On Multiplicative λ-Approximations and Some Geometric Applications*

Ilan Newman[†] Yuri Rabinovich[‡]

Abstract

Let $\mathcal{F}$ be a set system over an underlying finite set X, and let μ be a nonnegative measure over X. I.e., for every $S \subseteq X$, $\mu(S) = \sum_{x \in S} \mu(x)$. A measure μ^* on X is called a *multiplicative λ-approximation* of μ on $(\mathcal{F}, X)$ if for every $S \in \mathcal{F}$ it holds that $a\mu(S) \leq \mu^*(S) \leq b\mu(S)$, and $b/a = \lambda \geq 1$. The central question raised and partially answered in the present paper is about the existence of meaningful structural properties of $\mathcal{F}$ implying that for any μ on X there exists an $\frac{1+\epsilon}{1-\epsilon}$-approximation μ^* supported on a small subset of X.

It turns out that the parameter that governs the support size of a multiplicative approximation is the *triangular rank* of $\mathcal{F}$, $\mathrm{trk}(\mathcal{F})$. It is defined as the maximal length of a sequence of sets $\{S_i\}_{i=1}^t$ in $\mathcal{F}$ such that for all $1 < i \leq t$, $S_i \not\subseteq \cup_{j<i} S_j$. We show that for any μ on X and $0 < \epsilon < 1$, there is measure μ^* that $\frac{1+\epsilon}{1-\epsilon}$-approximates μ on $(X, \mathcal{F})$, and has support of size $O(\mathrm{trk}(\mathcal{F})^2 \log(\mathrm{trk}(\mathcal{F}))/\mathrm{poly}(\epsilon))$. We also present two alternative constructions which in some cases improve upon this bound. Conversely, we show that for any $0 \leq \epsilon < 1$ there exists a μ on X that cannot be $\frac{1+\epsilon}{1-\epsilon}$-approximated on $(\mathcal{F}, X)$ by any μ^* with support of size $< \mathrm{trk}(\mathcal{F})$. For special families $\mathcal{F}$ this bound can be improved to $\Omega(\mathrm{trk}(\mathcal{F})/\epsilon)$.

As an application we show a new dimension-reduction result for ℓ_1 metrics: Any ℓ_1-metric on n points can be (efficiently) embedded with $\frac{1+\epsilon}{1-\epsilon}$-distortion into $\mathbb{R}^{O(n/\epsilon^2)}$ equipped with the ℓ_1 norm. This improves over the best previously known bound of $O(n \log n/\mathrm{poly}(\epsilon))$ on dimension, due to Schechtman.

We obtain also some new results on efficient sampling of Euclidean volumes. In order to make the general framework applicable to this setting, we develop the basic theory of finite volumes, analogous to the theory of finite metrics, and get results of independent interest in this direction. To do so, we use basic combinatorial/topological facts about simplicial complexes, and study the naturally arising questions.

1 Introduction

Initially motivated by a problem in finite metric spaces, we pose the following general question.

Let $\mathcal{F}$ be a set system over an underlying finite set X. What are the structural properties of $\mathcal{F}$ that would ensure that for any nonnegative weighting of X, there exists a small weighted sample of X such that for every $S \in \mathcal{F}$, the original weight and the sampled weight differ by a small multiplicative *factor?*

An additive counterpart of this question has been extensively studied, and has turned out to be extraordinary fruitful. A rich theory that emerged has numerous applications in divers areas, e.g., Learning Theory, Discrete Geometry, randomness extraction etc. The key parameter in the additive setting is the Vapnik-Chervonenkis dimension of $\mathcal{F}$, defined as the size of the largest subset $Y \subseteq X$ shattered by $\mathcal{F}$, i.e., $\mathcal{F}|_Y = 2^Y$.

The multiplicative setting has achieved so far relatively less attention. It has been considered mainly by the Computational Geometry community in the framework of constructing efficient *core-sets*. In our multiplicative approximation problem, $\mathcal{F}$ is a collection of subsets of X, and it can be identified with a collection of 0/1 functions on X (the corresponding characteristic functions). In the more general setting addressed in Computational Geometry, $\mathcal{F}$ is a collection of functions from X to $\mathbb{R}^+$. A core-set is then a small weighted sample $X^* \subset X$ such that for every $f \in \mathcal{F}$, the average of f on X is multiplicatively approximated by its weighted average on X^*. [1] Most relevant works are dedicated to study of specific problems e.g., k-median, clustering etc., (cf., Chapt. 19 in [16]). A recent paper [13] introduces a general method, and shows how it can be applied to a long list of specific problems. It seems, however, that the methods and results of [13] are different, and even incomparable, with those of the present paper.

Two important works that are more relevant to this paper are the results of Benczúr and Karger [8], and Batson, Spielman and Srivastava [7]. These papers study a specific multiplicative approximation problem concerning cuts in undirected graphs. However, their deep and elegant findings indicate that there might exist a rich general theory. In this paper, using their achievements, we develop the foundations of such a theory.

The *triangular rank* of $\mathcal{F}$, $\mathrm{trk}(\mathcal{F})$, is perhaps the key parameter in the multiplicative setting. It has a number of equivalent definitions, the shortest $-$ being the size of the largest square lower-triangular submatrix of the

*This Research was supported by The Israel Science Foundation (grant number 862/10.)

[†]Department of Computer Science, University of Haifa, Haifa, Israel. E-mails: ilan@cs.haifa.ac.il.

[‡]Department of Computer Science, University of Haifa, Haifa, Israel. E-mails: yuri@cs.haifa.ac.il.

[1]Often a dual definition is used, i.e., a small weighted sample $\mathcal{F}^* \subset \mathcal{F}$ is sought, such that for every $x \in X$, the weighted average value of $f(x)$'s over $\mathcal{F}^*$ multiplicatively approximates the average value of $f(x)$'s over the entire $\mathcal{F}$.

incidence matrix of $\mathcal{F}$ vs. X with 1's on the diagonal. Another important parameter is $\text{rank}^*(\mathcal{F})$, the *soft rank* of $\mathcal{F}$, being the minimal possible rank (over $\mathbb{R}$) of the incidence matrix under all possible sign choices for its entries. Clearly, $\text{trk}(\mathcal{F}) \leq \text{rank}^*(\mathcal{F})$.

To formulate our findings, we introduce some definitions. A *measure* μ on $(X, \mathcal{F})$ is a nonnegative weighing $\{\mu(x)\}$ of X, that is extended to members of $\mathcal{F}$ in the standard way, namely, for $S \in \mathcal{F}$, $\mu(S) = \sum_{x \in S} \mu(x)$. A measure μ^* on X is called a *multiplicative* $\frac{1+\epsilon}{1-\epsilon}$-*approximation* of μ on $(\mathcal{F}, X)$ if for all $S \in \mathcal{F}$ it holds that $(1-\epsilon)\mu(S) \leq \mu^*(S) \leq (1+\epsilon)\mu(S)$. We shall be mostly concerned with constructing good approximations of small support.

The main positive result involving $\text{trk}(\mathcal{F})$ is that any measure μ on $(X, \mathcal{F})$ can be $\frac{1+\epsilon}{1-\epsilon}$-approximated by μ^* with support of size $O(\text{trk}(\mathcal{F})^2 \log(\text{trk}(\mathcal{F}))/\text{poly}(\epsilon))$. A complementary negative result is that there exist measures on $\mathcal{F}$ which are approximated arbitrarily badly by any μ^* on X with support of size $< \text{trk}(\mathcal{F})$. It is also easy to construct specific $(\mathcal{F}, X)$'s such that any $\frac{1+\epsilon}{1-\epsilon}$-approximation of the counting measure on X must have support of size $\Omega(\text{trk}(\mathcal{F})/\epsilon)$.

Since there is, however, a quadratic gap between the upper and the lower bounds on the size of the support of a $\frac{1+\epsilon}{1-\epsilon}$-approximation μ^*, we present two additional constructions that in some cases yield better upper bounds. The first construction produces μ^* with support of size $O(\text{trk}(\mathcal{F}) \cdot \log|\mathcal{F}|/\epsilon^2)$. In the second construction the support is of size $O(\text{rank}^*(\mathcal{F})/\epsilon^2)$.

All our constructions are randomized, but unlike in the additive setting, the sampling is not (and typically cannot be) uniform. The efficiency of the constructions crucially depends on the representation of $\mathcal{F}$ (sometimes an implicit representation is of interest, e.g., in [8] where $\mathcal{F}$ is the set of cuts of a given graph), and on the complexity of finding a set $S \in \mathcal{F}$ of a minimum (or at least approximately minimum) weight $\mu(S)$ for a given weight function μ.

We present a number of applications of the general theory. The most interesting among them is a new dimension-reduction result for ℓ_1 metrics. We show that any ℓ_1-metric on n points can be (efficiently) embedded with $\frac{1+\epsilon}{1-\epsilon}$ distortion into $\mathbb{R}^{O(n/\epsilon^2)}$ equipped with the ℓ_1 norm. This improves over the best previously known bound on dimension of $O(n \log n/\text{poly}(\epsilon))$ due to Schechtman [30], and comes close to the almost linear recent lower bound of Andoni, Charikar and Neiman [2].

Another application has to do with Euclidean volumes. Assume that X is embedded in an Euclidean space, and the goal is to estimate the average of the volumes of the d-simplices spanned by the $(d+1)$-tuples of X. For $d = 1$ the problem was addressed in [4], where it was shown that the average can be $(1+\epsilon)$-approximated by non-adaptively sampling a predefined (universal, efficiently constructible) set of $O(n/\epsilon^2)$ pairs of vertices, and outputting the average of the observed distances. We show that a similar result holds for any d, the number of $(d+1)$-tuples sampled being $O(n^d/\epsilon^2)$. This is a gain of an $\Omega(n)$ factor over the trivial $O(n^{d+1})$.

An examination of the range of applicability of our ϵ-approximation techniques, and an effort to gain a better understanding of a deceivingly simple argument behind the above ℓ_1 dimension reduction result, both lead us to a consideration of abstract finite volume spaces, a high-dimensional analogue of the finite metric spaces.

Finite volumes make a sporadic appearance in CS, e.g., in the classical algorithm of Feige [12] for approximating the bandwidth, or in [22], where a strong dimension-reduction result for Euclidean volumes is established. They are well suited to represent quantitative d-ary relations that naturally appear both in the theory of CS and in applications. However, we are aware of no formal treatment of finite volumes analogous to that of finite metrics.

We define a d-volume space (X, ν) as a volume function from (unordered) $(d+1)$-tuples of X to $\mathbb{R}^+$ satisfying a high-dimensional analogue of the triangle inequality, namely, that for any d-cycle C (this term will be clarified later), and any $(d+1)$-tuple $\sigma \in C$, it holds that $\nu(\sigma) \leq \sum_{\tau \in C,\ \tau \neq \sigma} \nu(\tau)$. Using basic ideas of Matroid Theory and Combinatorial Topology, we introduce the notions of hypertrees, hypercuts, etc. This, in turn, allows us to define ℓ_1- and negative-type volumes. In this framework we obtain a generalization of the above ℓ_1 dimension reduction result for metrics, with an application to geometrical sampling, and develop the tools needed to establish the application to Euclidean volumes mentioned above. We then proceed a little further than needed for the above applications, and establish upper and lower bounds on the approximation of general finite volumes by ℓ_1-volumes.

Moving to higher dimensions is not without difficulties. Even on the level of basic definitions, one has to wisely choose among many possible extensions of the 1-dimensional case. The underlying combinatorics becomes significantly more involved, and even the most natural questions become computationally difficult. And yet, as we hope to demonstrate in this paper, it is possible to construct a meaningful, tractable, and potentially useful theory of finite volume spaces. Furthermore, in analogy to the interplay between the theory of finite metric spaces and graph theory, there is there is a close interplay between the theory of finite volumes and the combinatorics of simplicial complexes. We believe, and partially demonstrate in this paper, that the study of finite volumes could provide a unique perspective on combinatorics of simplicial complexes, a fascinating area that rapidly gains popularity (see, e.g.,[20, 27, 25, 26, 36, 29]), helping in the choice of "right" definitions, leading to interesting natural questions, and suggesting tools for approaching some of these questions.

Note: Some proofs are omitted from this extended abstract. They will appear in the full version.

2 Multiplicative approximation via triangular rank

The *triangular rank* $\mathrm{trk}(\mathcal{F})$ of a set system $\mathcal{F}$ over X is defined as the maximal length of a sequence of sets $\{S_i\}_{i=1}^{t}$ in $\mathcal{F}$ such that for all $1 < i \le t$, $S_i \not\subseteq \cup_{j<i} S_j$. Equivalently, it is the size of the largest square lower-triangular submatrix (with 1's on the diagonal) in the the incidence matrix of $(\mathcal{F}, X)$. Although this will not be used in the present paper, let us mention that trk correlates well with the operation of union, but not with the intersection and the complement. In particular, it is easy to verify that $\mathrm{trk}(\mathcal{F} \cup \mathcal{H}) \le \mathrm{trk}(\mathcal{F}) + \mathrm{trk}(\mathcal{H})$, and also that $\mathrm{trk}(\mathcal{F}^{\cup}) = \mathrm{trk}(\mathcal{F})$, where $\mathcal{F}^{\cup}$ is the closure of $\mathcal{F}$ under taking unions.

Given a nonnegative measure μ over X, the goal is to construct a small-support measure μ^* on X such that for every $S \in \mathcal{F}$ it holds that $(1-\epsilon)\mu(S) \le \mu^*(S) \le (1+\epsilon)\mu(S)$. Such μ will be called a multiplicative $\frac{1+\epsilon}{1-\epsilon}$-approximation of μ with respect to $(\mathcal{F}, X)$.

In comparison, μ^* *additively* ϵ-approximates μ if $|\mu(S) - \mu^*(S)| \le \epsilon \cdot \mu(X)$ for every $S \in \mathcal{F}$. The Vapnik-Chervonenkis dimension of $\mathcal{F}$, $\mathrm{VCdim}(\mathcal{F})$, is the size of the maximum subset $Y \subseteq X$ shattered by $\mathcal{F}$. Clearly, $\mathrm{trk}(\mathcal{F}) \ge \mathrm{VCdim}(\mathcal{F})$, and, moreover, $\mathrm{trk}(\mathcal{F})$ can be arbitrarily large even when $\mathrm{VCdim}(\mathcal{F}) = 1$, e.g., consider $\mathcal{F} = \{x\}_{x \in X}$. The additive approximation is closely related to the VC-dimension:

Theorem 2.1. *(a special case of [19, 34]): Let $\mathcal{F}$ be a set system on the underlying set X. Then, for any measure μ on X and $0 < \delta \le 1$, there exists an additive δ-approximation μ^* on $(X, \mathcal{F})$ of size at most $O(\mathrm{VCdim}(\mathcal{F})/\delta^2)$.*

The main goal of this section, as well as one of the key contributions of this paper, is to show that the multiplicative approximation is closely related to the triangular rank.

We start with presenting a claim linking the triangular rank of $\mathcal{F}$ to the distribution of values of μ on $(\mathcal{F}, X)$.

Claim 2.1. *Let $\mathcal{F}$ be a set system on the underlying set X. Then, for any measure μ on X, the values $\{\mu(S) \mid S \in \mathcal{F}\}$ can be bucketed into at most $\mathrm{trk}(\mathcal{F})$ buckets, such that in any bucket the values differ by at most a multiplicative factor of 2.*

Proof. Consider the longest chain of sets $S_1, S_2, \ldots, S_t$ such that for all $1 \le i < t$, $2\mu(S_i) < \mu(S_{i+1})$. It follows that $\mu(S_{i+1}) > \mu(\cup_{j<i} S_j)$, and in particular $S_{i+1} \not\subseteq \cup_{j<i} S_j$, implying that $t \le \mathrm{trk}(\mathcal{F})$. Thus, $\mathcal{F}$ can be covered by at most $\mathrm{trk}(\mathcal{F})$ antichains with respect to this partial order, while any antichain defines a bucket with the desired property. ∎

Corollary 2.1. *Let $\mathcal{F}$ be a set system on the underlying set X. Then for every $\lambda \ge 1$ there exists a measure μ on X which cannot be λ-approximated by any μ^* with support of size $< \mathrm{trk}(\mathcal{F})$.*

Proof. Let $\{S_i\}_{i=1}^{\mathrm{trk}(\mathcal{F})}$ be a sequence of sets in $\mathcal{F}$ as in the definition of the triangular rank, and let $\{x_i\}_{i=1}^{\mathrm{trk}(\mathcal{F})}$ a sequence of corresponding elements of X such that $x_i \in S_i$, but $x_i \notin S_j$ for $j < i$. Define the measure μ on X by assigning $\mu(x_i) = (3\lambda)^i$ for every x_i, and assigning 0 to rest of X. Observe that $\mu(S_i) \ge (3\lambda)^i$ while $\mu(S_{i-1}) \le (3\lambda)^i/2\lambda$.

Assume by contradiction that μ^* as above exists. By Claim 2.1, the sets $S \in \mathcal{F}$ can be 2-bucketed according to their μ^*-values in less than $\mathrm{trk}(\mathcal{F}) - 1$ buckets. By the pigeonhole principle, there will be a bucket containing some two sets S_i and S_j as above. However, the ratio between their μ-values exceed 2λ, contrary to our assumptions that their μ^* values are within a factor of 2, and that μ^* λ-approximates μ. ∎

To get a lower bound on the support of μ^* in terms of λ for a specific family, assume that $|X|$ is large with respect to t, and $\mathcal{F}$ consists of all subsets of X of size larger than $|X| - t$. It is easy to verify that that $\mathrm{trk}(\mathcal{F}) = \mathrm{t}$, and that in order to $\frac{1}{1-\epsilon}$-approximate the counting measure on X one need a sample of size $\ge (t-1)/\epsilon$.

Next, we address the technically more demanding upper bounds, and present the two central results of this section. The first theorem is used to establish the second, but it is also of an independent value.

Theorem 2.2. *Let $\mathcal{F}$ be a set system on the underlying set X. Then, for any measure μ on X and $0 < \epsilon < 1$, there exists a multiplicative $\frac{1+\epsilon}{1-\epsilon}$-approximation μ^* on $(X, \mathcal{F})$ of size at most $O(\mathrm{trk}(\mathcal{F}) \cdot \log|\mathcal{F}|/\epsilon^2)$.*

Theorem 2.3. *Under the same assumptions as in Theorem 2.2, and for ϵ bounded from above* *by some* $c_0 < 1$, *there exists a multiplicative* $\frac{1+\epsilon}{1-\epsilon}$*-approximation* μ^* *on* $(X, \mathcal{F})$ *of size at most* $O(\,\mathrm{trk}^2(\mathcal{F}) \log \mathrm{trk}(\mathcal{F})/\epsilon^2 \;+\; \frac{1}{\epsilon^2} \log \frac{1}{\epsilon}\,)$.

2.1 Sketch of the Proof of Theorem 2.2 The method of proof generalizes the method of Karger and Benczúr from [8]. The existence of μ^* will be established using a probabilistic argument. We start with some preliminary observations.

For the proof, it suffices to address the case when μ is a counting measure, i.e., $\mu(S) = |S|$, since any other measure can be reduced to it: we first reduce to the case where μ is integral by scaling, and then take $\mu(x)$ copies of each element $x \in X$ and updating $\mathcal{F}$ accordingly[2]. Note that the triangular rank is not affected by any of the two steps and hence the result can be stated in terms of the triangular rank of the old system. However, the representation size may drastically change, hence algorithmic issues arise, and will be addressed later.

As we are about to sample the elements of X, notice that some elements are more essential than the others, and thus the sampling is necessarily non-uniform. For example, if a set $\{x\}$ belongs to $\mathcal{F}$, then the element x must necessarily be chosen. More generally, if $S \in \mathcal{F}$ is small, the elements $x \in S$ should be sampled with relatively hight probability. Thus, it makes sense to assign to each element $i \in X$ a *fragility* parameter indicating how carefully should it be sampled. We assume that $\cup_{S \in F} = X$ as we may remove elements that do not appear in any set.

Definition 2.1. *Define a partition of* X *with respect to* $\mathcal{F}$ *in the following manner:*
$i = 0$.
While $X \neq \emptyset$, *repeat:*
    $i = i + 1$;
    Let B_i *be the (currently) smallest nonempty set in* $\mathcal{F}$;
    Let $\mathcal{F} = \mathcal{F}|_{X - B_i} = \{S - B_i \,|\, S \in \mathcal{F}\}$, *and let* $X = X - B_i$.

Clearly, X *is a disjoint union of* B_i*'s created in the above process. The* strength $s(x)$ *of an element* $x \in B_k$ *is defined as* $s(x) = \max_{i \le k} |B_i|$. *The* fragility *of* x *is the inverse of its strength,* $f(x) = \frac{1}{s(x)}$.

The following lemma describes the basic properties of these notions.

Lemma 2.1.
1. *For any set* $S \in \mathcal{F}$ *it holds that* $|S| \geq \max_{x \in S} s(x)$.
2. $\sum_{x \in X} f(x) \;\le\; N$, *where* N *is the number of blocks in the above partition.*
3. $N \le \mathrm{trk}(\mathcal{F})$.

Proof. [of Lemma 2.1] For the first statement: Let i be the smallest number such that $|B_i| = s(x)$, and let k be smallest number such that $S \subseteq \cup_{j=1}^{k} B_j$. Then, by definition of the strength, it holds that $i \le k$, and thus the set $S - \cup_{j=1}^{i-1} B_j$ is not empty, and therefore was a candidate to be chosen at step i of the process. The fact that B_i was chosen instead means that $|S - \cup_{j=1}^{i-1} B_j| \geq |B_i|$, and the statement on $|S|$ follows.

For the second statement, observe that all elements in the block B_i have (the same) strength $s_i \geq |B_i|$. Thus,

$$\sum_{x \in X} f(x) \;=\; \sum_{i=1}^{N} \sum_{x \in B_i} f(x) \;=\; \sum_{i=1}^{N} |B_i|/s_i \;\le\; N\,.$$

For the third statement, let S_i be the set used in the definition of the block B_i. Then, no S_i is contained in the union of its predecessors $\cup_{j<i} S_i$. Since there can be at most $\mathrm{trk}(\mathcal{F})$ such sets in $\mathcal{F}$, the statement follows. ∎

We presently define the sampling procedure to be used in the proof of Theorem 2.2:

Definition 2.2. *Let* $\rho > 1$ *be a parameter to be defined later. For each element* $x \in X$, *define* $p_x = \min\{\rho f(x), 1\}$, *and let* Y_x *be the random variable (indicating whether* x *is chosen) defined by* $\Pr(Y_x = 1) = p_x$, *and* $\Pr(Y_x = 0) = 1 - p_x$. *Setting* $\alpha_x = 1/p_x$, *we define a random measure* μ^* *on* X *by* $\mu^*(S) = \sum_{x \in S} \alpha_x Y_x$.

In the remaining part of the proof the goal is to show that this μ^* almost surely has the required properties. The technical details will appear in the full version.

2.2 Proof of Theorem 2.3 We start with the same partition of X into N blocks B_i as in Definition 2.1. For each $(B_i, \mathcal{F}|_{B_i})$, we apply Theorem 2.1 with $\delta = \epsilon/2N$ to produce a measure μ_i on B_i so that for every $S \in \mathcal{F}$,

$$|\, |S \cap B_i| - \mu_i(S \cap B_i)\,| \;\le\; \delta |B_i|\,.$$

Since $\mathrm{VCdim}(\mathcal{F}|_{\mathrm{B_i}}) \le \mathrm{trk}(\mathcal{F}|_{\mathrm{B_i}}) \le \mathrm{trk}(\mathcal{F})$, and $N \le \mathrm{trk}(\mathcal{F})$, the size of the support of μ_i is at most $O(\,\mathrm{trk}(\mathcal{F}) \,/\, \delta^2\,) \;=\; \mathrm{O}(\,\mathrm{trk}^3(\mathcal{F}) \,/\, \epsilon^2\,)$.

Define a measure μ^* on X by $\mu^*(S) = \sum_{i=0}^{N} \mu_i(S \cap B_i)$. We claim that μ^* is a multiplicative $\frac{1+\epsilon/2}{1-\epsilon/2}$-approximation for the counting measure on $(X, \mathcal{F})$. Indeed, let S be a set in $\mathcal{F}$, and let t be the maximal

[2] Note that, in particular, elements of weight zero are removed. We may also remove elements that do not appear in any set - this does not affect the triangular rank, as well.

index such that $S \cap B_t$ is not empty. By definition of B_i's, $|B_i| \le |S|$ for $i \le t$. Therefore,

$$|\mu^*(S) - |S|| \;\le\; \sum_{i=1}^{t} |\mu_i(S) - |S|| \;\le\; \sum_{i=1}^{t} \delta |B_i| \;\le\; \sum_{i=i}^{t} \delta |S| \;\le\; \epsilon/2 \cdot |S|\,,$$

and the claim follows. The support X^* of μ^* is the union of supports of μ_i's, and thus its size is at most $O(\,\mathrm{trk}^4(\mathcal{F})\,/\,\epsilon^2\,)$. This already establishes a dependence solely in terms of $\mathrm{trk}(\mathcal{F})$ and ϵ, but it can be further strengthened in the following manner.

Applying Theorem 2.2 to $(X^*, \mathcal{F}^*)$ with precision $\epsilon/2$, we obtain an $\frac{1+\epsilon/2}{1-\epsilon/2}$-approximation μ^{**} of μ^*. Keeping in mind that μ^* is an $\frac{1+\epsilon/2}{1-\epsilon/2}$-approximation for μ, we conclude that μ^{**} is an $\frac{1+\epsilon}{1-\epsilon}$-approximation of μ, as $(\frac{1+\epsilon/2}{1-\epsilon/2})^2 \le \frac{1+\epsilon}{1-\epsilon}$. The size of the support of μ^{**} is $O(\mathrm{trk}(\mathcal{F}^*) \log |\mathcal{F}^*| \,/\, \epsilon^2)$. Since $\mathrm{VCdim}(\mathcal{F}^*) \le \mathrm{trk}(\mathcal{F}^*) \le \mathrm{trk}(\mathcal{F})$, the Sauer Lemma (see, e.g., [24], Lemma 5.9) implies that $\mathcal{F}^*$ contains at most $\sum_{i=0}^{\mathrm{trk}(\mathcal{F}^*)} \binom{|X^*|}{i}$ distinct sets, and thus $|\mathcal{F}^*| \le |X^*|^{\mathrm{trk}(\mathcal{F})}$. Combining the estimations for $|\mathcal{F}^*|$, $|X^*|$, and for the size of the support of μ^{**}, we conclude that μ^{**} is the desired approximation of μ. ∎

2.3 Algorithmic considerations

Recall that for simplicity of presentation the original measure μ was replaced by a counting measure by passing to infinitesimal units of weight, and duplicating each element according to its weight. This corresponds to sampling each element according to the Poisson distribution with parameter $\mu(x)$. A detailed discussion of this standard issue can be found e.g., in [8].

How efficient are the above procedures? The bottleneck is the partitioning process of Definition 2.1, the key issue is the ability to find a set $S \in \mathcal{F}$ of a minimum (or even approximately minimum) weight according to the (changing) measure μ. This poses no algorithmic difficulties, e.g., when $|\mathcal{F}|$ is polynomial, as in the forthcoming applications, or when $\mathcal{F}$ is the family of cuts in graphs as in [8].

3 Multiplicative approximation via soft rank

So far we were concerned solely with approximating measures on $(\mathcal{F}, X)$. In what follows it will be more convenient work in the extended setting, when $\mathcal{F}$ is a family of functions from X to $\mathbb{R}^+$ (rather then subsets of X). The problem is, given a nonnegative linear form $L(F) = \sum_{x\in X} w_x F(x)$, to produce a sparse linear form $L^*(F) = \sum_{x\in X} w^*_x F(x)$ that for every $f \in \mathcal{F}$ it holds that $(1-\epsilon)L(f) \le L^*(f) \le (1+\epsilon)L(f)$. Clearly, when f's are restricted to take values in $\{0,1\}$, one gets the original setting.

As it was brought to our attention, the extended setting was extensively studied in the context of concrete geometrical applications, and we refer the reader to the recent [13], presenting a unifying approach and stronger results for a long list of these applications. The methods and the applications of the present paper differ significantly from those of [13], yet both work only in special cases, and the extended setting is still not well understood.

It will be convenient to restate the problem using matrix terminology. Let M be an $|\mathcal{F}| \times |X|$ (i.e., $m \times n$) real nonnegative incidence matrix of $\mathcal{F}$ vs. X, i.e., $M(S,x) = 1$ is $x \in S$ in the original setting, and $M(f,x) = f(x)$ in the extended setting. The goal is produce a nonnegative vector $w^* \in \mathbb{R}^n$ of small support, such that for every $1 \le i \le m$, it holds that $(1-\epsilon)(Mw)_i \le (Mw^*)_i \le (1+\epsilon)(Mw)_i$. We call such w^* a multiplicative $\frac{1+\epsilon}{1-\epsilon}$-approximation of w with respect to M.

The key parameter of M discussed in this section is the minimum possible rank of the (Hadamard) square root of M.

Definition 3.1. *For $0 \le \delta < 1$, define $\mathrm{rank}^*_\delta(\mathrm{M})$ as the minimum rank over all matrices X such that for all i,j, it holds that $(1-\delta)M(i,j) \le X(i,j)^2 \le (1+\delta)M(i,j)$. In particular, let $\mathrm{rank}^*(\mathrm{M}) = \mathrm{rank}^*_0(\mathrm{M})$.*

Theorem 3.1. *Let M and w be as before. Then, for any $0 < \epsilon < 1$, there exists $w^* \in \mathbb{R}^n$ that $\frac{1+\epsilon}{1-\epsilon}$-approximates w with respect to M, and has support of size $O(\,\mathrm{rank}^*(\mathrm{M})/\epsilon^2)$. Moreover, the support of w^* is contained in that of w.*

This can be easily extended to $\frac{1+\epsilon+\delta}{1-\epsilon-\delta}$-approximation of w with support of size $O(\,\mathrm{rank}^*_\delta(\mathrm{M})/\epsilon^2)$.

Observe that $\mathrm{rank}^*_\delta(\mathrm{M}) \ge \mathrm{trk}(\mathrm{M})$ for any δ. However, the soft rank may be (and typically is) very far from the "right" triangular rank. A standard tensor product argument implies a lower bound on $\mathrm{rank}^*(\mathrm{M})$ in terms of rank(M): it holds that $\mathrm{rank}^*(\mathrm{M}) \ge \sqrt{\mathrm{rank}(\mathrm{M})}$. (This bound is sometimes tight, as shown by an application in Section 4.) Moreover, it is easy to see that $\mathrm{rank}^*(\mathrm{M}) \ge \mathrm{rank}_{\mathbb{F}_2}(\mathrm{M})$. In the case of $M = J - I$, where J is the all-1 matrix, this means that $\mathrm{rank}^*(\mathrm{M}) \ge \mathrm{n} - 1$, while $\mathrm{trk}(\mathrm{M}) = 2$.

The powerful technical tool we are going to employ, (implicitly) appears in its strongest form in an important paper of Batson, Spielman and Srivastava [7]:

Theorem 3.2. *[7] Let $B_{r\times n}$ be a real valued matrix,* *and let $Q_{r\times r}$ be $Q = BB^T$. Then, for every $0 < \epsilon < 1$ there exists (and can be efficiently constructed) a non-negative diagonal matrix $A_{n\times n}$ with at most $O(r/\epsilon^2)$ positive entries, and with following property. Let $Q^* = BAB^T$. Then, for* every $x \in \mathbb{R}^n$ *it holds that:*

$$(1-\epsilon)\cdot x^TQx \;\le\; x^TQ^*x \;\le\; (1+\epsilon)\cdot x^TQx\ .$$

Actually, [7] is solely interested in the Laplacian matrices of positively weighted graphs, and the above theorem is stated there only for such Q's. However, a close examination of the proof reveals that with a minor change (related to the rank of Q) it works also for general positive semidefinite symmetric Q's.

Proof. **(of Theorem 3.1):** Let $r = \mathrm{rank}^*(\mathrm{M})$. Then, there exist real matrices $Y_{m\times r}$ and $B_{r\times n}$ such that $(YB)(i,j) = \pm M(i,j)^{\frac{1}{2}}$ for all ij,. For any non-negative $v = (v_1,\ldots,v_n) \in \mathbb{R}^n$, define $B(v)$ as the matrix obtained from B by multiplying each column j of B by $\sqrt{v_j}$. Let y_i denote the i'th row of Y, a $1\times r$ real vector. Then, for any v and each $1 \le i \le m$, it holds that

$$y_i\,B(v)B(v)^T\,y_i^T \;=\; \sum_{j=1}^{n}\Big(M(i,j)^{\frac{1}{2}}\Big)^2 v_j \;=$$

$$\sum_{j=1}^{n} M(i,j)v_j \;=\; (Mv)_i\ . \tag{3.1}$$

Applying Theorem 3.2 to the matrix $B(w)B(w)^T$, we conclude that there exists a nonnegative diagonal $n\times n$ matrix A support of size $O(\mathrm{rank}^*(\mathrm{M})/\epsilon^2)$ such that for all $1 \le i \le m$,

$$(1-\epsilon)\cdot y_i\,B(v)B(w)^T\,y_i^T \;\le\; y_i\,B(w)AB(w)^T\,y_i^T \;\le\; (1+\epsilon)\cdot y_i\,B(w)B(v)^T\,y_i^T\ .$$

However, $B(w)AB(w)^T = B(w^*)B(w^*)^T$ where $w^* \in \mathbb{R}^n$ is defined by $w^*_j = w_j\cdot A(j,j)$ for $1 \le j \le n$. Applying once more the Equation (3.1), for $v = w^*$, we conclude that w^* is the required approximation of w with respect to M. ∎

4 Dimension reduction for ℓ_1-metrics

Since ℓ_1-norm and ℓ_1-metrics naturally arise in various contexts in various algorithmic applications, there is a need for a better understanding of their structure. Unfortunately, most questions concerning ℓ_1-metrics are considerably harder than their Euclidean counterparts, and the progress is slow. One such example, that is well understood for Euclidean metrics, but much less for ℓ_1-metrics is about the dimension of their realization: Given an ℓ_1-metric μ on n points together with its geometrical ℓ_1-norm realization, one seeks to find a low-dimension geometrical realization of μ, possibly at the price of introducing a small multiplicative distortion.

The exact problem (no distortion) was resolved (for the worst case) in [3], the answer being $\Theta(n^2)$. The approximate problem is more intricate. The elegant papers [10, 18] establish polynomial lower bounds for a concrete family of hard metrics, and the recent significant improvement of [2] strengthens this to $n^{1+O(1/\log 1/\epsilon)}$. The best upper bound so far, due to Schechtman [30] (extended by Talagrand [34]) asserts that $c_\epsilon n\log n$ dimensions always suffice for $1+\epsilon$ distortion.

Recall that an ℓ_1-metric μ allows two representations: one, geometrical, is an explicit embedding of the underlying space into an ℓ_1 space. The other, combinatorial, is as a sum of *cut metrics*, i.e., $\mu = \sum_{C\in\mathcal{C}} w_C\delta_C$, where $w_C \in \mathbb{R}^+$, $\mathcal{C}$ ranges over the partitions of the underlying space, and δ_C is the cut-metric (actually, semi-metric) corresponding to C, i.e., $\delta_C(x,y) = 1$ when x,y are partitioned by C, and 0 otherwise. In what follows, will shall be interested in the latter representation.

It is natural to define a *cut dimension* of an ℓ_1-metric μ as the minimal possible number of terms in the combinatorial representation of μ. Since every cut metric can be realized in one dimension, the cut dimension is never greater than the geometric dimension. As we shall see later (see Claim 6.4), the cut dimension of an ℓ_1-metric on n points is typically $\binom{n}{2}$, which is also an upper bound.

The following theorem, the main result of this section, improves upon [30] in two directions: the upper bound is smaller, and it bounds the cut dimension rather than the geometrical dimension.

Theorem 4.1. *Let d be an ℓ_1-metric on n points, and let $0 < \epsilon < 1$ be a constant. Then there exists (and is efficiently constructible) an ℓ_1 metric d^* that distorts d by at most a multiplicative factor of $\frac{1+\epsilon}{1-\epsilon}$, and the cut-dimension of d^* is at most $O(n/\epsilon^2)$.*

Proof. We shall work with the representation of d as a weighted sum of cut metrics, i.e., $d = \sum_{C\in\mathcal{C}} w_C\delta_C$.

Let M be a $\binom{n}{2}\times(2^{n-1}-1)$ Boolean matrix whose rows are indexed by edges of K_n, the columns are indexed by nonempty cuts of K_n, and $M(e,C) = 1$ if σ belongs to the cut C, and 0 otherwise. The key observation is that although M has a full rank (as we shall see later in greater generality), its soft rank and triangular rank are significantly smaller:

Claim 4.1. $\mathrm{trk(M)} = \mathrm{rank^*(M)} = \mathrm{n}-1.$

Proof. Consider the $n-1$ edges $\{e(v_1, v_j)\}_{j=2}^n$, and the $n-1$ corresponding cuts defined by the sets $\{v_2, \dots v_j\}_{j=2}^n$. Clearly, the corresponding minor of M is lower-triangular with 1's on the diagonal. Thus, $\mathrm{trk(M)} \geq \mathrm{n-1}$. For the other direction, since $\mathrm{trk(M)} \leq \mathrm{rank^*(M)}$, it suffices to show that $\mathrm{rank^*(M)} \leq \mathrm{n-1}$. Let Y be a matrix whose rows and indexed by edges, the columns are indexed by vertices, and for an arbitrarily oriented edge $e = (v, u)$, let $Y(e, v) = 0.5$, $Y(e, u) = -0.5$ and $Y(e, w) = 0$ otherwise. The matrix B is indexed by vertices vs. cuts, an for an arbitrarily oriented cut $C = (U, V-U)$, $B(v, C) = 1$ if $v \in U$, and $B(v, C) = -1$ otherwise. It is easily checked that $YB = \pm M$, and that $\mathrm{rank(YB)} \leq \mathrm{rank(Y)} = \mathrm{n-1}$. ∎

Interpreting each column of M as a cut-metric, Mw stands for a weighted sum of cut metrics, and $(Mw)(u, v) = d(u, v)$. The problem thus reduces to finding a multiplicative $\frac{1+\epsilon}{1-\epsilon}$-approximation w^* with a small support, which is readily done using the general tools developed in the previous sections. In particular, employing Theorem 2.2, we get an approximation of support $O(\mathrm{trk(M)} \log \mathrm{n}/\epsilon^2) = \mathrm{O}(\mathrm{n} \log \mathrm{n}/\epsilon^2)$ which matches the bound of [30]. Employing Theorem 3.1 leads to an improved bound of $O(\mathrm{rank^*(M)}/\epsilon^2) = \mathrm{O}(\mathrm{n}/\epsilon^2)$.

Both procedures take as an input the original representation of d as $d = \sum_{C \in \mathcal{C}} w_C \delta_C$, and work with $M|_{\mathcal{C}}$, i.e., only with the relevant columns. Thus, the running time is polynomial in the length of the input representation. ∎

One may wonder how tight is the bound of Theorem 4.1. As the following theorem shows, in terms of the dependence in n it is best possible. The dependence on ϵ is left for the future study.

THEOREM 4.2. *Let d be the metric on $\{1, 2, \dots, n\}$ defined by $d_n(i, j) = |10^i - 10^j|$. This is certainly an ℓ_1 metric. However, any metric $d' = \sum_{C \in \mathcal{C}'} \lambda_C \cdot \delta_C$ where $|\mathcal{C}'| \leq n-2$ distorts d by at least a factor of 9.*

5 Half-way discussion

Our starting point in the forthcoming study is Claim 4.1. Simple as it is, it was crucial in making possible the application of the approximation techniques of the previous sections. We start a generalization of this claim, aiming at extending the range of applicability of the approximation techniques. This generalization will be the basis for things to come.

5.1 co-Circuits in Matroids.

Keeping in mind that cuts are the co-circuits of the graphic matroids, we present here the generalization of Claim 4.1 to arbitrary matroids. The basic notions needed for the discussion are bases, circuits and co-circuits, all subsets of the underlying space X equipped with an abstract dependence structure satisfying some axioms. The bases are the maximal independent (equivalently, minimum spanning) sets over X. They all have the same size, called the *rank* of the matroid. The subsets of the bases (and only they) are also independent. Circuits are the minimal dependent sets. Co-circuits intersect every base, and are minimal with respect to this property.

In graphic matroids the elements are the edges, the bases are the trees, the circuits are the simple cycles, and the co-circuits are the cuts.

In linear matroids (the only type of matroids to be used) the elements are a subset $X \subseteq \mathbb{F}^n$, the bases are the maximal linearly independent subsets of X, the circuits are the minimal linearly dependent subsets of X, and the co-circuits C are the complements of the maximal non-spanning subsets in X. I.e., $X \not\subseteq \mathrm{span}(\overline{\mathrm{C}})$, but adding any element from C to $\overline{C}$ makes it spanning X. Equivalently, $C \subset X$ is a co-circuit iff there exists a linear function $L : \mathbb{F}^n \mapsto \mathbb{F}$ such that $\ker(L) \cap X = \overline{C}$, and $L|_X$ is uniquely defined (up to scaling).

LEMMA 5.1. *Let M be the incidence matrix of the underlying space X vs. the co-circuits of a matroid $\mathcal{M}$. Then,* $\mathrm{trk(M)} \leq \mathrm{rank}(\mathcal{M})$.

Proof. The only standard fact (and a simple exercise) about matroids to be exploited is that the size of the intersection of any circuit and co-circuit cannot be 1. Let Q be a square $N \times N$ lower triangular minor of M with 1's over the diagonal. Let the rows be indexed by $\{x_i\}_{i=1}^N$ and let the columns be indexed by $\{C_i\}_{i=1}^N$ respectively. It means, in particular, that $x_i \in C_i$, but $x_i \notin C_j$ for $j > i$. We claim that the set of elements $\{x_i\}_{i=1}^N$ does not contain cycles. Indeed, assume by contradiction that it does contain a cycle Z, and let r be the largest index such that $x_r \in Z$. Consider the corresponding co-circuit C_r. Since $x_r \in Z \cap C_r$, by the above fact, C_r must contain another element x_i from Z, $i < r$, contrary to the observation that $x_i \notin C_r$ for every $i < r$. Thus, $\{x_i\}_{i=1}^N$ is acyclic, implying that $N \leq \mathrm{rank}(\mathcal{M})$. ∎

In the second part of the paper we will study *simplicial matroids*, a natural generalization of the graphical matroids to higher dimensions. Using the most basic tools and notions of Combinatorial Topology, we shall study the structure of co-circuits in this setting, develop the theory of ℓ_1-volumes, and obtain the higher-dimensional analogue of Th. 4.1. This, in turn, will be used to obtain new results about Euclidean volumes. Let us describe these results here.

5.2 Additional applications For readers reluctant to delve into Combinatorial Topology, we state here the two purely geometrical implications of the forthcoming discussion. The proofs are delayed until Section 6.

The first application deals with estimating the average of an Euclidean d-dimensional volume. Assume that a finite set of 'points', V, is embedded in an arbitrary Euclidean space, and we are given access (either by getting the actual embedding, or by a black box) to the values of the induced Euclidean volumes of d-simplices over V. The goal is to estimate the average of volumes of all d-simplices on X. The trivial algorithm runs in time $O(n^{d+1})$. For $d = 1$ this was improved in [4] (in turn improving upon an earlier result of P. Indyk) to a linear time deterministic nonadaptive algorithm that makes queries about the distances between $O(n)$ predefined pairs, and outputs their average. We design a similar type $O(n^d)$-algorithm for Euclidean d-volumes.

THEOREM 5.1. *In order to $\frac{1+\epsilon}{1-\epsilon}$-approximate the average value of an Euclidean volume on V, it suffices to query $O(n^d/\epsilon^2)$ predefined $(d+1)$-tuples, and output a (predefined) linear combination of these queries.*

The second application is as follows.

THEOREM 5.2. *Let S be a set of n points in the plane. There exists a weighted sampling set Q containing at most $O(n^2/\epsilon^2)$ points of $\mathbb{R}^2$, such that the area enclosed by any non self-intersecting polygon P with vertices in S is $\frac{1+\epsilon}{1-\epsilon}$-approximated by the sum of weights of the points of Q enclosed by P.*

6 Finite Volume Spaces

6.1 Preliminaries

6.1.1 Basic definitions and facts from Combinatorial Topology Fix $\mathbb{F}$ to be a field; we shall always be working either over $\mathbb{R}$ or over $\mathbb{F}_2$. Let V be an underlying set of size n and let $K_n^{(d)} = \{\sigma \subseteq V \mid |\sigma| = d+1\}$ be the set of all d-dimensional simplices on V. It will be convenient to associate $\mathbb{F}$-weighted (formal) sums of d-simplices (called *d-chains*) with $\binom{n}{d+1}$-dimensional vectors of the corresponding weights.

Each simplex is either positively or negatively oriented (over $\mathbb{R}$; there are no orientations over $\mathbb{F}_2$), and it induces (in a standard manner described in details e.g., in [28]) the orientation of its subsimplices. The key feature of these orientations, and in fact its defining property, reflecting the structure of the underlying topological space, can be formulated as follows. Let M_d be the $\binom{n}{d} \times \binom{n}{d+1}$ incidence matrix, whose rows are indexed by (arbitrarily oriented) $(d-1)$-simplices, the columns are indexed by (arbitrarily oriented) d-simplices, and $M_d(\tau, \sigma) = 1$ if $\tau \subset \sigma$ and its orientation is consistent with the orientation induced by σ on its boundary, $M_d(\tau, \sigma) = -1$ if $\tau \subset \sigma$ but the orientations are inconsistent, and $M_d(\tau, \sigma) = 0$ if $\tau \not\subset \sigma$. Then, provided that the d-simplices indexing the two matrices are oriented consistently, it miraculously holds that $M_{d-1}M_d = 0$.

The right action of M_d can be interpreted as a mapping of weighted sums of d-simplices (i.e., d-chains) to weighted sums of $(d-1)$-simplices (i.e., $(d-1)$-chains), and is called ∂, the *boundary operator*. The left action of M_d can be interpreted as a mapping of weighted sums of $(d-1)$-chains to d-chains, and is called ∂^*, the *coboundary operator*.

A d-chain Z in the kernel of ∂_d is called a *d-cycle*. A d-chain B in the image of ∂^*_{d-1} is called a *d-coboundary*. The definitions immediately imply that the space of d-coboundaries (seen as vectors of weights) is precisely the dual of the space of d-cycles, i.e., every vector corresponding to a d-coboundary sums up to 0 on every d-cycle. A deeper fact is that $M_{d-1}M_d = 0$, and hence $\partial_{d-1}\partial_d = 0$ and $\partial^*_{d-1}\partial^*_d = 0$. Moreover, $\operatorname{Ker}\partial_{d-1} = \operatorname{Im}\partial_d$, and $\operatorname{Ker}\partial^*_d = \operatorname{Im}\partial^*_{d-1}$.

6.1.2 More Combinatorial definitions and facts We introduce here some non-standard suggestive terminology, aimed at stressing the analogy between graphs and high dimensional simplicial complexed, and facilitating the discussion. The underlying field $\mathbb{F}$ is, as before, either $\mathbb{F}_2$ or $\mathbb{R}$.

DEFINITION 6.1. *A set $S \subset K_n^{(d)}$ of d-simplices will be called* independent *if the corresponding $(d-1)$-chains $\{\partial_d\sigma \mid \sigma \in S\}$ are linearly independent over $\mathbb{F}$. Again, it might be more convenient to view the chains as the coefficient vectors.*
A (spanning) hypertree *$T \subset K_n^{(d)}$ is a maximal independent set.*
A set $K \subset K_n^{(d)}$ is homologically connected [3] *or just* connected *if it contains a hypertree.*
A hypercycle *$Z \subset K_n^{(d)}$ is a minimal dependent set.*
A hypercut *$C \subset K_n^{(d)}$ is a complement of a maximal nonspanning set, i.e.,* $\operatorname{span}\{\partial_d\sigma \mid \sigma \notin C\}$ *has codimension 1 with respect to* $\operatorname{Im}\partial_d$, *but moving any simplex from C to $\overline{C}$ makes the latter fully dimensional.*

The above definitions are fully consistent with the corresponding notions of the linear matroid whose elements

[3] From the topological perspective, such K should be treated as a simplicial complex containing, in addition to its part in $K_n^{(d)}$, all the lower dimensional simplices over V.

are associated with the columns of M_d. In particular, the hypertrees correspond to bases, and the hypercuts correspond to cocircuits. All d-hypertrees are of the same size, the rank of M_d. Since the set of all d-simplices containing a fixed vertex v of V is independent (as each corresponding boundary contains a $(d-1)$-simplex unique to it), and any other simplex is obviously spanned by these simplices, the size of any d-hypertree is $\binom{n-1}{d}$.

The following lemma summarizes the relations between hypercuts, hypertrees and coboundaries. The proofs are purely technical and use only elementary linear algebra (or, alternatively, for the first two items, elementary matroid theory). They are omitted from this preliminary version.

Lemma 6.1.

1. *Let T be a d-hypertree, and $\sigma \in T$. Then there exists a* unique *d-hypercut $C_{T,\sigma}$ such that $T \cap C_{T,\sigma} = \sigma$. More explicitly, $C_{T,\sigma}$ is the set of all the d-simplices τ such that the* unique *cycle Z created by adding τ to T, contains σ.*

2. *Let $\mathcal{C}$ be the set of d-hypercuts and let $\mathcal{T}$ be the set of d-hypertrees. Then, $\mathcal{C}$ is the* blocker *of $\mathcal{T}$, $\mathcal{C} = \mathcal{T}^B$. That is, every hypercut intersect every hypertree (and hence any connected set), and any set $S \subseteq K_n^{(d)}$ with this property that is minimal (with respect to containment) is a hypercut.*

3. *Over $\mathbb{F}_2$, any d-hypercut C is a d-coboundary* [4] *minimal with respect to containment.*

4. *Over $\mathbb{R}$, a set $C \subset K_n^{(d)}$ is a d-hypercut if and only if there exists a unique (up to scaling) real d-coboundary $\sum_{\sigma\in C} \beta_\sigma \sigma$ whose support is precisely C. The β's are defined by the equations $\forall\ i, j,\ \partial(\beta_{\sigma_i}\sigma_j - \beta_{\sigma_j}\sigma_i) \in \mathrm{span}\{\partial\tau,\ \tau \notin \mathrm{C}\}$.*

Here is perhaps the place to observe that every $\mathbb{F}_2$-hypertree is necessarily a $\mathbb{R}$-hypertree, but not the other way around. This is because the equation involved in definition of these terms are equation over vectors with entries $0, \pm1$, and thus a dependence over $\mathbb{R}$ implies a dependence over $\mathbb{F}_2$.

A very special class of d-hypercuts are the *geometrical hypercuts*, defined as follows:

Definition 6.2. *Let S^{d-1} be the unit sphere of dimension $d-1$ and $\phi : V \mapsto \mathrm{S}^{d-1}$ be a mapping such that the points in the image are in the general position (no nontrivial linear and affine dependencies). The set of d-simplices whose image under ϕ contains the origin will be called a geometric d-hypercut*[5].

A familiar, due to their use in the celebrate Sperner Lemma (see, e.g., [28]), subfamily of geometrical hypercuts are the *partition hypercuts*. They are a weaker generalization of the graphical cuts. Let $\mathcal{P} = \{V_1\cup, \ldots, V_{d+1}\}$ be a partition of V to $(d+1)$ disjoint nonempty parts. The corresponding partition hypercut is defined by $C_{\mathcal{P}} = \{\sigma \in K_n^{(d)}\ |; |\sigma \cap A_i| = 1, i = 1, 2, \ldots, d+1\}$.

Lemma 6.2. *Geometric d-hypercuts are indeed d-hypercuts both over $\mathbb{R}$ and over $\mathbb{F}_2$.*

Proof. **(Sketch)**: Let C be a geometric hypercut, and let ϕ be the corresponding geometric realization. Orient all d-simplices in $K_n^{(\leq d)}$ according to the positive orientation of $\mathbb{R}^d$. I.e., left to right for $d = 1$, counter-clockwise for $d = 2$, etc. (See e.g., [28] for the precise definition.) It will suffice to show that the chain B_C that assigns weight 1 to $\sigma \in C$, and 0 otherwise, is a minimal d-coboundary.

Showing that B_C is a real coboundary can be done either directly, using some calculus, or by a cleaner but less intuitive dual argument. For example, for $d = 2$, it is well known that for a 2-dimensional triangle D, the path-integral over the counter-clockwise oriented 1-dimensional boundary of D, $\frac{1}{2\pi}\int_{\partial D} x/(x^2+y^2)\,dy - y/(x^2+y^2)\,dx$ is 1 if D contains the origin, and 0 otherwise. Assigning to the oriented 1-simplices in $K_n^{(1)}$ the values of the above path integral on their realization, we obtain a 1-chain that under the action of ∂^* yields B_C.

In the dual setting, applying both to $\mathbb{R}$ and to $\mathbb{F}_2$, to conclude that B_C is coboundary, it suffices to show that it sums up to 0 on the oriented (as above) boundary of any $(d+1)$-simplex ζ. Observe that the origin is contained in either zero or two d-simplices belonging to the boundary of ζ. [6] In the former case we are done; in latter case, one of these simplices is necessarily oriented in a manner consistent with our orientation of $\mathbb{R}^d$, while the other is oriented inconsistently. Therefore, B_C sums up to 0 on $\partial\zeta$.

To show that B_C is minimal, is will suffice to show that for any $\sigma_{,}\sigma_j \in C$, it holds that $\partial(\sigma_i - \sigma_j) \in \mathrm{span}\{\partial\sigma,\ \sigma \notin \mathrm{C}\}$. Assume first that the two simplices are disjoint. We use the following cylindric construction. Consider two parallel copies of $\mathbb{R}^d$ in $\mathbb{R}^{d+1}$, each containing S^{d-1} with the ϕ-image of V.

[4] A d-chain is formally not a set of simplices, but over $\mathbb{F}_2$ there is an obvious bijection.

[5] That is, a simplex σ is in the hypercut defined by ϕ if $conv(\phi(V(\sigma)))$ contains the origin.

[6] For an interesting characterization of subsets of $K_n^{(2)}$ with this property over $\mathbb{F}_2$, see [14].

Choose σ_i from first copy, and σ_i from the second copy. Then, by the general position argument, the boundary of the $\text{conv}(\sigma_i \cup \sigma_i) \subset \mathbb{R}^{d+1}$ is triangulated by d-simplexes. For every d-simplex in this triangulation, consider the corresponding abstract simplex in $K_n^{(d)}$. An easy projection argument implies that all the simplices resulting from the lateral d-simplices in the above triangulation (i.e., all but σ_i and σ_j) are in $\overline{C}$. Since the chain corresponding to the boundary of convex polytope is a real (and hence an $\mathbb{F}_2$-) d-cycle, the statement follows.

If the two simplices σ_i and σ_j are not disjoint, we make the two copies of $\mathbb{R}^d$ intersect, such that all the common vertices (and only them) lie in the intersection, and proceed in same manner. ■

6.2 Volumes over $\mathbb{R}$ We start with the concise exposition of the more analytic theory of real volumes in order to get faster to the application. In the next section we shall treat the more combinatorial theory of volumes over $\mathbb{F}_2$ at a more leisurely pace, putting stress on its structural and combinatorial aspects. Every real volume will be an $\mathbb{F}_2$-volume, but not vice versa.

Let $K_n^{(\leq d)}$ be the simplicial complex on the underlying set V of size n containing all the simplices of dimension $\leq d$ on V. An abstract d-dimensional volume function $\mathrm{vol^{(d)}} : \mathrm{K_n^{(\leq d)}} \mapsto \mathbb{R}^+$ is a real nonnegative function with the following properties: (*) the simplices of dimension $< d$ have value 0; (**) the values of d-simplices satisfy the following generalization of the triangle inequality:

For every d-simplex σ and real d-cycle $Z = \sigma + \sum \alpha_i \sigma_i$, it holds that,

$$\mathrm{vol^{(d)}}(\sigma) \leq \sum |\alpha_\mathrm{i}| \, \mathrm{vol^{(d)}}(\sigma_\mathrm{i}). \tag{6.2}$$

We note that for $d > 1$, unlike the one-dimensional case when a triangle inequality suffices, condition (**) cannot be replaced by a requirement on cycles of bounded size.

The most natural class of the volume functions are the *Euclidean* volumes: given an embedding ϕ of V into an Euclidean space, the volume of a d-simplex σ, is the Euclidean d-volume of $\text{conv}(\phi(\sigma))$. As we shall see soon, Euclidean volumes are indeed volumes according to the above definition.

The notion of real *integral*[7] d-volume is a natural higher-dimensional generalization of a line metric:

Definition 6.3. *1. An absolute values of the weights of a real d-coboundary of $K_n^{(d)}$ will be called an* integral *volume.*

2. The absolute values of the weights of a real d-hypercut (as in the 4th item in Lemma 6.1) will be called a hypercut *volume, a special case of an integral volume.*

3. The convex combinations of integral d-volumes will be called (by analogy with the 1-dimensional case) the ℓ_1-volumes, and the convex cone formed by them will be called the (real) d-hypercut cone.

4. Sums of (pointwise) squares of integral volumes will be called functions of negative type.

Integral d-volumes are indeed d-volumes. Let $B = \beta\sigma + \sum_i \beta_i \sigma_i$ be a real d-coboundary. Let $Z = \sigma + \sum \alpha_i \sigma_i$ be a d-cycle. Since the weights of coboundaries sum up to 0 on cycles, it holds that $\beta + \sum \beta_i \alpha_i = 0$, and hence $|\beta| \leq \sum |\beta_i||\alpha_i|$.

Theorem 6.1. *The extremal rays of the d-hypercut cone are precisely the d-hypercut volumes.*

Integral d-volumes are intimately related to integrals of $d-1$-forms on simplices embedded in $\mathbb{R}^d$, hence their name.

Lemma 6.3.
1. *Euclidean d-volumes realizable in $\mathbb{R}^d$ are integral d-volumes.*
2. *Euclidean d-volumes (not necessarily realizable in $\mathbb{R}^d$) are ℓ_1, and moreover, they are nonnegative combinations of geometric d-hypercut volumes.*
3. *Geometric d-hypercut volumes and (pointwise) squares of Euclidean d-volumes are of negative type.*

Proof. [of Lemma 6.3] **(Sketch):** We start with the first statement. Consider a realization of $K_n^{(d)}$ in $\mathbb{R}^d$ defining an Euclidean d-volume vol. Orient all d-simplices according to the positive orientation of $\mathbb{R}^d$. We claim that vol, seen as a real-valued weighting of thus oriented d-simplices is a real d-boundary. Observe that vol is the integral over $\mathbb{R}^d$ of $B_{C(p)}$'s, where $C(p)$ is the geometrical hypercut defined by treating the point $p \in \mathbb{R}^d$ as the origin with respect to the above realization of $K_n^{(d)}$, and $B_{C(p)}$ is the corresponding coboundary, as in the proof of Lemma 6.2. Since the d-coboundaries are closed under addition, the statement follows.

For the second statement, it suffices to take the realization of $K_n^{(d)}$ in $\mathbb{R}^n$, and consider its projection on the random d-dimensional subspace. Clearly, the expected Euclidean volume of the projection of any d-dimensional simplex is proportional to its original volume. Thus, any Euclidean d-volume is a weighted sum of Euclidean d-volumes realizable in $\mathbb{R}^d$, and by

[7] From "integration", not from "integer".

the arguments used to prove the first statement we are done.

For the third statement, the correctness for the geometric d-hypercut volumes follows at once. For Euclidean volumes, by a corollary to Cauchy-Binet formula, the square of the volume of a d-simplex $\sigma \in \mathbb{R}^n$ is the sum of the squares of d-volumes of σ's projections on all subsets of k coordinates. Thus, the statement reduced to $n = d$, where it follows directly from the first statement. ∎

6.3 Applications We start with proving the following lemma:

LEMMA 6.4. *Let N be a real $\binom{n}{d+1} \times |\mathcal{F}|$ matrix with rows indexed by d-simplices $\sigma \in K_n^{(d)}$, and the columns indexed by functions $f \in \mathcal{F}$ of nonnegative type on $\sigma \in K_n^{(d)}$. Then,* $\mathrm{rank}^*(\mathrm{N}) \leq \binom{\mathrm{n}-1}{\mathrm{d}}$.

Proof. By definition of a function of nonnegative type, for every $f \in \mathcal{F}$ there exists a vector $x_f \in \mathbb{R}^{\binom{n}{d}}$ such that the f-column of N is equal to the (pointwise) square of the vector $x_f^T M_d$. Forming the matrix X whose raws are $\{x_f^T\}_{f \in \mathcal{F}}$, we conclude that (pointwise) $\sqrt{N} = \pm X^T M_d$. Hence, $\mathrm{rank}^*(\mathrm{N}) \leq \mathrm{rank}(\mathrm{M_d}) = \binom{\mathrm{n}-1}{\mathrm{d}}$. ∎

Let vol be a weighted sum of geometric hypercut d-volumes. Lemma 6.4 together with Theorem 3.1 imply that in this case vol can be multiplicatively $\frac{1+\epsilon}{1-\epsilon}$-approximated by a weighted sum of at most $\binom{n-1}{d}$ geometric hypercut d-volumes from the original sum.

We are now ready to address Theorem 5.2 stated in Section 5.2.

Proof. (**of Theorem 5.2**): Since planar polygons can be triangulated, it suffices to address the case when P is a triangle. Putting in the interior of each of $O(n^4)$ cells created by the line segments connecting the vertices of P, a point p with the associated weight w_p being the area of the cell, we get an initial sampling set that does not produce any errors, but it is too big. Associating with each such point p a geometrical hypercut 2-volume, the above discussion implies the desired construction. ∎

Next, we proceed towards establishing Theorem 5.1 stated in Section 5.2.

THEOREM 6.2. *For every nonnegative weighting w of $K_n^{(d)}$ define a linear form $F_w(f) = \sum_{\sigma \in K} w(\sigma) f(\sigma)$. Then, there exists (and is efficiently computable) a weighting w^* with support of size at most $O(\binom{n-1}{d}/\epsilon^2)$ such that for any (!) function f of nonnegative type on $K_n^{(d)}$, it holds that $(1-\epsilon)F_w(f) \leq F_{w^*}(f) \leq (1+\epsilon)F_w(f)$.*

Observe that the bound on the support is essentially tight in this generality, as sampling less than $\binom{n-1}{d}$ simplices (potentially) allows to predict only the values of the simplices spanned by this set, the rest remaining completely free. Thus, if the support of w is connected, the value of $F_w(f)$ cannot be approximated at all by a sample of non-connected subset of simplices, and hence by any small sample.

Proof. A proof based on Lemma 6.4 applied to N^T is quite natural here, but we prefer the original argument of [7] on which the latter theorem is based. Keeping in mind that the functions of nonnegative type are nonnegative combinations of (entrywise) squares of real d-coboundaries, it suffices to establish the statement for squares of real d-coboundaries.

Recall that a real d-coboundary $B_x \in \mathbb{R}^{\binom{n}{d+1}}$ is defined by a vector $x \in \mathbb{R}^{\binom{n}{d}}$ by $B_x^T = x^T M_d$, where M_d is the real incidence matrix as in Section 3. Thus, $F_w(B_x^2) = x^T(M_d W M_d^T)x$, where W is a diagonal $\binom{n}{d+1} \times \binom{n}{d+1}$ matrix indexed by d-simplices, in which $W(\sigma,\sigma) = w(\sigma)$. Applying Theorem 3.2 to the matrix $M_d W M_d^T = (M_d\sqrt{W}) \cdot (\sqrt{W} M_d^T)$ we conclude that there is another weighting w^* such that $|\mathrm{supp}(\mathrm{w}^*)| = \mathrm{O}(\mathrm{rank}(\mathrm{M_d})/\epsilon^2)$, and $x^T(M_d W M_d^T)x$ and $x^T(M_d W^* M_d^T)x$ differ by at most $(1 \pm \epsilon)$ multiplicative factor. Since $\mathrm{rank}(\mathrm{M_d}) = \binom{\mathrm{n}-1}{\mathrm{d}}$, we arrive at the desired conclusion. ∎

Since Euclidean volumes are by Lemma 6.3 of nonnegative type, this implies Theorem 5.1.

In the case of uniform weights, the constructed approximation weightings are high dimensional analogues of the *sparsifiers* from [7], which in turn are a slightly relaxed version of expanders. We feel that the structure of these special weighting is quite intriguing, they are potentially useful, and certainly they deserve further study. The expansion in simplicial complexes will reoccur also in the next section in a different context.

6.4 Volumes over $\mathbb{F}_2$ In this section we develop the theory of finite $\mathbb{F}_2$-volumes over $\mathbb{F}_2$. While for $d = 1$ the finite real and $\mathbb{F}_2$-volumes coincide, for larger dimensions the latter form a strictly larger family of functions. It appears that in many respects, the theory of hypercuts over $\mathbb{F}_2$, and in particular, of d-volumes over $\mathbb{F}_2$ provides a meaningful and even exciting generalization of the corresponding graph theoretic and metric theoretic notions. For example, one naturally arrives at a meaningful generalization of graph expansion, and, moreover, using this notion together with the generalization of Poincare forms, one can prove that certain volumes are hard to approximate by ℓ_1 volumes, much as in the metric theoretic case. We shall discuss this and other issues, without restricting ourselves to the matters directly related to sparsification.

Working over $\mathbb{F}_2$, it will be convenient to treat the d-chains simply as sub-complexes of $K_n^{(d-1)}$. This approach will be adopted throughout this section. Often, the situation for $d = 2$ is clearer than for higher dimensions, and the discussion will focus mostly on this case.

6.4.1 The Structure of Hypercuts Here we present some non-standard combinatorial notions and results to be used later in this section. The proofs are mostly omitted from this version.

In the $\mathbb{F}_2$ framework, the general theory directly implies the following facts.

Claim 6.1. *The incident vector of a d-coboundary B is of the form* $1_B^T = 1_G^T M_d$, *where* $G \subseteq K_n^{(d-1)}$. *The intersection of any d-cycle Z and d-coboundary B is always odd. Finally, for any d-hypercut C and any* $\sigma, \sigma' \in C$, *there is a d-cycle Z such that* $Z \cap C = \{\sigma, \sigma'\}$.

For $X \subseteq K_n^{(d)}$ and v a vertex of X, define the *link* of X with respect to v to be the following $(d-1)$-dimensional subcomplex of X:

$$\mathrm{link_v(X)} = \{\tau \in \mathrm{K_n^{(d-1)}} \mid \mathrm{v} \notin \tau \text{ and } \{\tau \cup \mathrm{v}\} \in \mathrm{X}\}.$$

Claim 6.2. *Let B be d-coboundary. Then, B is induced by* $\mathrm{link_v(B)}$, *namely* $1_B^T = 1_{\mathrm{link_v(B)}}^T M_d$. *Consequently, there is a 1-1 correspondence between the* $(d-1)$-*dimensional* G_{d-1}*'s on* $V - \{v\}$, *and the d-coboundaries* $B \subseteq K_n^{(d)}$.

Next, we characterize these 1-dimensional complexes (that is, graphs) G, so that $1_G^T M_2$ corresponds to a 2-hypercut (rather then just a coboundary).

Let $G = (V, E)$ be a graph. Call two adjacent edges (u, v), $(u, w) \in E(G)$ $\wedge$-*equivalent* if $(v, w) \notin E(G)$. I.e., the restriction of G to $\{u, v, w\}$ is a path of length 2 (namely a $\wedge$) with u at the middle. Taking the transitive closure of this relation, we call G $\wedge$-*connected* if any two edges of G are $\wedge$-equivalent.

Theorem 6.3. *Let B be a 2-coboundary, and let* $G = \mathrm{link_v(B)}$ *be its link with respect to an arbitrary vertex* v. *Then, B is a 2-hypercut iff G is* $\wedge$-*connected.*

Let us comment that a random graph G on $n-1$ vertices is almost surely $\wedge$-connected. (This is an easy exercise and we leave it to the reader.) Thus, in view of the above theorem, there are $2^{\Theta(n^2)}$ different 2-hypercuts.

Having characterized the hypercuts, we turn to the study of the distribution of their sizes. The first question is how large/small can a d-hypercut be? A partial answer is provided by the following claim.

Claim 6.3. *The size of the minimum (nonempty) d-hypercut in* $K_n^{(d)}$ *is* $n - d$. *The size of the maximum 2-hypercut is* $\binom{n}{3} - O(n^2)$.

Next result is about lower tail of the distribution of the sizes of d-hypercuts in $K_n^{(d)}$, in particular when $d = 2$. It should be noted that a similar but weaker result was shown earlier in [20] employing a somewhat more involved argument.

Theorem 6.4. *The number of d-hypercuts of size* αn *is at most* $n^{c_d \cdot \alpha}$ *where* c_d *can be (very roughly) upper-bounded by* $d(d+1)$. *For* $d = 2$ *we show a better upper bound of* $(4n)^{3\alpha+1}$.

Proof. Since $|C| = \alpha n$, the average size of $|\mathrm{link_v(C)}|$ is $(d+1)\alpha$, and therefore there exists a vertex v such that $|\mathrm{link_v(C)}| \leq (\mathrm{d}+1)\alpha$. Thus, $|C|$ is induced by G of size at most $(d+1)\alpha$. However, setting $m = \binom{n}{d}$, the number of such G's is at most $\binom{m}{(d+1)\alpha} = O(n^{d(d+1)\alpha})$. For $d = 2$ we know that G is $\wedge$-connected, hence it has at most one non trivial component containing at most 3α edges and $3\alpha + 1$ vertices. Thus, the number of such G's is at most

$$\binom{n}{3\alpha+1}\binom{\binom{3\alpha+1}{2}}{3\alpha} \leq$$

$$\left(\frac{en}{3\alpha+1}\right)^{3\alpha+1} \cdot \left(\frac{e \cdot 3\alpha(3\alpha+1)}{2 \cdot 3\alpha}\right)^{3\alpha} \leq (4n)^{3\alpha+1}$$

∎

6.4.2 Volumes: Basic Definitions Volumes over $\mathbb{F}_2$ are defined as non-negative real functions on $K_n^{(\leq d)}$, analogously to the definition of real volumes in Section 6.2:

For every d-simplex σ and d-cycle $Z = \sigma + \sum \sigma_i$ over $\mathbb{F}_2$ it holds that,

$$\mathrm{vol}(\sigma) \leq \sum \mathrm{vol}(\sigma_\mathrm{i}) \tag{6.3}$$

An important example of a volume function is the generalization of the shortest-path metric. Let $X \subseteq K_n^{(d)}$ be a connected (i.e., containing a d-hypertree) subcomplex with nonnegative weights on its d-simplices. The *lightest-cap* (called also *minimum filling*) volume $\mathrm{vol_X}$ induced by X on $K_n^{(d)}$ is defined by $\mathrm{vol_X} =$

$\min_{D_\sigma \subseteq X} \sum_{\sigma' \in D_\sigma} w_{\sigma'}$, where D_σ is a *σ-cap*, i.e., $\sigma \cup D_\sigma$ is a cycle. (In particular, σ itself is σ-cap.)

Another example are *hypercut volumes*, which in analogy to hypercut volumes over $\mathbb{R}$, and to cut metrics, are defined as follows. Let C be a d-hypercut in $K_n^{(d)}$ over $\mathbb{F}_2$. The corresponding volume function $\mathrm{vol}_C^{(d)}$ assigns 1 to every $\sigma \in C$, and 0 to every $\sigma \notin C$. To see that a hypercut volume is indeed a volume, it suffices to notice that a 0/1 function on d-simplices may fail to be a volume function iff there exists a cycle Z were all but one $\sigma \in Z$ have value 0. By Claim 6.1, such Z does not exist for $\mathrm{vol}_C^{(d)}$.

Since real d-cycles are a subset of $\mathbb{F}_2$ d-cycles, we conclude that the $\mathbb{F}_2$ d-volumes are contained in the real ones.

As with real volumes, volume functions over $\mathbb{F}_2$ on V are closed under addition and multiplication by a constant, and thus form a cone in $\mathbb{R}_+^{\binom{n}{d+1}}$. The extremal volumes in this cone are, as always, of particular interest. The following theorem provides a full characterization of 0/1 extremal volumes. Perhaps more important, it also establishes their inapproximability by any other metric.

The *multiplicative distortion* between two d-volume functions vol_1 and vol_2 on V is defined similarly to the metric distortion, i.e.,

$$\mathrm{dist}(\mathrm{vol}_1, \mathrm{vol}_2) \;=\; \max_\sigma \frac{\mathrm{vol}_1(\sigma)}{\mathrm{vol}_2(\sigma)} \cdot \max_\sigma \frac{\mathrm{vol}_2(\sigma)}{\mathrm{vol}_1(\sigma)}\,.$$

THEOREM 6.5. *A 0/1 volume function $\mathrm{vol}^{(d)}$ is extremal iff it is a hypercut volume. Moreover, the distortion between such $\mathrm{vol}^{(d)}$ and any other volume function $\mathrm{vol}_1^{(d)}$ is infinite unless $\mathrm{vol}_1^{(d)} = \alpha \cdot \mathrm{vol}^{(d)}$ for some positive constant α.*

Much of the modern theory of finite metric spaces is devoted to the study of special metric classes that constitute a sub-cone of the metric cone, notably ℓ_1 metrics and *NEG*-type metrics. Crucially for applications, any metric on n points can be approximated by a special metric with a bounded distortion c_n. E.g., for ℓ_1 the rough bound of $O(n)$ on distortion follows from the minimum spanning tree argument, and the much better $O(\log n)$ bound is implied by Bourgain's Theorem [9]. Theorem 6.5 implies that any (closed) sub-cone of volume functions with the approximation property *must* contain the cone spanned by the hypercut volumes. Moreover, as we shall soon see, this cone already has the required property. This justifies the following definition.

DEFINITION 6.4. *Analogously to the one dimensional case, we define ℓ_1 d-volumes to be the nonnegative combinations of hypercut d-volumes.*

Clearly, ℓ_1 d-volumes constitute a sub-cone of d-volumes.

6.4.3 ℓ_1 Volumes The most basic properties of ℓ_1 metrics are that they contain the class of tree-metrics and the class of Euclidean metrics. The situation with ℓ_1 d-volumes turns out to be fully analogous.

Euclidean d-volumes are sums of the geometrical hypercut volumes. Since geometrical hypercuts are $\mathbb{F}_2$ hypercuts, we conclude that the Euclidean volumes are ℓ_1 over $\mathbb{F}_2$.

For hypertrees, we have the following:

THEOREM 6.6. *Let T be a (spanning) d-hypertree with nonnegative weights on the d-simplices. Then, the lightest-cap d-volume $\mathrm{vol}_T^{(d)}$ is ℓ_1.*

Proof. Recall the definition of $C_{T,\sigma}$ from Lemma 6.1. We claim that $\mathrm{vol}_T^{(d)} = \sum_{\sigma \in T} \mathrm{vol}_{C^{(d)}_{T,\sigma}}$. For $\tau \in T$ this follows by definition, while for $\tau \notin T$, $\sum_{\sigma \in S} \mathrm{vol}^{(d)}_{C_{T,\sigma}}(\tau)$ is equal to the sum of weights of all the σ's in S belonging to the cycle created by adding τ to T, as it should be. ∎

This implies the following approximability result.

THEOREM 6.7. *Any d-volume on V can be approximated by an ℓ_1 d-volume with distortion at most $\binom{n-1}{d}$.*

Proof. Let $\mathrm{vol}^{(d)}$ be a d-volume function on $K_n^{(d)}$, and let T be the minimum (spanning) hypertree with respect to $\mathrm{vol}^{(d)}$. Then, for $\sigma \in T$, $\mathrm{vol}_T^{(d)}(\sigma) = \mathrm{vol}^{(d)}(\sigma)$. For $\sigma \notin S$, much like the MST in graphs, σ must be the heaviest d-simplex in the cycle $|Z|$ created by adding σ to T. Since the size of Z is at most $1+|T| \le 1+\binom{n-1}{d}$, the statement follows. ∎

While the upper bound on the distortion in Theorem 6.7 is probably too rough and the true exponent of n is probably smaller, we shall see in what follows that even for $d = 2$ the distortion can be as large as $\Omega(n^{\frac{1}{5}})$. Thus, in general it is polynomial, and not logarithmic as in the case for $d = 1$ (Bourgain's Theorem [9]).

The main negative result of this section is the following lower bound on distortion of approximating general 2-volumes by ℓ_1 2-volumes. On the way we define a d-dimensional analog of the graphical edge-expansion, which is of independent interest.

THEOREM 6.8. *There exists a 2-volume function such that any ℓ_1 volume distorts it by at least $\tilde{\Omega}(n^{1/5})$.*

Let us first outline the proof. Using the methods originally developed for the one-dimensional case, we show the existence of a connected 2-dimensional simplicial complex K with unit weights on its 2-simplices, and use a Poincare-type form to bound its approximability by a hypercut metric (this turns to be enough). A key feature that K must have is that $\mathrm{vol_K}$ has large average value (which will be guaranteed by the sparseness of K), and the other is that it intersect every hypercut significantly. This later feature suggests the definition of expansion which is interesting on its own.

Formally, given any connected complex K, consider the following Poincare-type form over the 2-volumes:

$$F_K(\mathrm{vol}) = \frac{\sum_{\sigma\in \mathrm{K}} \mathrm{vol}(\sigma)}{\mathrm{av}(\mathrm{vol})}, \tag{6.4}$$

where $\mathrm{av}(\mathrm{vol}) = \frac{1}{\binom{n}{3}} \cdot \sum_{\sigma\in \mathrm{K_n^{(2)}}} \mathrm{vol}(\sigma)$. By a standard argument frequently used in the theory of metric spaces, the distortion of embedding $\mathrm{vol_K}$ into ℓ_1 is lower-bounded by

$$\mathrm{dist}(\mathrm{vol_K} \hookrightarrow \ell_1) \ \geq\ \frac{\min_{\mathrm{vol}\in\ell_1} \mathrm{F_K}(\mathrm{vol})}{\mathrm{F_K}(\mathrm{vol_K})}. \tag{6.5}$$

Keeping in mind that K is unit-weighted, and that any $\mathrm{vol} \in \ell_1$ is a nonnegative combination of hypercut-volumes, we conclude that the above minimum is necessarily attained on a hypercut-volume. Thus Equation (6.4) becomes:

$$\mathrm{dist}(\mathrm{vol_K} \hookrightarrow \ell_1) \ \geq\ \mathrm{av}(\mathrm{vol_K})\cdot \min_{\mathrm{C}:\ \text{2-hypercut}} \frac{|\mathrm{K}\cap \mathrm{C}|/|\mathrm{C}|}{|\mathrm{K}|/\binom{\mathrm{n}}{3}} \tag{6.6}$$

Observe that for a graph G, the analogous expression of the second term in Equation (6.6) is,

$$\min_{C=E(A,\overline{A}):\ \text{cut}} \frac{|E(G)\cap C|/|C|}{|E(G)|/\binom{n}{2}} =$$

$$\min_{A\subset V, |A|\leq n/2} \left\{ \frac{|E(A,\overline{A})|}{|A|} \cdot \frac{1}{\text{average degree of } G} \right\} \cdot \frac{n-1}{n-|A|},$$

which is the normalized edge expansion of G up to a factor of 2. By analogy, we define [8]

Definition 6.5. *Let the* normalized (face) expansion *of $K \subseteq K_n^{(2)}$ be the value of*

$$\min_{C:\ \textit{2-hypercut}} \frac{|K\cap C|/|C|}{|K|/\binom{n}{3}}.$$

[8] A similar definition of face expansion was independently used in [26, 36]. See also the references therein.

I.e., the normalized expansion of K is the ratio between the minimum density of K with respect to a hypercut, and the density of K with respect to $K_n^{(2)}$.

Let $K_n^{(2)}(n,p)$ be the 2-dimensional analog of the Erdös-Rényi $G(n,p)$, where each $\sigma \in K_n^{(2)}$ is selected with probability $p = 25\log n/n$, independently from the others. Theorem 6.8 follows from the following two Lemmas.

Lemma 6.5. *For $K \in K^{(2)}(n,p)$ as above,* $\mathrm{av}(\mathrm{vol_K}) \geq \tilde{\Omega}(\mathrm{n}^{1/5})$ *with probability* $1-o(1)$.

Lemma 6.6. *The face expansion of $K \in K^{(2)}(n,p)$ is almost surely ≥ 0.5.*

Observe that Lemma 6.6 implies that K is connected, since if all 2-hypercuts meet K, then K must contain a blocker for the set of 2-hypercuts, namely a (spanning) 2-hypertree (Lemma 6.1). Thus, it strengthens the main result of [20] at the price of getting worst constants.

Before starting with the proof of Lemma 6.5, we need the following preparatory result.

Lemma 6.7. *Let Z be a 2-cycle, then $|V(Z)| \leq |Z|/2 + 2$.* [9]

Proof. Clearly, $\mathrm{link_v(Z)}$ is an Eulerian (1-dimensional) graph. As long as there is a vertex $v \in V(Z)$ for which $\mathrm{link_v(Z)}$ is not a simple cycle, do the following. Let $A_1, \ldots, A_r$ be the decomposition of $\mathrm{link_v(Z)}$ into edge-disjoint cycles. We introduce a new copy of v, $v_i, i = 1, \ldots r$ for each A_i, and replace each original 2-simplex $\{v, x, y\}$ containing v with a new 2-simplex $\{v_i, x, y\}$ where $(x, y) \in A_i$. This yields a new simple cycle Z'. Carry on with the this process on Z' etc. Since each time we produce a new 2-cycle with the same number of faces, but less vertices whose link is not a simple cycle, the process must terminate with a 2-cycle Z^* with all links being simple cycles. Such Z^*, using the language of algebraic topology, is a (vertex-) disjoint union of triangulations of 2-dimensional surfaces without boundary. Without loss of generality, assume that there is a single surface. It is known [23] that its Euler characteristics satisfies

$$\chi(Z^*) \ =\ |V(Z^*)| - |E(Z^*)| + |Z^*| \ \leq\ 2 \tag{6.7}$$

Observe that every edge e in Z^* appears in exactly two faces, and thus $2|E(Z^*)| = 3|Z^*|$. Plugging this into Equation (6.7) implies the Lemma for $|V(Z^*)|$, and hence for $|V(Z)|$. ■

[9] We were informed by Uli Wagner that the general version of this lemma is known as the Lower Bound Theorem. It states that for any d it holds that $d|V(Z)| - \binom{d+1}{2} \leq |Z|$, which is attended on the d-skeleton of the stacked $(d+1)$-polytope.

We are now ready to address Lemma 6.5.

Proof. **(of Lemma 6.5)** By Markov inequality K almost surely contains $o(n^3)$ 2-simplices, and thus $\mathrm{av}(\mathrm{vol_K})$ is determined by the 2-simplices $\sigma \notin K$. For each such σ, $\mathrm{vol_K}(\sigma)$ is the size of the smallest K-cap of σ, i.e., the minimum subset of simplices in K that together with σ form a simple cycle. Let us denote this cap by $\mathrm{Cap_K}(\sigma)$. Thus, to show that $\mathrm{av}(\mathrm{vol_K}) \geq \Omega(\lambda)$ (w.h.p.), it suffices to argue that the number of $\sigma \notin K$ for which the corresponding $\mathrm{Cap_K}(\sigma)$ has size less than λ, is $o(n^3)$ (w.h.p). Let N_λ be this number. Let n_k be the number of simple cycles of size exactly k in $K_n^{(2)}$. Then,

$$E[N_\lambda] \;=\; \sum_{k=4}^{\lambda} k \cdot n_k \cdot p^{k-1}(1-p) \tag{6.8}$$

Now, by Lemma 6.7, a cycle of size k has at most $k/2+2$ vertices. Fixing $t = k/2+2$ vertices, the number of size-k cycles on these vertices is clearly bounded by t^{3k}. Hence $n_k \leq (k/2+2)^{3k} \cdot \binom{n}{(k/2+2)} \;\leq\; n^2 \cdot (k^{2.5}\sqrt{n})^k$. Plugging this bound on n_k, and the value of p into Equation (6.8), we get,

$$E[N_\lambda] \;\leq\; n^2 \sum_{k=4}^{\lambda} (k^{2.5} \cdot \sqrt{n})^k \cdot k \cdot \left(\frac{25 \log n}{n}\right)^{k-1} \;\leq$$

$$\frac{n^3}{25 \log n} \cdot \sum_{k=4}^{\lambda} k \left(\frac{k^{2.5} \cdot 25 \log n}{\sqrt{n}}\right)^k$$

Choosing $\lambda = \frac{n^{1/5}}{50 \log n}$, we conclude that $E[N_\lambda] = O(n \log^3 n) = \tilde{O}(n)$, and by the Markov inequality we are done. ∎

Next, we turn to Lemma 6.6, the expansion Lemma.

Proof. **(of Lemma 6.6)** For a hypercut C, let $\gamma_K(C) = \frac{|K \cap C|/|C|}{|K|/\binom{n}{3}}$. We shall first estimate the probability that $\gamma_K(C) < 0.5$ for any *fixed* hypercut C, and then use the union bound to conclude that almost surely no such hypercut exists.

Observe first that $|K|$ is almost surely tightly concentrated around its mean which is $E[K] = p \cdot \binom{n}{3}$. Thus instead of discussing $\frac{|K \cap C|/|C|}{|K|/\binom{n}{3}}$, we may safely discuss $\frac{|K \cap C|/|C|}{|E[K]|/\binom{n}{3}} = \frac{|K \cap C|}{p \cdot |C|}$. Next, observe that $|K \cap C|$ is a sum of $|C|$ i.i.d Bernoulli variables, and its expectation is precisely $p|C|$. Thus, by Chernoff bound,

$$\Pr(\gamma_K(C) < 0.5) \;=\; \Pr(|K \cap C| < p \cdot |C|/2) \;\leq$$

$$e^{-p \cdot |C|/8}.$$

Let m_s be the number of 2-hypercuts of size s in $K_n^{(2)}$. By Theorem 6.4, $m_s \leq (4n)^{1+3s/n}$. Thus, the union bound implies that the probability that a bad C exists is at most

$$\sum_{s \geq n-2} m_s \cdot e^{-p \cdot s/8} \;\leq$$

$$4n \sum_{s \geq n-2} e^{\left(-\frac{25}{8}\frac{\log n}{n} + \frac{3\log(4n)}{n}\right) \cdot s} \;=\; o(1).$$

∎

6.4.4 Dimension Reduction for ℓ_1 Volumes

Given an ℓ_1 d-volume $\mathrm{vol} = \sum_{\mathrm{C} \in \mathcal{C}} \lambda_\mathrm{C} \cdot \mathrm{v_C}$, where $\mathcal{C}$ is a collection of d-hypercuts over $\mathbb{F}_2$, v_C is the cut volume associated with C, and λ_C are positive reals, we call $|\mathcal{C}|$ the *hypercut-dimension* of this particular representation of vol. We define the *hypercut-dimension* of vol as the minimum possible hypercut-dimension of any representation of it.

Let the *hypercut cone* be the convex cone formed by all ℓ_1 d-volumes on $K_n^{(d)}$. The extremal rays of this cone are the hypercut-d-volumes.

Claim 6.4. *The hypercut cone has full dimension.*

Proof. Assume that a function $f : K_n^{(d)} \mapsto \mathbb{R}$ sums up to 0 on every hypercut and therefore on any d-coboundary of $K_n^{(d)}$). It suffices to show that f is identically 0. Let σ be any d-simplex in $K_n^{(d)}$, and let τ_1, τ_2 be distinct $(d-1)$-dimensional faces of σ. Let B_1, B_2 and B_{12} be the d-coboundaries in $K_n^{(d)}$ induced by τ_1, τ_2 and $\{\tau_1, \tau_2\}$ respectively. Then, $0 = f(B_1) + f(B_2) - f(B_{12}) = 2f(\sigma)$, and the claim follows. ∎

Since the hypercut cone is a subset of $\mathbb{R}^{\binom{n}{d+1}}$, Caratheodory Theorem implies that the hypercut-dimension of any $\mathrm{vol^d}$ is at most $\binom{n}{d+1}$. However, we seek a multiplicative approximations of a much smaller hypercut-dimension. For volumes over the reals, Theorem 6.2 states that the hypercut dimension of an approximating metric can indeed be dropped down by a factor of n, with respect to the above Caratheodory bound. For the larger class of ℓ_1 d-volumes over $\mathbb{F}_2$, there is a similar phenomenon.

Theorem 6.9. *Let* vol *be an ℓ_1 d-volume on n points, and let $0 < \epsilon < 1$ be a constant. Then there exists an ℓ_1 d-volume* vol′ *that distorts* vol *by at most a multiplicative factor of $\frac{1+\epsilon}{1-\epsilon}$, and the hypercut-dimension of* vol′ *is at most $O(n^d \log n/\epsilon^2)$, thus improving the trivial $O(n^{d+1})$. Furthermore,* vol′ *is efficiently constructible.*

Proof. Let M be a $\binom{n}{d+1} \times |\mathcal{C}|$ Boolean matrix whose rows are indexed by d-simplices, the columns are indexed by d-hypercuts, and $M(\sigma, C) = 1$ if σ belongs to the hypercut C and 0 otherwise. Observe that $M\lambda$'s correspond to ℓ_1 d-volumes on $K_n^{(d)}$, and $|\text{supp}(\lambda)|$ is an upper bound on the hypercut-dimension of the respective d-volume. Thus, Theorem 2.2 applies, yielding an upper bound of $O(\,\mathrm{trk(M)} \cdot \mathrm{d} \log \mathrm{n} \,/\, \epsilon^2)$ on the hypercut dimension. It remains to upper-bound trk(M). It turns out be at most $\binom{n-1}{d}$.

Indeed, let Q be a square $N \times N$ lower triangular nonsingular minor of M. Let the rows be indexed by $\{\sigma_i\}_{i=1}^N$, and the columns be indexed by $\{C_i\}_{i=1}^N$ in this order. It means, in particular, that $\sigma_i \in C_i$, but $\sigma_i \notin C_j$ for $j > i$. We claim that the set of d-simplices $\{\sigma_i |\ i = 1, \ldots, N\}$ does not contain d-cycles. Indeed, assume by contradiction that it does contain a cycle Z, and r be the largest index such that $\sigma_r \in Z$. Consider the corresponding d-cut C_r. Since $\sigma_r \in Z \cap C_r$, by Claim 6.1, C_r must contain another d-simplex from Z, contrary to the fact $\sigma_i \notin C_r$ for every $i < r$.

Thus, $\{\sigma_i |\ i = 1, \ldots, N\}$ is acyclic, and N is bounded by the size of the maximum acyclic subcomplex, i.e., d-tree, which is $\binom{n-1}{d}$. ∎

6.5 Some Additional Remarks

6.5.1 Another example of an ℓ_1 volume. As mentioned in the Introduction, d-volumes are well suited and are potentially useful for representing quantitative d-ary relations. Here is an example to demonstrate what we mean.

Let $\mathcal{H}$ be a family of hyperplanes in $\mathbb{R}^d$ in the general positions. For every $(d+1)$-tuple of $\mathcal{H}$, define the measure of its non-collinearity as the Euclidean volume of the (unique) bounded cell formed by these hyperplanes.

Claim 6.5. *The above measure on the $(d+1)$-tuples is in fact an ℓ_1 d-volume (both real and $\mathbb{F}_2$) over $\mathcal{H}$.*

6.5.2 Sparse Spanners. It is well known that the average degree in a graph H with n vertices and girth g is $n^{O(\frac{1}{g})}$. Since (see [5]) the shortest-path metric d_G of a weighted graph G can be $(g-1)$-approximated by that of its subgraph H of girth g, there exists a g-spanner of G with at most $n^{1+O(\frac{1}{g})}$ edges. The construction naturally carries on to volumes, which brings us to a question: What is the maximal number of d-simplices in a simplicial complex K on n vertices, such that the smallest d-cycle of K is of size $\geq g$? Taking the field to be $\mathbb{F}_2$, the probabilistic construction of Lemma 6.6 (with small local amendments) shows that for $d = 2$ there exists K of average degree $O(\log n)$, and the smallest cycle of size $\tilde{\Omega}(n^{0.2})$. (By degree of a 1-simplex e we mean the number of 2-simplices in K that contain e.) Thus, the situation for $d = 2$ significantly differs from the graph theoretic case. It would be interesting to get tighter bounds for this problem. See also [21] for a somewhat related discussion.

6.5.3 On $\mathbf{c_1(K)}$. Like in graphs, given a d-complex K over $\mathbb{F}_2$, one may ask what is the worst possible distortion of approximating $\mathrm{vol_K}$, a lightest-cap volume of K (over all choices of nonnegative weights of its simplices), by an ℓ_1 volume. This important numerical parameter is called (by analogy with graphs) $c_1(K)$. One of the most important open questions in the theory of finite metric spaces is whether any graph G lacking a fixed minor has a constant $c_1(G)$ (see e.g., [15] for a related discussion and partial results). It is natural to ask a similar question about d-complexes: what properties of K would imply a nontrivial upper bound on $c_1(K)$? The techniques of [15] imply this: $c_1(K) \leq 2^{\chi(K)}$, where K (as usual) is assumed to have a complete $(d-1)$ skeleton and $\chi(K)$ is the Euler characteristic of K. The construction proceeds via repeatedly picking a minimal cycle, and removing a random d-simplex in it with probability proportional to its volume. The lightest-cap volume of the random (sub-)hypertree of K obtained in this manner dominates $\mathrm{vol_K}$, yet stretches it (in expectation) by only a constant factor.

Acknowledgments We would like to thank for valuable comments and discussions to Vladimir Hinich, Gil Kalai, Nati Linial, Avner Magen, Roy Meshulam, Tasos Sidiropoulos, Uli Wagner and the anonymous referees of this version.

References

[1] N. Alon, J. Spencer, The Probabilistic Method, John Wiley, 1992.

[2] A. Andoni, M.S. Charikar, O. Neiman and H.L. Nguyen. Near Linear Lower Bounds for Dimension Reduction in ℓ_1. to appear at *FOCS 2011*, October 2011.

[3] K. Ball. Isometric embedding in ℓ_p-spaces. *European J. Combin.* 11(4) 305-311, 1990.

[4] K. Barhum, O. Goldreich and A. Shraibman, On Approximating the Average Distance between points, in Proc. of APPRPX'07, 296-310, 2007.

[5] I. Althöfer, G. Das, D.P. Dobkin, D. Joseph and J. Soares, On Sparse Spanners of Weighted Graphs, *Discrete & Computational Geometry*, 9, 81–100, 1993.

[6] E. Babson, C. Hoffman, M. Kahle, The fundamental group of random 2-complexes, arXiv0711.2704.pdf.

[7] J. D. Batson, D. A. Spielman and N. Srivastava, Twice-Ramanujan Sparsifiers, *STOC'09*, 255-262, 2009.

[8] A. Benczúr and D. R. Karger. Randomized Approximation Schemes for Cuts and Flows in Capacitated Graphs. *CoRR*, cs.DS/0207078, 2002.

[9] J. Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52(1-2):46–52, 1985.

[10] B. Brinkman and M. Charikar. On the Impossibility of Dimension Reduction in ℓ_1. *J. ACM*, 52(5):766–788, 2005.

[11] R. Cordovil, B. Lindström Simplicial Matroids. Section 6 in *Combinatorial Geometries, N. White, editor*, Cambridge University Press, 1987.

[12] U. Feige. Approximating the bandwidth via volume respecting embeddings. Journal of Comput. Syst. Sci. 6(3), 510–539, 2000.

[13] D. Feldman, M. Langberg. A Unified Framework for Approximating and Clustering Data. *Proc. 43st Annu. ACM Symposium on Theory of Computing, STOC'11*, 569–578, 2011

[14] P. Frankl and Z. Füredi. An exact result for 3-graphs. *Discrete Mathematics*, 50:323–328, 1984.

[15] A. Gupta, I. Newman, Y. Rabinovich, A. Sinclair. Cuts, Trees and l_1-embeddings of graphs. *Combinatorica* 24(2):233–270, 2004.

[16] S. Har Peled. Geometric Approximation Algorithms (Mathematical Surveys and Monographs) *American Mathematical Society* June 15, 2011.

[17] W. Johnson and J. Lindenstrauss. Extensions of Lipschitz Maps into a Hilbert space. *Contemporary Mathematics*, 26:189-206, 1984.

[18] J. Lee and A. Naor, Embedding of the Diamond graph in L_p and Dimension Reduction in ℓ_1, *Geometric and functional analysis (GAFA)*, 14(4), 745-747 (2004).

[19] Y. Li, P.M. Long, A. Srinivasan, Improved Bounds on the Sample Complexity of Learning. *Journal of Comput. Syst. Sci.* 62(3), 516-527, (2001).

[20] N. Linial and R. Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006.

[21] A. Lubotzky and R. Meshulam. A Moore bound for simplicial complexes. *Bull. London Math. Soc.*, 39, 353-358, (2007).

[22] A. Magen and A. Zouzias Near Optimal Dimensionality Reductions That Preserve Volumes In *Proc. of APPROX-RANDOM 2008*, 523–534.

[23] W.S. Massey. Algebraic Topology: An Introduction *Springer Verlag, 1990*

[24] J. Matousek. Geometric Discrepancy. *Springer, 2nd printing edition, 2009.*

[25] J. Matousek, M. Tancer, U. Wagner. Hardness of Embedding Simplicial Complexes in $\mathbb{R}^d$ In *Proc. of SODA 2009*, 855–864.

[26] J. Matousek, U. Wagner. On Gromov's Method of Selecting Heavily Covered Points *CoRR*, abs/1102.3515, 2011

[27] R. Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Structures and Algorithms* 34(2009) 408–417.

[28] J.R. Munkres. Elements of Algebraic Topology *Addison-Wesley, 1984*

[29] N. Pippenger and K. Schleich. Topological characteristics of random triangulated surfaces. *Random Struct. Algorithms*, 28(3):247–288, 2006.

[30] G. Schechtman, More on Embedding Subspaces of l_p in l_1^r. *Compositio Mathematica*, 61(2): 159–169, 1987.

[31] J.J. Seidel A survey of two-graphs. In *Colloquio Internazionale sulle Teorie Combinatorie, Procedings* Vol. I, pp. 481-511. Atti dei Convegni Lincei, No. 17. Accademia Nazionale dei Lincei, Rome, 1976.

[32] D. A. Spielman, N. Srivastava, Graph Sparsification by Effective Resistances, *STOC*, 563-568, 2008.

[33] D. A. Spielman and S. Teng, Nearly-Linear Time Algorithms for Graph Partitioning, Graph Sparsification, and Solving Linear Systems, *arXiv:0809.3232, arXiv:0808.4134, arXiv:cs/0607105*, 2003

[34] M. Talagrand, Sharper Bounds for Gaussian and Empirical Processes. *Annals of Probability* 22: 28-76, 1994.

[35] M. Talagrand, Embedding subspaces of ℓ_1 into ℓ_1^n. *Proceedings of the American Mathematical Society*, 108(2): 363–369.

[36] U. Wagner. Minors in random and expanding hypergraphs. *27th SoCG'*, 351-360, 2011.