Explicit lossless expanders

Understanding the construction of Capalbo, Reingold, Vadhan, Wigderson

Péter Gács

Computer Science Department
Boston University

Spring 06

- Expanders, extractors condensers, etc.
- The requirement of explicitness.
- Algebraic constructions via an eigenvalue argument.
- The best eigenvalue may yield an expansion that is less by a factor of 2 than the best expansion.
- Some applications need expansion exceeding the eigenvalue bound or even near-maximal expansion (say, reliable memories).

- New ideas: “zigzag product” of Capalbo, Reingold, Vadhan, Wigderson and Capalbo.
- Relies only on an intuition of “entropy flow” and straighforward probability estimates instead of algebra or geometry.
- We will introduce to the details of the technique, but will not survey the many applications and related constructions.
- A surprising new application of the product technique, due to Reingold, will be presented next week by David: an algorithm for passing through an undirected maze in logarithmic space.

Notation, conventions

$[K] = \{1, \ldots, K\}$, $(k) = [2^k]$.
In a bipartite graph, we will talk about a left set (typically $[N]$) and a right set (typically $[M]$).

Definition

A bipartite graph $G = ([N], [M], E)$ is a (K, A)-expander if every subset $X \subseteq [N]$ of at most K vertices is connected to at least $A \cdot [N]$ neighbors (We need $M/K \leqslant A \leqslant N$).

Typically, the graph in question will be bi-regular: all points of the left set have the same degree D and also all points of the right set have the same degree. On the other hand, the graph is allowed to be a multigraph (with possibly parallel edges). Such a graph be described by a function

$$E : [N] \times [D] \to [M].$$

The bipartite graph can also be viewed as a Markov transition from left to right. Then the expander property seems to increase the uniformity from input distribution to output distribution. A successful measure of uniformity follows.

Definition

Let X be a discrete random variable with distribution $P = P_X$. The quantity

$$H_\infty(X) = H_\infty(P) = -\log \max_a P(a)$$

is the min-entropy of X. We way that X is a k-source if $H_\infty(X) \geqslant k$ (makes sense for non-integer k, too).

Example

Flat (uniform) distribution over a set of size K has min-entropy $\log K$.
In particular, U_d is a random variable uniformly distributed over (d).

Definition

Real number k is a log integer if 2^k is integer.

The min-entropy of flat distributions is always a log integer.

Proposition

The min-entropy is concave (just as ordinary entropy), that is for any distributions P_i over the same set, with (λ_i) with $\lambda_i \geqslant 0$, $\sum_i \lambda_i = 1$, with $P = \sum_i \lambda_i P_i$, we have $H_\infty(P) \geqslant \sum_i \lambda_i H_\infty(P_i)$.

This follows immediately from the concavity of logarithm.
A kind of "converse":

Theorem

If k is a log integer then every k-source is the convex combination of a finite number of flat k-sources.

Proof.

View the set of k-sources as a bounded convex polyhedron, and characterize its extremal points. □

Definition

For discrete distributions P, Q over the same set S, their statistical difference is

$$|P - Q| = \sup_{A \subseteq S} |P(A) - Q(A)| = \tfrac{1}{2} \sum_{x \in S} |P(x) - Q(x)|.$$

(This is a metric: it is $\frac{1}{2}$ the L_1 norm of the difference of the two functions.) If statistical the difference is smaller than ε the distributions are called ε-close.

The following is easy:

Proposition

Suppose that P_i is ε-close to P'_i for $i = 1, \ldots, k$ and $\lambda_i \geqslant 0$, $\sum_i \lambda_i = 1$. Then $\sum_i \lambda_i P_i$ is ε-close to $\sum_i \lambda_i P'_i$.

Definition

A distribution is a (k, ε)-source if it is ε-close to a k-source.

An example concept using all of the above:

Definition

A function $E : (n) \times (d) \to (m)$ is a (k, ε)-extractor if for any k-source X over (n), the distribution of $E(X, U_d)$ is ε-close to U_m.

Motivation: adapting "quasi-random" sources in randomized algorithms.

We need a kind of object generalizing expanders, extractors and the like, which is suitable to various operations of combination.

Definition

A function $E : (n) \times (d) \to (m)$ is a $(k_{\max}, \varepsilon, a)$ simple conductor with input size n, seed size d, output size m, ceiling $k_{\max}$, increment a, error ε if for all log integer $0 \leqslant k \leqslant k_{\max}$, for any k-source X over (n), the distribution of $E(X, U_d)$ is a $(k + a, \varepsilon)$-source.

(Here a can be negative but we will not use this possibility in the lecture.)
Conductors could be called entropy pumps.

Special case:

Definition

A function $E : (n) \times (d) \to (m)$ is an (ε, a) extracting conductor if it is an $(m - a, \varepsilon, a)$ simple conductor (with increment a and maximal ceiling).

In particular this must be a $(m - a, \varepsilon)$ extractor.

Another special case:

Definition

A function $E : (n) \times (d) \to (m)$ is a $(k_{\max}, \varepsilon)$ lossless conductor if it is a $(k_{\max}, \varepsilon, d)$ simple conductor (the increment equals the seed size).

Theorem

The function E is a $(k_{\max}, \varepsilon)$ lossless conductor if and only if the corresponding bipartite graph with left degree 2^d is a $(2^{k_{\max}}, (1-\varepsilon)2^d)$ expander.

So our goal is to construct large lossless conductors.

Proof of the theorem:

For any input random variable X, let Q be the distribution of $E(X, U_d)$, and let Γ be the support of Q.

Assume first that E is a (k, ε) lossless conductor. Let $S \subseteq (n)$ with $|S| = 2^k$, $k < k_{\max}$. Let X be uniform on S, then Q is ε-close to a distribution Q' with $H_\infty(Q') \geqslant k + d$. Then

$$\varepsilon \geqslant Q(\Gamma) - Q'(\Gamma) = 1 - Q'(\Gamma) \geqslant 1 - |\Gamma| \cdot 2^{-(k+d)},$$
$$|\Gamma| \geqslant (1 - \varepsilon) 2^{k+d}.$$

Assume now that E is a $(2^{k_{\max}}, (1-\varepsilon)2^d)$-expander. Let X be a (k,ε)-source with log integer $k \leqslant k_{\max}$. Using the convex combination theorem 6, it is sufficient to assume that X is flat. Let Q' be a distribution assigning probability $2^{-(k+d)}$ to all elements of Γ and also to some arbitrary elements outside Γ. Let Γ' be the support of Q'. We know $Q(y) \geqslant 2^{-(k+d)}$ for all $y \in \Gamma$, and also $|\Gamma| \geqslant (1-\varepsilon)2^{k+d}$, so we have

$$\begin{aligned} 2|Q - Q'| &\geqslant \sum_y |Q(y) - Q'(y)| \\ &= 1 - Q'(\Gamma) + Q'(\Gamma' \setminus \Gamma) \\ &\leqslant 1 - (1-\varepsilon) + \varepsilon = 2\varepsilon. \end{aligned}$$

This completes the proof of the theorem.

Combining extraction with losslessness by retaining some more input entropy in a buffer:

Definition

A pair of functions $\langle E, C\rangle : (n) \times (d) \to (m) \times (b)$ is a $(k_{\max}, \varepsilon, a)$ buffer conductor with buffer size b if E is an (ε, a) extracting conductor (with increment a) and $\langle E, C\rangle$ is a $(k_{\max}, \varepsilon)$ lossless conductor (with ceiling $k_{\max}$).

Special case:

Definition

A $(k_{\max}, \varepsilon, a)$ buffer conductor is an (ε, a) permutation conductor if $n + d = m + b$ and $\langle E, C\rangle$ is a permutation.

Theorem (Nonconstructive extracting)

For every $m \leqslant n$ and $\varepsilon > 0$ there is (ε, a) extracting conductor $E : (n) \times (d) \to (m)$ with

$$d = \log(n - m + 1) + 2\log(1/\varepsilon) + O(1),$$
$$a = d - \log(1/\varepsilon) - O(1).$$

In particular, for fixed ε and $m = n$ the "degree" d is constant.
This, and the following existence theorems are not proved in the paper, only promised in the final version. But it seems indeed that standard probabilistic technique combined with the convex combination theorem will yield them.
In the final construction, only constant-size versions of these nonconstructive objects will be used.

Theorem (Nonconstructive lossless)

For every $m \leqslant n$ *and* $\varepsilon > 0$ *there is* $(k_{\max}, \varepsilon)$ *lossless conductor* $E : (n) \times (d) \to (m)$ *with*

$$d = \log(n - m + 1) + \log(1/\varepsilon) + O(1),$$
$$k_{\max} = m - d - \log(1/\varepsilon) - O(1).$$

Combination into a buffer conductor:

Theorem (Nonconstructive buffer)

For every $m \leqslant n$ *and* $b, \varepsilon > 0$ *there is* $(k_{\max}, \varepsilon, a)$ *buffer conductor* $E : (n) \times (d) \to (m) \times (b)$ *with*

$$\begin{aligned} d &= \log(n - m + 1) + 2\log(1/\varepsilon) + O(1), \\ a &= d - 2\log(1/\varepsilon) - O(1), \\ k_{\max} &= m + b - d - \log(1/\varepsilon) - O(1). \end{aligned}$$

Known explicit objects

The known constructive expancers (say, the Ramanujan graphs) yield this:

Theorem (Explicit expander)

There are constants $a, d, c > 0$ such that for all n there is an explicit $(2^{n-a-c}, 2^a)$ expander $E : (n) \times (d) \to (n)$.

The well-known method of performing a d-step random walk on the above expander and remembering how to walk back, yields a buffer conductor:

Theorem (Explicit permutation)

For every $a < n$ and $\varepsilon > 0$ there is an explicit (ε, a) permutation conductor $\langle E, C \rangle : (n) \times (d) \to (n) \times (d)$, with $d = O(a + \log(1/\varepsilon))$.

The extractor part E alone here is not lossless, because of the big-O in $d = O(a + \log(1/\varepsilon))$.

Definition (Zigzag product)

Let

$$\langle E_1, C_1\rangle : (n_1) \times (d_1) \to (m_1) \times (b_1),$$
$$\langle E_2, C_2\rangle : (n_2) \times (d_2) \to (d_1) \times (b_2),$$
$$E_3 : (b_1 + b_2) \times (d_3) \to (m_3)$$

be three functions. For $(x_1) \in (n_1)$, $x_2 \in (n_2)$, $r_2 \in (d_2)$, $r_3 \in (d_3)$ let $n = n_1 + n_2$, $d = d_2 + d_3$, $m = m_1 + m_3$, and $E : (n) \times (d) \to (m)$ is defined as

$$E(x_1 \circ x_2, r_2 \circ r_3) = y_1 \circ y_2 \qquad \text{where}$$
$$\langle r_1, z_1\rangle = E_2(x_2, r_2),$$
$$\langle y_1, z_2\rangle = E_1(x_1, r_1),$$
$$y_2 = E_3(z_1 \circ z_2, r_3).$$

See drawing on the board. Let $a = 1000 \log(1/\varepsilon)$,

$$\begin{aligned} n_1 = m_1 &= n - 20a, \\ d_1 = b_1 &= 14a, \\ n_2 &= 20a, \\ d_2 = d_3 &= a, \\ b_2 &= 21a, \\ m_3 &= 17a. \end{aligned}$$

Assumptions

- $\langle E_1, C_1 \rangle$ is a $(\varepsilon, 6a)$ permutation conductor (explicit, big).
- $\langle E_2, C_2 \rangle$ is a (small) $(n_2, \varepsilon, 0)$ buffer conductor (nonconstructive, small).
- E_3 is a $(15a, \varepsilon)$ lossless conductor (nonconstructive, small).

Theorem

The resulting conductor $E : (n) \times (d) \to (b)$ is an $(n - 30a, 4\varepsilon)$ lossless conductor.

Lemma (Conditional decomposition)

Let (X_1, X_2) have a probability distribution on a finite product space. Given $\varepsilon > 0$ and a, there is a distribution for some (Y_1, Y_2) on the same space such that

(a) *The distributions of (X_1, X_2) and (Y_1, Y_2) are ε-close.*

(b) *The distribution of (Y_1, Y_2) is a convex combination of two other distributions $(\hat{Y}_1, \hat{Y}_2)$ and $(\check{Y}_1, \check{Y}_2)$ each having min-entropy at least $H_{\inf}(X_1, X_2) - \log(1/\varepsilon)$.*

(c) *For $x \in \mathrm{Supp}(X_1)$ we have $H_\infty(\hat{Y}_2 | \hat{Y}_1 = x) \geqslant a$.*

(d) *For $x \in \mathrm{Supp}(X_1)$ we have $H_\infty(\check{Y}_2 | \check{Y}_1 = x) < a$.*

The proof is simple: break up into two disjoint supports according to whether $H_\infty(X_2 | X_1 = x) \geqslant a$ or $< a$, then omit one of them if its probability is $< \varepsilon$.

Proof sketch

Ignoring ε: statistical differences just add up across conductors.
Assume $H_\infty(X_1, X_2) = k \leqslant n - 30a$.
We will prove $H_\infty(Y_1) \geqslant k - 14a$. Two cases, using the decomposition theorem (and losing $\log(1/\varepsilon) < a$).

1. $H_\infty(X_2|X_1 = x_1) \geqslant 14a$ for all x_1. Then $Y_2 = R_1$ becomes uniform and the needed entropy pumps into Y_1.
2. $H_\infty(X_2|X_1 = x_1) < 14a$ for all x_1. Then the needed entropy is already there, and will not get lost.

Now we now $H_\infty(Y_1) \geqslant k - 14a$. Due to the buffer and permutation property, $H_\infty(Y_1, Z_1, Z_2) \geqslant k + a$. From the two inequalities follows $H_\infty(Z_1, Z_2|Y_1 = y_1) \leqslant 15a$, and so E_3 can pump.