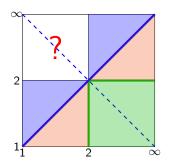
(In)Approximability of Matrix Norms





Madhur Tulsiani

Joint work with Vijay Bhattiprolu, Mrinalkanti Ghosh, Venkatesan Guruswami and Euiwoong Lee

Matrix Norms

- Given $A \in \mathbb{R}^{m \times n}$, find (or approximate)

$$\begin{split} \left\|A\right\|_{p\to q} &:= \max_{x\neq 0} \frac{\left\|Ax\right\|_q}{\left\|x\right\|_p} &= \max_{\left\|x\right\|_p = 1} \left\|Ax\right\|_q \end{split}$$
 where, $\left\|x\right\|_p = \left(\mathbb{E}_{i\in[n]}\left|x_i\right|^p\right)^{1/p}.$ $(p,q\geq 1).$

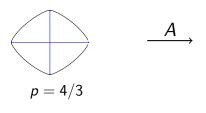
Matrix Norms

- Given $A \in \mathbb{R}^{m \times n}$, find (or approximate)

$$\|A\|_{p \to q} := \max_{x \neq 0} \frac{\|Ax\|_q}{\|x\|_p} = \max_{\|x\|_p = 1} \|Ax\|_q$$

where,
$$\left\|x\right\|_{p} = \left(\mathbb{E}_{i \in [n]} \left|x_{i}\right|^{p}\right)^{1/p}$$
. $(p, q \geq 1)$.

- How A maps $\|\cdot\|_p$ into $\|\cdot\|_q$.





 $-2 \rightarrow 2$:

- $2 \rightarrow 2$: Spectral Norm (we all know, love and compute.)

- 2 \rightarrow 2: Spectral Norm (we all know, love and compute.)
- $\infty \to 1$: Grothendieck problem. Constant-factor K_G approximation [Grothendieck 53] ... [Braverman (Makarychev)² Naor 11].

- 2 \rightarrow 2: Spectral Norm (we all know, love and compute.)
- $\infty \to 1$: Grothendieck problem. Constant-factor K_G approximation [Grothendieck 53] ... [Braverman (Makarychev)² Naor 11].
- $p \ge q$: Widely studied in terms of algorithms and hardness.

- 2 \rightarrow 2: Spectral Norm (we all know, love and compute.)
- $\infty \to 1$: Grothendieck problem. Constant-factor K_G approximation [Grothendieck 53] ... [Braverman (Makarychev)² Naor 11].
- $p \ge q$: Widely studied in terms of algorithms and hardness.
- p < q: Hypercontractive norms. Relevant for small-set expansion like problems.

- p and p^* define dual norms if $\frac{1}{p} + \frac{1}{p^*} = 1$.

- p and p^* define dual norms if $\frac{1}{p} + \frac{1}{p^*} = 1$.
- $\|x\|_{p^*} = \max_{\|y\|_p \le 1} \langle x, y \rangle.$

- p and p^* define dual norms if $\frac{1}{p} + \frac{1}{p^*} = 1$.

$$- \|x\|_{p^*} = \max_{\|y\|_p \le 1} \langle x, y \rangle.$$

$$- \|A\|_{p \to q} = \|A^T\|_{q^* \to p^*}$$

- p and p^* define dual norms if $\frac{1}{p} + \frac{1}{p^*} = 1$.

$$- \|x\|_{p^*} = \max_{\|y\|_p \le 1} \langle x, y \rangle.$$

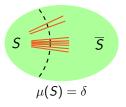
$$\begin{aligned} - & \|A\|_{p \to q} &= \left\|A^{T}\right\|_{q^{*} \to p^{*}} \\ & \|A\|_{p \to q} &= \max_{\|x\|_{p} \le 1} \|Ax\|_{q} \\ &= \max_{\|x\|_{p} \le 1} \max_{\|y\|_{q^{*}} \le 1} \langle y, Ax \rangle \\ &= \max_{\|x\|_{p} \le 1} \max_{\|y\|_{q^{*}} \le 1} \left\langle A^{T}y, x \right\rangle &= \left\|A^{T}\right\|_{q^{*} \to p^{*}} \end{aligned}$$

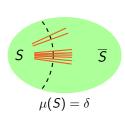
- p and p^* define dual norms if $\frac{1}{p} + \frac{1}{p^*} = 1$.

$$- \|x\|_{p^*} = \max_{\|y\|_p \le 1} \langle x, y \rangle.$$

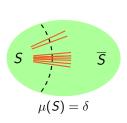
$$\begin{aligned} - & \|A\|_{p \to q} &= \left\|A^{T}\right\|_{q^{*} \to p^{*}} \\ & \|A\|_{p \to q} &= \max_{\|x\|_{p} \le 1} \|Ax\|_{q} \\ &= \max_{\|x\|_{p} \le 1} \max_{\|y\|_{q^{*}} \le 1} \langle y, Ax \rangle \\ &= \max_{\|x\|_{p} \le 1} \max_{\|y\|_{q^{*}} \le 1} \left\langle A^{T}y, x \right\rangle &= \left\|A^{T}\right\|_{q^{*} \to p^{*}} \end{aligned}$$

Hypercontractive remains hypercontractive $(p \le q \implies q^* \le p^*)$.



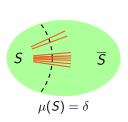


$$1 - \Phi_G(S) = \frac{\langle \mathbb{1}_S, A_G \mathbb{1}_S \rangle}{\delta}$$



$$1 - \Phi_{G}(S) = \frac{\langle \mathbb{1}_{S}, A_{G} \mathbb{1}_{S} \rangle}{\delta}$$

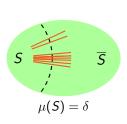
$$\leq \frac{\|\mathbb{1}_{S}\|_{q^{*}} \|A_{G} \mathbb{1}_{S}\|_{q}}{\delta}$$



$$1 - \Phi_{G}(S) = \frac{\langle \mathbb{1}_{S}, A_{G} \mathbb{1}_{S} \rangle}{\delta}$$

$$\leq \frac{\|\mathbb{1}_{S}\|_{q^{*}} \|A_{G} \mathbb{1}_{S}\|_{q}}{\delta}$$

$$\leq \|A_{G}\|_{p \to q} \cdot \frac{\|\mathbb{1}_{S}\|_{q^{*}} \|\mathbb{1}_{S}\|_{p}}{\delta}$$



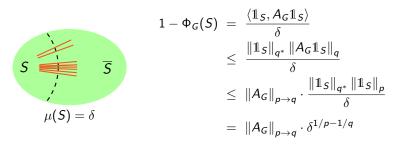
$$1 - \Phi_{G}(S) = \frac{\langle \mathbb{1}_{S}, A_{G} \mathbb{1}_{S} \rangle}{\delta}$$

$$\leq \frac{\|\mathbb{1}_{S}\|_{q^{*}} \|A_{G} \mathbb{1}_{S}\|_{q}}{\delta}$$

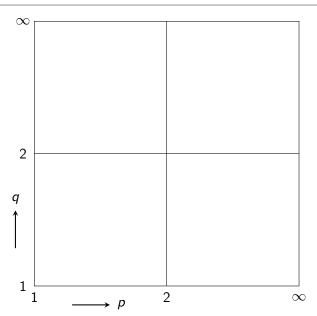
$$\leq \|A_{G}\|_{p \to q} \cdot \frac{\|\mathbb{1}_{S}\|_{q^{*}} \|\mathbb{1}_{S}\|_{p}}{\delta}$$

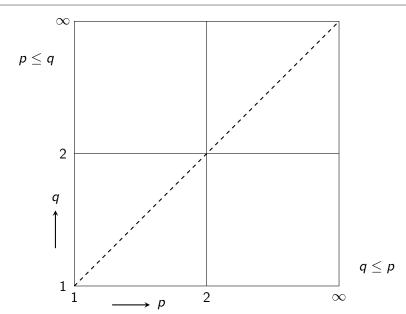
$$= \|A_{G}\|_{p \to q} \cdot \delta^{1/p - 1/q}$$

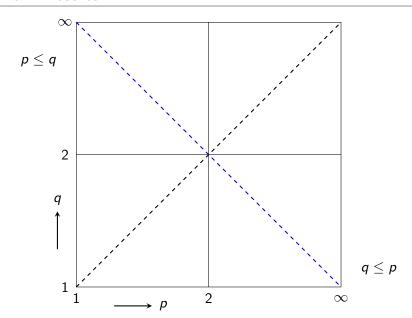
- Bound on any hypercontractive norm of A_G implies small-set expansion of graph G.

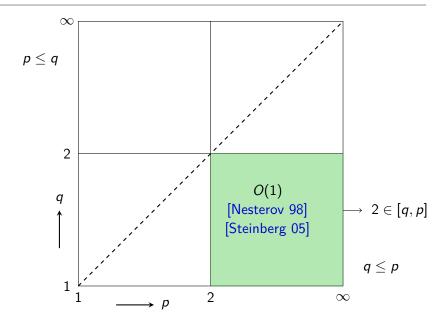


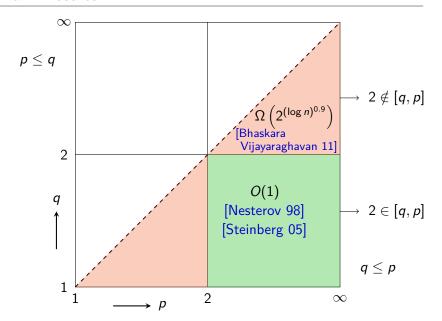
- [BBHKSZ 12]: Two-sided connection for 2 ightarrow q norm of related matrix A_G' .

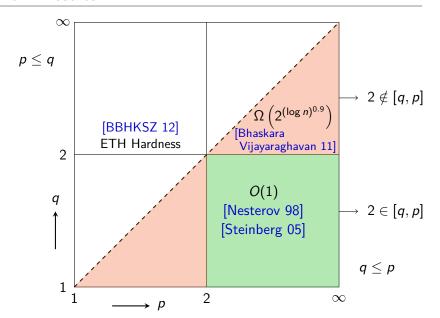


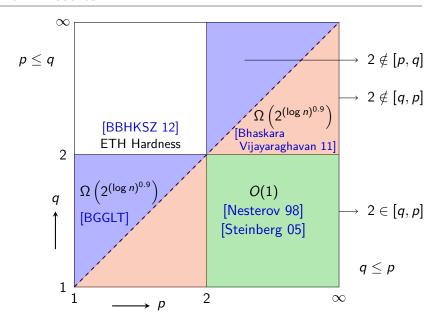


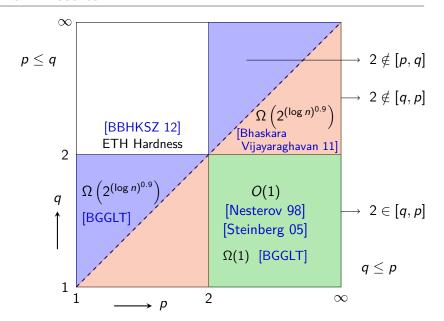


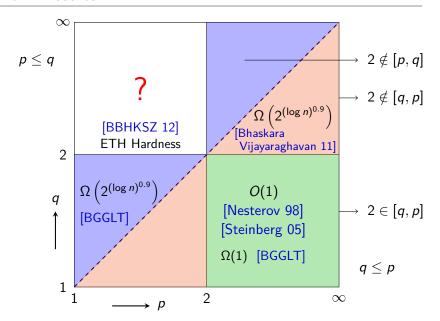












- $\|A\|_{p o q}$ is hard to approximate within $2^{(\log n)^{1-\epsilon}}$ for any $\epsilon > 0$ when $2 \notin [p,q]$ (assuming NP not in randomized quasipolynomial time). Hardness of $(2+\delta) \to q$ norm for any $\delta > 0$.

- $\|A\|_{p o q}$ is hard to approximate within $2^{(\log n)^{1-\epsilon}}$ for any $\epsilon > 0$ when $2 \notin [p,q]$ (assuming NP not in randomized quasipolynomial time). Hardness of $(2+\delta) \to q$ norm for any $\delta > 0$.
- $\|A\|_{p \to q}$ is hard to approximate within $1/(\gamma_{p^*} \cdot \gamma_q)$ when $2 \in [q, p]$, where $\gamma_r = \|g\|_r$ for $g \sim N(0, 1)$.

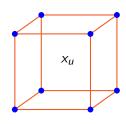
- $\|A\|_{p o q}$ is hard to approximate within $2^{(\log n)^{1-\epsilon}}$ for any $\epsilon > 0$ when $2 \notin [p,q]$ (assuming NP not in randomized quasipolynomial time). Hardness of $(2+\delta) \to q$ norm for any $\delta > 0$.
- $\|A\|_{p \to q}$ is hard to approximate within $1/(\gamma_{p^*} \cdot \gamma_q)$ when $2 \in [q, p]$, where $\gamma_r = \|g\|_r$ for $g \sim N(0, 1)$.
- Gadget reductions had similar bottlenecks for all $p \le q$. Can overcome when $2 \notin [p,q]$.

- $\|A\|_{p o q}$ is hard to approximate within $2^{(\log n)^{1-\epsilon}}$ for any $\epsilon > 0$ when $2 \notin [p,q]$ (assuming NP not in randomized quasipolynomial time). Hardness of $(2+\delta) \to q$ norm for any $\delta > 0$.
- $\|A\|_{p \to q}$ is hard to approximate within $1/(\gamma_{p^*} \cdot \gamma_q)$ when $2 \in [q, p]$, where $\gamma_r = \|g\|_r$ for $g \sim N(0, 1)$.
- Gadget reductions had similar bottlenecks for all $p \le q$. Can overcome when $2 \notin [p, q]$.
- Some evidence that the case when $2 \in [p, q]$ may be different from $2 \notin [p, q]$?

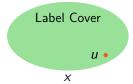
Hardness of proving hypercontractive hardness (via gadgets)



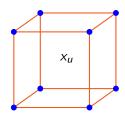
Hardness of proving hypercontractive hardness (via gadgets)

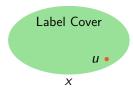


- Usually $x_u = \left(\widehat{f_u}(\{1\}), \dots, \widehat{f_u}(\{R\})\right)$ for $f_u : \{-1,1\}^R \to \{-1,1\}$ defining encoding of label-cover solution.



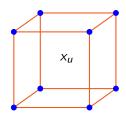
Hardness of proving hypercontractive hardness (via gadgets)

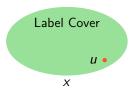




- Usually $x_u = \left(\widehat{f_u}(\{1\}), \dots, \widehat{f_u}(\{R\})\right)$ for $f_u : \{-1, 1\}^R \to \{-1, 1\}$ defining encoding of label-cover solution.
- Want global solution to be spread (nonzero mass on most blocks u).
- Need to prove soundness by viewing x as Label-Cover solution.

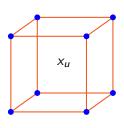
Hardness of proving hypercontractive hardness (via gadgets)



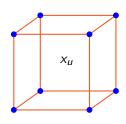


- Usually $x_u = \left(\widehat{f_u}(\{1\}), \dots, \widehat{f_u}(\{R\})\right)$ for $f_u : \{-1,1\}^R \to \{-1,1\}$ defining encoding of label-cover solution.
- Want global solution to be spread (nonzero mass on most blocks u).
- Need to prove soundness by viewing x as Label-Cover solution.
- $\frac{\|x\|_q}{\|x\|_p}$ maximized when x is sparse (for p < q). No mass on most blocks.

Hardness of $2 \rightarrow r$ for r < 2 (extending [Briët Regev Saket 15])



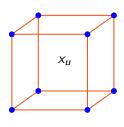
-
$$x_u = (\widehat{f}_u(\{1\}), \dots, \widehat{f}_u(\{R\}))$$



$$-x_u = (\widehat{f}_u(\{1\}), \ldots, \widehat{f}_u(\{R\}))$$

- Rademacher gadget:

$$x_u \stackrel{A}{\longrightarrow} \left\{ \sum_{i=1}^R \epsilon_i \cdot x_{u,i} \right\}_{\epsilon_1,\dots,\epsilon_R \in \{-1,1\}}$$

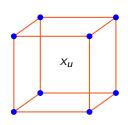


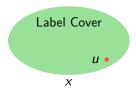
$$-x_u = (\widehat{f}_u(\{1\}), \ldots, \widehat{f}_u(\{R\}))$$

- Rademacher gadget:

$$x_u \stackrel{A}{\rightarrow} \left\{ \sum_{i=1}^R \epsilon_i \cdot x_{u,i} \right\}_{\epsilon_1,\dots,\epsilon_R \in \{-1,1\}}$$

- Enforces local sparsity. For $\|x_u\|_2 \le 1$:
 - $||Ax_u||_r = 1$ for $x_u = (1, 0, \dots, 0)$.
 - $\|Ax_u\|_r \le \gamma_r < 1$ for spread x_u .



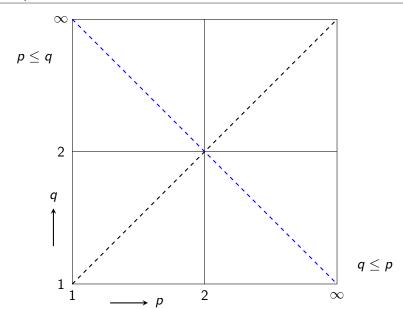


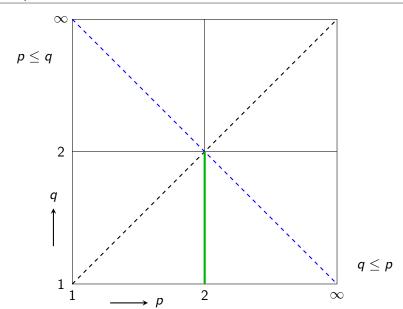
$$-x_u = (\widehat{f}_u(\{1\}), \ldots, \widehat{f}_u(\{R\}))$$

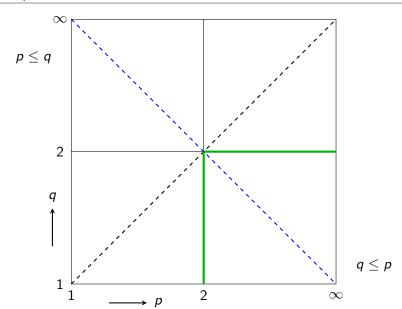
- Rademacher gadget:

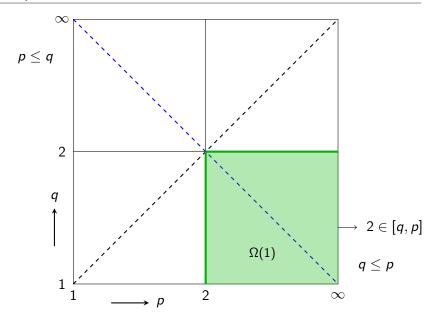
$$x_u \xrightarrow{A} \left\{ \sum_{i=1}^R \epsilon_i \cdot x_{u,i} \right\}_{\epsilon_1,\dots,\epsilon_R \in \{-1,1\}}$$

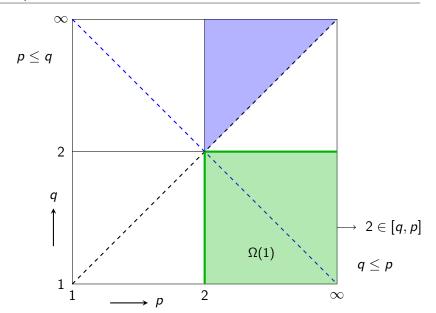
- Enforces local sparsity. For $||x_u||_2 \le 1$:
 - $||Ax_u||_r = 1$ for $x_u = (1, 0, \dots, 0)$.
 - $\|Ax_u\|_r \le \gamma_r < 1$ for spread x_u .
 - Globally project to linear space implied by (smooth) label cover constraints.
 r < 2 ensures global spread.





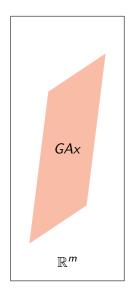






- $1/\gamma_r$ hardness for $\|\cdot\|_{2\to r}$ when r<2.

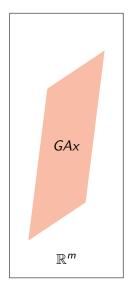
- $1/\gamma_r$ hardness for $\left\|\cdot\right\|_{2 \to r}$ when r < 2.
- $1/\gamma_{p^*}$ hardness for $\|A\|_{p\to 2} = \|A^T\|_{2\to p^*}$. $(p>2 \implies p^*<2)$



- $1/\gamma_r$ hardness for $\|\cdot\|_{2\to r}$ when r<2.
- $1/\gamma_{p^*}$ hardness for $\|A\|_{p\to 2} = \|A^T\|_{2\to p^*}$. $(p>2 \implies p^*<2)$
- For random Gaussian matrix $G \in \mathbb{R}^{m \times n}$ and $z \in \mathbb{R}^n$

$$\|Gz\|_q \approx \|z\|_2$$

[Dvoretzky]: Simultaneously $\forall z$ if $m \geq n^{q/2}$.



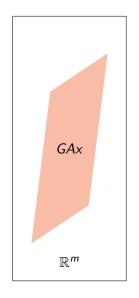
- $1/\gamma_r$ hardness for $\|\cdot\|_{2\to r}$ when r<2.
- $1/\gamma_{p^*}$ hardness for $\|A\|_{p\to 2} = \|A^T\|_{2\to p^*}$. $(p>2 \implies p^*<2)$
- For random Gaussian matrix $G \in \mathbb{R}^{m \times n}$ and $z \in \mathbb{R}^n$

$$\|Gz\|_q \approx \|z\|_2$$

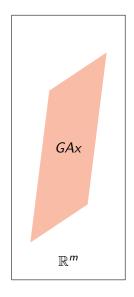
[Dvoretzky]: Simultaneously $\forall z$ if $m \geq n^{q/2}$.

- Reduction from p o 2 norm, to p o q norm

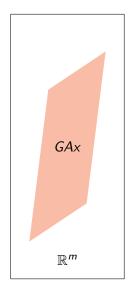
$$\|\mathit{GAx}\|_q \approx \|\mathit{Ax}\|_2$$
 for all x
 $\|\mathit{GA}\|_{p \to q} \approx \|\mathit{A}\|_{p \to 2}$



- $\|GA\|_{p\to q} \approx \|A\|_{p\to 2}$

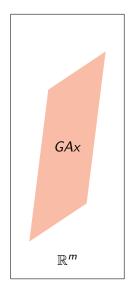


- $\|GA\|_{p \to q} \approx \|A\|_{p \to 2}$
- Matrix A derived from Label Cover via gadgets. G represents (simple) global transformation.



-
$$\|GA\|_{p \to q} \approx \|A\|_{p \to 2}$$

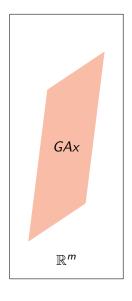
- Matrix A derived from Label Cover via gadgets. G represents (simple) global transformation.
- Constant factor $(1/\gamma_{p^*})$ hardness for $\|\cdot\|_{p \to q}$.



$$- \|GA\|_{p \to q} \approx \|A\|_{p \to 2}$$

- Matrix A derived from Label Cover via gadgets. G represents (simple) global transformation.
- Constant factor $(1/\gamma_{p^*})$ hardness for $\left\|\cdot\right\|_{p \to q}$.
- Hardness amplification for p o q norm (via tensoring). For all $p \le q$

$$\|A^{\otimes t}\|_{p \to q} = (\|A\|_{p \to q})^t$$

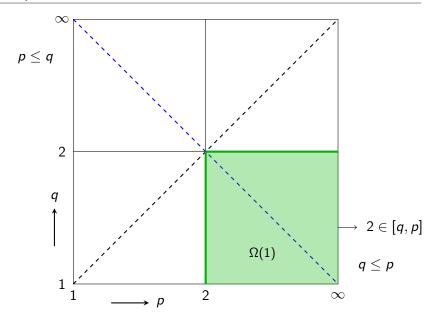


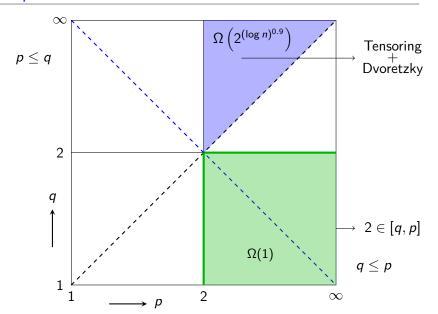
$$- \|GA\|_{p\to q} \approx \|A\|_{p\to 2}$$

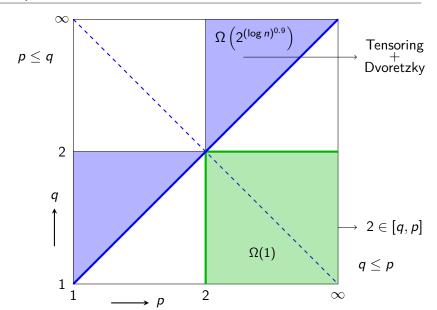
- Matrix A derived from Label Cover via gadgets. G represents (simple) global transformation.
- Constant factor $(1/\gamma_{p^*})$ hardness for $\left\|\cdot\right\|_{p \to q}$.
- Hardness amplification for p o q norm (via tensoring). For all $p \le q$

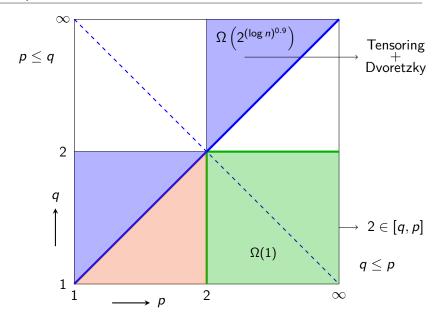
$$\|A^{\otimes t}\|_{p \to q} = (\|A\|_{p \to q})^t$$

- c^t hardness for instances of size $N = n^{O(tq/2)}$. $2^{(\log N)^{1-\epsilon}}$ hardness for $t = (\log n)^{O(1/\epsilon)}$.

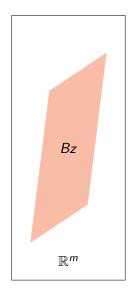








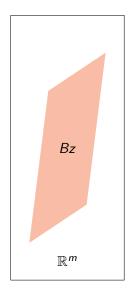
- Previous result gives strong hardness for $\|\cdot\|_{p \to p}$ when p < 2, since $\|A\|_{p \to p} = \|A^T\|_{p^* \to p^*}$ and $p^* > 2$.



- Previous result gives strong hardness for $\|\cdot\|_{p\to p}$ when p<2, since $\|A\|_{p\to p}=\|A^T\|_{p^*\to p^*}$ and $p^*>2$.
- [Schechtman 87]: For $q , there exists samplable distribution of matrices <math>B \in \mathbb{R}^{m \times n}$ (with $m = O(n^3)$) such that w.h.p.

$$\|Bz\|_q \approx \|z\|_p \quad \forall z \in \mathbb{R}^n.$$

(slight modification of Schechtman's result).



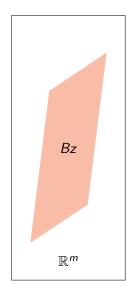
- Previous result gives strong hardness for $\|\cdot\|_{p\to p}$ when p<2, since $\|A\|_{p\to p}=\|A^T\|_{p^*\to p^*}$ and $p^*>2$.
- [Schechtman 87]: For $q , there exists samplable distribution of matrices <math>B \in \mathbb{R}^{m \times n}$ (with $m = O(n^3)$) such that w.h.p.

$$\|Bz\|_q \approx \|z\|_p \quad \forall z \in \mathbb{R}^n.$$

(slight modification of Schechtman's result).

- Reduction from p o p norm, to p o q norm

$$\|BAx\|_q \approx \|Ax\|_p$$
 for all x
 $\|BA\|_{p \to q} \approx \|A\|_{p \to p}$



- Previous result gives strong hardness for $\|\cdot\|_{p\to p}$ when p<2, since $\|A\|_{p\to p}=\|A^T\|_{p^*\to p^*}$ and $p^*>2$.
- [Schechtman 87]: For $q , there exists samplable distribution of matrices <math>B \in \mathbb{R}^{m \times n}$ (with $m = O(n^3)$) such that w.h.p.

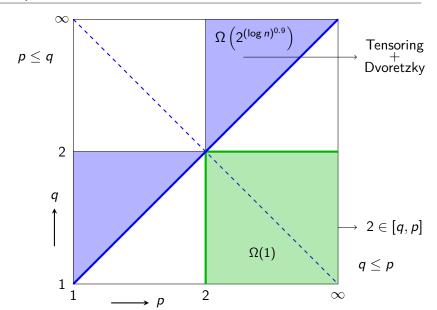
$$\|Bz\|_q \approx \|z\|_p \quad \forall z \in \mathbb{R}^n.$$

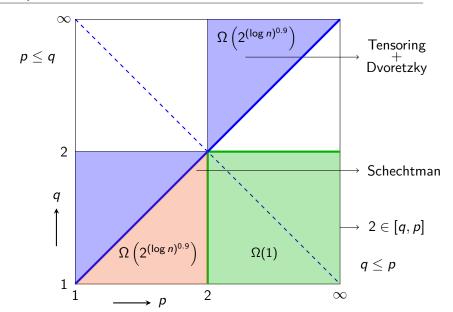
(slight modification of Schechtman's result).

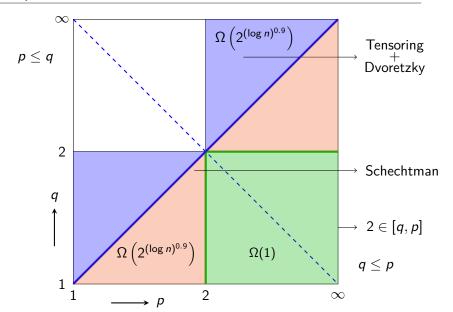
- Reduction from p o p norm, to p o q norm

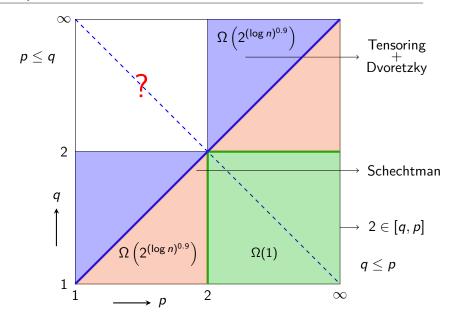
$$\|BAx\|_q \approx \|Ax\|_p$$
 for all x
 $\|BA\|_{p \to q} \approx \|A\|_{p \to p}$

- Simplified (but randomized) proof of [BV 11].









- The case when $p \le 2 \le q$. Hardness of $2 \to 4$ norm? Of any degree-4 polynomial over ℓ_2 ball.

- The case when $p \le 2 \le q$. Hardness of $2 \to 4$ norm? Of any degree-4 polynomial over ℓ_2 ball.
- Hardness of $(2 + \delta) \rightarrow 4$ norm. How small can δ be?

- The case when $p \le 2 \le q$. Hardness of $2 \to 4$ norm? Of any degree-4 polynomial over ℓ_2 ball.
- Hardness of $(2 + \delta) \rightarrow$ 4 norm. How small can δ be?
- Known algorithms for Unique Games, SSE imply sub-exponential algorithms for $2 \to q$ norms. Can this be extended to all $p \le q$ when $2 \in [p,q]$?

- The case when $p \le 2 \le q$. Hardness of $2 \to 4$ norm? Of any degree-4 polynomial over ℓ_2 ball.
- Hardness of $(2 + \delta) \rightarrow$ 4 norm. How small can δ be?
- Known algorithms for Unique Games, SSE imply sub-exponential algorithms for $2 \to q$ norms. Can this be extended to all $p \le q$ when $2 \in [p,q]$?
- Right (form of) approximation ratio when $p \ge q$ and $2 \in [q, p]$.
 - Hardness result is tight when *p* or *q* equals 2.
 - [BGGLT 18]: Matching approximation ratio (up to uncertainty in value of K_G).

Thank You

Questions?