera. 2008. Quantitative property A, Poincaré inequalities, L^p-compression and L^p-distortion for metric measure spaces. *Geom. Dedicata* 136 (2008), 203–220. DOI:http://dx.doi.org/10.1007/s10711-008-9286-5

[88] Luca Trevisan. 2012. On Khot's unique games conjecture. *Bull. Amer. Math. Soc. (N.S.)* 49, 1 (2012), 91–111. DOI:http://dx.doi.org/10.1090/S0273-0979-2011-01361-1

[89] Hans S. Witsenhausen. 1986. Minimum dimension embedding of finite metric spaces. *J. Combin. Theory Ser. A* 42, 2 (1986), 184–199. DOI:http://dx.doi.org/10.1016/0097-3165(86)90089-0