1 ℓ_1 diameter in small dimension

The diameter of a pointset P is defined as $\text{diam}(P) = \max_{u,v \in P} d(u, v)$. The goal of the exercise is to show an algorithm for *exact* computation of diameter in ℓ_1^d metric, when d is small (i.e. $o(\log n)$).

1. Show an exact embedding $\phi : \ell_1^d \to \ell_\infty^{2^d}$.
2. Show an algorithm that finds ℓ_1^d diameter of poinset P, $|P| = n$, in $\mathcal{O}(n \cdot 2^d)$ time.

2 Johnson-Lindenstrauss

Consider following useful result by Johnson–Lindenstrauss:

Lemma 1 *For $P \subset \mathbb{R}^n$, $|P| = n$, and for $0 < \varepsilon \leq 1$, there is $k = \mathcal{O}(\log n/\varepsilon^2)$ and map $\varphi : \ell_2^n \to \ell_2^k$, such that:*

$$\forall_{p,q \in P} \quad \|p - q\|_2 \leq \|\phi(p) - \phi(q)\|_2 \leq (1 + \varepsilon) \cdot \|p - q\|_2$$

More surprisingly, such mapping can be easily constructed. It can be shown, that for appropriate k, and picked u.a.r. matrix $A \in \{-1, 1\}^{k \times n}$, a map $\phi(p) = A \cdot p$ works (w.h.p., and up to some fixed scaling factor).

1. Show that random matrix A preserves squared norm *in expectation*, that is

   $$\mathbb{E}[\|\phi(p)\|_2^2] = k\|p\|_2^2$$

   where the expectation is taken across choices of A.
2. For the harder part, that is *concentration*: weaker bound (i.e. $\mathcal{O}(\text{poly}(\log n)/\varepsilon^2)$) can be derived from known concentration bounds
   OR
   find and read the proof of online.

3 Sketching for Hamming distance

Consider a following sketching problem: Alice and Bob hold each text, i.e. $T_A, T_B \in \Sigma^n$. They want to send relatively short messages (that is, *sketches*) to Charlie, so that Charlie can compute *Hamming distance* of T_A and T_B.

1. Consider case of $\Sigma = \{0, 1\}$. Show that Alice and Bob can encode messages of $\mathcal{O}(\log n/\varepsilon^2)$ *words* so that Charlie decodes $1 + \varepsilon$ approximation of Hamming distance.
2. Consider random projection $\varphi : \Sigma \to \{0, 1\}$. Show that it preserves Hamming distance *in expectation* (up to factor 1/2).
3. Show that averaging over many projections $\varphi_1, \varphi_2, \ldots, \varphi_k$ for some $k = \mathcal{O}(\log n/\varepsilon^2)$ preserves Hamming distance up to $1 + \varepsilon$ factor, w.h.p.

4. Combine all of above ideas to obtain a communication protocol:
    - Alice and Bob send messages of length $\mathcal{O}(\log^2 n/\varepsilon^4)$ words,
    - Charlie can decode, with high probability, an $(1+\varepsilon)$ approximation of the Hamming distance.
5. (*) Show that by smarter combining of J-L lemma and random projections, one can reduce the size of messages to $\mathcal{O}(\log n/\varepsilon^2)$ words.

4 Applications of the sparsest cut

4.1 Approximating edge expansion

For undirected, unweighted graph G and set $S \subset V$ s.t. $|S| \leq |V|/2$, the *edge expansion* of S is defined to be $|\delta S|/|S|$. Show an $\mathcal{O}(\log n)$ approximation algorithm for finding *minimal expansion* in graph G.

4.2 Minimum bisection cut

A *bisection cut* is a cut (S, S') such that $|S| = |S'| = n/2$. A r-*balanced cut* is a cut where $r \cdot n \leq |S| \leq (1-r) \cdot n$. Our goal is to (pseudo) approximate *minimal bisection cut.* Consider following procedure:

- Initialize $U \leftarrow \emptyset$, $W \leftarrow V$.
- Until $|U| < n/3$, repeat:
- X is the approximate small edge expansion cut in G_W
- $U \leftarrow U \cup X$, $W \leftarrow W \setminus X$.
- return (U, W)

Show, that this procedure returns an 1/3-balanced cut, with size being $\mathcal{O}(\log n)$ factor from minimal bisection cut.