[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[Mat02] Jiri Matousek. *Lectures on Discrete Geometry*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.

[Mat07] J. Matoušek. Collection of open problems on low-distortion embeddings of finite metric spaces, March 2007. Available online. Last access in August, 2007.

[OR05] R. Ostrovsky and Y. Rabani. Low distortion embeddings for edit distance. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, pages 218–224, 2005.

[Raz92] A. A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

A Hellinger distance and Communication Protocols

PROPOSITION A.1. ([BJKS04]) *Let $\mathcal{P}$ be a randomized private-coin protocol on $\mathcal{X} \times \mathcal{Y}$. Let $(u_1, v_1), (u_2, v_2) \in \mathcal{X} \times \mathcal{Y}$ be two distinct inputs whose transcript wave functions in $\mathcal{P}$ are denoted by $\psi(u_1, v_1)$ and $\psi(u_2, v_2)$, respectively.*

Mutual information to Hellinger distance:
Suppose $(U, V) \in_R \{(u_1, v_1), (u_2, v_2)\}$. If Π denotes the transcript random variable, then

$$\mathrm{I}(U, V : \Pi) \geq \hat{\|}\psi(u_1, v_1) - \psi(u_2, v_2)\hat{\|}.$$

Soundness:
Suppose $\mathcal{P}$ is a correct protocol for a decision problem g defined on $\mathcal{L} \subseteq \mathcal{X} \times \mathcal{Y}$. Suppose $(u_1, v_1), (u_2, v_2) \in \mathcal{L}$ such that $g(u_1, v_1) \neq g(u_2, v_2)$. Then,

$$\hat{\|}\psi(u_1, v_1) - \psi(u_2, v_2)\hat{\|} \geq 1 - 2\sqrt{\delta}.$$

Pythagorean property:
Consider the combinatorial rectangle of 4 inputs $\{u_1, u_2\} \times \{v_1, v_2\}$ and label them as $A = (u_1, v_1)$, $B = (u_1, v_2)$, $C = (u_2, v_1)$ and $D = (u_2, v_2)$. Then,

$$\hat{\|}\psi(A) - \psi(D)\hat{\|} \geq \begin{cases} \frac{1}{2} \cdot (\hat{\|}\psi(A) - \psi(B)\hat{\|} + \hat{\|}\psi(C) - \psi(D\hat{\|}) \\ \frac{1}{2} \cdot (\hat{\|}\psi(A) - \psi(C)\hat{\|} + \hat{\|}\psi(B) - \psi(D)\hat{\|}) \end{cases}$$

□

The first property in the above proposition is just a restatement of the fact that the Jensen-Shannon distance between $\psi(u)$ and $\psi(v)$ is bounded from below by their Hellinger distance. The next property follows by relating Hellinger to variational distance and then invoking the correctness of the protocol. The last property relies on the structure of *deterministic* communication protocols, namely, that the transcripts partition the space of inputs into *combinatorial rectangles*. The property itself can be seen as one generalization to randomized protocols. (In [BJKS04], another property is shown which generalizes the cut-and-paste property of deterministic communication protocols. This is not needed for our results.)