Hardy's inequalities: unification and application to an isoperimetric inequality

Jon A. Wellner

University of Washington, Seattle

High Dimensional Probability 9

June 17, 2020

Based on joint work with: Chris Klaassen and Adrien Saumard

Outline

- 1: Hardy's inequalities and some relatives
- 2: Hardy's two inequalities united via probability
- 3: Isoperimetric inequalities and covariance inequalities
- 4: Summary: questions and open problems.

1. Hardy's inequalities

First the continuous (or integral form) inequality: For a non-negative function ψ on $(0, \infty)$ and $p > 1$,

$$\int_0^\infty \left(\frac{1}{x} \int_0^x \psi(y) dy \right)^p dx \le \left(\frac{p}{p-1} \right)^p \int_0^\infty \psi^p(x) dx. \tag{1}$$

On the other hand, the the discrete (or series form) inequality is: for a sequence $\{a_n\}$ of non-negative real numbers and $p > 1$

$$\sum_{n=1}^\infty \left(\frac{1}{n} \sum_{k=1}^n a_k \right)^p \le \left(\frac{p}{p-1} \right)^p \sum_{n=1}^\infty a_n^p. \tag{2}$$

When G. H. Hardy developed these inequalities during the period 1915 - 1925, probability theory was non-existent or nearly so, and hence it is not too surprising that neither of these statements involves probability theory. Averages appear on the left sides, but this almost seems to be an accident. As Persi Diaconis has written in his (2002) London Math. Soc. paper "G. H. Hardy and Probability???", *I want to argue that Hardy had no knowledge of probability theory, and indeed had a genuine antipathy to the subject. ... At the time when he was working, the mathematical underpinnings of probability were a vague mess; ....*

Unification via probability

Here we propose the following unification of Hardy's inequalities (1) and (2):

Theorem 1: For any distribution function F on $\mathbb{R}$, non-negative function ψ on $\mathbb{R}$, and $p > 1$

$$\int_{\mathbb{R}} \left(\frac{1}{F(x)} \int_{(-\infty,x]} \psi dF \right)^p dF(x) \;\leq\; \left(\frac{p}{p-1} \right)^p \int_{\mathbb{R}} \psi^p(x) dF(x). \quad (3)$$

Corollary:
(i) For any $p > 1$ and non-negative $\psi \in L^p$, (1) holds.
(ii) For any $p > 1$ and non-negative sequence $\{a_n\} \in \ell^p$, (2) holds.

Proof of the Corollary:

(i) Follows from Theorem 1: take F to be the distribution function corresponding to the uniform probability measure on $[0, K]$, multiply by K, and let $K \to \infty$.

(ii) Follows from Theorem 1: take F to be the distribution function corresponding to the uniform probability measure on $\{1, \ldots, K\}$, multiply by K, and let $K \to \infty$.

We can rewrite Theorem 1 in terms of random variables: let X and Y be independent random variables with right-continuous distribution function F on $(\mathbb{R}, \mathcal{B})$, and let ψ be a non-negative, measurable function on $(\mathbb{R}, \mathcal{B})$. Then if $p > 1$

$$E\left(\left[\frac{E\left(\psi(Y)\mathbf{1}_{[Y \leq X]} \mid X\right)}{F(X)}\right]^p\right) \leq \left(\frac{p}{p-1}\right)^p E\left(\psi^p(Y)\right). \tag{4}$$

For absolutely continuous distribution functions F the constant $(p/(p-1))^p$ is the smallest possible.

Hardy's inequalities (1), (2), and our new inequality (4) are formulated in terms of "left tail averages". By replacing X by $-X$ and F by $1-F_-(-x)$, we obtain an analogue of (3) expressed in terms of "right tail averages" as follows:

Corollary 2: For $p>1$ and $\psi \geq 0$, $\psi \in L_p(F)$,

$$E\left(\left[\frac{E\left(\psi(Y)\mathbf{1}_{[Y\geq X]} \mid X\right)}{1-F(X-)}\right]^p\right) \leq \left(\frac{p}{p-1}\right)^p E\left(\psi^p(Y)\right). \tag{5}$$

Q: Is there a version of Hardy's inequality with both left and right tail averaging? To this end consider the averaging operators T_c defined as follows: for $c \in \mathbb{R}$ and $\psi \geq 0$ in $L_p(F)$

$$T_c\psi(x) = \frac{\mathbf{1}_{(-\infty,c]}(x)}{F(x)}\int_{(-\infty,x]}\psi dF + \frac{\mathbf{1}_{(c,\infty)}(x)}{1-F(x)}\int_{(x,\infty)}\psi dF.$$

Corollary 3: For $c \in \mathbb{R}$, $p > 1$ and $\psi \geq 0$, $\psi \in L_p(F)$

$$E\left[T_c\psi(X)\right]^p \leq \left(\frac{p}{p-1}\right)^p E\left(\psi^p(Y)\right).$$

Proof: Since $(X|X \leq c)$ has d.f. $F(\cdot)/F(c)$ and similarly for the right tail $(X|X > c)$ has survival function $(1 - F(\cdot))/(1 - F(c))$, we can compute and use (4) and (5) conditionally: thus we find that

$$\begin{aligned} I &\equiv E\left(\left[\frac{E\left(\psi(Y)\mathbf{1}_{[Y \leq X]} \mid X\ \right)}{F(X)}\right]^p \middle| X \leq c, \right) \\ &= E\left(\left[\frac{E\left(\psi(Y)\mathbf{1}_{[Y \leq X]} \mid X, Y \leq c\right)}{F(X)/F(c)}\right]^p \middle| X \leq c \right) \\ &\leq \left(\frac{p}{p-1}\right)^p E\left(\psi^p(Y)|X \leq c\right). \end{aligned}$$

Similarly,

$$\begin{aligned} II &\equiv E\left(\left[\frac{E\left(\psi(Y)\mathbf{1}_{[Y\geq X]} \mid X\right)}{1-F(X-)}\right]^p \Bigg| X>c, \right) \\ &= E\left(\left[\frac{E\left(\psi(Y)\mathbf{1}_{[Y\leq X]} \mid X, Y>c\right)}{(1-F(X-))/(1-F(c))}\right]^p \Bigg| X>c \right) \\ &\leq \left(\frac{p}{p-1}\right)^p E\left(\psi^p(Y) | X>c\right) \end{aligned}$$

Thus, noting that $E\left[T_c\psi(X)\right]^p = I + II$, yields

$$E\left[T_c\psi(X)\right]^p = I + II \leq \left(\frac{p}{p-1}\right)^p E\left(\psi^p(Y)\right),$$

and this completes the proof of Corollary 3. □

Isoperimetric inequalities and covariance inequalities

Let μ be a probability measure on $\mathbb{R}^d$ and define μ^+ by

$$\mu^+(A) \equiv \liminf_{h \searrow 0} \frac{\mu(A^h) - \mu(A)}{h}$$

where

$$A^h \equiv \{x \in \mathbb{R}^d : \ |x - a| < h \ \text{ for some } a \in A\}.$$

If we can show that

$$\mu^+(A) \geq c \min\{\mu(A), \ 1 - \mu(A)\} \quad \text{for all} \quad A \in \mathcal{B}^d,$$

then the optimal (largest value of) $c \equiv Is(\mu)$ is the *Cheeger constant* of μ. It turns out that $Is(\mu)$ is related to the best constant in the classical Poincaré inequality: if $\lambda > 0$ satisfies

$$\lambda Var_\mu(g) \leq \int |\nabla g|^2 d\mu$$

for all locally Lipschitz g, then

$$\lambda \geq Is(\mu)^2/4.$$

For a measure μ on $\mathbb{R}$ it turns out that

$$Is(\mu) = \text{essinf}_{a<x<b} \frac{f(x)}{\min\{F(x), 1 - F(x)\}}$$

where $f =$ density of F_{ac}; here $F = F_c + F_d = F_{ac} + F_s + F_d$ is the distribution function of μ.

$Is(\mu)$ is also the optimal constant satisfying the following L^1–Poincaré inequality

$$c \int |g - \text{med}_\mu(g)| d\mu \leq \int |\nabla g| d\mu.$$

This last inequality can also be viewed as a covariance inequality: since $\text{sign}(x) = 2 \cdot 1_{[x \geq 0]} - 1$,the definition of the median yields

$$E_\mu \left[\text{sign}(g - \text{med}_\mu(g)) d\mu \right] = \int \text{sign}(g - \text{med}_\mu(g)) d\mu = 0$$

and hence ...

$$\int |g - \text{med}_\mu(g)| d\mu = Cov_\mu(g - \text{med}_\mu(g), \text{sign}(g - \text{med}_\mu(g)))$$

where

$$Cov_\mu(g,h) = \int g(h - E_\mu h) d\mu = \int gh d\mu - \int g d\mu \int h d\mu.$$

Menz and Otto (2013) established the following asymmetric Brascamp-Lieb inequality: if μ is a strictly log-concave measure on $\mathbb{R}$, then for smooth and square integrable g, h on $\mathbb{R}$,

$$\left| Cov_\mu(g,h) \right| \leq \left\| g' \right\|_1 \cdot \left\| \frac{h'}{\varphi''} \right\|_\infty$$

where φ is the potential of μ: $d\mu = e^{-\varphi} dx$.

This has been generalized to $\mathbb{R}^d$ and to L_p–L_q by Carlen, Cordero-Erausquin, and Lieb (2013): for a strictly log-concave measure μ on $\mathbb{R}^d$, $d\mu = \exp(-\varphi) dx$, and square integrable locally Lipschitz functions g and H, and $p \in [2, \infty)$ with $1/p + 1/q = 1$,

$$\left|Cov_\mu(g,h)\right| \leq \left\|\lambda_{min}^{(2-p)/p} \mathsf{Hess}_\varphi^{-1/p} \nabla g\right\|_q \left\|\mathsf{Hess}_\varphi^{-1/p} \nabla h\right\|_p$$

where $\lambda_{min}(x)$ is the least eigenvalue of $\mathsf{Hess}_\varphi(x)$.

But this type of covariance inequality relies strongly on the log-concavity of ν. If $\lambda_{min}(x) \geq \rho > 0$ (which implies that μ is strongly log-concave), the best inequalities relating $Is^2(\mu)$ to λ_P (the optimal Poincaré constant) and to λ_{min} are

$$\rho \leq \lambda_P \quad \text{and} \quad \frac{Is(\mu)^2}{4} \leq \lambda_P \leq 36 Is(\mu)^2$$

To understand the type of inequality we are aiming for, suppose that μ satisfies an L_1–Poincaré inequality (with centering by a median): i.e. for every smooth integrable g we have

$$c_1 \|g - \text{med}_\mu g\|_1 \leq \|g'\|_1.$$

Then if $h \in L_\infty$ we get

$$\begin{aligned} |Cov(g,h)| &= |E[(g - \text{med} g)(h - E(h))]| \\ &\leq \|g - \text{med}_\mu g\|_1 \cdot \|h - E(h)\|_\infty \\ &\leq c_1^{-1} \|g'\|_1 \cdot \|h_0\|_\infty. \end{aligned}$$

The optimal $c_1 = Is(\mu)$. The following inequality from Saumard and W (2019) begins the connection to Hardy's inequalities.

Theorem 1. (Saumard and W, 2019)
Let μ be a probability measure with a positive density f on $\mathbb{R}$, d.f. F, and median $m \in \mathbb{R}$. Let $g \in L_\infty(F)$ and $h \in L_1(F)$. Assume also that g and h are absolutely continuous. Then

$$
\begin{aligned}
& Is(\mu)|Cov(g,h)| \\
& \quad \le \ \max\left\{ \sup_{a<x\le m} \left| \frac{\int_a^x h dF}{F(x)} - Eh \right|, \sup_{m<x<b} \left| \frac{\int_x^b h dF}{1-F(x)} - Eh \right| \right\} \cdot \int_{\mathbb{R}} |g'| dF
\end{aligned}
$$

where $a \equiv \inf\{x: \ F(x) > 0\}$, $\equiv \sup\{x: \ F(x) < 1\}$.

Proof: From Hoeffding (1940) it follows that

$$
Cov(g(X), h(X)) = \int_{\mathbb{R}^2} g'(x) K_\mu(x,y) h'(y) dx dy
$$

where $K_\mu(x,y) = F(x \wedge y) - F(x)F(y)$. Then use Fubini's theorem.

Theorem 2. (Saumard and W, 2019)
Let μ be a probability measure with a positive density f on $\mathbb{R}$, cumulative distribution function F, and median $m \in \mathbb{R}$. Let $g \in L_p(F)$ and $h \in L_q(F)$ with $p^{-1} + q^{-1} = 1$. Assume also that g and h are absolutely continuous. Then

$$\begin{aligned} Is(\mu)|Cov(g,h)| &\leq \|g'\|_p \|T_m(h_0)\|_q \\ &\leq \|g'\|_p \frac{q}{q-1} \|h_0\|_q = p\|g'\|_p \|h_0\|_q \end{aligned} \quad (6)$$

where, with $h_0 \equiv h - \int h dF$,

$$\begin{aligned} T_m(h_0)(x) &\equiv T_m h_0(x) \\ &= \frac{1_{(a,m]}(x)}{F(x)} \int_{(a,x]} h_0 dF + \frac{1_{(m,b)}(x)}{1-F(x)} \int_{(x,b)} h_0 dF \end{aligned}$$

and where $a \equiv \inf\{x: \ F(x) > 0\}$, $b \equiv \sup\{x: \ F(x) < 1\}$ as before.

The second inequality in (6) follows from a "two-sided" Hardy inequality as follows: for $k \in \overline{\mathbb{R}}$

$$\|T_k(h)\|_q^q \leq \left(\frac{q}{q-1}\right)^q \|h\|_q^q.$$

In Saumard and W (2019) we proved this for any continuous distribution F. From the results in Klaassen and W (2020), we now know that this holds for an arbitrary distribution function.

Corollary: Let μ be a probability measure with positive density f on $\mathbb{R}$ and d.f. F. Let $g \in L_2(F)$ be absolutely continuous. Then

$$Var_\mu(g) \leq \frac{4}{Is(\mu)^2}\|g'\|_2^2$$

and hence the optimal constant λ_1 in the Poincaré inequality satisfies $\lambda_1 \geq (Is(\mu))^2/4$.

Proof: Take $g = h$ and $p = 2$ in Theorem 2. $\square$

5. Open questions and further problems:

- Is there a simpler approach to unification of Hardy's two inequalities via probability theory?
- What are the right inequalities for improving the isoperimetric constants for distributions on $\mathbb{R}^d$? (Improve on the results of Ledoux for log-concave distributions?)
- Martingales and Hardy's inequalities with $p = 2$?
- Other generalizations of Hardy's inequalities:
  reverse Hardy inequalities; Copson's inequality; reverse Copson's inequality (yes!)

References:

- Kufner, A., Maligranda, L., and Persson, L-E. (2006). The Prehistory of the Hardy Inequality. *The American Math. Monthly*, **113**, 715-732.

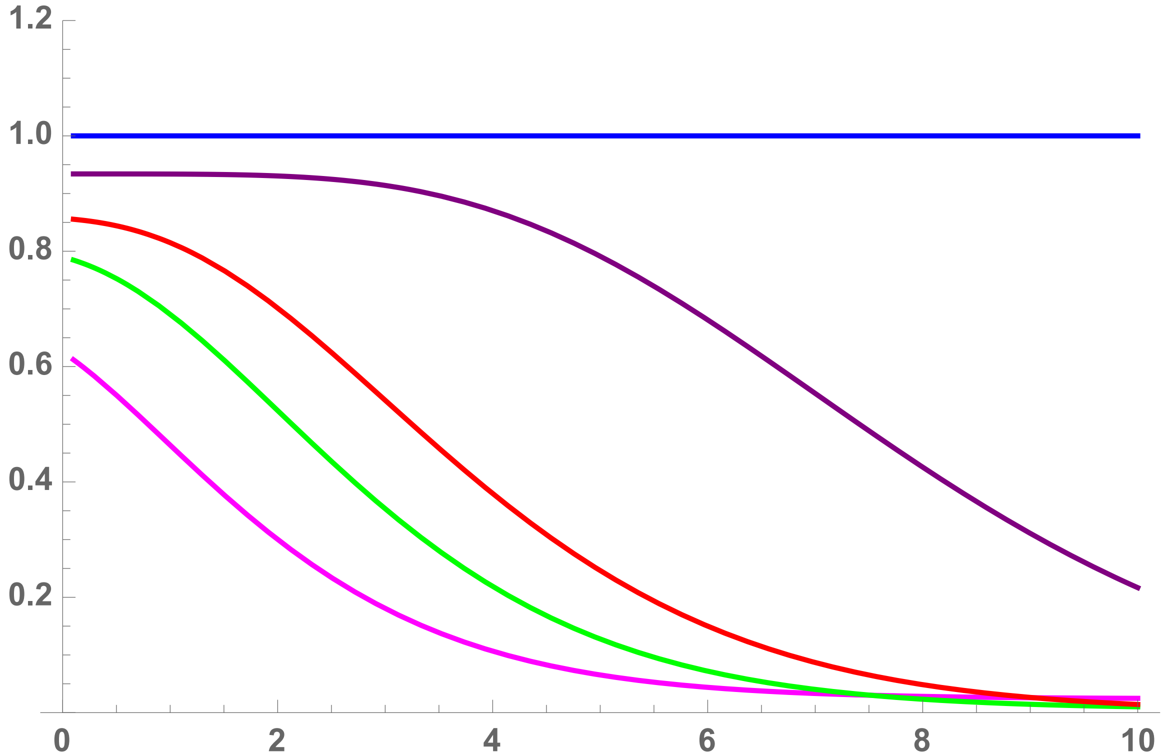
- Diaconis, P. (2002). *Bull. London Math. Soc.* **34**, 385-402.

- Saumard, A. and Wellner, J. A. (2019). On the isoperimetric constant, covariance inequalities and L_p-Poincaré inequalities in dimension one. *Bernoulli* **25**, 1794 - 1815.

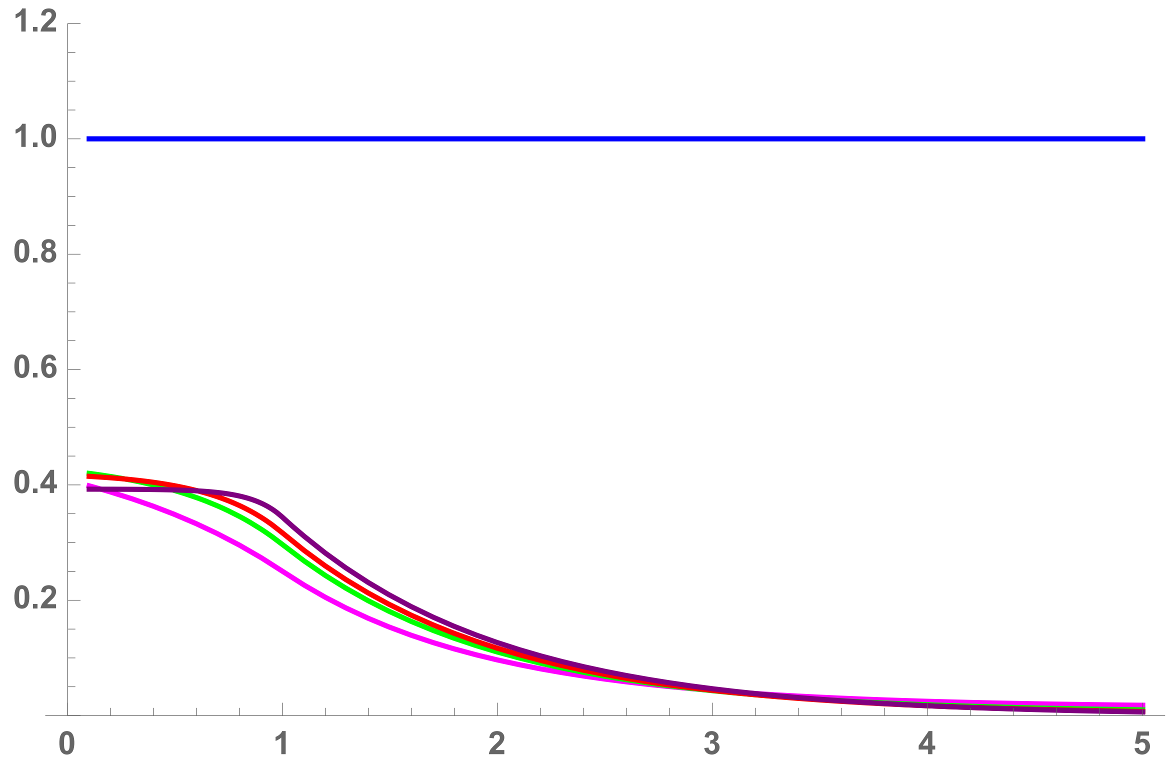
- Bobkov, S.G. and Houdre, C. (1997). Isoperimetric constants for product probability measures. *Ann. Probability* **25**, 184 - 205.

- Klaassen, C. A. J. and W (2020). Hardy's inequality in terms of random variables. *Submitted*.

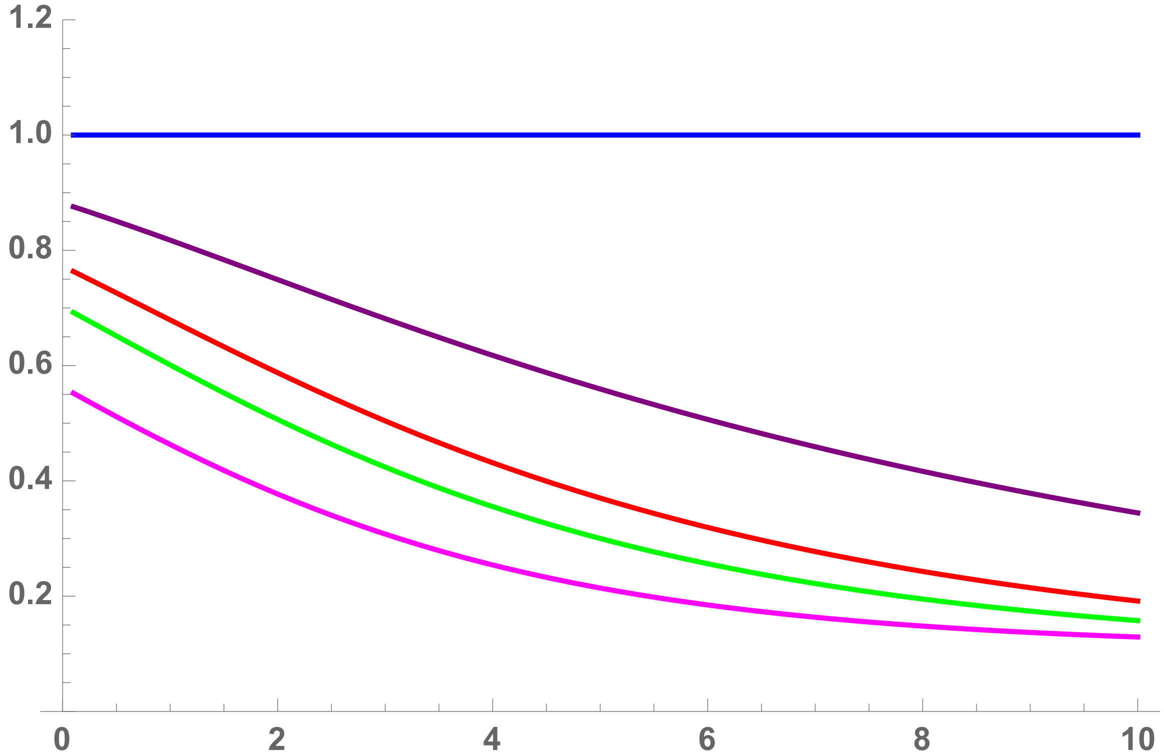
- Ledoux, M. (2004). Spectral gap, logarithmic Sobolev constant, and geometric bounds. *Surveys in Differ. Geom.*, vol. IX, 219 - 240. Int. Press, Somerville, MA.
- Bobkov, S. G. (2007). On isoperimetric constants for log-concave probability distributions. Geometric aspects of functional analysis, 81 - 88, Lecture Notes in Math., 1910, Springer, Berlin, 2007.
- Bobkov, Sergey G. (2009). On the isoperimetric constants for product measures. (2009). Problems in mathematical analysis. No. 40. J. Math. Sci. (N.Y.) 159 (2009), no. 1, 47?53.



Example 1: $f(x) = e^{-x}$, $\psi(x) = x$, $p \in \{2, 3, 4, 8\}$ with 2 =magenta, 3 =green, 4 =red, 8 =purple. The plots show $\|T_c(\psi)\|_p^p / (p/(p-1))^p \|\psi\|_p^p$ as a function of c.



Example 2: $f(x) = e^{-x}$, $\psi(x) = 1_{[1,\infty)}(x)$, $p \in \{2, 3, 4, 8\}$ with 2 =magenta, 3 =green, 4 =red, 8 =purple. The plots show $\|T_c(\psi)\|_p^p / (p/(p-1))^p \|\psi\|_p^p$ as a function of c.



Example 3: $f(x) = e^{-x}$, $\psi(x) = \exp(x/t)$, $(p,t) \in \{(2,3),(3,4),(4,5),(8,9)\}$ with $(2,3) =$magenta, $(3,4) =$green, $(4,5) =$red, $(8,9) =$purple. The plots show $\|T_c(\psi)\|_p^p/(p/(p-1))^p\|\psi\|_p^p$ as a function of c.

Thank You!