



Inequalities for L^p -Norms that Sharpen the Triangle Inequality and Complement Hanner's Inequality

Eric A. Carlen¹ · Rupert L. Frank^{2,3} · Paata Ivanisvili⁴ · Elliott H. Lieb⁵

Received: 6 November 2019 / Published online: 25 May 2020
© The Author(s) 2020

Abstract

In 2006 Carbery raised a question about an improvement on the naïve norm inequality $\|f + g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p)$ for two functions f and g in L^p of any measure space. When $f = g$ this is an equality, but when the supports of f and g are disjoint the factor 2^{p-1} is not needed. Carbery's question concerns a proposed interpolation between the two situations for $p > 2$ with the interpolation parameter measuring the overlap being $\|fg\|_{p/2}$. Carbery proved that his proposed inequality holds in a special case. Here, we prove the inequality for all functions and, in fact, we prove an inequality of this type that is stronger than the one Carbery proposed. Moreover, our stronger inequalities are valid for *all* real $p \neq 0$.

Keywords L^p space · Minkowski's inequality · Convexity

This paper may be reproduced, in its entirety, for non-commercial purposes. Work partially supported by NSF grants DMS-1501007 and DMS-1764254 (E.A.C.), DMS-1363432 (R.L.F.), DMS-1856486 (P.I.), PHY-1265118 (E.H.L.).

✉ Rupert L. Frank
rlfrank@caltech.edu

Eric A. Carlen
carlen@math.rutgers.edu

Paata Ivanisvili
pivanisv@uci.edu

Elliott H. Lieb
lieb@princeton.edu

¹ Department of Mathematics, Hill Center, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA

² Mathematisches Institut, Ludwig-Maximilians Universität München, Theresienstr. 39, 80333 Munich, Germany

³ Mathematics 253-37, Caltech, Pasadena, CA 91125, USA

⁴ Department of Mathematics, University of California, Irvine, CA 92617, USA

⁵ Departments of Mathematics and Physics, Princeton University, Princeton, NJ 08544, USA

1 Introduction and Main Theorem

1.1 Main Result

Since $|z|^p$ is a convex function of z for $p \geq 1$, for any measure space, the L^p unit ball, $\{f : \int |f|^p \leq 1\}$, is convex. One way to express this is with Minkowski’s triangle inequality $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. Another is with the inequality

$$\|f + g\|_p^p \leq 2^{p-1} (\|f\|_p^p + \|g\|_p^p), \tag{1.1}$$

valid for any functions f and g on any measure space. There is equality in (1.1) if and only if $f = g$ and, in our main result (Theorem 1.1), we improve (1.1) substantially when f and g are far from equal.

In 2006 Carbery [3] proposed several plausible refinements of (1.1) for $p \geq 2$, of which the strongest was

$$\int |f + g|^p \leq \left(1 + \frac{\|fg\|_{p/2}}{\|f\|_p \|g\|_p}\right)^{p-1} \int (|f|^p + |g|^p). \tag{1.2}$$

He proved that this inequality holds when f and g are characteristic functions of sets, but left the general case open. Our result provides the first proof of the inequality proposed by Carbery.

The ratio $\Gamma = \frac{\|fg\|_{p/2}}{\|f\|_p \|g\|_p}$, that appears in (1.2), varies between 0 and 1 and, therefore, the factor of $(1 + \Gamma)^{p-1}$ varies between 1 and 2^{p-1} . Thus, (1.2), which we show to be true, is a refinement of (1.1).

In contrast to (1.1), there is equality in (1.2) not only when $f = g$, but also when $fg = 0$. The extreme values 2^{p-1} and 1 of the factor of $(1 + \Gamma)^{p-1}$ correspond to these two cases of equality in (1.2).

Here, we propose and prove a strengthening of (1.2) in which Γ is replaced by the quantity

$$\tilde{\Gamma} := \|fg\|_{p/2} \left(\frac{\|f\|_p^p + \|g\|_p^p}{2}\right)^{-2/p}. \tag{1.3}$$

By virtue of the arithmetic-geometric mean inequality we have $\tilde{\Gamma} \leq \Gamma$ and therefore (1.2) with Γ replaced by $\tilde{\Gamma}$ is a stronger inequality than (1.2).

Moreover, our improved inequalities are not restricted to $p > 2$, but are valid for all $p \in \mathbb{R}$. We write

$$\|f\|_p := \left(\int |f|^p\right)^{1/p} \text{ for all } p \neq 0.$$

We now state our main result. It has three parts. The first part concerns the validity of (1.2) with Γ replaced by $\tilde{\Gamma}$ and its analogue for $p < 2$. This part, in particular, shows

that (1.2) is valid. The second part of the theorem states that, within a natural class of related inequalities, our inequality is best possible. The third part of the theorem settles the cases of equality in our inequality.

Theorem 1.1 (Main Theorem) *For all $p \in (0, 1] \cup [2, \infty)$ and functions f and g on any measure space,*

$$\int |f + g|^p \leq \left(1 + \frac{2^{2/p} \|fg\|_{p/2}}{(\|f\|_p^p + \|g\|_p^p)^{2/p}} \right)^{p-1} \int (|f|^p + |g|^p). \tag{1.4}$$

The inequality reverses if $p \in (-\infty, 0) \cup [1, 2]$, where, for $p \in [1, 2]$, it is assumed that f and g are nonnegative almost everywhere.

For $p \in [2, \infty)$ (resp. for $p \in (0, 1)$), the inequality is false if $\tilde{\Gamma}$ is raised to any power $q > 1$ (resp. $q < 1$).

For $p \in (-\infty, 0)$ (resp. for $p \in (1, 2]$), the reversed inequality is false if $\tilde{\Gamma}$ is raised to any power $q > 1$ (resp. $q < 1$).

For $p \in (0, \infty) \setminus \{1, 2\}$ and $\|f\|_p, \|g\|_p < \infty$, there is equality in (1.4) if and only if f and g have disjoint supports, up to a null set, or are equal almost everywhere.

For $p \in (-\infty, 0)$ and $\|f\|_p, \|g\|_p < \infty$, there is equality in (1.4) if and only if f and g are equal almost everywhere.

We note that Carbery’s proposed inequality (1.2) involves three kinds of quantities on the right side (namely, $\|fg\|_{p/2}$, $\|f\|_p^p + \|g\|_p^p$ and $\|f\|_p \|g\|_p$), while our inequality (1.4) involves only two (namely, $\|fg\|_{p/2}$ and $\|f\|_p^p + \|g\|_p^p$). This both strengthens the result and simplifies the proof.

We note that (1.4) is an equality for $p = 1, 2$ and any nonnegative f and g .

As we already mentioned, Carbery proved that his proposed inequality (1.2) is valid when f and g are characteristic functions. Our theorem can also be easily proved in this special case. We do not see how to use this fact in the proof of the general case.

Another important special case is when f and g are proportional to each other. Even in this special case, inequality (1.4) is quite nontrivial and, in fact, constitutes the core of the proof of Theorem 1.1. We will discuss this momentarily.

1.2 Outline of Our Proof

Our proof of Theorem 1.1 consists of three parts:

Part A: We show how to reduce the inequality to a simpler one involving only one function, namely $\alpha := f/(f + g)$ for $f, g \geq 0$, which takes values in $[0, 1]$, and a reference measure that is a probability measure. The inequality in question is

$$1 \leq \left(1 + \frac{2^{2/p} \|\alpha(1 - \alpha)\|_{p/2}}{(\|\alpha\|_p^p + \|1 - \alpha\|_p^p)^{2/p}} \right)^{p-1} (\|\alpha\|_p^p + \|1 - \alpha\|_p^p) \tag{1.5}$$

for $p \in (0, 1] \cup [2, \infty)$ and its reverse for $p \in (-\infty, 0) \cup [1, 2]$. The reduction exploits the fact that the only important quantity is the ratio of f to g . This part is very easy. The details are presented in Sect. 2.

Part B: In the second part, which is more difficult than Part A, we show that inequality (1.5) (and therefore (1.4) in Theorem 1.1) is true if it is true when the function α is constant. This is the same as saying f and g are proportional to each other on the set where both are nonzero. Our proof of this fact is based on a convexity argument and is presented in Sect. 3.

Part C: With Parts A and B complete, the proof of (1.4) and its reverse reduces to a seemingly elementary inequality, parametrized by p , for a number $\alpha \in [0, 1]$, namely

$$1 \leq \left(1 + \left(\frac{2\alpha^{p/2}(1-\alpha)^{p/2}}{\alpha^p + (1-\alpha)^p} \right)^{2/p} \right)^{p-1} (\alpha^p + (1-\alpha)^p) . \tag{1.6}$$

for $p \in (0, 1] \cup [2, \infty)$ and its reverse for $p \in (-\infty, 0) \cup [1, 2]$. The proof of this is Part C.

While the validity of (1.6) appears to be a consequence of Theorem 1.1, one can also view Theorem 1.1 as a consequence of (1.6).

In order to deal with the optimality statement in the second part of Theorem 1.1, we consider a generalization of (1.6), namely,

$$1 \leq \left(1 + \left(\frac{2\alpha^{p/2}(1-\alpha)^{p/2}}{\alpha^p + (1-\alpha)^p} \right)^q \right)^{p-1} (\alpha^p + (1-\alpha)^p) , \tag{1.7}$$

with a parameter q . Note that the quantity $R := \frac{2\alpha^{p/2}(1-\alpha)^{p/2}}{\alpha^p + (1-\alpha)^p}$ lies in $[0, 1]$ for all α and p and therefore, R^q decreases as q increases. Thus, for $p \in [2, \infty)$, the inequality (1.7) strengthens as q increases, and for $p \in (0, 1]$, it strengthens as q decreases. Likewise, for $p \in [1, 2]$ the reverse of (1.7) is stronger for smaller q , and for $p \in (-\infty, 0)$ it is stronger for larger q .

In Sect. 4 we shall prove the following facts about inequalities (1.6) and (1.7).

Theorem 1.2 *For $p \in (0, 1] \cup [2, \infty)$ and all numbers $\alpha \in [0, 1]$, inequality (1.6) is valid.*

For $p \in (-\infty, 0) \cup [1, 2]$ and all numbers $\alpha \in [0, 1]$, the reversed inequality in (1.6) is valid (where $\alpha \in (0, 1)$ for $p < 0$).

For $p \in [2, \infty)$, (resp. for $p \in (0, 1)$) inequality (1.7) is false if $q > 2/p$, (resp. if $q < 2/p$).

For $p \in (-\infty, 0)$, (resp. for $p \in (1, 2]$) the reversed inequality in (1.7) is false if $q > 2/p$, (resp. if $q < 2/p$).

For $p \in (0, \infty) \setminus \{1, 2\}$, there is equality in (1.6) if and only if $\alpha \in \{0, 1/2, 1\}$.

For $p \in (-\infty, 0)$, there is equality in (1.6) if and only if $\alpha = 1/2$.

Our proof of Theorem 1.2 is elementary, but rather lengthy. We leave it as a challenge to simplify and shorten this proof.

This concludes our outline of the proof of Theorem 1.1. Details will be provided in Sect. 4.3.

1.3 Relation to Other Convexity Inequalities

Theorem 1.1 may be viewed as a refinement of Minkowski’s triangle inequality. Since (1.1), like Minkowski’s inequality, is a direct expression of the convexity of the L^p unit ball, it is equivalent to Minkowski’s inequality. We recall the simple argument: For any unit vectors $u, v \in L^p$, (1.1) says that $\|(u + v)/2\|_p \leq 1$, and then by continuity, $\|\lambda u + (1 - \lambda)v\|_p \leq 1$ for all $\lambda \in (0, 1)$. Suppose $0 < \|f\|_p, \|g\|_p < \infty$, and define $\lambda = \|f\|_p / (\|f\|_p + \|g\|_p)$, $u = \|f\|_p^{-1} f$, and $v = \|g\|_p^{-1} g$. Then

$$\|f + g\|_p^p = (\|f\|_p + \|g\|_p)^p \|\lambda u + (1 - \lambda)v\|_p^p \leq (\|f\|_p + \|g\|_p)^p,$$

which is Minkowski’s inequality.

When $p = 1$ and $f, g \geq 0$, (1.1) is an identity; otherwise when $p > 1$, there is equality in (1.1) if and only if $f = g$. When the supports of f and g are disjoint, however, (1.1) is far from an equality and the factor 2^{p-1} is not needed. There is equality in Minkowski’s inequality whenever f is a multiple of g or *vice-versa*. Hence although (1.1) is equivalent to Minkowski’s inequality, it becomes an equality in fewer circumstances.

There is another well-known refinement of Minkowski’s inequality for $1 < p < \infty$, namely *Hanner’s inequality*, [2,6,9] which gives the exact modulus of convexity of the unit ball in L^p , $B_p := \{f : \int |f|^p \leq 1\}$. For $p \geq 2$, and unit vectors u and v , Hanner’s inequality says that

$$\left\| \frac{u + v}{2} \right\|_p^p + \left\| \frac{u - v}{2} \right\|_p^p \leq 1, \tag{1.8}$$

which is also a consequence of one of Clarkson’s inequalities [1]. When u and v have disjoint supports, $\|u + v\|_p^p = \|u - v\|_p^p = 2$, and then the left hand side is 2^{2-p} , so that for unit vectors u and v , the condition $uv = 0$, which yields equality in the inequality of Theorem 1.1, does not yield equality in Hanner’s inequality. On the other hand, while one can derive a bound on the modulus of convexity in L^p from (1.4), one does not obtain the sharp exact result provided by Hanner’s inequality. Both inequalities express a quantitative strict convexity property of B_p , but neither implies the other; they provide complimentary information, with the information provided by Theorem 1.1 being especially strong when f and g have small overlap as measured by $\|fg\|_{p/2}$.

We also refer to a recent sharpening of Hölder’s inequality in [4].

1.4 Restatement of Theorem 1.2 in Terms of Means

Inequality (1.6) can be restated in terms of q th power means [7]: For $x, y > 0$, define

$$M_q(x, y) = ((x^q + y^q)/2)^{1/q} \text{ if } q \in \mathbb{R} \setminus \{0\} \quad \text{and} \quad M_0(x, y) = \sqrt{xy}.$$

Note that $M_0(x, y)$ is the *geometric mean* of x and y and $M_{-1}(x, y)$ is their *harmonic mean*.

Corollary 1.3 For all $x, y > 0$, and all $p \in (0, 1] \cup [2, \infty)$

$$M_1^p(x, y) \leq \left(\frac{M_p(x, y) + M_{-p}(x, y)}{2} \right)^{p-1} M_p(x, y), \tag{1.9}$$

while the reverse inequality is valid for all $p \in (-\infty, 0) \cup [1, 2]$.

Proof A simple calculation shows that for all $p > 0$, $\frac{M_{-p}(x, y)}{M_p(x, y)} = \frac{2^{2/p}xy}{(x^p + y^p)^{2/p}}$.

Thus, taking $x = \alpha$ and $y = 1 - \alpha$, the inequality (1.6) can be written as

$$\frac{1}{2} \leq \left(1 + \frac{M_{-p}(\alpha, 1 - \alpha)}{M_p(\alpha, 1 - \alpha)} \right)^{p-1} M_p^p(\alpha, 1 - \alpha),$$

Then by homogeneity and the fact that $M_1(\alpha, 1 - \alpha) = 1/2$, (1.6) is equivalent to (1.9) □

The following way to write our inequality sharpens and complements the arithmetic-geometric mean inequality for any two numbers $x, y > 0$, provided one has information on $M_p(x, y)$.

Corollary 1.4 (Improved and complemented AGM inequality) For all $x, y > 0$, and all $p > 2$,

$$1 - \left(\frac{A}{M_p} \right)^{p'} \geq \frac{1}{2} \left(1 - \left(\frac{G}{M_p} \right)^2 \right) \geq \frac{1}{2} \left(1 - \left(\frac{G}{M_{p'}} \right)^2 \right) \geq 1 - \left(\frac{A}{M_{p'}} \right)^p \tag{1.10}$$

where $p' = p/(p - 1)$, $A = (x + y)/2$ and $G = \sqrt{xy}$.

Remark 1.5 Since $p, p' \geq 1$, all of the quantities being compared in these inequalities are nonnegative.

Despite the classical appearance of (1.9), we have not been able to find it in the literature, most of which concerns inequalities for means $M_q(x_1, \dots, x_n) = (\frac{1}{n} \sum_{j=1}^n x_j^q)^{1/q}$ of an n -tuple of nonnegative numbers, often with more general weights. The obvious generalization of (1.9) from two to three nonnegative numbers x, y , and z is false as one sees by taking $z = 0$: Then there is no help from $M_{-p}(x, y, z)$ on the right. A valid generalization to more variables probably involves means over $M_{-p}(x_j, x_k)$ for the various pairs. In any case, as far as we know, (1.9) is new.

1.5 Further Discussion of Inequality (1.6)

A truly remarkable feature of the inequality (1.6) (or, equivalently, (1.9)) is that it is *surprisingly close to equality uniformly in the arguments*. To see this, let $f(\alpha, p)$ denote the right hand side of (1.6). Contour plots of this function for various ranges of p are shown in Figs. 1, 2 and 3.

Figure 1 is a contour plot of this function in $[1/2, 1] \times [2, 4]$. The contours shown in Fig. 1 range from 1.00001 to 1.018. Note that the function f is identically 1 along three sides of plot: $\alpha = 1/2, 1$, and $p = 2$. The maximum value for $2 \leq p \leq 4$, near 1.018, occurs towards the middle of the segment at $p = 4$.

Figure 2 is a contour plot of f on $[1/2, 1] \times [1, 2]$. The contours range from 0.9961 (the small closed contour) to 0.99999999 (close to the boundary). *Amazingly, the function in (1.6) is quite close – within two percent – to the constant 1 over the range $p \geq 1$ and $\alpha \in [0, 1]$* . Moreover, the “landscape” is quite flat: The gradient has a small norm over the whole domain.

Fig. 1 Level plots of $f(\alpha, p)$ on $[1/2, 1] \times [2, 4]$

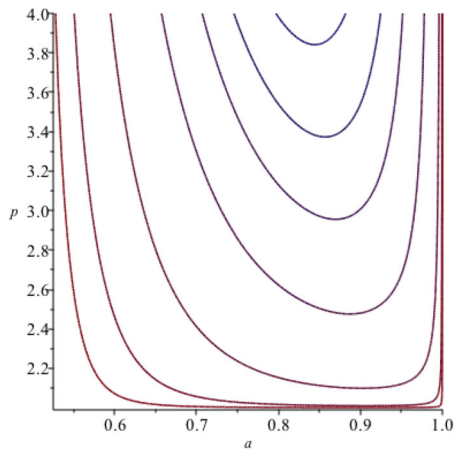


Fig. 2 Level plots of $f(\alpha, p)$ on $[1/2, 1] \times [1, 2]$

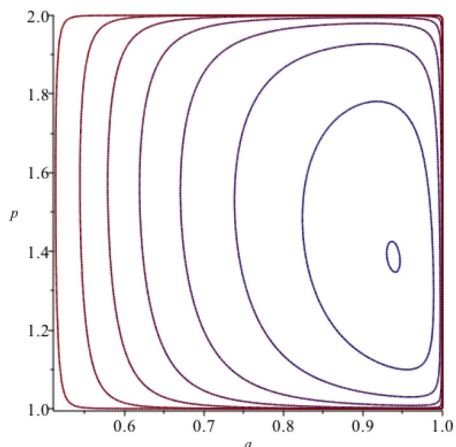


Fig. 3 Level plots of $f(\alpha, p)$ on $[0, 1/2] \times [0, 1]$

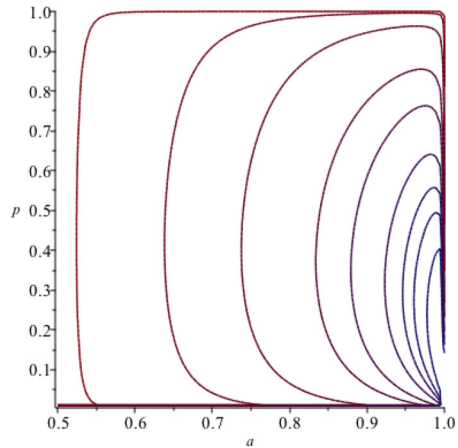


Figure 3 is a contour plot of f in the domain $[0, 1/2] \times [0, 1]$. The contours in Fig. 3 range from 1.0000001 to 1.06. Higher values are to the right. For p in this range, the maximum is not so large – about 1.06 – but the landscape gets very “steep” near $\alpha = 1$ and $p = 0$. The proof of the inequality is especially delicate in this case.

For $p < 0$, there is equality only at $\alpha = 1/2$, and the inequality is not so uniformly close to an identity. The contour plot is less informative, and hence is not recorded here. This is the case in which the inequality is easiest to prove.

It is possible to give a simple direct proof of inequality (1.6) for certain integer values of p , as we discuss in Sect. 5. We also give a simple proof that for $p > 2$ and for $p < 0$, validity of the inequality at p implies validity of the inequality at $2p$, and we briefly discuss an application of this to the problem in which functions are replaced by operators and integrals are replaced by traces.

2 Part A: Reduction from Two Functions to One

Our first observation is that in proving the inequality in Theorem 1.1, we may always assume that f and g are nonnegative. In fact, the right side of (1.4) only depends on $|f|$ and $|g|$, and the left side does not decrease for $p > 0$ and does not increase for $p < 0$ if f and g are replaced by $|f|$ and $|g|$. The latter follows since $|f + g| \leq |f| + |g|$ implies $|f + g|^p \leq (|f| + |g|)^p$ for $p > 0$ and $|f + g|^p \geq (|f| + |g|)^p$ for $p < 0$.

While Theorem 1.1 involves two functions f and g one can use the arbitrariness of the measure to reduce the question to a *single* function defined on a *probability space* (that is, $\int 1 = 1$). We have already observed that it suffices to prove the inequality in the case where f and g are both nonnegative. For nonnegative functions f and g , set

$$\alpha = f/(f + g), \quad 1 - \alpha = g/(f + g)$$

on the set where $f + g > 0$. Replacing the underlying measure dx by the new measure $(f + g)^p dx / \|f + g\|_p^p$ we see that it suffices to prove the following inequality for

$p \in (0, 1] \cup [2, \infty)$, and also to prove the reverse inequalities for $p \in (-\infty, 0) \cup [1, 2]$:

$$1 \leq \left(1 + \frac{2^{2/p} \|\alpha(1 - \alpha)\|_{p/2}}{(\|\alpha\|_p^p + \|1 - \alpha\|_p^p)^{2/p}} \right)^{p-1} (\|\alpha\|_p^p + \|1 - \alpha\|_p^p) \tag{2.1}$$

for a single function $0 \leq \alpha \leq 1$ on a probability space, i.e., $\int 1 = 1$.

3 Part B: Reduction to a Constant Function

In this section we prove the following.

Proposition 3.1 *If $p \in (0, 1] \cup [2, \infty)$, then inequality (2.1) is true for all functions α (which is equivalent to (1.4) for all f, g) if and only if it is true for all constant functions, that is, for all numbers $\alpha \in [0, 1]$,*

$$1 \leq \left(1 + \left(\frac{2\alpha^{p/2}(1 - \alpha)^{p/2}}{\alpha^p + (1 - \alpha)^p} \right)^{2/p} \right)^{p-1} (\alpha^p + (1 - \alpha)^p). \tag{3.1}$$

If $p \in (-\infty, 0) \cup [1, 2]$, then the reverse of inequality (2.1) is true for all functions α (which is equivalent to the reverse of (1.4) for all f, g) if and only if it is true for all constant functions, that is, for all numbers $\alpha \in [0, 1]$, the reverse of (3.1) holds.

Moreover, for $p \in \mathbb{R} \setminus \{0, 1, 2\}$ there is equality in (2.1) if and only if $\max\{\alpha(x), 1 - \alpha(x)\}$ is constant almost everywhere and for this constant equality holds in (3.1).

To prove this proposition we need a definition and a lemma.

Definition 3.2 Fix $p \in \mathbb{R} \setminus \{0\}$ and for $0 \leq a \leq 1$, let

$$h(a) := a^{p/2}(1 - a)^{p/2} \quad \text{and} \quad b(a) := a^p + (1 - a)^p.$$

Clearly, b determines the unordered pair a and $1 - a$ and, therefore, b determines h . Thus, we can consider the function

$$b \mapsto H(b) := h(a^{-1}(b))$$

(in which the dependence on p is suppressed in the notation).

Lemma 3.3 (convex/concave H) *The function $b \mapsto H(b)$ is strictly convex when $p \in (2, \infty)$ and strictly concave when $p \in (-\infty, 2) \setminus \{0, 1\}$.*

Proof To prove this lemma we use the chain rule to compute the second derivative of H . As a first step we define a useful reparametrization as follows: $e^{2x} := a/(1 - a)$. A quick computation shows that $h = (2 \cosh x)^{-p}$ and $b = 2 \cosh(px)(2 \cosh x)^{-p}$. Thus, $h = b/(2 \cosh(px))$. By symmetry, we can restrict our attention to the half-line $x \geq 0$.

We now compute the first two derivatives:

$$db/dx = 2^{1-p} p \frac{\sinh((p-1)x)}{(\cosh x)^{p+1}} \tag{3.2}$$

$$dh/dx = -p \frac{\tanh x}{(2 \cosh x)^p} \tag{3.3}$$

$$(dH/db)(x) = \frac{dh/dx}{db/dx} = -\frac{\sinh x}{2 \sinh((p-1)x)} \tag{3.4}$$

$$(d/dx)(dH/db)(x) = \cosh(x) \frac{(p-1) \tanh x - \tanh((p-1)x)}{2 \sinh((p-1)x) \tanh((p-1)x)} \tag{3.5}$$

$$(d^2 H/db^2)(x) = \frac{(d/dx)(dH/db)(x)}{db/dx} \tag{3.6}$$

Our goal is to show that (3.6) has the correct sign (depending on p) for all $x \geq 0$.

Clearly, the quantity (3.2) is nonpositive for $p \in (0, 1]$ and nonnegative elsewhere. We claim that the quantity (3.5) is nonpositive for $p \in (-\infty, 0] \cup [1, 2]$ and nonnegative elsewhere. In fact, the denominator is always positive. For the numerator we write $t = p - 1$ and use the fact that for all $x > 0$, $t \tanh(x) - \tanh(tx) > 0$ for $t > 1$ and for $-1 < t < 0$, while the inequality reverses, and is strict for other values of t except $t = 0$ and $t = \pm 1$.

To see this, fix $x > 0$, and define $f(t) := t \tanh(x) - \tanh(tx)$. Evidently $f(t) = 0$ for $t = -1, 0, 1$. Then since $f''(t) = 2x^2 \sinh(tx) / \cosh^3(tx)$, $f''(t) > 0$ for $t > 0$, and $f''(t) < 0$ for $t < 0$. It follows that $f(t) > 0$ for $-1 < t < 0$ and $t > 1$, while $f(t) < 0$ for $0 < t < 1$ and $t < -1$.

According to (3.6) the products of the signs of (3.2) and (3.5) yield the strict convexity/concavity properties of $H(b)$ shown in rows 2 to 4 of the table (Fig. 4). □

	$p < 0$	$0 < p < 1$	$1 < p < 2$	$p > 2$
$\frac{db}{dx}$	≥ 0	≤ 0	≥ 0	≥ 0
$\frac{d}{dx} \left(\frac{dH}{db} \right)$	≤ 0	≥ 0	≤ 0	≥ 0
$H(b)$	concave	concave	concave	convex
$p(p-1)$	≥ 0	≤ 0	≥ 0	≥ 0
Direction	\geq	\leq	\geq	\leq

Fig. 4 Table of signs determining the direction of the main inequality (1.4)

Proof of Proposition 3.1 We consider the ratio $\|\alpha(1 - \alpha)\|_{p/2}^{p/2} / (\|\alpha\|_p^p + \|1 - \alpha\|_p^p)$ in (2.1). The denominator is

$$B := \int (\alpha^p + (1 - \alpha)^p) = \int b(\alpha(x)) ,$$

and the numerator is the integral $\int H(b(\alpha(x)))$. By Jensen’s inequality (recalling that the underlying measure is a probability measure) and the convexity/concavity of H in Lemma 3.3, this latter integral is bounded from below by $H(B)$ in the convex case and from above in the concave case. That is,

$$\frac{\|\alpha(1 - \alpha)\|_{p/2}^{p/2}}{\|\alpha\|_p^p + \|1 - \alpha\|_p^p} = \frac{\int H(b(\alpha(x)))}{B} \geq \frac{H(B)}{B} \tag{3.7}$$

for $p \geq 2$, while the reverse is true for $p \leq 2$.

Moreover, when $p \notin \{0, 1, 2\}$, by the strict convexity/concavity of H , the inequality in (3.7) is strict unless $b(\alpha(x))$ is almost everywhere equal to a constant. It is easy to see that $b(\alpha(x))$ is almost everywhere equal to a constant if and only if $\max\{\alpha(x), 1 - \alpha(x)\}$ is almost everywhere equal to a constant.

Then, taking into account the signs of $2/p$ and $p - 1$ in the various ranges,

$$\left(1 + \left(\frac{2\|\alpha(1 - \alpha)\|_{p/2}^{p/2}}{\|\alpha\|_p^p + \|1 - \alpha\|_p^p} \right)^{2/p} \right)^{p-1} \geq \left(1 + \left(\frac{2H(B)}{B} \right)^{2/p} \right)^{p-1}$$

for $p \in (0, 1] \cup [2, \infty)$, with the reverse in equality for $p \in (-\infty, 0) \cup [1, 2]$. The last two rows in Fig. 4 summarize the interaction of the convexity/concavity properties of $H(b)$ and the signs of the exponents $p/2$ and $p - 1$ in the direction of the inequality in (3.1) for the different ranges of p .

To complete the proof of the theorem, we note that the range of the function $b(\alpha(x))$ lies in the interval $[2^{1-p}, 1]$ if $p > 1$ and in the interval $[1, 2^{1-p}]$ if $p < 1$. Therefore, its average value B lies in this same interval. Consequently, there is a number $\alpha \in [0, 1]$ such that $B = \alpha^p + (1 - \alpha)^p$. (Note that it is not claimed that this number α is related in any particular way to the function $\alpha(x)$. However, if $b(\alpha(x))$ is constant, then, as we have already mentioned, $\max\{\alpha(x), 1 - \alpha(x)\}$ is constant, and the value of this constant coincides either with the number α or the number $1 - \alpha$.) Therefore, if inequality (3.1) or its reverse holds for all numbers α , inequality (2.1) or its reverse holds for all functions $\alpha(x)$. Taking into account the cases of equality discussed above, this yields the result as stated. □

The proof of Proposition 3.1 is based on Jensen’s inequality, but the reduction to constants is *not* obtained by applying Jensen’s inequality to show that α must be constant in cases of equality. In fact, this is false, since there is equality in (1.4) when f and g have disjoint support. In this case, α is the indicator function of the support of f , while $1 - \alpha$ is the indicator function of the support of g . In the proof of Proposition 3.1, Jensen’s inequality is applied to show that in cases of equality, $\alpha^p + (1 - \alpha)^p$ must

be constant, and this is true almost everywhere with respect to the relevant probability measure, when f and g have disjoint support.

4 Part C: Proof of Theorem 1.2

4.1 Proof of the Inequality

Our goal in this subsection is to prove the first part of Theorem 1.2, that is, the inequality

$$(\alpha^p + (1 - \alpha)^p) \left(1 + \left(\frac{2\alpha^{p/2}(1 - \alpha)^{p/2}}{\alpha^p + (1 - \alpha)^p} \right)^{2/p} \right)^{p-1} \geq 1 \quad \text{for all } \alpha \in [0, 1] \quad (4.1)$$

if $p \in (0, 1] \cup [2, \infty)$, and the reverse inequality if $p \in (-\infty, 0) \cup [1, 2]$ (where $\alpha \in (0, 1)$ if $p < 0$). We will also characterize the cases of equality stated in the third part of Theorem 1.2.

For $p > 0$, there is evidently equality in (4.1) for $\alpha \in \{0, 1/2, 1\}$, and for $p < 0$, there is equality for $\alpha = 1/2$. Moreover, the inequality is invariant under exchanging α and $1 - \alpha$. Thus, for the proof of (4.1) it suffices to consider $\alpha \in (1/2, 1)$ for $p > 0$, and $\alpha \in (0, 1/2)$ if $p < 0$. It is convenient to change variables

$$t := \left(\frac{1 - \alpha}{\alpha} \right)^p \in (0, 1) \quad \text{and} \quad c := 1/p.$$

Moreover, for fixed c we introduce the function

$$f(t) := -\frac{1}{c} \ln(1 + t^c) + \ln(1 + t) + \frac{1 - c}{c} \ln \left(1 + \left(\frac{4t}{(t + 1)^2} \right)^c \right).$$

By taking logarithms we see that the claimed inequality (4.1) is equivalent to

$$f(t) \geq 0 \quad \text{for } t \in (0, 1)$$

if $p \in (0, 1] \cup [2, \infty)$ (that is, $c \in (0, 1/2] \cup [1, \infty)$), and the reverse inequality in (4.1) is equivalent to the reverse inequality if $p \in (-\infty, 0) \cup [1, 2]$ (that is, $c \in (-\infty, 0) \cup [1/2, 1]$). We shall show that for $c > 0$ the derivative f' has a unique sign change in $(0, 1)$ and it changes sign from $+$ to $-$ if $c \in (0, 1/2) \cup (1, \infty)$ and from $-$ to $+$ if $c \in (1/2, 1)$. Moreover, for $c < 0$ we shall show that the derivative f' is positive on $(0, 1)$.

Since $f(0) = f(1) = 0$ for $c > 0$, this proves that $f \geq 0$ if $c \in (0, 1/2) \cup (1, \infty)$ and that $f \leq 0$ if $c \in (1/2, 1)$. Moreover, since $f(1) = 0$ for $c < 0$, this proves that $f \leq 0$ if $c < 0$. Moreover, this argument shows that $f(t) \neq 0$ for all $t \in (0, 1)$ and all $c \in \mathbb{R} \setminus \{0, 1/2, 1\}$. Thus, we have reduced the proof of the first and the third part of Theorem 1.2 to proving the above sign change properties of f' .

In order to discuss the sign changes of f' we compute

$$f'(t) = \frac{(1-c)(1-t)}{t(1+t)} \left(\frac{1}{\left(\frac{(1+t)^2}{4t}\right)^c + 1} - \frac{t^c - t}{(1-c)(t^c + 1)(1-t)} \right). \tag{4.2}$$

Clearly, it suffices to consider the sign changes of the second factor and therefore to consider the sign changes of

$$g(t) := \left(\frac{(1+t)^2}{4t}\right)^c - \left(\frac{(1-c)(t^c + 1)(1-t)}{t^c - t} - 1\right). \tag{4.3}$$

We shall show that for $c > 0$, g has a unique sign change in $(0, 1)$ and it changes sign from $-$ to $+$ if $c \in (0, 1/2)$ and from $+$ to $-$ if $c \in (1/2, \infty)$. Moreover, for $c < 0$ we shall show that g is negative on $(0, 1)$. Clearly, these properties of g imply the claimed properties of f' and therefore will conclude the proof.

We next observe that the second term in (4.3) is positive.

Lemma 4.1 *For any $c \in \mathbb{R} \setminus \{1\}$ and $t \in (0, 1)$,*

$$\frac{(1-c)(t^c + 1)(1-t)}{t^c - t} > 1.$$

Proof First, consider the case $c \in [0, 1)$. Then concavity of the map $t \mapsto t^c$ implies $1 - c + ct - t^c \geq 0$, therefore $\frac{(1-c)(1-t)}{t^c - t} \geq 1$, and the claim follows from $t^c + 1 > 1$.

Next, for $c > 1$ the argument is similar using convexity of the map $t \mapsto t^c$.

Finally, for $c < 0$ convexity of $t \mapsto t^{1-c}$ implies that

$$\frac{(1-c)(t^c + 1)(1-t)}{t^c - t} - 1 = \frac{(1-c)(1+t^{-c})(1-t)}{1-t^{1-c}} - 1 > \frac{(1-c)(1-t)}{1-t^{1-c}} - 1 \geq 0.$$

This concludes the proof of the lemma. □

Because of Lemma 4.1, we can define

$$h(t) := c \ln \left(\frac{(1+t)^2}{4t}\right) - \ln \left(\frac{(1-c)(t^c + 1)(1-t)}{t^c - t} - 1\right). \tag{4.4}$$

We shall show that for $c > 0$, h has a unique sign change in $(0, 1)$ and it changes sign from $-$ to $+$ if $c \in (0, 1/2)$ and from $+$ to $-$ if $c \in (1/2, \infty)$. Moreover, for $c < 0$ we shall show that h is negative on $(0, 1)$. Clearly, these properties of h imply the claimed properties of g and therefore will conclude the proof.

We will prove this by investigating sign changes of h' . Namely, we shall show that for $c > 0$, h' has a unique sign change in $(0, 1)$ and it changes sign from $+$ to $-$ if $c \in (0, 1/2)$ and from $-$ to $+$ if $c \in (1/2, \infty)$. Moreover, for $c < 0$ we shall show that h' is positive on $(0, 1)$.

Let us show that this implies the claimed properties of h . Indeed, an elementary limiting argument shows that

$$h(0) = \begin{cases} -\infty & \text{if } c < 0, \\ -2c \ln 2 - \ln(1 - c) & \text{if } c \in (0, 1), \\ +\infty & \text{if } c > 1. \end{cases}$$

and

$$h(1) = 0 \quad \text{for all } c.$$

The function $-2c \ln 2 - \ln(1 - c)$ is convex on $(0, 1)$ and vanishes at $c = 0$ and $c = 1/2$. From this we conclude that

$$h(0) < 0 \text{ if } c < 1/2, \quad h(0) = 0 \text{ if } c = 1/2, \quad h(0) > 0 \text{ if } c > 1/2.$$

Because of this behavior of $h(0)$ and $h(1)$, the claimed properties of h' imply the claimed properties of h .

Therefore in order to complete the proof of Theorem 1.2 we need to discuss the sign changes of h' . We compute

$$h'(t) = \frac{v(t)}{(1+t)(t^c - t)^2 \left(\frac{(1-c)(t^c+1)(1-t)}{t^c-t} - 1 \right)}$$

with

$$v(t) := t(2c^2 - 1) - t^2c^2 + 2c(1 - 2c)(t^c - t^{c+1}) + t^{2c}(1 - 2c^2) + (t^{1+2c} - 1)(1 - c)^2 + t^{2c-1}c^2.$$

We shall show that for $c > 0$, v has a unique sign change in $(0, 1)$ and it changes sign from $+$ to $-$ if $c \in (0, 1/2)$ and from $-$ to $+$ if $c \in (1/2, \infty)$. Moreover, for $c < 0$ we shall show that v is positive on $(0, 1)$.

Since, by Lemma 4.1 the denominator in the above expression for h' is positive, these properties of v clearly imply those of h' and therefore complete the proof of the theorem.

In order to prove the claimed properties of v we shall study the sign changes of v'' . We shall show that for $c > 0$, v'' has a unique sign change in $(0, 1)$ and it changes sign from $+$ to $-$ if $c \in (0, 1/2)$ and from $-$ to $+$ if $c \in (1/2, \infty)$. Moreover, for $c < 0$ we shall show that v'' is positive.

Let us now argue that these properties of v'' indeed imply the claimed properties of v . We compute

$$v'(t) = 2c^2 - 1 - 2c^2t + 2c(1 - 2c)(ct^{c-1} - (c+1)t^c) + 2c(1 - 2c^2)t^{2c-1} + (1 - c)^2(1 + 2c)t^{2c} + c^2(2c - 1)t^{2c-2},$$

$$\begin{aligned}
 v''(t) = & 2c \cdot [-c + c(1 - 2c)(c - 1)t^{c-2} - c(1 - 2c)(c + 1)t^{c-1} \\
 & + (2c - 1)(1 - 2c^2)t^{2c-2} \\
 & + (1 - c)^2(1 + 2c)t^{2c-1} + c(2c - 1)(c - 1)t^{2c-3}], \tag{4.5}
 \end{aligned}$$

and finally

$$v'''(t) = 2c(1 - 2c)(c - 1)t^{c-3}w(t) \tag{4.6}$$

with

$$w(t) := c(c - 2) - (c + 1)ct - 2t^c(1 - 2c^2) + (1 - c)(1 + 2c)t^{c+1} - c(2c - 3)t^{c-1}.$$

From these formulas we easily infer that

$$v(1) = v'(1) = v''(1) = 0, \quad v'''(1) = 2c(1 - 2c)(c - 1)^2.$$

In particular, $v'''(1) > 0$ if $c \in (0, 1/2)$ and $v'''(1) < 0$ if $c \in (-\infty, 0) \cup (1/2, 1) \cup (1, \infty)$. This means that v is convex near $t = 1$ if $c \in (-\infty, 0) \cup (1/2, \infty)$ and concave near $t = 1$ if $c \in (0, 1/2)$.

Let us discuss the behavior near $t = 0$. If $c < 1/2$, then $v(t)$ behaves like $t^{2c-1}c^2$, so $v(0) = +\infty$, and $v''(0) > 0$. If $c > 1/2$, then $v(0) = -(1 - c)^2$ and $v''(0) < 0$.

This behavior of v near 0 and 1, together with the claimed sign change properties of v'' , implies the claimed sign change properties of v and will therefore complete the proof of Theorem 1.2. This is because, for example, if v is convex near $t = 1$ with $v(1) = v'(1) = 0$, and v has a single inflection point $t_0 \in (0, 1)$, then v is positive on $[t_0, 0)$, and v is concave on $(0, t_0)$.

Thus, we are left with studying the sign changes of v'' . In order to do so, we need to distinguish several cases. For $c < 1$ we will argue via the sign changes of v''' , while for $c > 1$ we will argue directly.

Case $c \in (0, 1)$. We want to show that v'' changes sign from + to - if $c \in (0, 1/2)$ and from - to + if $c \in (1/2, 1)$.

Since $v''(0) > 0$ if $c \in (0, 1/2)$, $v''(0) < 0$ if $c \in (1/2, 1)$, $v''(1) = 0$, and $v'''(1) > 0$, it suffices to show that v''' changes sign only once on $(0, 1)$. Because of (4.6) this is the same as showing that w changes sign only once on $(0, 1)$. Notice that $w(0) = +\infty$, and $w(1) = c - 1 < 0$. Moreover,

$$w''(t) = c(1 - c)t^{c-3}p(t)$$

with

$$p(t) := t^2(c + 1)(1 + 2c) + 2t(1 - 2c^2) + 2c^2 - 7c + 6.$$

The quadratic polynomial p is positive. Indeed, when $c \in (0, 1/2)$ this follows from the fact that all its coefficients are positive. When $c \in (1/2, 1)$ we observe that the parabola

p is minimized on \mathbb{R} at $t = \frac{2c^2-1}{(c+1)(1+2c)}$, and its minimal value is $\frac{(5-3c^2)+c(11-8c^2)}{(c+1)(1+2c)}$, which is positive for $c \in (1/2, 1)$.

The fact that p is positive means that w is convex. Since $w(0) = +\infty$ and $w(1) < 0$, we conclude that w has only one root.

Case $c \in (-\infty, 0)$. We want to show that v'' is positive.

Since $v''(1) = 0$, it suffices to show that v''' is negative which, by (4.6), is the same as showing that w is negative. Clearly, $w(0) = -\infty$, $w''(0) < 0$, and $w(1) = c - 1 < 0$, and $w'(1) = (3c - 1)(c - 1) > 0$, so it suffices to show that $w'' < 0$ on $(0, 1)$. For this it suffices to show that $p > 0$ on $(0, 1)$. We have $p(0) > 0$, and $p(1) = 9 - 4c > 0$. Thus if $(1 + c)(1 + 2c) \leq 0$ we have proved the claim. Consider the case when $(1 + c)(1 + 2c) > 0$. The vertex of the parabola is $t_0 = \frac{2c^2-1}{(c+1)(1+2c)}$. If $c < -1$ then clearly $\frac{2c^2-1}{(c+1)(1+2c)} > 1$. If $c \in (-1/2, 0)$, then clearly $\frac{2c^2-1}{(c+1)(1+2c)} < 0$.

Case $c \in (1, \infty)$. We want to show that v'' changes sign from $-$ to $+$.

We begin with the case $c \in (1, 2)$. We write (4.5) as $v''(t) = 2cq(t)$ with

$$q(t) := -c + c(1 - 2c)(c - 1)t^{c-2} - c(1 - 2c)(c + 1)t^{c-1} + (2c - 1)(1 - 2c^2)t^{2c-2} + (1 - c)^2(1 + 2c)t^{2c-1} + c(2c - 1)(c - 1)t^{2c-3}.$$

Clearly $q(0) = -\infty$ and $q(1) = 0$. It is enough to show that q' changes sign from $+$ to $-$. We have

$$q'(t) = t^{2c-4}(2c - 1)(c - 1)m(t)$$

with

$$m(t) := c(2 - c)t^{1-c} + c(c+1)t^{2-c} + 2(1 - 2c^2)t + (c - 1)(1+2c)t^2 + c(2c - 3).$$

We shall show that $m(t)$ changes sign only once from $+$ to $-$. Clearly $m(0) = +\infty$ and $m''(0) > 0$. Next, $m(1) = 1 - c < 0$, and $m''(1) = (c - 1)(c^2 + 2c + 2) > 0$. Thus it suffices to show $m'' > 0$ on $(0, 1)$. Since $m''(0) > 0$, $m''(1) > 0$, then $m'' > 0$ will follow from m''' having the constant sign. We have

$$m'''(t) = t^{-c-2}c^2(c - 1)(c - 2)(c + 1)(1 - t) < 0.$$

This finishes the case $c \in (1, 2)$.

If $c = 2$, then $q(t) = (t - 1)(5t^2 - 16t + 8)$, and we see that it changes sign only once.

In what follows we assume $c > 2$. Let us rewrite (4.5) as $v''(t) = 2ct^{2c-3}u(t)$ with

$$u(t) := -ct^{3-2c} + c(1 - 2c)(c - 1)t^{1-c} - c(1 - 2c)(c + 1)t^{2-c} + (2c - 1)(1 - 2c^2)t + (1 - c)^2(1 + 2c)t^2 + c(2c - 1)(c - 1).$$

We need to show that u changes sign only once. We have $u(0) = -\infty$, and $u''(0) < 0$. At the point $t = 1$, we have $u(1) = 0$, $u'(1) = -(2c - 1)(c - 1)^2 < 0$, $u''(1) =$

$-2(2c - 1)(c - 1)^2 < 0$. It suffices to show that $u'' < 0$ on $(0, 1)$. Since $u''(0) < 0$, $u''(1) < 0$, the latter claim will follow from showing that u''' has a constant sign. We have

$$u'''(t) = t^{-2-c}c(2c - 1)(c - 1)b(t)$$

with

$$b(t) := c^3 - c - tc(c + 1)(c - 2) + 2t^{2-c}(2c - 3).$$

The factor b has the property that $b(0) = +\infty$, $b(1) = (c + 6)(c - 1) > 0$. On the other hand,

$$b'(t) = -c(c + 1)(c - 2)t^{1-c} \left(t^{c-1} + \frac{2(2c - 3)}{c(c + 1)} \right)$$

is negative, so b is positive.

This concludes the proof of the inequality of Theorem 1.2.

4.2 Sharpness of the Exponent $2/p$

Our goal in this subsection is to prove the optimality statement in Theorem 1.2 corresponding the exponent q in inequality (1.7).

We begin by discussing the case $q = 2/p$ and present an alternative way of writing inequality (1.6). Introduce a new variable $s \in (0, 1)$ through

$$\alpha = \frac{1 + \sqrt{s}}{2}$$

Rewriting (1.6), and taking the $\frac{1}{p-1}$ root of both sides, we may rearrange terms to obtain

$$2 \leq \eta^{\frac{1}{p-1}}(s) \left(1 + \frac{1-s}{\eta^{\frac{2}{p}}(s)} \right) = \eta^{\frac{1}{p-1}}(s) + (1-s)\eta^{\frac{2-p}{p(p-1)}}(s) \tag{4.7}$$

for $0 \leq s \leq 1$, where

$$\eta(s) := \frac{(1 + \sqrt{s})^p + (1 - \sqrt{s})^p}{2}. \tag{4.8}$$

Taking the $\frac{1}{p-1}$ root eliminates the change of direction in the inequality at $p = 1$, and it now takes on a nontrivial form at $p = 1$: Define

$$f_p(s) := \eta^{\frac{1}{p-1}}(s) + (1-s)\eta^{\frac{2-p}{p(p-1)}}(s) - 2, \tag{4.9}$$

for $p \neq 1$, and one easily computes the limit at $p = 1$:

$$f_1(s) := (2 - s)(1 - \sqrt{s})^{\frac{1-\sqrt{s}}{2}} (1 + \sqrt{s})^{\frac{1+\sqrt{s}}{2}} - 2 .$$

The first assertion in Theorem 1.2 is equivalent to the assertion that for all $s \in (0, 1)$,

$$\boxed{f_p(s) \geq 0 \text{ for } p \in (-\infty, 0) \cup [2, \infty) \quad \text{and} \quad f_p(s) \leq 0 \text{ for } p \in (0, 2] .}$$

(4.10)

In this form, the inequality is easy to check for some values of p . For example, for $p = -1$, $\eta(s) = \frac{1}{1-s}$ and $f_{-1}(s) = (1 - s)^{1/2} + (1 - s)^{-1/2} - 2$, which is clearly positive. One can give simple proofs of (4.10) for other integer values of p , e.g., $p = 3$ and $p = 4$ along these lines.

We now turn our attention to inequality (1.7) with a general power q . If one makes the transformations described above and sets $r = qp/2$, one is led to the function

$$g_{r,p}(s) := \eta^{\frac{1}{p-1}}(s) \left(1 + \left(\frac{1-s}{\eta^{\frac{2}{p}}(s)} \right)^r \right) - 2 \tag{4.11}$$

instead of $f_p(s)$. Inequality (1.7) for $p \in (0, 1] \cup [2, \infty)$ and its reverse for $p \in (-\infty, 0) \cup [1, 2]$ are equivalent to the assertion that for all $s \in (0, 1)$

$$g_{r,p}(s) \geq 0 \text{ for } p \in (-\infty, 0) \cup [2, \infty) \quad \text{and} \quad g_{r,p}(s) \leq 0 \text{ for } p \in (0, 2]. \tag{4.12}$$

A motivation for this reparametrization is that for fixed p , the function on the right hand side of (1.6) is equal to 1 up to order $\mathcal{O}((\alpha - 1/2)^4)$ at $\alpha = 1/2$. In the variable s , the leading term in Taylor expansion in s will be second order, and we prove the sharpness of the power $q = 2/p$ by an expansion at this point.

Proof of the second paragraph of Theorem 1.2 For fixed $r > 0$, define the function $g_{r,p}(s)$ by (4.11). By the arithmetic-geometric mean inequality, $(1 - s)^{p/2} \leq \eta(s)$ for all p , and hence $(1 - s)/\eta^{2/p}(s) < 1$ for $p > 0$, while $(1 - s)/\eta^{2/p}(s) > 1$ for $p < 0$. Therefore, for fixed s and p , $g_{r,p}(s)$ decreases as r increases for $p > 0$, and does the opposite for $p < 0$.

A Taylor expansion shows that, as $s \rightarrow 0$,

$$g_{r,p}(s) = p(1 - r)s + o(s) .$$

It follows that $g_{r,p}(s) \geq 0$ on $[0, 1]$ is false (near $s = 0$) for $p \geq 2$ and $r > 1$, and for $p < 0$ and $r < 1$. Likewise, it follows that $g_{r,p}(s) \leq 0$ on $[0, 1]$ is false for $p \in (0, 2]$ and $r < 1$. Since the exponent q in (1.7) corresponds to $r(2/p)$, this, together with the remarks leading to (4.12), justifies the statements referring to q in Theorem 1.2. \square

4.3 Proof of Theorem 1.1

The proof of Theorem 1.1 is now essentially complete. For the sake of clarity, let us summarize the whole argument.

The first part of Theorem 1.2, which was proved in Sect. 4.1, establishes the validity of inequality (1.6) for $p \in (0, 1] \cup [2, \infty)$ and its reverse for $p \in (-\infty, 0) \cup [1, 2]$. According to Proposition 3.1, this yields inequality (2.1) for $p \in (0, 1] \cup [2, \infty)$ and its reverse for $p \in (-\infty, 0) \cup [1, 2]$. Finally, according to the discussion in Sect. 2, this proves inequality (1.4) for $p \in (0, 1] \cup [2, \infty)$ and its reverse for $p \in (-\infty, 0) \cup [1, 2]$.

Next, we turn to the second part of Theorem 1.1 referring to q . Inequality (1.4) with $\tilde{\Gamma}$ replaced by $\tilde{\Gamma}^q$ reduces for f and g which are positive multiples of each other to inequality (1.7), where $\alpha = f/(f + g)$. In the second part of Theorem 1.2, which was proved in Sect. 4.2, we have determined precisely under which conditions on q this inequality holds. This proves the second part of Theorem 1.1.

Finally, we discuss the third part of Theorem 1.1 concerning the cases of equality. The last part of Theorem 1.2, which was proved in Sect. 4.1, establishes that equality in inequality (1.6) holds if and only if $\alpha \in \{0, 1/2, 1\}$ for $p \in (0, \infty) \setminus \{1, 2\}$ and if and only if $\alpha = 1/2$ if $p \in (-\infty, 0)$. According to the second part of Proposition 3.1, this means that equality holds in (2.1) if and only if the function $\max\{\alpha(x), 1 - \alpha(x)\}$ is almost everywhere equal to a constant, which has values in $\{0, 1/2, 1\}$ for $p \in (0, \infty) \setminus \{1, 2\}$ and has the value $\alpha = 1/2$ if $p \in (-\infty, 0)$. Clearly, $\max\{\alpha(x), 1 - \alpha(x)\} \equiv 1/2$ if and only if $\alpha(x) \equiv 1/2$, and $\max\{\alpha(x), 1 - \alpha(x)\} \equiv 1$ if and only if α is the characteristic function of a set. Now given $f, g \geq 0$ set $\alpha = f/(f + g)$ on the set where $f + g > 0$. Then $\alpha(x) \equiv 1/2$ almost everywhere if and only if $f = g$ almost everywhere, and α is the characteristic function of a set if and only if f and g have disjoint supports, up to a null set. This proves the statement of equality in Theorem 1.1 in case of nonnegative functions f and g .

We now show that this implies the statement of equality for general functions f and g . Indeed, by the argument at the beginning of Sect. 2, equality in (1.4) implies that $|f + g| = |f| + |g|$ almost everywhere and that equality in (1.4) holds for $|f|$ and $|g|$. As we have just shown, the latter fact implies that either $|f| = |g|$ almost everywhere or, if $p \in (0, \infty) \setminus \{1, 2\}$, $|f|$ and $|g|$ have disjoint supports, up to a null set. This, together with the former fact implies that either $f = g$ almost everywhere or, if $p \in (0, \infty) \setminus \{1, 2\}$, f and g have disjoint supports, up to a null set. This completes the proof of Theorem 1.1. □

5 Doubling Arguments and a Generalization to Schatten Norms

5.1 Doubling Arguments

We begin this section with a simple proof showing that if the inequality (1.4) is valid for some $p \geq 2$ or some $p < 0$, then it is also valid for 2. Since the inequality (1.4) holds as an identity for $p = 2$, and is simple to prove for $p = -1$ (see Sect. 4.2), this yields a simple proof of infinitely many cases of the inequality (1.4). The proof is not only simple and elegant; it applies to certain noncommutative generalizations of (1.4)

for which the reductions in parts A and B of the proof we have just presented are not applicable, as we discuss.

To introduce the doubling argument we present a direct proof of Theorem 1.1 for $p = 4$.

Proof Suppose $f, g \geq 0$. By homogeneity, we may suppose that $\|f\|_4^4 + \|g\|_4^4 = 2$. Define

$$X := fg, \quad Y := f^2 + g^2, \quad \alpha := \|X\|_2 \quad \text{and} \quad \beta := \|Y\|_2. \tag{5.1}$$

By the arithmetic-geometric mean inequality, $X \leq \frac{1}{2}Y$, and hence

$$\int X^2 d\mu \leq \frac{1}{4} \int Y^2 d\mu = \frac{1}{4} \int (f^4 + g^4 + 2f^2g^2) d\mu = \frac{1}{2} + \frac{1}{2} \int X^2 d\mu.$$

This yields $\alpha \leq 1$ and $\beta \leq 2$. Then $(f + g)^2 = Y + 2X$ and hence

$$\|f + g\|_4^2 = \|Y + 2X\|_2 \leq \|Y\|_2 + 2\|X\|_2 = \beta + 2\alpha. \tag{5.2}$$

It suffices to prove that $\beta + 2\alpha \leq 2^{1/2}(1 + \alpha)^{3/2}$. Note that $\beta^2 = \int (f^2 + g^2)^2 d\mu = 2 + 2\alpha^2$, and then since $\alpha \in [0, 1]$. Thus it suffices to show that

$$(1 + \alpha^2)^{1/2} \leq (1 + \alpha)^{3/2} - 2^{1/2}\alpha \quad \text{for all } 0 \leq \alpha \leq 1. \tag{5.3}$$

Squaring both sides, this is equivalent to $1 + \alpha^2 \leq (1 + \alpha)^3 + 2\alpha^2 - 2^{3/2}\alpha(1 + \alpha)^{3/2}$. This reduces to $2^{3/2}(1 + \alpha)^{3/2} \leq 3 + 4\alpha + \alpha^2$. Squaring both sides again, this reduces to $(\alpha^2 - 1)^2 \geq 0$, completing the proof. \square

What made this proof work is the fact that the inequality holds for $p = 2$ – as an identity, but that is unimportant. Then, using Minkowski’s inequality, as in (5.2), together with the numerical inequality (5.3) we arrive at the inequality for $p = 4$. This is a first instance of the general doubling proposition, to be proved next. The inequality (5.3) is a special case of the general inequality (5.4) proved below.

This strategy can be adapted to give direct proof of the inequality for other integer values of p ; e.g., $p = 3$. When p is an integer, and f and g are nonnegative, one has the binomial expansion of $(f + g)^p = f^p + g^p + \text{mixed terms}$. Under the assumption that $\int (f^p + g^p) = 2$, one is left with estimating the mixed terms, and one can use Hölder for this. When p is not an integer, there is no useful expression for $(f + g)^p - f^p - g^p$.

Proposition 5.1 *Suppose that for some $p \geq 2$, (1.4) is valid for all $f, g \geq 0$. Then (1.4) is valid with p replaced by $2p$ for all $f, g \geq 0$. Likewise, if for some $p < 0$ the reverse of (1.4) is valid for all $f, g > 0$, then the reverse of (1.4) is valid with p replaced by $2p$ for all $f, g > 0$.*

The proof of Proposition 5.1 relies on the following lemma.

Lemma 5.2 For $t \in \mathbb{R}$, define ψ_t on $[0, \infty)$ by

$$\psi_t(\alpha) = (1 + \alpha)^{1+t} - (1 + \alpha^2)^t - 2^t \alpha. \tag{5.4}$$

Then for $t \in [0, 1]$, $\psi_t(\alpha) \geq 0$ on $[0, \infty)$, while for $t > 1$, $\psi_t(\alpha) \leq 0$ on $[0, \infty)$.

Proof We write $\psi_t(\alpha) = (1 + \alpha)^t - (1 + \alpha^2)^t - (2^t - (1 + \alpha)^t)\alpha$. Therefore,

$$\frac{\psi_t(\alpha)}{\alpha(1 - \alpha)} = \frac{(1 + \alpha)^t - (1 + \alpha^2)^t}{\alpha(1 - \alpha)} - \frac{2^t - (1 + \alpha)^t}{1 - \alpha}.$$

Defining $a := 1 + \alpha^2$, $b := 1 + \alpha$ and $c := 2$, and defining $\varphi(\alpha) := x^t$, the right hand side is the same as

$$\frac{\varphi(b) - \varphi(a)}{b - a} - \frac{\varphi(c) - \varphi(b)}{c - b}.$$

For $\alpha \in [0, 1)$ we have $a < b < c$ and therefore this quantity is positive when φ is concave, and negative when φ is convex. For $\alpha \in (1, \infty)$ we have $a > b > c$ and therefore this quantity is negative when φ is concave, and positive when φ is convex. □

Proof of Proposition 5.1 Let $f, g \in L^{2p}$ with $\|f\|_{2p}^{2p} + \|g\|_{2p}^{2p} = 2$. Define $X := fg$ and $Y := f^2 + g^2$, and $\gamma := \|X\|_p$ and $\beta := \|Y\|_p$. By the triangle inequality we have

$$\|f + g\|_{2p}^2 = \|Y + 2X\|_p \begin{cases} \leq \|Y\|_p + 2\|X\|_p = \beta + 2\gamma & \text{if } p \geq 2, \\ \geq \|Y\|_p + 2\|X\|_p = \beta + 2\gamma & \text{if } p < 0. \end{cases}$$

(Note that the triangle inequality reverses for $p < 0$.) We now use the assumption that the inequality (1.4) is valid for p . Applying the inequality with exponent p to the functions f^2 and g^2 , which satisfy $\|f^2\|_p^p + \|g^2\|_p^p = \|f\|_{2p}^{2p} + \|g\|_{2p}^{2p} = 2$, we obtain for $p \geq 2$,

$$\beta^p = \|f^2 + g^2\|_p^p \leq 2 \left(1 + \|f^2 g^2\|_{p/2}\right)^{p-1} = 2 \left(1 + \gamma^2\right)^{p-1}$$

and similarly $\beta^p \geq 2 \left(1 + \gamma^2\right)^{p-1}$ for $p < 0$. To summarize, we have shown that

$$\|f + g\|_{2p}^2 \begin{cases} \leq 2^{1/p}(1 + \gamma^2)^{1-1/p} + 2\gamma & \text{if } p \geq 2, \\ \geq 2^{1/p}(1 + \gamma^2)^{1-1/p} + 2\gamma & \text{if } p < 0. \end{cases}$$

According to Lemma 5.2 (with $t = 1 - 1/p$ and $\alpha = \gamma$) this is bounded from above for $p \geq 2$ and from below for $p < 0$ by $2^{1/p}(1 + \gamma)^{2-1/p} = 2^{1/p}(1 + \|fg\|_p)^{2-1/p}$, which is the claimed inequality. □

5.2 A Generalization to Schatten Norms

For $p \in [1, \infty)$, an operator A on some Hilbert space belongs to the Schatten p -class \mathcal{S}_p in case $(A^*A)^{p/2}$ is trace class, and the Schatten p norm on \mathcal{S}_p is defined by $\|A\|_p = (\text{Tr}[(A^*A)^{p/2}])^{1/p}$. One possible noncommutative analog of (part of) Theorem 1.1 would assert that for positive $A, B \in \mathcal{S}_p, p > 2$.

$$\text{Tr}(A + B)^p \leq \left(1 + \left(\frac{\text{Tr}[B^{p/4} A^{p/2} B^{p/4}]}{\frac{1}{2}\|A\|_p^p + \frac{1}{2}\|B\|_p^p} \right)^{2/p} \right)^{p-1} \text{Tr}(A^p + B^p). \tag{5.5}$$

Note that for $p = 2$, (5.5) holds as an identity.

In this setting, it is not clear how to implement analogs of Parts A and B of our proof for functions. However, the direct proofs sketched at the beginning of this section do allow us to prove the validity of (5.5) for all $p = 2^k, k \in \mathbb{N}$.

Theorem 5.3 *If (5.5) is valid for some $p \geq 2$ and all positive $A, B \in \mathcal{S}_p$, then it is valid for $2p$ and all $A, B \in \mathcal{S}_{2p}$. In particular, since (5.5) holds as an identity for $p = 2$, it is valid for $p = 2^k$ for all $k \in \mathbb{N}$.*

Proof Let A and B be positive operators in \mathcal{S}_{2p} , and assume that $\|A\|_{2p}^{2p} + \|B\|_{2p}^{2p} = 2$, which, by homogeneity, entails no loss of generality. Define

$$X := \frac{1}{2}(AB + BA) \quad \text{and} \quad Y = A^2 + B^2.$$

Note that

$$\|X\|_p \leq \frac{1}{2}(\|AB\|_p + \|BA\|_p).$$

By definition, the Lieb–Thirring inequality [10], and cyclicity of the trace,

$$\|AB\|_p^p = \text{Tr}[(BA^2B)^{p/2}] \leq \text{Tr}[B^{p/2}A^pB^{p/2}] = \text{Tr}[A^{p/2}B^pA^{p/2}].$$

Define

$$\beta := \|Y\|_p \quad \text{and} \quad \gamma := (\text{Tr}[B^{p/2}A^pB^{p/2}])^{1/p}.$$

Therefore, $\|A + B\|_{2p}^2 = \|Y + 2X\|_p \leq \|Y\|_p + 2\|X\|_p \leq \beta + 2\gamma$. Since $\|A^2\|_p^p + \|B^2\|_p^p = 2$, we can apply (5.5) to deduce that

$$\beta^p = \|A^2 + B^2\|_p^p \leq 2 \left(1 + (\text{Tr}[B^{2p/4} A^p B^{2p/4}])^{2/p} \right)^{p-1} = 2 \left(1 + \gamma^2 \right)^{p-1}.$$

Altogether

$$\|A + B\|_{2p}^2 \leq 2^{1/p}(1 + \gamma^2)^{1-1/p} + 2\gamma$$

and, by Lemma 5.2, the right side is bounded above by $2^{1/p}(1+\gamma)^{2-1/p}$, which proves the inequality, \square

Note Since this paper was submitted, two further papers by the authors [5,8] have appeared which, in particular, explore extensions of the inequalities discussed here to more than three functions. The results are less complete than for two functions.

Acknowledgements Open Access funding provided by Projekt DEAL. We thank Anthony Carbery for useful correspondence.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Clarkson, J.A.: Uniformly convex spaces. *Trans. Am. Math. Soc.* **40**, 396–414 (1936)
2. Ball, K., Carlen, E.A., Lieb, E.H.: Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.* **115**(3), 463–482 (1994)
3. Carbery, A.: Almost orthogonality in the Schatten-von Neumann classes. *J. Oper. Theory* **62**(1), 151–158 (2009)
4. Carlen, E.A., Frank, R.L., Lieb, E.H.: Stability estimates for the lowest eigenvalue of a Schrödinger operator. *Geom. Funct. Anal.* **24**(1), 63–84 (2014)
5. Carlen, E.A., Frank, R.L., Lieb, E.H.: Inequalities that sharpen the triangle inequality for sums of N functions in L^p . *Ark. Math.* **58**(1), 57–69 (2020)
6. Hanner, O.: On the uniform convexity of L^p and ℓ^p . *Ark. Math.* **3**, 239–244 (1956)
7. Hardy, G., Littlewood, J.E., Polya, G.: *Inequalities*. Cambridge Univ Press, Cambridge (1934)
8. Mooney, C., Ivanisvili, P.: Sharpening the triangle inequality: envelopes between L^2 and L^p spaces. *Anal. PDE (to appear)*
9. Lieb, E.H., Loss, M.: *Analysis*, 2nd edn. American Mathematical Society, Providence, RI (2014)
10. Lieb, E.H., Thirring, W.: Inequalities for the moments of the eigenvalues of the Schrödinger hamiltonian and their relation to Sobolev inequalities. In: Lieb, E.H., Simon, B., Wightman, A. (eds.) *Studies in Mathematical Physics*, pp. 269–303. Princeton University Press (1976)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.