Discrepancy, Graphs, and the Kadison–Singer Problem

N Srivastava

Discrepancy theory seeks to understand how well a continuous object can be approximated by a discrete one, with respect to some measure of uniformity. For instance, a celebrated result due to Spencer says that given any set family $S_1, \ldots, S_n \subset [n]$, it is possible to colour the elements of $[n]$ Red and Blue in a manner that:

$$\forall S_i \qquad \left| |S_i \cap R| - \frac{|S_i|}{2} \right| \leq 3\sqrt{n},$$

where $R \subset [n]$ denotes the set of red elements. In other words, it is possible to partition $[n]$ into two subsets so that this partition is very close to balanced on *each one* of the test sets S_i. Note that a "continuous" partition which splits each element exactly in half will be exactly balanced on each S_i; the content of Spencer's theorem is that we can get very close to this ideal situation with an actual, discrete partition which respects the wholeness of each element.

Spencer's theorem and its variants have had applications in approximation algorithms, numerical integration, and many other areas. In this post I will describe a new discrepancy theorem [1] due to Adam Marcus, Dan Spielman, and myself, which also seems to have many applications. The theorem is about "uniformly" partitioning sets of vectors in $\mathbb{R}^n$ and says the following:

Theorem 1. (*implied by Corollary 1.3 in [1]*) *Given vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ satisfying $\|v_i\|^2 \leq \alpha$ and*

$$\sum_{i=1}^{m} \langle v_i, x\rangle^2 = 1 \qquad \forall \|x\| = 1, \tag{1}$$

there exists a partition $T_1 \cup T_2 = [m]$ satisfying

$$\left| \sum_{i \in T_j} \langle v_i, x\rangle^2 - \frac{1}{2} \right| \leq 5\sqrt{\alpha} \qquad \forall \|x\| = 1.$$

Thus, instead of being nearly balanced with respect to a finite set family as in Spencer's setting, we require our partition of $v_1, \ldots, v_m$ to be nearly balanced with respect to the infinite set of test vectors $\|x\| = 1$. In this context, "nearly balanced" means that about half of the quadratic form ("energy") of the $v_1, \ldots, v_m$ in direction x comes from T_1 (and the rest, which must also be about half, comes from T_2). We will henceforth refer to the maximum deviation from perfect balance (i.e. $1/2$) over all x as the *discrepancy* of a partition. Note that every partition has discrepancy at most $1/2$, so the guarantee of the theorem is nontrivial whenever $5\sqrt{\alpha} < 1/2$.

This type of theorem was conjectured to hold by Nik Weaver [2], with any constant strictly less than $1/2$ (independent of m and n) in place of $5\sqrt{\alpha}$. The reason he was interested in it is that he showed it implies a positive solution to the so-called Kadison–Singer (KS) problem, a central question in operator theory which had been open since 1959. KS was itself motivated by a basic question about the mathematical foundations of quantum mechanics — check out the blog **soulphysics** [3] for an intuitive description of its physical significance. If you want to know exactly what the statement of KS is and how it can be reduced to finite-dimensional vector discrepancy statements similar to Theorem 1, I highly recommend the accessible and self-contained survey article written recently by Nick Harvey [4].

In the rest of the post I will try to demystify what Theorem 1 is about, say a bit about the proof, and describe a simple application to graph theory.

1. What the Theorem Says

Let's examine how restrictive the hypotheses of Theorem 1 are. To see that some bound on the norms of the v_i is necessary for the conclusion of the theorem to hold, consider an example where one vector has large norm, say $\|v_1\|^2 = 3/4$. In any partition of $v_1, \ldots, v_m$, one of the sets, say T_1, will contain v_1. If we now examine the quadratic form in the direction $x = v_1/\|v_1\|$, we see that

$$\sum_{i \in T_1} \langle v_i, x\rangle^2 \geq \|v_1\|^2 = 3/4,$$

so this partition has discrepancy at least $1/4$. The problem is that v_1 by itself accounts for significantly more than half of the quadratic form in direction x, and there is no way to get closer to half without splitting the vector.

Another instructive example is the one-dimensional instance $v_1, \ldots, v_m \in \mathbb{R}^1$, with $v_i^2 = 1/m = \alpha$ for all i and m odd. Here, the larger side of any partition must have $\sum_{i \in T_j} \langle v_i, e_1 \rangle^2 = \sum_{i \in T_j} v_i^2 \geq 1/2 + \alpha/2$, leading to a discrepancy of at least $\alpha/2$.

In general, the above examples show that the presence of large vectors is an obstruction to the existence of a low discrepancy partition. Theorem 1 shows that this is the only obstruction, and if all the vectors have sufficiently small norm then an appropriately low discrepancy partition must exist. It is worth mentioning that by a more sophisticated example than the ones above, Weaver has shown that the $O(\sqrt{\alpha})$ dependence in Theorem 1 cannot be improved.

Let us now consider the "isotropy" condition (1). This may seem like a very strong requirement at first, but it is in fact best viewed as a normalisation condition. To see why, let us first write the theorem using matrix notation. It says that given vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ with $\|v_i\|^2 \leq \alpha$ and

$$\sum_{i=1}^{m} v_i v_i^T = I,$$

there is a partition $T_1 \cup T_2 = [m]$ satisfying

$$\left(\frac{1}{2} - 5\sqrt{\alpha}\right) I \preceq \sum_{i \in T_j} v_i v_i^T \preceq \left(\frac{1}{2} + 5\sqrt{\alpha}\right) I,$$

where $A \preceq B$ means that

$$x^T A x \leq x^T B x \qquad \forall x \in \mathbb{R}^n,$$

or equivalently that $B - A$ is positive semidefinite.

Now suppose I am given some arbitrary vectors $w_1, \ldots, w_m \in \mathbb{R}^n$, which are not necessarily isotropic. Assume that the span of the w_i is $\mathbb{R}^n$ (otherwise, change the basis and write them as vectors in some lower-dimensional $\mathbb{R}^k$). This implies that the positive semidefinite matrix $W := \sum_{i=1}^{m} w_i w_i^T$ is invertible, and therefore has a negative square root $W^{-1/2}$. Now consider the "normalised" vectors

$$v_i = W^{-1/2} w_i, \qquad i = 1, \ldots, m$$

and observe that

$$\sum_{i=1}^{m} v_i v_i^T = W^{-1/2} \left(\sum_{i=1}^{m} w_i w_i^T \right) W^{-1/2} = I,$$

so these vectors are isotropic. The normalised vectors have norms

$$\|v_i\|^2 = \|W^{-1/2} w_i\|^2.$$

To better grasp what these norms mean, we can write:

$$\begin{aligned} \|W^{-1/2} w_i\|^2 &= \sup_{x \neq 0} \frac{\langle x, W^{-1/2} w_i \rangle^2}{x^T x} \qquad (*) \\ &= \sup_{y = W^{1/2} x \neq 0} \frac{\langle W^{1/2} y, W^{-1/2} w_i \rangle^2}{y^T W y} \\ &= \sup_{y \neq 0} \frac{\langle y, w_i \rangle^2}{\sum_i \langle y, w_i \rangle^2}. \end{aligned}$$

Thus, the norms $\|v_i\|^2$ measure the maximum fraction of the quadratic form of W that a single vector w_i can be responsible for — exactly the critical quantity in the example at the beginning of this section.

These numbers are sometimes called "leverage scores" in numerical linear algebra and statistics. As long as the leverage scores are bounded by α, we can apply Theorem 1 to $v_1, \ldots, v_m$ to obtain a partition satisfying

$$\begin{aligned} \left(\frac{1}{2} - 5\sqrt{\alpha}\right) I \preceq \sum_{i \in T_j} v_i v_i^T &= W^{-1/2} \left(\sum_{i \in T_j} w_i w_i^T \right) W^{-1/2} \\ &\preceq \left(\frac{1}{2} + 5\sqrt{\alpha}\right) I. \end{aligned}$$

We now appeal to the fact that $A \preceq B$ iff $MAM \preceq MBM$, for any invertible M (this amounts to a simple change of variables similar to what we did in (*)). Multiplying by $W^{1/2}$ on both sides, we find that the partition $T_1 \cup T_2$ guaranteed by Theorem 1 satisfies:

$$\begin{aligned} \left(\frac{1}{2} - 5\sqrt{\alpha}\right) \left(\sum_{i=1}^{m} w_i w_i^T \right) &\preceq \sum_{i \in T_j} w_i w_i^T \\ &\preceq \left(\frac{1}{2} + 5\sqrt{\alpha}\right) \left(\sum_{i=1}^{m} w_i w_i^T \right). \qquad (2) \end{aligned}$$

Thus, we have the following restatement of Theorem 1:

Theorem 2. *Given any vectors $w_1, \ldots, w_m \in \mathbb{R}^n$, there is a partition $T_1 \cup T_2 = [m]$ such that* (2) *holds with $\alpha = \max_i w_i^T (\sum_{i=1}^{m} w_i w_i^T)^+ w_i$.*

Note that we have used the **pseudoinverse** instead of the usual inverse to handle the case where the vectors do not span $\mathbb{R}^n$.

For those who do not like to think about sums of rank one matrices (I know you're out there), Theorem 2 may be restated very concretely as:

Theorem 3. *Any matrix $B_{m\times n}$ whose rows w_i^T have leverage scores $w_i^T(B^TB)^+w_i$ bounded by α can be partitioned into two row submatrices B_1 and B_2 so that for all $x \in \mathbb{R}^n$:*

$$(1/2 - 5\sqrt{\alpha})\|Bx\|^2 \le \|B_jx\|^2 \le (1/2 + 5\sqrt{\alpha})\|Bx\|^2.$$

The reason this theorem is powerful is that lots of diverse objects can be encoded as quadratic forms of matrices. We will see one such application later in the post.

2. Matrix Chernoff Bounds and Interlacing Polynomials

Let me quickly say a bit about the proof of Theorem 1. One reasonable way to try to find a good partition $T_1 \cup T_2$ is randomly, and indeed this strategy is successful to a certain extent. The tool that we use to analyse a random partition is the so-called "Matrix Chernoff Bound", developed and refined by Lust-Piquard, Rudelson, Ahlswede-Winter, Tropp, and others. The variant that is most convenient for our application is the following:

Theorem 4. (*Theorem 4.1 in [5]*)
Given symmetric matrices $A_1 \dots, A_m \in \mathbb{R}^{n\times n}$ and independent random Bernoulli signs $\epsilon_1, \dots, \epsilon_m$, we have

$$\mathbb{P}\left[\left\|\sum_{i=1}^m \epsilon_i A_i\right\| \ge t\right] \le 2n \cdot \exp\left(-\frac{t^2}{2\|\sum_{i=1}^m A_i^2\|}\right).$$

Applying the theorem to $A_i = v_iv_i^T$ and taking $T_1 = \{i : \epsilon_i = +1\}$ yields Theorem 1 with a discrepancy of $O(\sqrt{\alpha \log n})$, which is nontrivial when $\alpha \le O(1/\log n)$. This bound is interesting and useful in some settings, but it is not sufficient to prove the Kadison–Singer conjecture (which requires a uniform bound as $n \to \infty$), or for the application in the next section. It may be seen as analogous to the discrepancy of $O(\sqrt{n \log n})$ achieved by a random colouring of a set family $S_1, \dots, S_n \subset [n]$, which is easily analysed using the usual Chernoff bound and a union bound.

In order to remove the logarithmic factor and obtain Theorem 1, we prove the following stronger but less general inequality, which controls the deviation of a sum of independent rank-one matrices at a constant rather than logarithmic scale, but only with nonzero (rather than high) probability:

Theorem 5. (*Theorem 1.2 in [1]*)
If $\epsilon > 0$ and $v_1, \dots, v_m$ are independent random vectors in $\mathbf{R}^n$ with finite support such that

$$\sum_{i=1}^m \mathbb{E} v_iv_i^T = I,$$

and

$$\mathbb{E}\|v_i\|^2 \le \alpha$$

for all i, then

$$\mathbb{P}\left[\left\|\sum_{i=1}^m v_iv_i^T\right\| \le (1+\sqrt{\alpha})^2\right] > 0.$$

The conclusion of the theorem is equivalent to the following existence statement: there is a point $\omega \in \Omega$ in the probability space implicitly defined by the v_is such that

$$\left\|\sum_{i\le m} v_i(\omega)v_i(\omega)^T\right\| \le (1+\sqrt{\alpha})^2.$$

To prove the theorem, we begin by considering for every ω the univariate polynomial

$$P[\omega](x) := \det\left(xI - \sum_{i\le m} v_i(\omega)v_i(\omega)^T\right).$$

The roots of $P[\omega]$ are real since it is the characteristic polynomial of a symmetric matrix, and in particular the largest root is equal to the spectral norm of $\sum_{i\le m} v_i(\omega)v_i(\omega)^T$.

The proof now proceeds in two steps. First, we show that there must exist an ω such that

$$\lambda_{\max}(P[\omega]) \le \lambda_{\max}(\mathbb{E}P), \tag{3}$$

where $\lambda_{\max}$ denotes the largest root of a polynomial. This type of statement may be seen as a generalisation of the probabilistic method to polynomial-valued random variables, and was introduced in the paper [6], where we used it to show the existence of bipartite Ramanujan graphs of every degree. Note that (3) does *not* hold for general polynomial-valued random variables — in general, the roots of a sum of polynomials do not have much to do with the roots of the individual polynomials. The reason it holds in this particular case is that the $P[\omega]$ are generated by sums of rank-one matrices (which by Cauchy's theorem produce interlacing characteristic polynomials) and form what we call an "interlacing family".

The second step is to upper bound the roots of the expected polynomial $\mu(x) := \mathbb{E}P(x)$. It turns out that the right way to do this is to write $\mu(x)$ as a linear transformation of a certain m-variate polynomial $Q(z_1, \ldots, z_m)$, and show that Q does not have any roots in a certain region of $\mathbb{R}^m$. This is achieved by a new multivariate generalisation of the "barrier function" argument used in [7] to construct spectral sparsifiers of graphs. The multivariate barrier argument relies heavily on the theory of real stable polynomials, which are a multivariate generalisation of real-rooted polynomials.

Rather than giving any further details, I encourage you to read the paper. The proof is not difficult to follow, and from what I have heard quite "readable".

3. Partitioning a Graph into Sparsifiers

One very fruitful setting in which to apply Theorem 2 is that of Laplacian matrices of graphs. Recall that for an undirected graph $G = (V, E)$ on n vertices, the Laplacian is the $n \times n$ symmetric matrix defined by:

$$L_G = \sum_{ij \in E} b_{ij} b_{ij}^T,$$

where $b_{ij} := (e_i - e_j)$ is the incidence vector of the edge ij. The Laplacian quadratic form

$$x^T L_G x = \sum_{ij \in E} (x(i) - x(j))^2$$

encodes a lot of useful information about a graph. For instance, it is easy to check that given any cut $S \subset V$, the quadratic form $x_S^T L_G x_S$ of the indicator vector $x_S(i) = \mathbf{1}_{\{i \in S\}}$ is equal to the number of edges between S and $\overline{S}$. Thus, the values of $x^T L_G x$ completely determine the cut structure of G. (We mention in passing that the extremisers of the quadratic form are eigenvalues and are related to various other properties of G — this is the subject of **spectral graph theory**.)

Now consider $G = K_n$, the complete graph on n vertices, which has Laplacian

$$L_{K_n} = \sum_{ij} b_{ij} b_{ij}^T.$$

An elementary calculation reveals that the leverage scores in this graph are all very small:

$$b_{ij}^T L_{K_n}^+ b_{ij} = \frac{2}{n}.$$

This is a good time to mention that the leverage scores of the incidence vectors b_{ij} in any graph G have a natural interpretation — they are simply the effective resistances of the edges ij when the graph is viewed as an electrical network (this happens because inverting L_G is equivalent to computing an electrical flow, and the quantity $x^T L_G^+ x$ is equal to the energy dissipated by the flow.) In any case, for the complete graph, all of the edges have effective resistances equal to $2/n$, so we may apply Theorem 2 with $\alpha = 2/n$ to conclude that there is a partition of the edges into two sets, T_1 and T_2, each satisfying

$$\left(1/2 - O(1/\sqrt{n})\right) L_{K_n} \preceq \sum_{ij \in T_k} b_{ij} b_{ij}^T \preceq \left(1/2 + O(1/\sqrt{n})\right) L_{K_n}. \quad (4)$$

Now observe that each sum over T_k is the Laplacian L_{G_k} of a *subgraph* G_k of K_n. By recalling the connection to cuts, this implies that K_n can be partitioned into two subgraphs, G_1 and G_2, each of which approximates its cuts up to a $1/2 \pm O(1/\sqrt{n})$ factor.

This seems like a cute result, but we can go a lot further. As long as the effective resistances of edges in G_1 and G_2 are sufficiently small, we can apply Theorem 2 *again* to each of them to obtain four subgraphs. And then again to obtain eight subgraphs, and so on.

How long can we keep doing this? The answer depends on how fast the effective resistances grow as we keep partitioning the graph. The following simple calculation reveals that they grow geometrically at a favourable rate. Initially, all of the effective resistances are equal to $\ell_0 = 2/n$. After one partition, the maximum effective resistance of an edge in G_k is at most

$$\ell_1 := \max_{ij \in G_k} b_{ij}^T L_{G_k}^+ b_{ij} \le (1/2 - O(1/\sqrt{n}))^{-1} b_{ij} L_{K_n}^+ b_{ij} = (1/2 - O(1/\sqrt{n}))^{-1} \cdot (2/n).$$

In general, after i levels of partitioning, we have the inequalities:

$$\begin{aligned} 2\exp(O(\sqrt{\ell_{i-1}}))\ell_{i-1} &\ge (1/2 - O(\sqrt{\ell_{i-1}}))^{-1}\ell_{i-1} \\ &\ge \ell_i \\ &\ge (1/2 + O(\sqrt{\ell_{i-1}}))^{-1}\ell_{i-1} \\ &\ge (3/2)\ell_{i-1}, \end{aligned}$$

as long as ℓ_{i-1} is bounded by some sufficiently small absolute constant δ. Applying these inequalities iteratively we find that after t levels:

$$\begin{aligned}\ell_t &\le 2^t \exp\left(O\left(\sum_{i=0}^{t-1}\sqrt{\ell_i}\right)\right)\ell_0 \\ &\le 2^t\cdot\exp\left(O\left(\sum_{i=0}^{t-1}\sqrt{(2/3)^{t-1-i}\ell_{t-1}}\right)\right)\cdot(2/n) \\ &\le \exp\left(O(\sqrt{\delta})\right)\cdot 2^t(2/n),\end{aligned}$$

and the inequalities are valid as long as we maintain that $\ell_{t-1} \le \delta$. Taking binary logs, we find that these conditions are satisfied as long as

$$O(\sqrt{\delta}) + t + \log(2/n) \le \log(\delta),$$

which means we can continue the recursion for

$$t = \log n - 1 + \log(\delta) - O(\sqrt{\delta}) = \log n - O(1)$$

levels. This yields a partition of K_n into $O(n)$ subgraphs, each of which is an $O(1)$-factor spectral approximation of $(1/2^t)K_n$, in the sense of (4). This latter approximation property implies that each of the graphs must have constant degree (by considering that the degree cuts must approximate those of $(1/2^t)K_n$) and constant spectral gap; thus we have shown that K_n can be partitioned into $O(n)$ constant degree expander graphs.

The real punchline, however, is that we did not use anything special about the structure of K_n other than the fact that its effective resistances are bounded by $O(1/n)$. In fact, the above proof works exactly the same way on *any* graph on n vertices with m edges, whose effective resistances are bounded by $O(n/m)$ — for such a graph, the same calculations reveal that we can recursively partition the graph for $\log(m/n) - O(1)$ levels, while maintaining a constant factor approximation! Note that the total effective resistance of any unweighted graph on n vertices is $n - 1$, so the boundedness condition is just saying that every effective resistance is at most a constant times the average over all m edges.

For instance, the effective resistances of all edges in the hypercube Q_n on $N = 2^n$ vertices are very close to $1/2n = 1/2\log N$. Thus, repeatedly applying Theorem 2 implies that it can be partitioned into $O(\log N)$ constant degree subgraphs, each of which is an $O(1)$-factor spectral approximation of $1/\log N \cdot Q_n$. In fact, this type of conclusion holds for *any* edge-transitive graph, in which symmetry implies that each edge has exactly the same effective resistance.

The above result may be seen as a generalisation of the theorem of Frieze and Molloy [9], which says that up to a certain extent, any sufficiently good expander graph may be partitioned into sparser expander graphs. It may also be seen as an unweighted version of the spectral sparsification theorem of Batson, Spielman, and myself [7], which says that every graph has a *weighted* $O(1)$-factor spectral approximation with $O(n)$ edges. The recursive partitioning argument that we have used is quite natural and appears to have been observed a number of times in various contexts; see for instance paper of Rudelson [10], as well as the very recent work of Harvey and Olver [11].

4. Conclusion and Open Questions

Theorem 1 essentially shows that under the mildest possible conditions, a quadratic form/sum of outer products can be "split in two" while preserving its spectral properties. Since graphs can be encoded as quadratic forms/outer products, the theorem implies that they also can be "split into two" while preserving some properties. However, a lot of other objects can also be encoded this way. For instance, applying Theorem 1 to a submatrix of a Discrete Fourier Transform (it also holds over $\mathbb{C}^n$) or Hadamard matrix yields a strengthening of the "uncertainty principle" for Fourier matrices, which says that a signal cannot be localised both in the time domain and the frequency domain; see paper of Casazza and Weber [12] for details. This strengthening has implications in signal processing, and its infinite-dimensional analogue is useful in analytic number theory. For a thorough survey of the consequences of the Kadison–Singer conjecture and Theorem 1 in many diverse areas, check out [13].

To conclude, let me point out that the current proof of Theorem 1 is not algorithmic, since it involves reasoning about polynomials which are in general $\#P$-hard to compute. Finding a polynomial-time algorithm which delivers the low-discrepancy partition promised by the theorem is likely to yield further insights into the techniques used to prove it as well as more connections to other areas — just as the beautiful work of Moser-Tardos [14], Bansal [15], and Lovett-Meka [16] has done for the Lovasz Local Lemma and Spencer's theorem. It would also be nice to see if the methods used here can also be used to recover known results in discrepancy theory, such as Spencer's theorem itself.

Acknowledgements

Thanks to Nick Harvey, Daniel Spielman, and Nisheeth Vishnoi for helpful suggestions, comments, and corrections during the preparation of this article. Special thanks to my coauthors Adam Marcus and Daniel Spielman of Yale University.

References

[1] A. Marcus, D. A. Spielman and N. Srivastava, Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem, arXiv:1306.3969 (2013).
[2] N. Weaver, The Kadison–Singer problem in discrepancy theory, *Discrete Mathematics* **278** (2004) 227–239.
[3] B. Roberts, Philosophy and physics in the Kadison–Singer conjecture, http://www.soulphysics.org/2013/06/philosophy-and-physics-in-the-kadison-singer-conjecture/
[4] N. J. A. Harvey, An introduction to the Kadison–Singer problem and the paving conjecture (2013).
[5] J. A. Tropp, User-friendly tail bounds for sums of random matrices, arXiv:1004.4389 (2011).
[6] A. Marcus, D. A. Spielman and N. Srivastava, Interlacing families I: Bipartite Ramanujan graphs of all degrees, arXiv:1304.4132 (2013).
[7] J. Batson, D. A. Spielman and N. Srivastava, Twice-Ramanujan sparsifiers, arXiv:0808.0163 (2009).
[8] D. G. Wagner, Multivariate stable polynomials: theory and applications, arXiv:0911.3569 (2009).
[9] A. M. Frieze and M. Molloy, Splitting an expander graph, *J. Algorithms* **33**(1) (1999) 166–172.
[10] M. Rudelson, Almost orthogonal submatrices of an orthogonal matrix, arXiv:math/9606213 (1996).
[11] N. J. A. Harvey and N. Olver, Pipage rounding, pessimistic estimators and matrix concentration, arXiv:1307.2274 (2013).
[12] P. G. Casazza and E. Weber, The Kadison–Singer problem and the uncertainty principle, *Proc. Amer. Math. Soc.* **136**(12) (2008) 4235–4243.
[13] P. G. Casazza, M. Fickus, J. C. Tremain and E. Weber, The Kadison–Singer problem in mathematics and engineering: a detailed account, in D. Han, P. E. T. Jorgensen and D. R. Larson, eds. *Contemporary Mathematics*, Vol. 414 (2006), pp. 299–356.
[14] R. A. Moser and G. Tardos, A constructive proof of the general Lovasz Local Lemma, arXiv:0903.0544 (2009).
[15] N. Bansal, Constructive algorithms for discrepancy minimization, arXiv:1002.2259 (2010).
[16] S. Lovett and R. Meka, Constructive discrepancy minimization by walking on the edges, arXiv:1203.5747 (2012).

Nikhil Srivastava

Microsoft Research, Bangalore, India

niksri@microsoft.com

Nikhil Srivastava graduated in Mathematics and Computer Science from Union College, Schenectady, NY in June 2005. In May 2010 he obtained PhD from Yale University in Computer Science with the dissertation "Spectral Sparsification and Restricted Invertibility". He did his postdoc at Princeton, MSRI (Berkeley) and IAS (Princeton). Currently he is with Microsoft Research, Bangalore, India. His current research interests are theoretical computer science, linear algebra, random matrices, and convex geometry.