SIAM J. COMPUT.
Vol. 42, No. 3, pp. 855–883

ON MULTIPLICATIVE λ-APPROXIMATIONS AND SOME GEOMETRIC APPLICATIONS*

ILAN NEWMAN† AND 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 $\widetilde{O}(\mathrm{trk}(\mathcal{F}) \cdot \mathrm{VCdim}(\mathcal{F})/\epsilon^2)$, where $\mathrm{VCdim}(\mathcal{F})$, bounded from above by $\mathrm{trk}(\mathcal{F})$, is the VC-dimension of $\mathcal{F}$. We also present some 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 is 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 Schechtman, showing that the dimension is bounded by $O(n \log n/\mathrm{poly}(\epsilon))$. 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.

Key words. sparsification, dimension reduction, multiplicative approximation, core sets

AMS subject classifications. 51F, 60C, 68R

DOI. 10.1137/100801809

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 extraordinarily 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 so far achieved 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

*Received by the editors July 12, 2010; accepted for publication (in revised form) March 18, 2013; published electronically May 16, 2013. A preliminary version of this paper appeared in SODA2012. This Research was supported by The Israel Science Foundation, grant 862/10.
http://www.siam.org/journals/sicomp/42-3/80180.html

†Department of Computer Science, University of Haifa, Haifa, Israel (ilan@cs.haifa.ac.il, yuri@cs.haifa.ac.il).

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 a study of specific problems, e.g., k-median, clustering, etc. (cf. Chapter. 19 in [16]). A recent paper [12] introduces a general approach that can be related to our methods. However, this paper focuses solely on specific applications, and is in a complementary relation to 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 incidence matrix of $\mathcal{F}$ versus X with 1's on the diagonal. Another important parameter is $\mathrm{rank}^*(\mathcal{F})$, 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, $\mathrm{trk}(\mathcal{F}) \leq \mathrm{rank}^*(\mathcal{F})$.

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 $\mathrm{trk}(\mathcal{F})$ is that any measure μ on $(X, \mathcal{F})$ can be $\frac{1+\epsilon}{1-\epsilon}$-approximated by μ^* with support of size $\widetilde{O}(\mathrm{trk}^2(\mathcal{F})/\epsilon^2)$, or, to be more precise, $\widetilde{O}(\mathrm{trk}(\mathcal{F}) \cdot \mathrm{VCdim}(\mathcal{F})/\epsilon^2)$, where $\mathrm{VCdim}(\mathcal{F}) \leq \mathrm{trk}(\mathcal{F})$ is the VC-dimension of $\mathcal{F}$. 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 $< \mathrm{trk}(\mathcal{F})$. 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 $\Omega(\mathrm{trk}(\mathcal{F})/\epsilon)$.

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 $O(\mathrm{trk}(\mathcal{F}) \cdot \log|\mathcal{F}|/\epsilon^2)$. In the second construction the support is of size $O(\mathrm{rank}^*(\mathcal{F})/\epsilon^2)$.

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 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 $\mathbb{R}^{O(n/\epsilon^2)}$

[1] Usually a dual definition is used, i.e., a small weighted sample $\mathcal{F}^* \subset \mathcal{F}$ is sought, so 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 $\mathcal{F}$.

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 et al. [3].

Another application has to do with Euclidean volumes. Assume that X is embedded in a 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 [6], where it was shown that the average can be $(1+\epsilon)$-approximated by nonadaptively 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 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 computing science, e.g., in the classical algorithm of Feige [11] 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 computer science and in applications. However, we are aware of no formal treatment of finite volumes analogous to that of finite metrics.

For a finite set X, we say that (X, ν) is a d-volume space if ν is a 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 $\nu(\sigma) \le \sum_{\tau \in C, \tau \neq \sigma} \nu(\tau)$. 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 one-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 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 the combinatorics of simplicial complexes, a fascinating area that rapidly gains popularity (see, e.g., [20, 27, 25, 26, 34, 29, 4]), helping in the choice of "right" definitions, leading to interesting natural questions, and suggesting tools for approaching some of these questions.

2. Multiplicative approximation via triangular rank. The *triangular rank* $\text{trk}(\mathcal{F})$ 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

incidence matrix of $(\mathcal{F}, X)$.[2] 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 intersection or 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^a *additively* ϵ-approximates μ if $|\mu(S) - \mu^a(S)| \le \epsilon \cdot \mu(X)$ for every $S \in \mathcal{F}$. The quality of additive approximation is closely related to the Vapnik–Chervonenkis dimension of $\mathcal{F}$, $\mathrm{VCdim}(\mathcal{F})$, defined as the size of the maximum subset $Y \subseteq X$ shattered by $\mathcal{F}$.

THEOREM 2.1 (a special case of [19, 32]). *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^a on $(X, \mathcal{F})$ that has support size $O(\mathrm{VCdim}(\mathcal{F})/\delta^2)$.*

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 $X = \{1, \dots, n\}$ and $\mathcal{F} = \{\{1, \dots, i\} \mid 1 \le i \le n\}$.

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. Let "$<$" be a transitive antisymmetric relation on $\mathcal{F}$ defined by $S < T$ if $2\mu(S) < \mu(T)$. Let $(S_1, S_2, \dots, S_t)$ be the largest chain with respect to "$<$," that is, for all $1 \le i < t$, it holds that $2\mu(S_i) < \mu(S_{i+1})$. 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 "$<$," and any antichain defines a bucket with the desired property. ◻

THEOREM 2.2. *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 $\{x_i\}_{i=1}^{\mathrm{trk}(\mathcal{F})}$ be 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 $\mu(x_i) = (2\lambda+1)^i$ for every x_i, and assigning 0 to the rest of X. Observe that $\frac{\mu(S_i)}{\mu(S_{i-1})} \ge \frac{(2\lambda+1)^i}{\sum_1^{i-1}(2\lambda+1)^j} \ge 2\lambda$.

Assume by contradiction that μ^* is a λ-approximation of μ, and it has support X' of size $|X'| = t < \mathrm{trk}(\mathcal{F})$. By Claim 2.1 applied to $\mathcal{F}|_{X'}$, i.e., the sets restricted to the support of μ^*, the values of μ^* can be 2-bucketed into at most t buckets (as obviously the trk is bounded by the support size). Hence, by the pigeonhole principle, for some two sets $S_i, S_j, i < j$, $\mu^*(S_i), \mu^*(S_j)$ belong to the same bucket.

Hence, $\frac{\mu^*(S_j)}{\mu^*(S_i)} \le 2$, while, as noted above, $\frac{\mu(S_j)}{\mu(S_i)} > 2\lambda$. This contradicts the assumption that μ^* multiplicatively λ-approximates μ. ◻

[2]That is, after permuting the columns and rows of the matrix so as to obtain the largest possible value.

While the lower bound of Theorem 2.2 is universal and it holds for every family $\mathcal{F}$, the following claim presents a specific family for which the lower bound on the support, in terms of the dependence in the multiplicative error, is stronger.

CLAIM 2.2. *Let $n \geq t(1+1/\epsilon)$, and let $\mathcal{F}$ be the family of all subsets of $[n]$ of size $n-t+1$. Then $\mathrm{trk}(\mathcal{F}) = \mathrm{t}$, and every μ^* that $\frac{1+\epsilon}{1-\epsilon}$-approximates the counting measure on $[n]$ has support size $\geq t/2\epsilon$.*

Proof. To see that $\mathrm{trk}(\mathcal{F}) = \mathrm{t}$, note that on one hand any t distinct sets in $\mathcal{F}$ cover the entire $[n]$, and on the other hand, the sets $S_i = [n-t] \cup \{n-t+i\}$, $i = 1, \ldots, t$, define a triangular minor of size t.

For the lower bound on the support size, observe first that the symmetry of $\mathcal{F}$ implies, without loss of generality, that μ^* is uniform on its support, namely, $\mu^*(i) = \alpha > 0$ for every i in the support of μ^*. Let s be the size of the support of μ^*. If $s > n-t$, then $s > t/\epsilon$, and we are done; else, the maximal value of μ^* on $\mathcal{F}$ is $\alpha \cdot s$, while the minimum value is $\alpha \cdot (s-t)$. Therefore, it must hold that $\frac{s}{s-t} \leq \frac{1+\epsilon}{1-\epsilon}$, implying $s \geq (1+\epsilon)t/2\epsilon > t/2\epsilon$. $\square$

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. One should bear in mind that $\mathrm{VCdim}(\mathcal{F}) \leq \mathrm{trk}(\mathcal{F})$.

THEOREM 2.3. *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 2.4. *Under the same assumptions as in Theorem 2.3, and for ϵ bounded from above by some universal constant $0 < c_0 < 1$, there exists a multiplicative $\frac{1+\epsilon}{1-\epsilon}$-approximation μ^* on $(X, \mathcal{F})$ of size at most $O(\mathrm{trk}(\mathcal{F}) \cdot \mathrm{VCdim}(\mathcal{F})/\epsilon^2 \cdot (\log(\mathrm{trk}(\mathcal{F})) + \log\frac{1}{\epsilon}))$.*

2.1. Proof of Theorem 2.3. The method of proof generalizes the method of Benczúr and Karger from [8]. The existence of μ^* will be established using a probabilistic argument. We start with some preliminary observations.

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. 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 original system. However, the representation size may drastically change, hence algorithmic issues arise. They will be addressed in section 2.3.

As we are about to sample the elements of X, notice that some elements are more essential than others, and thus the sampling is necessarily nonuniform. 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 high probability. Thus, it makes sense to assign to each element $i \in X$ a *fragility* parameter indicating how carefully it should be sampled. Without loss of generality we assume any element of X appears in some set $S \in \mathcal{F}$; otherwise, it can be simply removed.

DEFINITION 2.5. *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.[3]

LEMMA 2.6.

1. *For any set* $S \in \mathcal{F}$ *it holds that* $|S| \ge \max_{x \in S} s(x)$.
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. For the first statement, let x be any element in S, let i be the smallest number such that $|B_i| = s(x)$, and let k be smallest number such that $S \subseteq \cup_{j=1}^{k} B_j$. 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. Therefore, it was a candidate to be chosen at the step i of the process. Since actually B_i was chosen at this step, it means that $|S - \cup_{j=1}^{i-1} B_j| \ge |B_i| = s(x)$, and the statement on $|S|$ follows.

For the second statement, observe that for every $i = 1, \dots, N$, all elements in the block B_i have (the same) strength $s(B_i) \ge |B_i|$. Thus,

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

For the third statement, for $i = 1, \dots, N$, let S_i be the original set involved in the definition of the block B_i, in other words, $B_i = S_i - \cup_{j=1}^{i-1} B_j$. Clearly, no S_i is contained in the union of its predecessors $\cup_{j<i} S_j$. Since there can be at most $\mathrm{trk}(\mathcal{F})$ such sets in $\mathcal{F}$, the statement follows. ∎

The main sampling procedure, i.e., the sparse approximation measure μ^*, is defined in the following manner.

DEFINITION 2.7. *Let* $\rho > 1$ *be a parameter to be specified later. For each element* $x \in X$, *let* $p_x = \min\{\rho f(x), 1\}$, *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* $\alpha_x = 1/p_x$, *the random measure* μ^* *on* $(\mathcal{F}, X)$ *is defined as* $\mu^*(S) = \sum_{x \in S} \alpha_x Y_x$.

The rest of the section is devoted to showing that μ^* almost surely has the required properties.

In what follows we shorten a statement of the type $x \notin [(1-\epsilon)a, (1+\epsilon)a]$ by $x \notin (1 \pm \epsilon)a$.

We shall require the following version of the Chernoff bound.

THEOREM 2.8 (see [1]). *Let* $Y_1, \dots Y_n$ *be independent Poisson trials such that* $\Pr(Y_i = 1) = p_i$, *and let* $\nu = \sum p_i$. *Then, for any* $0 < \beta < 1$, $\Pr[\sum Y_i \notin (1 \pm \beta) \cdot \nu] \le 2e^{-\beta^2 \nu/3}$.

The first goal is to show that almost surely the size of the support of μ^*, namely of $X^* = \{x \in X |\ Y_x = 1\}$, is $O(\rho N)$. As before, N is the number of blocks B_i in the partition.

LEMMA 2.9. *With probability* $\ge 1 - 2e^{-8\rho N/3}$, *it holds that* $|X^*| \le 2\rho N$.

[3] The definitions of $s(x), f(x)$, and the number of blocks N, in the partition above, may appear to depend on a somewhat arbitrary choice of the next set B_i in the case of ties. It can be shown that these parameters are well defined and do not depend on these choices. However, this will be of no importance in what follows, and all claims hold for *every* possible partition.

Proof. Since $|X^*| = \sum_{x \in X} Y_x$, item 2 of Lemma 2.6 implies that

$$E[|X^*|] = \sum_{x \in X} p_x \leq \sum_{x \in X} \rho f(x) \leq \rho N\,.$$

Since the Y_x are independent, Theorem 2.8 applies, implying that $\Pr(|X^*| > 2\rho N) \leq 2e^{-8\rho N/3}$. ∎

Showing that μ^* is indeed an $\frac{1+\epsilon}{1-\epsilon}$-approximation of μ on $(X, \mathcal{F})$ is a harder task, to be achieved in a number of steps.

LEMMA 2.10. *Set $\rho = \frac{3}{\epsilon^2}(\ln(2|\mathcal{F}|) + \ln t + k)$, where $k > 0$ is any real number, and t is the number of distinct strengths s_i. Then, the probability that there exists $S \in \mathcal{F}$ such $\mu^*(S) \notin (1 \pm \epsilon)|S|$ is $\leq e^{-k}$.*

Proof. First, observe that the expectation of μ^* is precisely μ, i.e., for every $S \subseteq X$ it holds that $\mathrm{E}[\mu^*(S)] = |S|$. Indeed,

$$\mathrm{E}[\mu^*(S)] = \mathrm{E}\left[\sum_{x \in S} \alpha_x Y_x\right] = \sum_{x \in S} \alpha_x \cdot \mathrm{E}[Y_x] = \sum_{x \in S} \alpha_x p_x = |S|.$$

The next step is to show that almost surely for all $S \in \mathcal{F}$ (simultaneously), $\mu^*(S)$ is tightly concentrated around its mean. The appropriate tool for showing this is Theorem 2.8; however it does not directly apply to our case as we have *weighted* sums of Bernoulli random variables. The following construction is designed to overcome this difficulty.

DEFINITION 2.11. *Let $s_1 < s_2 < \cdots < s_t$ be the set of all distinct strength values produced by the process from Definition* 2.5. *For each i, let k_i be the smallest number such that $|B_{k_i}| = s_i$. Let $\alpha_i = \alpha_x$ where x is any element in B_{k_i}. Define the "layering" of X by $X^i = \cup_{j \geq k_i} B_j, i = 1, \ldots, t$. In particular, $X^1 = X$. Finally, define the following random measures on X, $\mu^i(S) = \sum_{x \in S \cap X^i} Y_x$, $i = 1, \ldots, t$, where the Y_x's are the random Bernoulli variables of Definition* 2.7.

Let $\Delta_i = \alpha_i - \alpha_{i-1}$ and $\Delta_1 = \alpha_1$. Observe that

$$\mu^* = \sum_i \Delta_i \cdot \mu^i. \tag{2.1}$$

The crucial observation is that for any $S \in \mathcal{F}$ and any $j \leq t$, $\mu^j(S)$ is either identically 0, or it's expectation is at least ρ. The argument is similar to the one used in the proof of the first item of Lemma 2.6. Let i be the maximal number such that $S \cap X^i \neq \emptyset$. By the definition of B_{k_i}, it holds that $|S \cap X^i| \geq |B_{k_i}| = s_i$, and that for any $x \in S \cap X^i$, $s(x) = s_i$. Thus, for all $j \leq i$,

$$\mathrm{E}[\mu^j(S)] \geq \mathrm{E}[\mu^i(S)] = \sum_{x \in S \cap X^i} \mathrm{E}[Y_x] = \sum_{x \in S \cap X^i} p_x \geq s_i \cdot \frac{\rho}{s_i} = \rho\,,$$

where the first inequality holds since $X_i \subset X_j$ for $j \leq i$. For the case $j > i$, it clearly holds that $\mu^i(S) = 0$.

We proceed to conclude the proof of Lemma 2.10. Consider the random variable $\mu^i(S) = \sum_{x \in S \cap X^i} Y_x$ for any $1 \leq i \leq t$ and $S \in \mathcal{F}$. Either it is identically 0, and thus is equal to its expectation, or its expectation is at least ρ. Applying to it Theorem 2.8, one obtains

$$\Pr[\,\mu^i(S) \notin (1 \pm \epsilon) \cdot \mathrm{E}[\mu^i(S)]\,] \leq 2e^{-\frac{\epsilon^2}{3} \cdot \mathrm{E}[\mu^i(S)]} \leq 2e^{-\frac{\epsilon^2}{3} \cdot \rho}\,. \tag{2.2}$$

Substituting the proposed value for ρ, we conclude that the above probability is at most $|\mathcal{F}|^{-1} \cdot t^{-1} \cdot e^{-k}$. Taking the union bound over all $1 \leq i \leq t$ and $S \in \mathcal{F}$, we conclude that the probability that there exist a bad pair (i, S) with $\mu^i(S) \notin (1 \pm \epsilon) \cdot \mathrm{E}[\mu^i(x,y)]$ is at most e^{-k}. Keeping in mind that $\mu^* = \sum_{i=1}^t \Delta_i \cdot \mu^i$, the statement follows. $\square$

To sum up, choosing k in $\rho = \frac{3}{\epsilon^2}(\ln(2|\mathcal{F}|) + \ln t + k)$ to be a large enough constant, and keeping in mind that $t \leq N \leq \mathrm{trk}(|\mathcal{F}|)$, Lemma 2.10 implies that μ^* is with high probability a multiplicative $\frac{1+\epsilon}{1-\epsilon}$-approximation of μ on $(X, \mathcal{F})$. On the other hand, Lemma 2.9 implies that X^*, the support of μ^*, is with high probability of size $\leq 2\rho N = O(\mathrm{trk}(\mathcal{F}) \log |\mathcal{F}|)/\epsilon^2$. This establishes Theorem 2.3.

As a concluding remark, let us mention that sometimes a better upper bound can be obtained, as in [8], by strengthening (2.2). Instead of using a uniform lower bound on the expected value of $\mu^i(S)$, one may exploit the distribution of these expectation over $\mathcal{F}$, and use, e.g., the property that small values are rare. Let us also mention without much elaboration the recent paper [14] that adds new understanding to the framework of [8], and may potentially contribute to the general framework as well.

2.2. Proof of Theorem 2.4. We start with the same partition of X into N blocks B_i as in Definition 2.5. 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 $S \in \mathcal{F}$,

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

Since $\mathrm{VCdim}(\mathcal{F}|_{\mathrm{B_i}}) \leq \mathrm{trk}(\mathcal{F}|_{\mathrm{B_i}}) \leq \mathrm{trk}(\mathcal{F})$, and $N \leq \mathrm{trk}(\mathcal{F})$, Theorem 2.1 implies that 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 index such that $S \cap B_t$ is not empty. By definition of B_i's, $|B_i| \leq |S|$ for $i \leq t$. Therefore,

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

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 2.3 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 $(\frac{1+\epsilon/2}{1-\epsilon/2})^2 \leq \frac{1+\epsilon}{1-\epsilon}$. The size of the support of μ^{**} is $O(\mathrm{trk}(\mathcal{F}^*) \log |\mathcal{F}^*| \,/\, \epsilon^2)$. Let $d = \mathrm{VCdim}(\mathcal{F})$. Since $\mathrm{VCdim}(\mathcal{F}^*) \leq$ d, the Sauer lemma (see, e.g., [24], Lemma 5.9) implies that $\mathcal{F}^*$ contains at most $\sum_{i=0}^d \binom{|X^*|}{i}$ distinct sets, and thus $|\mathcal{F}^*| \leq |X^*|^{d+1}$. Combining the estimates for $|\mathcal{F}^*|$, $|X^*|$, and for the size of the support of μ^{**}, we conclude that μ^{**} is the desired approximation of μ. $\square$

2.3. Algorithmic considerations. Recall that for simplicity of presentation the original measure μ was replaced by a counting measure. This was done 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 2.5; 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 with approximating measures on $(\mathcal{F}, X)$ where $\mathcal{F}$ is a family of subsets of X or, equivalently, a family of $\{0,1\}$-valued functions on X. In this section it will be just as convenient to work in an extended setting, where $\mathcal{F}$ is a family of functions from X to $\mathbb{R}^+$. The extended problem can be formulated as follows. Given a linear form with nonnegative coefficients $L(f) = \sum_{x\in X} w_x f(x)$ over functions $f : X \mapsto \mathbb{R}^+$, the goal is 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 obtains the original setting.

Recently it was brought to our attention, that the extended setting was extensively studied in the context of concrete geometrical applications, and we refer the reader to the recent [12], presenting a unifying approach and strongest known results for a long list of such applications.

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 3.2. *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 trivially extends 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 triangular rank. A standard tensor product argument implies a lower bound on $\mathrm{rank}^*(\mathrm{M})$ in terms of $\mathrm{rank}(\mathrm{M})$: 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 $\mathrm{trk}(\mathrm{M}) = 2$.[4]

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 3.3 (see [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 nonnegative diagonal matrix $A_{n\times n}$ with at most $O(r/\epsilon^2)$ positive entries, and with the following property; Let $Q^* = BAB^T$. Then, for every $x \in \mathbb{R}^n$ it holds that*

$$(1-\epsilon) \cdot x^T Q x \le x^T Q^* x \le (1+\epsilon) \cdot x^T Q x \ .$$

[4]Currently, we do not have similarly good lower bounds on rank^*_δ for $\delta > 0$.

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 also works for general positive semidefinite symmetric Q's.

Proof of Theorem 3.2. Let $r = \text{rank}^*(\text{M})$. 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 i,j. For any nonnegative $w = (w_1,\ldots,w_n) \in \mathbb{R}^n$, define $B(w)$ as the matrix obtained from B by multiplying each column j of B by $\sqrt{w_j}$. Let y_i denote the ith row of Y, a $1\times r$ real vector. Then, for any v and each $1 \le i \le m$, it holds that

$$y_i\, B(w)B(w)^T\, y_i^T = \sum_{j=1}^{n} \left(M(i,j)^{\frac{1}{2}}\right)^2 w_j = \sum_{j=1}^{n} M(i,j) w_j = (Mw)_i\,. \tag{3.1}$$

Applying Theorem 3.3 to the matrix $B(w)B(w)^T$, we conclude that there exists a nonnegative diagonal $n\times n$ matrix A of support size $O(\text{rank}^*(\text{M})/\epsilon^2)$, such that for all $1 \le i \le m$,

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

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 $1 \le j \le n$. Using once more (3.1), this time with 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, well understood for Euclidean metrics but much less understood for ℓ_1-metrics, is about the dimension of a faithful 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 [5], 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 [3] strengthens this to $n^{1+O(1/\log 1/\epsilon)}$. The best upper bound so far, due to Schechtman [30] (extended by Talagrand [33]) 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, semimetric) 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 smaller 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, Claim 6.4), its soft rank and its triangular rank are significantly smaller.

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

Proof. Consider the $n-1$ edges $\{e(v_j, v_{j+1})\}_{j=1}^{n-1}$, and the $n-1$ corresponding cuts defined by the sets $\{v_1, \ldots, v_j\}_{j=1}^{n-1}$. Clearly, the corresponding minor of M is lower triangular with 1's on the diagonal. Thus, $\mathrm{trk(M) \geq n-1}$. For the other direction, since $\mathrm{trk(M) \leq rank^*(M)}$, it suffices to show that $\mathrm{rank^*(M) \leq n-1}$. Let Y be a matrix whose rows are 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, indexed by vertices versus cuts, is defined as follows. For an arbitrarily oriented cut[5] $C = (U, V-U)$ and a vertex $v \in V$, let $B(v, C) = 1$ if $v \in U$, and let $B(v, C) = -1$ otherwise. It is easily checked that $YB = \pm M$, and that $\mathrm{rank(YB) \leq rank(Y) = n-1}$. ∎

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 2.3, we get an approximation of support $O(\mathrm{trk(M)} \log \mathrm{n}/\epsilon^2) = \mathrm{O(n \log n}/\epsilon^2)$ which matches the bound of [30]. Employing Theorem 3.2 leads to an improved bound of $O(\mathrm{rank^*(M)}/\epsilon^2) = \mathrm{O(n}/\epsilon^2)$.

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 claim shows, in terms of the dependence on n it is the best possible. The dependence on ϵ is left for future study.

CLAIM 4.2. *Consider a line with n points $\{p_1, p_2, \ldots, p_n\}$ on it (in this order), where the distance between p_{i+1} and p_i is $(2\lambda+1)^i$. Let d be the corresponding line metric. Then, 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 λ.*

Proof. It holds that $d = \sum_{i=1}^{n-1} (2\lambda+1)^i \cdot \delta_{C_i}$, where C_i is the cut defined by points indexed $\{1, 2, \ldots, i\}$. Associating cut metrics with cuts seen as subsets of edges, and observing that the system of cuts $\{C_i\}_{i=1}^{n-1}$ over E has triangular rank $n-1$ (see also the proof of Claim 4.1), one proceeds precisely as in the proof of Theorem 2.2, and the conclusion follows. ∎

5. Halfway discussion. So far, we have established some general tools for the multiplicative λ-approximation, and have shown their efficiency for the ℓ_1 dimension-reduction problem. In this section we partially address the naturally arising question

[5]By orientation of a cut $C(\{U, V-U\})$ we mean either $C(U, V-U)$ or $C(V-U, V)$.

about the applicability of these tools. Namely, what conditions on $\mathcal{F}$ might ensure that it has small triangular rank. This is followed (section 6) by an in-depth study of a certain class of such families $\mathcal{F}$ of a combinatorial-topological origin, and to finite volume spaces.

The starting point of the forthcoming discussion 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 with two generalizations of this claim, aiming at extending the range of applicability of the approximation techniques. The first generalization will not be further exploited in this paper; the second will be the basis for things to come.

5.1. Splitting set systems. The Boolean matrix M used in the proof of Claim 4.1 (i.e., the inclusion matrix of edges versus edge cuts) could be described somewhat differently using vertices instead of edges. In this representation the rows correspond to subsets e of V of size 2, the columns correspond to nontrivial subsets A of V, and $M(e, A) = 1$ iff $|e \cap A| = 1$. This situation is a special case of what we call a *splitting set system*, and the claim that $\mathrm{trk(M)} = |\mathrm{V}| - 1$ turns out to be a special case of a more general theorem.

Let $\mathcal{F}, \mathcal{C} \subseteq 2^V$ be any two families of subsets of V. For every $f \in \mathcal{F}$ and $c \in \mathcal{C}$, say that c splits f if $c \cap f \neq \emptyset$ and $\bar{c} \cap f \neq \emptyset$. Define the incidence matrix M by $M(f, c) = 1$ if c splits f, and $M(f, c) = 0$ otherwise.

CLAIM 5.1. *Let M be the incidence matrix as above. Then,* $\mathrm{trk(M)} \leq |\mathrm{V}| - 1$.

Proof. Let Q be a square $N \times N$ lower-triangular nonsingular minor of M. Let the rows be indexed by $\{f_i\}_{i=1}^N$, and the columns be indexed by $\{c_i\}_{i=1}^N$ in this order. It means, in particular, that c_i always splits f_i, but c_j with $j > i$, does not split f_i. Consider the partition of V, the underlying set induced by the family $\{c_{i+1}, \dots, c_N\}$. Since no c_j in it splits f_i, f_i must be contained in a single atom of the partition. Since c_i splits f_i, the partition induced by $\{c_i, c_{i+1}, \dots, c_N\}$ must strictly refine the previous partition. Therefore, the number of atoms in the partition induced by $\{c_1, c_2, \dots, c_N\}$ is at least $N + 1$. But then $N + 1 \leq |V|$, and the statement follows. ◻

5.2. Cocircuits in matroids. Keeping in mind that cuts are the cocircuits 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 cocircuits, 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 are precisely the independent sets. Circuits are the minimal dependent sets. Cocircuits intersect every base, and are minimal with respect to this property. Equivalently, a cocircuit is the complement of a maximal nonspanning subset of X.

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

In linear matroids, the only type of matroids used in this paper, $\mathbb{F}$ is a field (only $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{F}_2$ will be considered), and the elements are the members of a fixed subset $X \subseteq \mathbb{F}^n$, viewed as n-dimensional vectors over $\mathbb{F}$. The bases are the maximal linearly independent subsets of X, the circuits are the minimal linearly dependent subsets of X, and the cocircuits are the minimal subsets of X that intersect every base. Equivalently, $C \subseteq X$ is a cocircuit if it is the complement of a maximal nonspanning subset of X, that is, $X \not\subseteq \mathrm{span}(\overline{\mathrm{C}})$, but adding any element from C to $\overline{C}$ makes it spanning X. In the language of linear algebra, this can be (with a minor effort) restated as follows: C is a cocircuit 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 versus the cocircuits of a matroid $\mathcal{M}$. Then,* $\mathrm{trk}(\mathrm{M}) \leq \mathrm{rank}(\mathcal{M})$.

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 cocircuit cannot be 1.

Let Q be a square $N \times N$ lower-triangular minor of M with 1's on the diagonal. Let the rows be indexed by $\{x_i\}_{i=1}^N$ and let the columns be indexed by $\{C_i\}_{i=1}^N$. 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 circuits. Indeed, assume by contradiction that it does contain a circuit Z, and let r be the largest index such that $x_r \in Z$. Consider the corresponding cocircuit 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 shall study *simplicial matroids*, a natural generalization of the graphical matroids to higher dimensions. Using the basic tools and notions of combinatorial topology, we shall study the structure of cocircuits/hypercuts in this setting, develop the theory of ℓ_1-volumes, and obtain a higher-dimensional analogue of Theorem 4.1. This, in turn, will be used to arrive at new results about Euclidean volumes.

There are two parallel theories that we consider, differing in the definition of the hypercut volumes, a generalization of the cut metrics. While in both cases the support of a hypercut volume is a cocircuit of the corresponding simplicial matroid, in the former case a hypercut volume is an $\mathbb{F}$-valued function, while in the latter case it is a $\{0, 1\}$-valued function.[6] The former theory, more analytic in nature, will be explored for $\mathbb{F} = \mathbb{R}$. The results in this case require some acquaintance with the basic notions of algebraic topology, and they have direct and interesting applications to some simply stated geometrical problems (Theorems 5.2 and 5.3 below).

The latter (combinatorial) theory will be explored for $\mathbb{F} = \mathbb{F}_2$. While it is consistent with the algebraic-topological approach, basic linear algebra and matroid theory are sufficient for its understanding. It yields slightly weaker results in the same vein, but provides a somewhat different and a more general perspective.

5.3. 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 a Euclidean d-dimensional volume. Assume that a finite set V is embedded in an arbitrary Euclidean space, and we are given black box access to the values of the induced Euclidean volumes of d-simplices over V. The goal is to estimate the average volume of all d-simplices on V, making a sublinear, that is $o(n^{d+1})$, number of *nonadaptive* queries. For $d = 1$ this was achieved in [6] (improving upon an earlier result of Indyk [17]), by constructing 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 for any $d \geq 1$.

THEOREM 5.2. *In order to $\frac{1+\epsilon}{1-\epsilon}$-approximate the average value of a Euclidean volume on V, it suffices to nonadaptively query $O(n^d/\epsilon^2)$ predefined $(d + 1)$-tuples, and output a (predefined) linear combination of the obtained values.*

[6] For $\mathbb{F} = \mathbb{F}_2$ there is no distinction.

An additional value of this result stems from the observation that if one were to approximate *squares* of Euclidean distances (a larger class, strictly containing the Euclidean distances) in this manner, the solutions in the $d = 1$ case are precisely the linear size sparsifiers of a complete graph in the sense of [7]. As argued there, these objects are quite akin to the constant degree expanders. Since Theorem 5.2 works for squares of Euclidean volumes as well, the structure and the properties of the resulting objects for $d > 1$ are most intriguing in the context of expansion of simplicial complexes.[7]

The second application is as follows.

THEOREM 5.3. *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.*

We note that Theorem 5.2 and Theorem 5.3 can be viewed as results about *core sets*, as the multiplicative sparsifiers are called in the computational geometry community (cf. [16]). To our knowledge, these results are new, and do not follow directly from existing tools, e.g., of [12].

6. Finite volume spaces.

6.1. Preliminaries.

6.1.1. Basic definitions and facts from combinatorial topology. Fix $\mathbb{F}$ to be a field; in this paper it shall always be either $\mathbb{R}$ or $\mathbb{F}_2$. 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 $\mathbb{R}$; over $\mathbb{F}_2$ there is a unique orientation), and it induces (in a standard manner described in detail, e.g., in [28]) the orientation of its subsimplices. The key feature of this orientation, 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, it miraculously holds that $M_{d-1}M_d = 0$, provided that the d-simplices indexing the two matrices are identically oriented.

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), denoted ∂, and is called the *boundary operator*. The left action of M_d, denoted ∂^*, can be interpreted as a mapping of weighted sums of $(d-1)$-simplices (i.e., $(d-1)$-chains) to weighted sums of d-simplices (i.e., d-chains),[8] 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

[7] A different notion of high-dimensional expansion is studied in section 6.3.3.

[8] Alternatively, ∂^* can be interpreted as a mapping of the *weightings* of $(d-1)$-simplices (called $(d-1)$-*cochains*) to the weightings of d-simplices (d-cochains). Then, ∂^* assigns to a d-simplex the sum of weights of the $(d-1)$-simplices forming its boundary, and is akin to integration of differential forms from calculus.

d-coboundaries (seen as vectors of weights) is precisely the orthogonal complement of the space of d-cycles, i.e., any vector of weights corresponding to a d-coboundary sums up to 0 on every d-cycle.

Since $M_{d-1}M_d = 0$, it holds that $\partial_{d-1}\partial_d = 0$ and $\partial_d^*\partial_{d-1}^* = 0$. This, together with the (homological) connectivity of $K_n^{(d)}$, to be discussed in the next section, implies that $\operatorname{Ker}\partial_{d-1} = \operatorname{Im}\partial_d$, and $\operatorname{Ker}\partial_d^* = \operatorname{Im}\partial_{d-1}^*$. In particular, the space of $(d-1)$-cycles is spanned by the boundaries of d-simplices.

We conclude this section with a (nonstandard) definition of a minimal d-coboundary.

DEFINITION 6.1. *A d-coboundary $B = \sum_{\sigma\in C} b_\sigma\sigma$ supported on $C \subseteq K_n^{(d)}$ will be called* minimal *if there is no nontrivial d-coboundary whose support is strictly contained in C. Equivalently, B is a unique (up to scaling) nontrivial d-coboundary supported on C. This can be conveniently restated in the following form: B is minimal iff for all $\sigma, \tau, \partial(b_\sigma\tau - b_\tau\sigma) \in \operatorname{span}\{\partial\tau, \tau \notin \mathrm{C}\}$.*

Clearly, any d-coboundary is a sum of minimal d-coboundaries.

6.1.2. Basic combinatorial definitions and further facts. We introduce here some suggestive terminology from matroid theory, aimed at stressing the analogy between graphs and high-dimensional simplicial complices.

DEFINITION 6.2. *A set $S \subset K_n^{(d)}$ of d-simplices will be called* acyclic *if the corresponding $(d-1)$-chains $\{\partial\sigma \mid \sigma \in S\}$ are linearly independent over $\mathbb{F}$. A set that is not acyclic will be called* dependent.

* *A (spanning)* hypertree *$T \subset K_n^{(d)}$ is a maximal acyclic set.*

* *A set $K \subset K_n^{(d)}$ is* homologically connected[9] *or just* connected *if it contains a hypertree.*

* *A* hypercycle *$Z \subset K_n^{(d)}$ is a minimal dependent (i.e., nonacyclic) set.*

* *A* hypercut *$C \subset K_n^{(d)}$ is a complement of a maximal nonspanning set, i.e., $\operatorname{span}\{\partial\sigma \mid \sigma \notin \mathrm{C}\}$ 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 for the linear matroid whose elements are associated with the columns of M_d. In particular, the hypertrees correspond to bases, and the hypercuts correspond to cocircuits of this matroid. All d-hypertrees are of the same size, the rank of M_d. Since the set of all the d-simplices containing a fixed vertex v of V is acyclic (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 and hypertrees, and formalizes the intuitive correspondence between hypercuts and coboundaries. These facts are either standard in matroid theory, or are based on elementary linear algebra. The details are omitted.

LEMMA 6.3.

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 *hypercycle 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 intersects every hypertree (and*

[9] From the algebraic-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.

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. *A set $C \subset K_n^{(d)}$ is a d-hypercut iff it is the support of a minimal nontrivial d-coboundary.*

Here is perhaps the place to observe that every $\mathbb{F}_2$-hypertree is necessarily an $\mathbb{R}$-hypertree, but not the other way around. This is because the equation involved in the definition of these terms is an equation over vectors with entries $0, \pm 1$, and thus a dependence over $\mathbb{R}$ implies a dependence over $\mathbb{F}_2$.

A special class of d-hypercuts are the *geometric hypercuts*, a weaker generalization of the graph cuts, defined as follows.

DEFINITION 6.4. *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.*[10] *(An equivalent definition where ϕ maps V into $\mathbb{R}^d$ instead of S^{d-1} will also be frequently used.)*

A familiar example (and a proper subfamily) of the geometric hypercuts are the *partition hypercuts*. They are employed in the famous combinatorial proof[11] of the Sperner lemma (see, e.g., [28]). The definition is as follows. Let $\mathcal{P} = \{V_1 \cup, \dots, V_{d+1}\}$ be a partition of V to $(d+1)$ disjoint nonempty parts. The corresponding partition hypercut is defined by $C_{\mathcal{P}} = \{\sigma \in K_n^{(d)} \mid ; |\sigma \cap A_i| = 1, i = 1, 2, \dots, d+1\}$. Note that all graphical cuts are of this kind.

LEMMA 6.5. *Geometric d-hypercuts are indeed d-hypercuts both over $\mathbb{R}$ and over $\mathbb{F}_2$.*

Proof. 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 $d = 1$, counterclockwise for $d = 2$, etc. (See, e.g., [28] for the precise definition.) By Lemma 6.3, it will suffice to show that the chain B_C that assigns weight 1 to all $\sigma \in C$, and 0 to all $\sigma \notin C$, 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. The direct proof is based on the (well-known) existence of a $(d-1)$-form on $\mathbb{R}^d$ whose integral on the boundary of a (geometrical) d-simplex is 1 if it contains the origin, and 0 otherwise. Consider, for example, the case $d = 2$. Then, the path integral over the counterclockwise oriented one-dimensional boundary of a triangle 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, as required. Assigning to the oriented 1-simplices in $K_n^{(1)}$ the value of the above path integral on their geometrical realization, we obtain a 1-chain that under the action of ∂_2^* yields B_C.

Following the dual argument, applicable both to $\mathbb{R}$ and to $\mathbb{F}_2$, in order to show that B_C is coboundary it suffices to show that it sums up to 0 on the (oriented) boundary of any $(d+1)$-simplex ζ. Indeed, recall that the linear space of d-coboundaries is precisely the orthogonal complement of the space of d-cycles, and the latter is spanned by the boundaries of $(d+1)$-simplices. The key observation is that the origin is always contained in either zero or two d-simplices belonging to the boundary of ζ.[12] In the former case we are done; in the latter case, one of these simplices is necessarily oriented

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

[11] In our present terminology, the key argument in this proof is that over $\mathbb{F}_2$, a partition hypercut is orthogonal to a d-cycle.

[12] For an interesting characterization of subsets of $K_n^{(2)}$ with this property over $\mathbb{F}_2$, see [13].

in a manner consistent with our orientation (induced by the positive 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, one needs to show that that for any $\sigma_i, \sigma_j \in C$, 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 cylindrical 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. Choose σ_i from first copy, and σ_j from the second copy. Then, by the general position argument, the boundary of $\mathrm{conv}(\sigma_i \cup \sigma_j) \subset \mathbb{R}^{d+1}$ 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 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 a 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 the 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 to the applications faster. 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. The two theories (analytic versus combinatorial), differ in the choice of their basic objects. In the former, these are the (topological) d-cycles and minimal d-coboundaries. In the latter, these are the (combinatorial) d-hypercycles and d-hypercuts.

For $d = 1$ both theories coincide over any $\mathbb{F}$, forming the theory of metric spaces. They also coincide for any d over $\mathbb{F}_2$.

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 $Z = \sigma + \sum a_i\sigma_i$, it holds that

$$\mathrm{vol}^{(d)}(\sigma) \ \leq\ \sum |\mathrm{a_i}|\, \mathrm{vol}^{(d)}(\sigma_\mathrm{i}). \tag{6.1}$$

We note that for $d > 1$, unlike the one-dimensional case when the triangle inequality suffices, condition (**) cannot be replaced by a requirement on cycles of bounded size.[13]

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 $\mathrm{conv}(\phi(\sigma))$. As we shall see soon, Euclidean volumes are indeed volumes according to the above definition.

Definition 6.6.

1. *The absolute values of the weights of a real d-coboundary of $K_n^{(d)}$ will be called an* integral[14] *d-volume on V. For $d = 1$ these are precisely the line metrics.*

[13] For example, consider a realization of a large 4-connected triangulated planar graph on the sphere. Take all but one triangle σ of this realization to be of value 0, while all other triangles (including σ) to be of value 1. This is not a legitimate volume, as σ belongs to a cycle (the sphere triangulation) in which every other triangle has value 0. It can be argued that no shorter cycle that violates condition (**) exists.

[14] From "integration," not from "integer." Integral d-volumes are related to integrals of real differential $(d-1)$-forms.

2. *The absolute values of the weights of a* minimal *real d-coboundary will be called a* hypercut *d-volume on V, a special case of an integral volume.*

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

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

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

THEOREM 6.7. *The extremal rays of the d-hypercut cone are precisely the d-hypercut volumes.*

Proof. Consider an integral d-volume vol, let B be the corresponding d-coboundary, and let S be the set of all d-simplices σ such that $\mathrm{vol}(\sigma) = 0$. By a standard variational argument, vol is extremal in the d-hypercut cone (i.e., cannot be represented as a sum of two different integral volumes) iff B is a unique (up to scaling) nonzero d-coboundary that vanishes on S. But this, according to Definition 6.3, means that vol is a d-hypercut volume. $\square$

Next, we relate the abstract theory to geometric hypercut volumes and Euclidean volumes. To begin with, recall that the proof of Lemma 6.5 shows that the (minimal, real) d-coboundary B_C corresponding to a geometric hypercut C has weights 0/1 for a suitable orientation of simplices. Thus, the hypercut volume vol_C corresponding to C is an integral volume with $\mathrm{vol}(\sigma) = 1$ for $\sigma \in C$, and $\mathrm{vol}(\sigma) = 0$ otherwise.

LEMMA 6.8.

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, as well as (pointwise) squares of Euclidean d-volumes, are of negative type.*

Proof. 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-coboundary.

This could be proved directly by an argument sketched in the beginning of the proof of Lemma 6.5, by exhibiting a (well-known) $(d-1)$-differential form whose integral over ∂S is the Euclidean volume of S for any nice $S \subset \mathbb{R}^d$, in particular for any d-simplex realized in $\mathbb{R}^d$. We shall adopt a more transparent approach. Let $C(p)$ be the geometric 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 let $B_p = \sum_{\sigma\in C(p)} \sigma$ be the corresponding minimum coboundary, as in the proof of Lemma 6.5. Observe that the integral over $\mathbb{R}^d$ of the B_p's is $\sum_{\sigma\in K_n^{(d)}} \mathrm{vol}(\sigma)\sigma$. 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 the second statement reduces to the first one.

For the third statement, the correctness for the geometric d-hypercut volumes is obvious, as they only take values $\{0, 1\}$. For Euclidean volumes, by a well-known corollary to the 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 reduces to $n = d$, and further, in view of the first statement, to the geometrical d-hypercut volumes. $\square$

6.2.1. Applications. We start with proving the following lemma.

LEMMA 6.9. *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 rows are $\{x_f^T\}_{f\in\mathcal{F}}$, we conclude that (pointwise) $\sqrt{N} = \pm X^T M_d$. Hence, $\mathrm{rank}^*(\mathrm{N}) \leq \mathrm{rank}(\mathrm{M_d}) = \binom{\mathrm{n}-1}{\mathrm{d}}$. $\square$

Since by Lemma 6.8 geometric hypercut d-volumes are of negative type, an immediate corollary to Lemma 6.9, in view of Theorem 3.2, is the following.

COROLLARY 6.10. *Let* vol *be a weighted sum of geometric hypercut d-volumes. Then,* vol *can be (efficiently) multiplicatively $\frac{1+\epsilon}{1-\epsilon}$-approximated by a weighted sum of at most $\binom{n-1}{d}$ geometric hypercut d-volumes appearing in the the original sum.* We are now ready to prove Theorem 5.3 stated in section 5.3.

Proof of Theorem 5.3. Since planar polygons can be triangulated, it suffices to address the case when P is a triangle. Drawing all lines through the pairs of points in S, one gets $O(n^2)$ lines, which in turn partition the plane into $O(n^4)$ cells. Putting in the interior of each cell a single point p with the associated weight α_p being the area of this cell, one gets an initial sampling set that does not produce any errors, but it is too big. Associate with each such point p the corresponding geometric hypercut obtained by treating p as the origin. Let $\mathrm{vol_p}$ be the corresponding (geometric hypercut) 2-volume $\mathrm{vol_p}$. Now, the function that we aim to approximate, i.e., mapping each triangle over S to its Euclidean area, can be expressed in the form $\sum \alpha_p \mathrm{vol_p}$. Applying Corollary 6.10 to this sum yields the desired construction. $\square$

Next, we proceed towards establishing Theorem 5.2 stated in section 5.3.

THEOREM 6.11. *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 one 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 a nonconnected subset of simplices, and hence by any small sample.

Proof. A proof based on Lemma 6.9 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 3.3 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 $|\mathrm{supp(w^*)}| = \mathrm{O(rank(M_d)}/\epsilon^2)$, and $x^T(M_dWM_d^T)x$ and $x^T(M_dW^*M_d^T)x$ differ by at most $(1\pm\epsilon)$ multiplicative factor. Since $\mathrm{rank(M_d)} = \binom{\mathrm{n}-1}{\mathrm{d}}$, we arrive at the desired conclusion. □

Since Euclidean volumes are, by Lemma 6.8, of nonnegative type, this implies Theorem 5.2.

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 weightings is quite intriguing, potentially useful, and certainly deserves further study. The expansion in simplicial complexes will reoccur also in the next section in a different context.

6.3. Volumes over $\mathbb{F}_2$. The theory of finite d-volumes over $\mathbb{F}_2$ is more combinatorial in nature than the theory over $\mathbb{R}$. It provides a clean and often appealing generalization of related graph theoretic and metric theoretic concepts and methods. For example, it leads to a meaningful generalization of graph expansion, which, together with a generalization of Poincaré forms, allows one to 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, allowing occasional excursions to matters not immediately related to sparsification. It should also be mentioned that all the applications obtained so far can be proved by using the theory over $\mathbb{F}_2$ as well, up to multiplicative logarithmic factors.

Working over $\mathbb{F}_2$, it will be convenient to treat the d-chains simply as subcomplexes 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.

6.3.1. The structure of hypercuts. Here we present some nonstandard combinatorial notions and results to be used later in this section.

In the $\mathbb{F}_2$ framework, the general theory of section 6.1 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 even.*[15] *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 a 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 $B \subseteq K_n^{(d)}$.*

Proof. Let B' be the d-coboundary induced by $\mathrm{link_v(B)}$. Consider first a d-simplex σ that contains v. Since $\mathrm{link_v(B)}$ lacks all the $(d-1)$-faces of σ containing v, and contains the remaining $(d-1)$-face $\tau = \sigma - \{\mathrm{v}\}$ iff $\sigma \in B$, the definition of coboundary B' implies that $\sigma \in B'$ iff $\sigma \in B$. Consider next a d-simplex $\sigma = (v_1, v_2, \ldots, v_{d+1})$ that

[15]This is true for any binary matroid, and not only a simplicial one, where in the language of binary matroids a coboundary is just a cocycle.

does not contain v. Consider the d-boundary of the $(d+1)$-simplex $(v_1, v_2, \ldots, v_{d+1}, \mathrm{v})$. It is a cycle, and all its d-faces with exception of σ contain v. Since B' and B agree on all these faces, the parity argument from Claim 6.1 implies that they agree on σ as well. Thus, $B' = B$. $\square$

Next, we characterize these one-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 6.12. *Let B be a 2-coboundary, and let $G = \mathrm{link}_{\mathrm{v}}(\mathrm{B})$ be its link with respect to an arbitrary vertex* v*. Then, B is a 2-hypercut iff G is $\wedge$-connected.*

Proof. Let x be a vector with coordinates indexed by the edges of K_n. Consider the following system of equations in x. For each e containing the vertex v, $x_e = 0$; for each triangle $\sigma \notin B$, $\sum_{e\in\sigma} x_e = 0$. We claim that this system of equations has a unique nontrivial solution iff B is a hypercut. Indeed, by definition, $x = 1_{E(G)}$ is one nontrivial solution, as $1_{E(G)}$ induces B. The existence of another nontrivial solution x' is equivalent to existence of a nontrivial 2-coboundary B' (induced by x') strictly contained in B, as on every triangle $\sigma \in \bar{B}$, x' must sum to 0. Recall that different links define different coboundaries.

Assigning the forced value 0 to all x_e where e contains v, and to all $x_{(a,b)}$ where the triangle $\{a, b, \mathrm{v}\} \notin \mathrm{B}$, we arrive at the equivalent system of equations $x_{(a,b)} + x_{(b,c)} = 0$ whenever $a, b, c \in V - \{\mathrm{v}\}$, and (a, b), $(b, c) \in E(G)$; $(a, c) \notin E(G)$. Thus, the edges in the same $\wedge$-equivalence class must be assigned the same value, but there is no restrictions for edges in different $\wedge$-equivalence classes. We conclude that there is a unique solution iff there is one $\wedge$-equivalence class, i.e., G is $\wedge$-connected. $\square$

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 values. 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 $\binom{n}{3} - O(n^2)$.*

Proof. We start with the first statement, and prove it by induction on n, d. Since the minimum coboundary is a hypercut, it suffices to prove it for coboundaries. The statement clearly holds for $d = 1$ and for $n = d + 1$. Assume that the statement is true for all pairs (n', d') where $n' < n$, $d' \leq d$. Let C be a nonempty d-coboundary, and let v be a vertex. Consider $\mathrm{link}_{\mathrm{v}}(\mathrm{C})$. Then, $|C| = |C'| + |\mathrm{link}_{\mathrm{v}}(\mathrm{C})|$, where C' is the restriction of C on $V - \{\mathrm{v}\}$, clearly a d-coboundary of $K_{n-1}^{(d)}$. Recall that $\mathrm{link}_{\mathrm{v}}(\mathrm{C})$ cannot be empty. If $C' \neq \emptyset$, then by inductive hypothesis $|C| \geq (n-1-d)+1 = n-d$. Otherwise, by the previous discussion, $\mathrm{link}_{\mathrm{v}}(\mathrm{C})$ must be a $(d - 1)$-coboundary of $K_{n-1}^{(d-1)}$, and thus by inductive hypothesis $|C| = |\mathrm{link}_{\mathrm{v}}(\mathrm{C})| \geq (\mathrm{n}-1) - (\mathrm{d}-1) = \mathrm{n}-\mathrm{d}$. The bound is tight, as shown by a d-hypercut that consists of all the d-simplices containing a fixed $(d - 1)$-simplex τ.

For the second statement, consider the 2-coboundary B of $K_n^{(2)}$ whose link is a complete graph on $n-1$ points excluding a Hamiltonian cycle. It is easy to verify that the criterion of Theorem 6.12 holds, and thus B is a 2-hypercut. A simple calculation shows that for $n \geq 5$, $|B| = \binom{n}{3} - (n-1)(n-4)$. $\square$

We conclude this section with a result about 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 6.13. *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 $|\mathrm{link_v(C)}| \le \mathrm{(d+1)}\alpha$. 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 nontrivial component containing at most 3α edges and $3\alpha + 1$ vertices. Thus, the number of such G's is at most

$$\binom{n}{3\alpha+1}\binom{\binom{3\alpha+1}{2}}{3\alpha} \le \left(\frac{en}{3\alpha+1}\right)^{3\alpha+1} \cdot \left(\frac{e \cdot 3\alpha(3\alpha+1)}{2 \cdot 3\alpha}\right)^{3\alpha} \le (4n)^{3\alpha+1}. \qquad \square$$

To conclude this section, let us but mention without elaborating the two-graphs of Seidel [31], that are clearly related to the $\mathbb{F}_2$ 2-hypercuts, and may potentially contribute to the future study in this direction.

6.3.2. Volumes: Basics. Volumes over $\mathbb{F}_2$ are defined as nonnegative real functions on $K_n^{(\le d)}$, analogously to the definition of real volumes in section 6.2. The generalized triangle inequality is the following.

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

$$\mathrm{vol}(\sigma) \le \sum \mathrm{vol}(\sigma_\mathrm{i}). \tag{6.2}$$

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 $\mathrm{vol_X}(\sigma) = \min_{\mathrm{D}_\sigma \subseteq \mathrm{X}} \sum_{\sigma' \in \mathrm{D}_\sigma} \mathrm{w}_{\sigma'}$, 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 is of *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 where all but one $\sigma \in Z$ have value 0. By Claim 6.1, such Z does not exist.

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 volume.

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 6.14. *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.*

Proof. Let $\mathrm{vol}^{(\mathrm{d})}$ be a hypercut d-volume function defined by a hypercut C. Assume that $\mathrm{vol}^{(\mathrm{d})} = \mathrm{vol}_1^{(\mathrm{d})} + \mathrm{vol}_2^{(\mathrm{d})}$. Consider $\mathrm{vol}_1^{(\mathrm{d})}$. It must be 0 outside of C. By Claim 6.1, for any two $\sigma, \sigma' \in C$ there exists a cycle $Z = Z_{\sigma,\sigma'}$ such that $Z \cap C = \{\sigma, \sigma'\}$. Since all the d-simplices in $\overline{C}$ have volume 0, the generalized triangle inequality implies that $\mathrm{vol}_1^{(\mathrm{d})}(\sigma) = \mathrm{vol}_1^{(\mathrm{d})}(\sigma')$. Thus, $\mathrm{vol}_1^{(\mathrm{d})} = \alpha \cdot \mathrm{vol}^{(\mathrm{d})}$, as claimed.

For the other direction, consider an extremal $0/1$ d-volume function $\mathrm{vol}^{(\mathrm{d})}$. Define $C \subset K_n^{(d)}$ as $C = \{\sigma \mid \mathrm{vol}^{(\mathrm{d})}(\sigma) = 1\}$. Then C intersects every spanning tree T, as otherwise, if $C \cap T = \emptyset$ and $\sigma \in C$, then $T \cup \{\sigma\}$ contains a cycle that would intersect C in σ in contradiction to (6.2). Thus C contains a coboundary. Moreover, if C is not a minimal coboundary, let $C' \subsetneq C$ be a hypercut (minimal coboundary). Define $\mathrm{vol}_1^{(\mathrm{d})}$ and $\mathrm{vol}_2^{(\mathrm{d})}$ as follows. Outside of C both are 0. For $\sigma \in C \setminus C'$, $\mathrm{vol}_1^{(\mathrm{d})}(\sigma) = \mathrm{vol}_2^{(\mathrm{d})}(\sigma) = \frac{1}{2}$; for $\sigma \in C'$, $\mathrm{vol}_1^{(\mathrm{d})}(\sigma) = 0.4$, and $\mathrm{vol}_2^{(\mathrm{d})}(\sigma) = 0.6$. The definition of C' implies that both $\mathrm{vol}_1^{(\mathrm{d})}$ and $\mathrm{vol}_2^{(\mathrm{d})}$ are volume functions, contradicting the assumption that $\mathrm{vol}^{(\mathrm{d})}$ is extremal.

The second statement follows easily along the same line of reasoning. The support of any volume function approximating such $\mathrm{vol}^{(\mathrm{d})}$ must coincide with the support of $\mathrm{vol}^{(\mathrm{d})}$, and moreover, arguing as above, it must be constant on it. ∎

More can be said on the structure of $0/1$ volumes; e.g., every $0/1$ volume is a sum of hypercut volumes. A proof for this as well as further study of high-dimensional finite volumes will appear in a subsequent paper.

Much of the modern theory of finite metric spaces is devoted to the study of special metric classes that constitute a subcone 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 6.14 implies that any (closed) subcone 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 6.15. *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 subcone of d-volumes.

6.3.3. $\boldsymbol{\ell_1}$-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 geometric hypercut volumes. Since geometric hypercuts are $\mathbb{F}_2$ hypercuts (Lemma 6.5), we conclude that the Euclidean volumes are ℓ_1 over $\mathbb{F}_2$.

For hypertrees, we have the following theorem.

THEOREM 6.16. *Let T be a (spanning) d-hypertree with nonnegative weights on the d-simplices. Then, the lightest-cap d-volume $\mathrm{vol}_{\mathrm{T}}^{(\mathrm{d})}$ is ℓ_1.*

Proof. Recall the definition of $C_{T,\sigma}$ from Lemma 6.3. We claim that $\mathrm{vol}_{\mathrm{T}}^{(\mathrm{d})} = \sum_{\sigma \in \mathrm{T}} \mathrm{vol}_{\mathrm{C}_{\mathrm{T},\sigma}^{(\mathrm{d})}}$. For $\tau \in T$ this follows by definition, while for $\tau \notin T$, $\sum_{\sigma \in S} \mathrm{vol}_{\mathrm{C}_{\mathrm{T},\sigma}}^{(\mathrm{d})}(\tau)$ 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 6.17. *Any d-volume on V can be approximated by an ℓ_1 d-volume with distortion at most $\binom{n-1}{d}$.*

Proof. Let $\mathrm{vol}^{(\mathrm{d})}$ be a d-volume function on $K_n^{(d)}$, and let T be the minimum (spanning) hypertree with respect to $\mathrm{vol}^{(\mathrm{d})}$. Then, for $\sigma \in T$, $\mathrm{vol}_{\mathrm{T}}^{(\mathrm{d})}(\sigma) = \mathrm{vol}^{(\mathrm{d})}(\sigma)$. For $\sigma \notin S$, much like the minimum spanning tree 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 $1 + |T| \leq 1 + \binom{n-1}{d}$, the statement follows. ∎

While the upper bound on the distortion in Theorem 6.17 is probably too rough and the true exponent of n is probably smaller, we show, somewhat unexpectedly, that even for $d = 2$ the distortion can be as large as $\widetilde{\Omega}(n^{\frac{1}{5}})$. 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 6.18. *There exists a 2-volume function such that any ℓ_1-volume distorts it by at least $\widetilde{\Omega}(n^{1/5})$.*

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

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

$$F_K(\mathrm{vol}) = \frac{\sum_{\sigma \in \mathrm{K}} \mathrm{vol}(\sigma)}{\mathrm{av}(\mathrm{vol})}\,, \tag{6.3}$$

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

$$\mathrm{dist}(\mathrm{vol}_{\mathrm{K}} \hookrightarrow \ell_1) \;\geq\; \frac{\min_{\mathrm{vol} \in \ell_1} \mathrm{F}_{\mathrm{K}}(\mathrm{vol})}{\mathrm{F}_{\mathrm{K}}(\mathrm{vol}_{\mathrm{K}})}\,. \tag{6.4}$$

Since 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. Keeping in mind that K is unit weighted, (6.4) becomes:

$$\mathrm{dist}(\mathrm{vol}_{\mathrm{K}} \hookrightarrow \ell_1) \;\geq\; \mathrm{av}(\mathrm{vol}_{\mathrm{K}}) \cdot \min_{\mathrm{C:\ 2-hypercut}} \frac{|\mathrm{K} \cap \mathrm{C}|/|\mathrm{C}|}{|\mathrm{K}|/\binom{\mathrm{n}}{3}}. \tag{6.5}$$

Observe that for a graph G, the analogous expression of the second term in (6.5) is

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

which is the normalized edge expansion of G up to a factor of 2. By analogy, we have the following definition.[16]

Definition 6.19. *Let the* normalized (face) expansion *of* $K \subseteq K_n^{(2)}$ *be the value of*

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

Thus, 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 $K_n^{(2)}$.

We now formally proceed with the proof of Theorem 6.18.

Proof. Let $K_n^{(2)}(n,p)$ be the two-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 of the others. Theorem 6.18 follows from Lemmas 6.20, 6.21, and (6.5). ∎

Lemma 6.20. *For* $K \in K^{(2)}(n,p)$ *as above,* $\mathrm{av(vol_K)} \geq \tilde{\Omega}(\mathrm{n}^{1/5})$ *with probability* $1 - o(1)$.

Lemma 6.21. *The face expansion of* $K \in K^{(2)}(n,p)$ *is almost surely* ≥ 0.5.

Observe that Lemma 6.21 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 6.3). Thus, it strengthens the main result of [20] at the price of getting worse constants.

Before starting with the proof of Lemma 6.20, we need the following preparatory result.

Lemma 6.22. *Let* Z *be a* 2*-cycle; then* $|V(Z)| \leq |Z|/2 + 2$.[17]

Proof. Clearly, $\mathrm{link_v(Z)}$ is an Eulerian (one-dimensional) graph. As long as there is a vertex $\mathrm{v \in V(Z)}$ 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 $\mathrm{v, v_i, i = 1, \ldots r}$ for each A_i, and replace each original 2-simplex $\{\mathrm{v, x, y}\}$ containing v with a new 2-simplex $\{\mathrm{v_i, x, y}\}$ 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 fewer 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 satisfy

$$\chi(Z^*) = |V(Z^*)| - |E(Z^*)| + |Z^*| \ \leq \ 2. \tag{6.6}$$

Observe that every edge e in Z^* appears in exactly two faces, and thus $2|E(Z^*)| = 3|Z^*|$. Plugging this into (6.6) implies the lemma for $|V(Z^*)|$, and hence for $|V(Z)|$. ∎

We are now ready to address Lemma 6.20.

Proof of Lemma 6.20. By the Markov inequality K almost surely contains $o(n^3)$ 2-simplices, and thus $\mathrm{av(vol_K)}$ is determined by the 2-simplices $\sigma \notin K$. For each such

[16] A similar definition of face expansion was independently used in [26, 34]. See also the references therein.

[17] 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.

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 $\mathrm{av}(\mathrm{vol_K}) \geq \Omega(\lambda)$ (with high probability), 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 $o(n^3)$ (with high probability). Let N_λ be this number. Let n_k be the number of simple cycles of size exactly k in $K_n^{(2)}$. Then,

$$E[N_\lambda] = \sum_{k=4}^{\lambda} k \cdot n_k \cdot p^{k-1}(1-p). \tag{6.7}$$

Now, by Lemma 6.22, 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 (6.7), 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. $\square$

Next, we turn to Lemma 6.21, the expansion lemma.

Proof of Lemma 6.21. 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 $|C|$ independently and identically distributed Bernoulli variables, and its expectation is precisely $p|C|$. By Chernoff bound,

$$\Pr\left(\gamma_K(C) < 0.5\right) = \Pr\left(|K \cap C| < p \cdot |C|/2\right) \leq e^{-p \cdot |C|/8}.$$

Let m_s be the number of 2-hypercuts of size s in $K_n^{(2)}$. By Theorem 6.13, $m_s \leq (4n)^{1+3s/n}$. Thus, the union bound implies that the probability that a bad C exists is at most

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

6.3.4. 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 $\mathbb{R}^{\binom{n}{d+1}}$, the Caratheodory theorem implies that the hypercut dimension of any $\mathrm{vol}^{\mathrm{d}}$ is at most $\binom{n}{d+1}$. However, we seek a multiplicative approximation of a much smaller hypercut dimension. For volumes over the reals, Theorem 6.11 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 class of ℓ_1 d-volumes over $\mathbb{F}_2$, there is a similar phenomenon.

THEOREM 6.23. *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 $|\mathrm{supp}(\lambda)|$ is an upper bound on the hypercut-dimension of the respective d-volume. Thus, Theorem 2.3 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). Since the hypercuts $\mathcal{C}$ of $K_n^{(d)}$ are the co-circuits of the simplicial matroid corresponding to $K_n^{(d)}$ (see Definition 6.2 and the discussion immediately following it), Lemma 5.1 applies, implying that $\mathrm{trk(M)} \leq \binom{\mathrm{n}-1}{\mathrm{d}}$. ∎

6.4. Some additional remarks.

6.4.1. 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 an ℓ_1 d-volume (both real and $\mathbb{F}_2$) over $\mathcal{H}$.*

Proof. Assuming $|\mathcal{H}| = n$, we associate $H_i \in \mathcal{H}$ with the ith vertex of $K_n^{(d)}$. Arguing as in the proof of Lemma 6.8 (to be more specific, the second proof of item(1)), it suffices to show that for any generic point $p \in \mathbb{R}^d$, the subset $C_p \subseteq K_n^{(d)}$ of d-simplices corresponding to the $(d+1)$-tuples of $\mathcal{H}$ that contain p in their bounded cell, constitute a geometric hypercut. Indeed, given this, since each C_p defines a hypercut d-volume V_p, the above measure on the $(d+1)$-tuples of $\mathcal{H}$ corresponds to $\int_{p \in \mathbb{R}^d} V_p$, which is ℓ_1 by definition.

For each $H_i \in \mathcal{H}$, let $x_i^p \in H_i$ be the orthogonal projection of p on H_i. Obviously, p lies in the bounded cell of $\{H_{i_1}, \ldots, H_{i_{d+1}}\}$ iff p is in the convex hull of $\{x_{i_1}^p, \ldots, x_{i_{d+1}}^p\}$. Thus, C_p is precisely the geometric hypercut defined by $\{x_i^p\}_{i=1}^n \subset \mathbb{R}^d$ with respect to p. ∎

6.4.2. 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 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 (see [2]), 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 6.21 (with small local amendments) shows that for $d = 2$ there exist 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.

6.4.3. 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 $\mathrm{vol}_{\mathrm{K}}$, 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 $\mathrm{vol}_{\mathrm{K}}$, yet stretches it (in expectation) by only a constant factor.

Acknowledgments. We would like to thank for valuable comments and discussions Vladimir Hinich, Gil Kalai, Michael Langberg, Nati Linial, Avner Magen, Roy Meshulam, Tasos Sidiropoulos, Uli Wagner, and the anonymous referees of the SODA'12 version.

REFERENCES

[1] N. Alon and J. Spencer, *The Probabilistic Method*, John Wiley, New York, 1992.

[2] I. Althöfer, G. Das, D.P. Dobkin, D. Joseph, and J. Soares, *On sparse spanners of weighted graphs*, Discrete Comput. Geom., 9 (1993), pp. 81–100.

[3] A. Andoni, M.S. Charikar, O. Neiman, and H.L. Nguyen, *Near linear lower bounds for dimension reduction in* ℓ_1, FOCS, IEEE, Piscataway, NJ, 2011, pp. 315–323.

[4] E. Babson, C. Hoffman, and M. Kahle, *The Fundamental Group of Random* 2*-Complexes*, preprint, arXiv0711.2704, 2008.

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

[6] K. Barhum, O. Goldreich, and A. Shraibman, *On approximating the average distance between points*, in Proceedings of APPRPX'07, Springer-Verlag, Berlin, 2007, pp. 296–310.

[7] J.D. Batson, D.A. Spielman, and N. Srivastava, *Twice-Ramanujan sparsifiers*, in STOC'09, 2009, ACM, New York, pp. 255–262.

[8] A. Benczúr and D.R. Karger, *Augmenting undirected edge connectivity in* $\tilde{O}(n^2)$ *time*, J. Algorithms, 37 (2000), pp. 2–36.

[9] J. Bourgain, *On Lipschitz embeddings of finite metric spaces in Hilbert space*, Israel J. Math., 52 (1985), pp. 46–52.

[10] B. Brinkman and M. Charikar, *On the impossibility of dimension reduction in* ℓ_1, J. ACM, 52 (2005), pp. 766–788.

[11] U. Feige, *Approximating the bandwidth via volume respecting embeddings*, J. Comput. System Sci., 6 (2000), pp. 510–539.

[12] D. Feldman and M. Langberg, *A unified framework for approximating and clustering data*, in Proceedings of 43rd Annual ACM Symposium on Theory of Computing, ACM, New York, 2011, pp. 569–578.

[13] P. Frankl and Z. Füredi, *An exact result for* 3*-graphs*, Discrete Math., 50 (1984), pp. 323–328.

[14] W.S. Fung, R. Hariharan, N. Harvey, and D. Debmalya, *A general framework for graph sparsification*, in STOC'11, ACM, New York, 2011, pp. 71–80.

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

[16] S. Har Peled, *Geometric approximation algorithms*, Mathematical Surveys and Monographs 173, AMS, Providence, RI, 2011.

[17] P. Indyk, *Sublinear time algorithms for metric space problems*, in STOC'99, ACM, New York, 1999, pp. 428–434.

[18] J. Lee and A. Naor, *Embedding of the diamond graph in L_p and dimension reduction in ℓ_1*, Geom. Funct. Anal., 14 (2004), pp. 745–747.

[19] Y. Li, P.M. Long, and A. Srinivasan, *Improved bounds on the sample complexity of learning*, J. Comput. System Sci., 62 (2001), pp. 516–527.

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

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

[22] A. Magen and A. Zouzias, *Near optimal dimensionality reductions that preserve volumes*, in Approximation, Randomization and Combinatorial Optimization Algorithms and Techniques, Lecture Notes in Comput. Sci. 5171, 2008, pp. 523–534.

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

[24] J. Matousek, *Geometric Discrepancy*, 2nd ed., Berlin, Springer, 2009.

[25] J. Matousek, M. Tancer, and U. Wagner, *Hardness of embedding simplicial complexes in $\mathbb{R}^d$*, in Proceedings of SODA, ACM-SIAM, 2009, pp. 855–864.

[26] J. Matousek and U. Wagner, *On Gromov's Method of Selecting Heavily Covered Points*, preprint, arXiv:1102.3515, 2011.

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

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

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

[30] G. Schechtman, *More on embedding subspaces of l_p in l_1^r*, Compositio Math., 61 (1987), pp. 159–169.

[31] J.J. Seidel, *A survey of two-graphs*, in Colloquio Internazionale sulle Teorie Combinatorie, Vol. I, Atti dei Convegni Lincei 17, 1 Accademia Nazionale dei Lincei, Rome, 1976, pp. 481–511.

[32] M. Talagrand, *Sharper bounds for Gaussian and empirical processes*, Ann. Probab., 22 (1994), pp. 28–76.

[33] M. Talagrand, *Embedding subspaces of ℓ_1 into ℓ_1^n*, Proc. Amer. Math. Soc., 108 (1990), pp. 363–369.

[34] U. Wagner, *Minors in random and expanding hypergraphs*, in 27th SoCG', ACM, New York, 2011, pp. 351–360.