On the Optimality of Dictator functions and Isoperimetric Inequalities on Boolean Hypercubes

Zijie Chen and Chandra Nair
Department of Information Engineering
The Chinese University of Hong Kong
Sha Tin, NT, Hong Kong
{zijie,chandra}@ie.cuhk.edu.hk

Abstract

In this paper, we present a family of conjectures on the optimality of the dictator function among all Boolean functions for a new family of Φ-entropies. Our main theorem shows that there is an ordering of these conjectures, in that if the conjecture is established for a value α, then it holds for any $\beta : \beta \geq \alpha$. For the parameter range $\frac{1}{2} \leq \alpha, \beta < 1$, this family of conjectures is stronger than the mutual information conjecture, originally proposed in a paper by Courtade and Kumar. By considering a limiting value of the noise parameter ρ, we show how these conjectures relate to isoperimetric inequalities on the Boolean Hypercube of a flavor first considered by Talagrand and later by Kahn and Park. Finally, we obtain bounds to our conjecture using ideas from the proofs of isoperimetric inequalities. [1]

I. Introduction

A. Background

Denote $\mathbb{H}_n = \{-1,1\}^n$ to be the Boolean Hypercube. The following conjecture by Kumar, appearing in [1], has received much attention from the information theory community.

Conjecture 1 (Kumar, [1]). *Let $X \sim \text{Unif}(\mathbb{H}_n)$. Let Y be obtained by flipping each bit of X independently with probability $\frac{1-\rho}{2}$. Let $f(X)$ be a Boolean function. Then*

$$I(f(X);Y) \leq 1 - H\left(\frac{1-\rho}{2}\right). \tag{1}$$

Here $H(\frac{1-\rho}{2}) = -(\frac{1-\rho}{2})\log_2(\frac{1-\rho}{2}) - (\frac{1+\rho}{2})\log_2(\frac{1+\rho}{2})$ is the binary entropy function.

A weaker version of the conjecture

$$I(f(X);g(Y)) \leq 1 - H\left(\frac{1-\rho}{2}\right), \tag{2}$$

where $f(X)$ and $g(Y)$ are Boolean functions was established in [2]. Earlier, Andrej Bogdanov and the second author (to whom Kumar had mentioned his conjecture in 2011) had shown the above weaker form of the conjecture when $f = g$, as reported in [1]. This conjecture was extensively studied by the second author and collaborators during a Simons workshop in 2015. There (as reported in [8]), they reformulated Conjecture 1 using Φ-entropies (in particular, using the Jensen-Shannon divergence) and proposed the following conjecture based on the Helliger distance.

Conjecture 2 ([3]). *Let $X \sim \text{Unif}(\mathbb{H}_n)$. Let Y be obtained by flipping each bit of X independently with probability $\frac{1-\rho}{2}$. Let $f(X)$ be a Boolean function that takes values in $\{-1,1\}$. Then*

$$\sqrt{1-\mathbb{E}[f]^2} - \mathbb{E}\left[\sqrt{1-((T_\rho f)(Y))^2}\right] \leq 1 - \sqrt{1-\rho^2}. \tag{3}$$

Here, $T_\rho f$ is the standard Boolean operator $\mathbb{E}[f(X)|Y]$. Furthermore, it was shown (using ideas that motivated this article) that if Conjecture 2 holds, then Conjecture 1 holds.

Samorodnitsky, [4], showed that there exists a positive ρ_0 such that Conjecture 1 is true for balanced Boolean functions when $|\rho| \leq \rho_0$. This result was extended by Lei Yu, [5], who showed that Conjecture 1 is true for balanced Boolean functions for $|\rho| \leq 0.44$ (approximately).

Li and Médard, [6], investigated the Boolean function that maximizes $\mathbb{E}[|T_\rho f|^\alpha]$, for a fixed mean and $\alpha \in [1,2]$, and conjectured the optimality of the dictator function. Barnes and Özgür, [7], showed that for balanced Boolean functions, the Li-Médard conjecture is equivalent to the Courtade-Kumar conjecture.

Note: For the rest of the article, we assume that $f(X)$ takes values in $\{-1,1\}$.

[1]This work was supported by the following grants from the University Grants Committee of the Hong Kong Special Administrative Region: GRF14210120, GRF14221822.

B. Motivation and Isoperimetric Inequalities

Since the $(T_\rho)_{\rho\geq 0}$ operator forms a Markov semigroup, i.e. $T_{\rho_1}\circ T_{\rho_2} = T_{\rho_1\rho_2}$, it is intuitively expected that the above conjectures would be harder to prove as ρ becomes closer to one. (By symmetry, we assume that ρ takes values in $[0,1]$.) The results obtained so far by the community seem to corroborate this belief. However, one of the many indications of the validity of Conjecture 1 rests on the observation that as $\rho \uparrow 1$ (see Remark 1), Conjecture 1 would hold provided that for all balanced functions, i.e. $f : \mathbb{E}[f] = 0$, we have

$$\mathbb{E}[s_f(X)] \geq 1,$$

where $s_f(X)$ (or the sensitivity of the function at X) denotes the number of neighbours of X where the function takes the opposite sign. On the other hand, the above statement is a special case of Harper's isoperimetric inequality and immediately follows from Boolean Fourier Analysis.

Remark 1. When $\rho = 1$, the conjectured Boolean function inequalities are trivially true, and equality holds if and only if f is balanced. Then, one can consider the derivative (in ρ) of the conjectured inequality at $\rho = 1$ for balanced functions and deduce the inequality about the sensitivity of the function.

However, a similar statement for Conjecture 2 is unknown. It was known that as $\rho \uparrow 1$, Conjecture 2 would hold provided that for all balanced functions, i.e. $f : \mathbb{E}[f] = 0$, we have

$$\mathbb{E}\left[\sqrt{s_f(X)}\right] \geq 1.$$

It turns out that the above statement has received considerable attention from the probability community, and several interesting results have been established.

A well-known result is Talagrand's isoperimetric inequality on the Boolean Hypercube. Equip the vertices in $\mathbb{H}_n$ with the uniform probability measure. Let A be a subset of vertices of $\mathbb{H}_n$ and for $x \in A$, let $h_A(x)$ be defined as the number of neighbors of x that do not belong to A (and $h_A(x) = 0$ if $x \notin A$). Then Talagrand [8] established that

$$\mathbb{E}\left[\sqrt{h_A}\right] \geq \sqrt{2}\mu(A)(1-\mu(A)).$$

If one defines $f(x) = (-1)^{\mathbb{1}_{A^c}(x)}$, then observe that $s_f(x) = h_A(x) + h_{A^c}(x)$. Further as one of $h_A(x)$ or $h_{A^c}(x) = 0$ at every x, we also see that $s_f^\alpha(x) = h_A^\alpha(x) + h_{A^c}^\alpha(x)$ for any $\alpha > 0$. Therefore Talagrand's isoperimetric inequality implies that for balanced functions,

$$\mathbb{E}\left[\sqrt{s_f(X)}\right] \geq \frac{1}{\sqrt{2}}.$$

Bobkov improved this in [9] to show that for balanced Boolean functions,

$$\mathbb{E}\left[\sqrt{s_f(X)}\right] \geq \sqrt{\frac{2}{\pi}}.$$

Finally, a recent result, [10], implies that for balanced Boolean functions,

$$\mathbb{E}\left[\sqrt{s_f(X)}\right] \geq 0.91.$$

If one were to consider alternately $\mathbb{E}\left[s_f^\alpha(X)\right]$, note that for any $\alpha > 0$, the dictator function yields a value of one. It was known (see [8]) that for $\alpha < \frac{1}{2}$, the majority function would yield a value lower than one. On the other hand, a rather recent result by Kahn and Park [11] showed that

$$\mathbb{E}\left[h_A^{\log_2\left(\frac{3}{2}\right)}\right] \geq 2\mu(A)(1-\mu(A)).$$

This implies that for all balanced Boolean functions

$$\mathbb{E}\left[s_f^{\log_2\left(\frac{3}{2}\right)}(X)\right] \geq 1.$$

Building on this work, [10], proved that for $\forall \mu(A) \geq \frac{1}{2}$

$$\mathbb{E}\left[h_A^{0.53}\right] \geq 2\mu(A)(1-\mu(A)).$$

This implies that for all balanced Boolean functions

$$\mathbb{E}\left[s_f^{0.53}(X)\right] \geq 1.$$

The main contribution of this work is the following:

- we present a family of conjectures parameterized by $\alpha, \frac{1}{2} \le \alpha \le 2$, where $\alpha = \frac{1}{2}$ corresponds to Conjecture 2; $\alpha = 1$ corresponds to Conjecture 1, and $\alpha = 2$ corresponds to a known statement about the optimality of the dictator function. Further, we prove that establishing the conjecture for any α in this interval would imply the conjecture for any $\beta : \alpha \le \beta \le 2$.
- we obtain a lower bound for $\mathbb{E}\left[\sqrt{1-((T_\rho f)(Y))^2}\right]$, equivalently, an upper bound for Conjecture 2, by borrowing on ideas used by Bobkov [9] that improved on Talagrand's isoperimetric inequality.

II. Main

A. Preliminaries

As given by Chafaï [12], the Φ entropy of a function f is defined by

$$H_\Phi(f) := \mathbb{E}\left[\Phi(f)\right] - \Phi(\mathbb{E}\left[f\right]),$$

where $\Phi(x)$ is a convex function.

1) A new family of Φ-Entropies: We define a parameterized family of functions $\Phi_\alpha : [0,1] \mapsto [0,1]$, $\alpha \neq 1$ to be

$$\Phi_\alpha(p) = 1 - \frac{\alpha\left(\left(p^\alpha + (1-p)^\alpha\right)^{\frac{\log_2 \alpha}{\alpha-1}} - 1\right)}{1-\alpha}. \tag{4}$$

Note that $\lim_{\alpha \to 1} \Phi_\alpha(p) = 1 - H_2(p)$, where $H_2(p)$ is the binary entropy function, so we define $\Phi_1(p)$ accordingly.

Remark 2. Notice that $\Phi_\alpha(p)$ can also be written as the form $\Phi_\alpha(p) = 1 - \frac{\alpha\left(\alpha^{-H_\alpha(p)}-1\right)}{1-\alpha}$, where $H_\alpha(p) = \frac{1}{1-\alpha}(\log_2(p^\alpha + (1-p)^\alpha)$ is the Rényi entropy of a binary random variable taking values with probability $\{p, 1-p\}$, respectively.

The reader can verify the following immediate results about this parameterized family.

Lemma 1. *The parameterized family $\{\Phi_\alpha\}, \alpha > 0$ satisfies the following properties:*

(i) $\Phi_{\frac{1}{2}}(p) = 1 - 2\sqrt{p(1-p)}$,
(ii) $\Phi_2(p) = (2p-1)^2$,
(iii) $\Phi_\alpha(\frac{1}{2}) = 0$,
(iv) $\Phi_\alpha(0) = 1$,
(v) $\Phi_\alpha(p) = \Phi_\alpha(1-p), p \in [0,1]$,
(vi) $\Phi_\alpha(p)$ *is decreasing in the interval* $p \in [0, \frac{1}{2}]$ *and convex in the interval* $p \in [0,1]$.

A proof of the last part is provided in the full version [13].

A natural extension of the family of functions to generalized discrete probability distributions is the following

Denote $\mathbf{p} = (p_1, p_2, \cdots, p_n)$ be a probability vector, we define the following

$$\begin{aligned}\Phi_\alpha(\mathbf{p}) &= \frac{\alpha^{\log_2 n}\left(n^{-\log_2 \alpha} - \left(\sum_{i=1}^n p_i^\alpha\right)^{\frac{\log_2 \alpha}{\alpha-1}}\right)}{1-\alpha^{\log_2 n}}, \\ &= \frac{1-\alpha^{\log_2 n - H_\alpha(\mathbf{p})}}{1-\alpha^{\log_2 n}}.\end{aligned}$$

Denote Unif(n) to be the uniform distribution of a set with size n and $\pi(\mathbf{p})$ be any permutation of $\mathbf{p}$. The above function satisfies the following properties.

(i) $\Phi_{\frac{1}{2}}(\mathbf{p}) = 1 - \frac{2\sum_{i<j}\sqrt{p_i p_j}}{n-1}$,
(ii) $\Phi_1(\mathbf{p}) = 1 - \frac{H_2(\mathbf{p})}{\log_2 n}$
(iii) $\Phi_2(\mathbf{p}) = 1 - \frac{2n\sum_{i<j} p_i p_j}{n-1}$,
(iv) $\Phi_\alpha(\text{Unif}(n)) = 0$,
(v) $\Phi_\alpha(\pi((1, 0, \cdots, 0))) = 1$,
(vi) $\Phi_\alpha(\mathbf{p}) = \Phi_\alpha(\pi(\mathbf{p}))$,
(vii) $\Phi_\alpha(\mathbf{p})$ is jointly convex for $\forall \alpha \le 2$,
$(viii)$ $\Phi_\alpha(t\mathbf{p} + (1-t)\text{Unif}(n))$ is non-negative and monotonically increasing with respect to $t \in [0,1]$.

We attach the proof of the last two items in the appendix.

The main family of conjectures, in this article, is the following:

Conjecture 3. *Let $X \sim$ Unif$(\mathbb{H}_n)$. Let Y be obtained by flipping each bit of X independently with probability $\frac{1-\rho}{2}$. Let $f(X)$ be a Boolean function that takes values in $\{-1, 1\}$. Then, for any $\frac{1}{2} \le \alpha < 2$, we conjecture that*

$$H_{\Phi_\alpha}\left(\frac{1-(T_\rho f)(Y)}{2}\right) \le H_{\Phi_\alpha}\left(\frac{1-\rho Y_1}{2}\right), \tag{5}$$

where Y_1 is the first co-ordinate of Y.

As the right-hand-side corresponds to the evaluation of the left-hand-side for the dictator function, i.e. $f(X) = X_1$, the conjecture can be rephrased as saying that the dictator function maximizes the left-hand-side.

Remark 3. The following points are worth noting:

- When $\alpha = \frac{1}{2}$, the Conjecture 3 reduces to Conjecture 2.
- When $\alpha = 1$, the Conjecture 3 reduces to Conjecture 1.
- When $\rho \to 1$, the validity of the conjecture in the limit ($\rho \to 1$) is the same as verifying for all balanced Boolean functions
$$\mathbb{E}\left[s_f^{\alpha}(X)\right] \geq 1.$$
This is true, [10], if $\alpha \geq 0.53$. Furthermore, the above statement is false when $\alpha < \frac{1}{2}$. This implies that Conjecture 3 cannot be extended to $\alpha < \frac{1}{2}$.
- When $\alpha = 2$, the equivalent statement would be
$$\mathbb{E}\left[((T_\rho f)(Y))^2\right] - (\mathbb{E}[f])^2 \leq \rho^2,$$
which is easily verified by considering the Fourier representation and Parseval's theorem.

One of the main results of this paper is the following:

Theorem 1. *If Conjecture 3 holds for any $\alpha \in [\frac{1}{2}, 2)$, it also holds for any β satisfying $\alpha \leq \beta < 2$.*

Remark 4. This shows that proving Conjecture 2 will prove Conjecture 3; however proving Conjecture 3 for some alpha in $(\frac{1}{2}, 1]$ would be sufficient to imply Conjecture 1. This family is motivated by the fact that recent success has been made in the lower bound for $\mathbb{E}\left[s_f^{\alpha}(X)\right]$, for $\alpha = \log_2\left(\frac{3}{2}\right)$.

We will use the following standard Boolean Fourier representation,
$$f(\mathbf{x}) = \hat{f}_\emptyset + \sum_{S \subset [1:n], |S| \geq 1} \hat{f}_S \chi_S(\mathbf{x}).$$

Note that $\mathbb{E}[f] = \hat{f}_\emptyset$. Using this notation, Conjecture 2 is equivalent to requiring that all Boolean
$$\mathbb{E}\left[\sqrt{1 - ((T_\rho f)(Y))^2}\right] \geq \sqrt{1 - \hat{f}_\emptyset^2} + \sqrt{1 - \rho^2} - 1.$$
In Theorem 2 below, we show how the tools used in [9] for deriving isoperimetric inequalities can lead to a lower bound for the left-hand-side. In particular. we show that

Theorem 2. *Let $X \sim \mathrm{Unif}(\mathbb{H}_n)$. Let Y be obtained by flipping each bit of X independently with probability $\frac{1-\rho}{2}$. Let $f(X)$ be a Boolean function taking values in $\{-1, 1\}$. Let*
$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \hat{G}(x) = \int_{-\infty}^{x} g(t) dt.$$
Define $G(a) = 2g\left(\hat{G}^{-1}\left(\frac{1+a}{2}\right)\right)$. Then,
$$\mathbb{E}\left[\sqrt{1 - ((T_\rho f)(Y))^2}\right] \geq G\left(\hat{f}_\emptyset\right)\sqrt{1 - \rho^2}. \tag{6}$$

Remark 5. There are non-trivial regimes of parameters $(\mathbb{E}[f], \rho)$ where
$$\sqrt{1 - (\mathbb{E}[f])^2} + \sqrt{1 - \rho^2} - 1 \leq G(\mathbb{E}[f])\sqrt{1 - \rho^2}.$$
For this set of parameters, Theorem 2 gives a proof of Conjecture 2. However, we are more interested in demonstrating the link between proof ideas in isoperimetric inequalities and the family of Conjectures.

From Theorem 2 we obtain that
$$\sqrt{1 - \hat{f}_\emptyset^2} - \mathbb{E}\left[\sqrt{1 - ((T_\rho f)(Y))^2}\right] \leq \sqrt{1 - \hat{f}_\emptyset^2} - G(\hat{f}_\emptyset)\sqrt{1 - \rho^2},$$
or equivalent
$$H_{\Phi_{\frac{1}{2}}}\left(\frac{1 - (T_\rho f)(Y)}{2}\right) \leq \sqrt{1 - \hat{f}_\emptyset^2} - G(\hat{f}_\emptyset)\sqrt{1 - \rho^2}. \tag{7}$$

III. Proofs of the Results

A. Proof of Theorem 1

The idea of the proof is similar to that in [3] where it was shown that Conjecture 2 would imply Conjecture 1. Let us denote $\Psi_\alpha(p) = 1 - \Phi_\alpha(p)$ and for $\alpha > 0$, let us define $\Psi_\alpha^{-1} : [0,1] \mapsto [0, \frac{1}{2}]$ as the inverse mapping to the interval $[0, \frac{1}{2}]$. This is well-defined from part (vi) of Lemma 1. Further, again using the same result, if $0 \le y_1 < y_2 < 1$, we have $\Psi_\alpha^{-1}(y_1) < \Psi_\alpha^{-1}(y_2)$ and consequently $\Psi_\beta\left(\Psi_\alpha^{-1}(y_1)\right) < \Psi_\beta\left(\Psi_\alpha^{-1}(y_2)\right)$. Therefore $\Psi_\beta\left(\Psi_\alpha^{-1}(y)\right)$ is non-negative and increasing in $[0,1]$.

We divide the proof of Theorem 1 into two parts.

1) Part 1: $\frac{1}{2} \le \alpha \le \beta \le 1$:

Lemma 2. *For $\frac{1}{2} \le \alpha \le \beta \le 1$, the function $\Psi_\beta\left(\Psi_\alpha^{-1}(y)\right)$ is convex in y, for $y \in [0,1]$.*

Proof. The proof is in the full version [13]. □

Note that the inequality in Conjecture 3 can be equivalently written as

$$\Psi_\alpha\left(\frac{1-\mathbb{E}\left[(T_\rho f)(Y)\right]}{2}\right) - \mathbb{E}\left[\Psi_\alpha\left(\frac{1-(T_\rho f)(Y)}{2}\right)\right] \le \Psi_\alpha\left(\frac{1}{2}\right) - \Psi_\alpha\left(\frac{1-\rho}{2}\right).$$

To complete the proof of this part, we employ a special case of a well-known weak majorization inequality (Karamata's inequality).

Lemma 3 (Karamata)**.** *Let $f(x)$ is non-negative, increasing, and convex function on $[0,1]$, such that $x_1 + x_2 \le y_1 + y_2$ and $\max\{y_1, y_2\} \ge \max\{x_1, x_2\}$. Then*

$$f(x_1) + f(x_2) \le f(y_1) + f(y_2).$$

The proof of this part of Theorem 1, i.e. $\frac{1}{2} \le \alpha \le \beta \le 1$, is completed by the following proposition.

Proposition 1. *For $\frac{1}{2} \le \alpha \le \beta \le 1$, suppose the inequality*

$$\Psi_\alpha\left(\frac{1-\mathbb{E}\left[(T_\rho f)(Y)\right]}{2}\right) - \mathbb{E}\left[\Psi_\alpha\left(\frac{1-(T_\rho f)(Y)}{2}\right)\right] \le \Psi_\alpha\left(\frac{1}{2}\right) - \Psi_\alpha\left(\frac{1-\rho}{2}\right)$$

holds, then

$$\Psi_\beta\left(\frac{1-\mathbb{E}\left[(T_\rho f)(Y)\right]}{2}\right) - \mathbb{E}\left[\Psi_\beta\left(\frac{1-(T_\rho f)(Y)}{2}\right)\right] \le \Psi_\beta\left(\frac{1}{2}\right) - \Psi_\beta\left(\frac{1-\rho}{2}\right)$$

also holds.

Proof. The second statement is equivalently written as requiring to show

$$\Psi_\beta\left(\Psi_\alpha^{-1}\left(\Psi_\alpha\left(\frac{1-\mathbb{E}\left[(T_\rho f)(Y)\right]}{2}\right)\right)\right) - \mathbb{E}\left[\Psi_\beta\left(\Psi_\alpha^{-1}\left(\Psi_\alpha\left(\frac{1-(T_\rho f)(Y)}{2}\right)\right)\right)\right] \le \Psi_\beta\left(\Psi_\alpha^{-1}\left(\Psi_\alpha\left(\frac{1}{2}\right)\right)\right) - \Psi_\beta\left(\Psi_\alpha^{-1}\left(\Psi_\alpha\left(\frac{1-\rho}{2}\right)\right)\right).$$

In light of Lemma 2 and Lemma 3, by taking $f(x) = \Psi_\beta\left(\Psi_\alpha^{-1}(x)\right)$, $x_1 = \Psi_\alpha\left(\frac{1-\mathbb{E}[(T_\rho f)(Y)]}{2}\right)$, $x_2 = \Psi_\alpha\left(\frac{1-\rho}{2}\right)$, $y_1 = \mathbb{E}\left[\Psi_\alpha\left(\frac{1-(T_\rho f)(Y)}{2}\right)\right]$, and $y_2 = \Psi_\alpha\left(\frac{1}{2}\right) = 1$, it suffices to show that $x_1 + x_2 \le y_1 + y_2$ (as $y_2 = 1 \ge \max\{x_1, x_2\}$), or equivalently

$$\Psi_\alpha\left(\frac{1-\mathbb{E}\left[(T_\rho f)(Y)\right]}{2}\right) - \mathbb{E}\left[\Psi_\alpha\left(\frac{1-(T_\rho f)(Y)}{2}\right)\right] \le \Psi_\alpha\left(\frac{1}{2}\right) - \Psi_\alpha\left(\frac{1-\rho}{2}\right).$$

This shows that if Conjecture 3 folds for α, then it also holds for β for $\frac{1}{2} \le \alpha \le \beta \le 1$. □

2) Part 2: $1 \leq \alpha \leq \beta \leq 2$: The proof of this part is more or less the same as the previous part. We prove a similar convexity result as in Lemma 2.

Lemma 4. *For $1 \leq \alpha \leq \beta \leq 2$, the function $\Psi_\beta\left(\Psi_\alpha^{-1}(y)\right)$ is convex in y, for $y \in [0,1]$.*

Proof. The proof is given in the full version [13]. □

Given this convexity, we now can apply Proposition 1 (with $1 \leq \alpha \leq \beta \leq 2$) to complete the proof of Part 2.

Remark 6. Notice that the limit of Ψ_α at $\alpha = 1$ is well-defined and equals the Shannon entropy. Hence, establishing the conjecture for $\alpha < 1$, would imply Conjecture 1 (using Part 1), and Conjecture 1 would imply Conjecture 3 for any $1 < \beta \leq 2$ (using Part 2).

This completes the proof of Theorem 1.

B. Proof of Theorem 2

The techniques used in all the works to obtain isoperimetric inequalities on the Boolean Hypercube revolve around an inductive proof strategy to derive a functional inequality. Here, we adopt a similar strategy to get a lower bound for $\mathbb{E}\left[\sqrt{1-((T_\rho f)(Y))^2}\right]$ in terms of the mean of the Boolean function.

Given a Boolean function f, we define $f_+ = f(\tilde{\mathbf{x}}, +1)$ and $f_- = f(\tilde{\mathbf{x}}, -1)$ to be two child functions on a smaller dimensional Hypercube generated by $f(\mathbf{x})$, where $\mathbf{x}$ is a vector in the Hypercube of dimension $n+1$ and $\tilde{\mathbf{x}}$ is in the subcube of dimension n. Elementary calculations yield that

$$(T_\rho f)(\mathbf{x}) = \frac{1+\rho x_{n+1}}{2}(T_\rho f_+)(\tilde{\mathbf{x}}) + \frac{1-\rho x_{n+1}}{2}(T_\rho f_-)(\tilde{\mathbf{x}}).$$

With appropriate abuse of notation for simplicity, it follows that

$$\begin{aligned} 1 - T_\rho f^2 = &\left(\frac{1+\rho x_{n+1}}{2}\right)^2 (1 - T_\rho f_+^2) \\ &+ \left(\frac{1-\rho x_{n+1}}{2}\right)^2 (1 - T_\rho f_-^2) \\ &+ 2\left(\frac{1-\rho^2}{4}\right)(1 - T_\rho f_+ T_\rho f_-). \end{aligned}$$

Consequently,

$$\begin{aligned} \mathbb{E}\left[\sqrt{1-T_\rho f^2}\right] = &\frac{1}{2}\mathbb{E}\Bigg[\Bigg\{\left(\frac{1+\rho}{2}\right)^2 (1 - T_\rho f_+^2) \\ &+ \left(\frac{1-\rho}{2}\right)^2 (1 - T_\rho f_-^2) \\ &+ 2\left(\frac{1-\rho^2}{4}\right)(1 - T_\rho f_+ T_\rho f_-)\Bigg\}^{\frac{1}{2}}\Bigg] \\ &+ \frac{1}{2}\mathbb{E}\Bigg[\Bigg\{\left(\frac{1-\rho}{2}\right)^2 (1 - T_\rho f_+^2) \\ &+ \left(\frac{1+\rho}{2}\right)^2 (1 - T_\rho f_-^2) \\ &+ 2\left(\frac{1-\rho^2}{4}\right)(1 - T_\rho f_+ T_\rho f_-)\Bigg\}^{\frac{1}{2}}\Bigg]. \end{aligned}$$

Note that for $a,b,c \geq 0$, the function: $\Phi(x,y,z) := \sqrt{ax^2+by^2+cz^2}$ is jointly convex. Therefore, we have

$$\begin{aligned}
&\mathbb{E}\left[\sqrt{1-T_\rho f^2}\right] \\
&\geq \frac{1}{2}\Bigg\{\left(\frac{1+\rho}{2}\right)^2\left(\mathbb{E}\left[\sqrt{(1-T_\rho f_+^2)}\right]\right)^2 \\
&\quad+\left(\frac{1-\rho}{2}\right)^2\left(\mathbb{E}\left[\sqrt{(1-T_\rho f_-^2)}\right]\right)^2 \\
&\quad+2\left(\frac{1-\rho^2}{4}\right)\left(\mathbb{E}\left[\sqrt{(1-T_\rho f_+ T_\rho f_-)}\right]\right)^2\Bigg\}^{\frac{1}{2}} \\
&+\frac{1}{2}\Bigg\{\left(\frac{1-\rho}{2}\right)^2\left(\mathbb{E}\left[\sqrt{(1-T_\rho f_+^2)}\right]\right)^2 \\
&\quad+\left(\frac{1+\rho}{2}\right)^2\left(\mathbb{E}\left[\sqrt{(1-T_\rho f_-^2)}\right]\right)^2 \\
&\quad+2\left(\frac{1-\rho^2}{4}\right)\left(\mathbb{E}\left[\sqrt{(1-T_\rho f_+ T_\rho f_-)}\right]\right)^2\Bigg\}^{\frac{1}{2}}.
\end{aligned}$$

The third term can be expressed as follows:

$$\begin{aligned}
\sqrt{1-T_\rho f_+ T_\rho f_-} = \Bigg\{&\frac{1}{2}\sqrt{1-T_\rho f_+^2}^2+\frac{1}{2}\sqrt{1-T_\rho f_-^2}^2 \\
&+\frac{1}{2}(T_\rho f_+ - T_\rho f_-)^2\Bigg\}^{\frac{1}{2}}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{E}\left[\sqrt{1-T_\rho f_+ T_\rho f_-}\right] \geq \Bigg\{&\frac{1}{2}\mathbb{E}\left[\sqrt{1-T_\rho f_+^2}\right]^2 \\
&+\frac{1}{2}\mathbb{E}\left[\sqrt{1-T_\rho f_-^2}\right]^2 \\
&+\frac{1}{2}\mathbb{E}\left[T_\rho f_+ - T_\rho f_-\right]^2\Bigg\}^{\frac{1}{2}}.
\end{aligned}$$

Let $\varphi : [-1,1] \mapsto (0,1)$, satisfying $\varphi(-1)=\varphi(1)=0$, and

$$\begin{aligned}
&\frac{1}{2}\Bigg\{\left(\frac{1+\rho}{2}\right)^2\varphi^2(a)+\left(\frac{1-\rho}{2}\right)^2\varphi^2(b) \\
&\quad+2\left(\frac{1-\rho^2}{4}\right)\frac{1}{2}\left(\varphi^2(a)+\varphi^2(b)+(a-b)^2\right)\Bigg\}^{\frac{1}{2}} \\
&+\frac{1}{2}\Bigg\{\left(\frac{1-\rho}{2}\right)^2\varphi^2(a)+\left(\frac{1+\rho}{2}\right)^2\varphi^2(b) \\
&\quad+2\left(\frac{1-\rho^2}{4}\right)\frac{1}{2}\left(\varphi^2(a)+\varphi^2(b)+(a-b)^2\right)\Bigg\}^{\frac{1}{2}} \\
&\quad\geq \varphi\left(\frac{a+b}{2}\right),
\end{aligned}$$

for all $-1 \leq a, b \leq 1$. By induction, it is immediate from above any such $\varphi(a)$ provides a lower bound for $\mathbb{E}\left[\sqrt{1-((T_\rho f)(Y))^2}\right]$, with $\mathbb{E}[f] = a$. Simplifying, the above functional condition is equivalent to

$$\left\{\frac{1+\rho}{2}\varphi^2(a) + \frac{1-\rho}{2}\varphi^2(b) + \left(\frac{1-\rho^2}{4}\right)(a-b)^2\right\}^{\frac{1}{2}} + \left\{\frac{1-\rho}{2}\varphi^2(a) + \frac{1+\rho}{2}\varphi^2(b) + \left(\frac{1-\rho^2}{4}\right)(a-b)^2\right\}^{\frac{1}{2}} \geq 2\varphi\left(\frac{a+b}{2}\right).$$

Now, we assume that $\varphi(a) = \sqrt{1-\rho^2}\psi(a)$, among which $\psi(a)$ does not depend on ρ. (This restricts our choice and perhaps leads to a sub-optimal lower bound). Now have that

$$\left\{\frac{1+\rho}{2}\psi^2(a) + \frac{1-\rho}{2}\psi^2(b) + \frac{(a-b)^2}{4}\right\}^{\frac{1}{2}} + \left\{\frac{1-\rho}{2}\psi^2(a) + \frac{1+\rho}{2}\psi^2(b) + \frac{(a-b)^2}{4}\right\}^{\frac{1}{2}} \geq 2\psi\left(\frac{a+b}{2}\right).$$

The left-hand-side is concave and symmetric in ρ, for $\rho \in [-1, 1]$, therefore the minimum is attained at $\rho = 1$ (as value at $\rho = 1$ is same as $\rho = -1$). Therefore $\Psi(a)$ must satisfy

$$\left\{\psi^2(a) + \frac{(a-b)^2}{4}\right\}^{\frac{1}{2}} + \left\{\psi^2(b) + \frac{(a-b)^2}{4}\right\}^{\frac{1}{2}} \geq 2\psi\left(\frac{a+b}{2}\right),$$

for $-1 \leq a, b \leq 1$. Let $\hat{\psi}(x) = \frac{1}{2}\psi(2x-1)$. Therefore, we require $\hat{\psi}(0) = \hat{\psi}(1) = 0$ and satisfies

$$\frac{1}{2}\left\{\hat{\psi}^2(x) + \frac{(x-y)^2}{4}\right\}^{\frac{1}{2}} + \frac{1}{2}\left\{\hat{\psi}^2(y) + \frac{(x-y)^2}{4}\right\}^{\frac{1}{2}} \geq \hat{\psi}\left(\frac{x+y}{2}\right),$$

for $0 \leq x, y \leq 1$. This is precisely the functional recursion considered and solved by Bobkov [9] while studying the isoperimetric inequality, except $\hat{\psi}(x) \in (0, \frac{1}{2})$. The following was proved in [9]. Let

$$g(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad \hat{G}(x) = \int_{-\infty}^{x} g(t)dt.$$

Then $I(x) = g\left(\hat{G}^{-1}(x)\right), 0 \leq x \leq 1$ is the maximal non-negative function satisfying the functional recursion of $\hat{\psi}$. Consequently $G(a) = 2I\left(\frac{1+a}{2}\right)$ is a non-negative function satisfying the functional recursion of ψ.

Therefore, putting the above observations and noting that $\mathbb{E}[f] = \hat{f}_\emptyset$, we obtain that

$$\mathbb{E}\left[\sqrt{1-((T_\rho f)(Y))^2}\right] \geq G\left(\hat{f}_\emptyset\right)\sqrt{1-\rho^2},$$

completing the proof of Theorem 2.

IV. Future Work and Conclusion

In this paper, we proposed a parameterized family of phi=entropies and related conjectures on the optimality of Boolean functions and related a limiting case of this conjecture to isoperimetric inequalities on the Boolean Hypercube. Further, we showed that the family of conjectures is ordered (in the sense that establishing the conjecture for a particular value would imply the conjecture for higher values). Finally, we showed that mimicking proof techniques in the isoperimetric inequalities leads to non-trivial lower bounds to our conjectures.

As an immediate future work, it would be useful to see if one can recover the functional recursion used by Kahn and Park [11] (in the sense of the proof of Theorem 2) for some $\alpha \in (\frac{1}{2}, 1)$. An immediate consequence would be the proof of Conjecture 1 for all balanced functions.

References

[1] T. A. Courtade and G. R. Kumar, "Which boolean functions maximize mutual information on noisy inputs?," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4515–4525, 2014.

[2] G. Pichler, P. Piantanida, and G. Matz, "Dictator functions maximize mutual information," 2018.

[3] V. Anantharam, A. Bogdanov, A. Chakrabarti, T. Jayaram, and C. Nair, "A conjecture regarding optimality of the dictator function under Hellinger distance," *Information Theory and Applications Workshop*, 2017. Available at http://chandra.ie.cuhk.edu.hk/pub/papers/HC/hel-conj.pdf.

[4] A. Samorodnitsky, "On the entropy of a noisy function," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5446–5464, 2016.

[5] L. Yu, "On the Φ-stability and related conjectures," *Probability Theory and Related Fields*, vol. 186, pp. 1045–1080, Aug 2023.

[6] J. Li and M. Médard, "Boolean functions: Noise stability, non-interactive correlation distillation, and mutual information," *IEEE Transactions on Information Theory*, vol. 67, no. 2, pp. 778–789, 2021.

[7] L. P. Barnes and A. Özgür, "The courtade-kumar most informative boolean function conjecture and a symmetrized li-médard conjecture are equivalent," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 2205–2209, 2020.

[8] M. Talagrand, "Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem," *Geometric & Functional Analysis GAFA*, vol. 3, no. 3, pp. 295–314, 1993.

[9] S. G. Bobkov, "An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in gauss space," *Ann. Probab.*, vol. 25, pp. 206–214, 01 1997.

[10] D. Beltran, P. Ivanisvili, and J. Madrid, "On sharp isoperimetric inequalities on the hypercube," 2023.

[11] J. Kahn and J. Park, "An isoperimetric inequality for the hamming cube and some consequences," 2019.

[12] D. Chafaï, "Entropies, convexity, and functional inequalities, on ϕ -entropies and ϕ -Sobolev inequalities," *Kyoto Journal of Mathematics*, vol. 44, jan 2004.

[13] Z. Chen and C. Nair, "A parameterized family of Phi-entropies, on optimality of dictator functions and their relation to isoperimetric inequalities on the Boolean Hypercube." http://chandra.ie.cuhk.edu.hk/pub/publications.html, May 2024.

Appendix: On the Optimality of Dictator functions and Isoperimetric Inequalities on Boolean Hypercubes

A. Proof of Lemma 1

Lemma 5. *The following inequality holds for $a \neq b$*

$$e^{\frac{a+b}{2}} < \frac{e^a - e^b}{a-b} < \frac{e^a + e^b}{2}.$$

Proof. This is well-known, but we give an argument here for completeness. Let $a - b = u$ and w.l.o.g assume that $u > 0$. Now, multiplying by e^{-b} on both sides, we need to show

$$e^{\frac{u}{2}} < \frac{e^u - 1}{u} < \frac{e^u + 1}{2}.$$

Now, by performing Taylor-Series expansion and comparing coefficients, the inequality is immediate. □

Remark 7. Observing that by setting $t^a = e^{a \ln t}, t > 1$, Lemma 5 implies that

$$t^{\frac{a+b}{2}} < \frac{t^a - t^b}{(a-b)\ln t} < \frac{t^a + t^b}{2}.$$

Proof of Lemma 1. We first prove the monotonicity,

$$\begin{aligned}\Phi'_\alpha(p) = \frac{\alpha^2 \log_2 \alpha}{(1-\alpha)^2}&(p^\alpha + (1-p)^\alpha)^{\frac{\log_2 \alpha}{\alpha-1}-1} \\ &\times (p^{\alpha-1} - (1-p)^{\alpha-1}).\end{aligned}$$

When $\alpha < 1, p \in [0, \frac{1}{2}]$, we have that $\log_2 \alpha < 0$ and $p^{\alpha-1} \geq (1-p)^{\alpha-1}$, the function is monotonically decreasing.

When $\alpha > 1, p \in [0, \frac{1}{2}]$, we have that $\log_2 \alpha > 0$ and $p^{\alpha-1} \leq (1-p)^{\alpha-1}$, the function is also monotonically decreasing.

Lastly, when $\alpha = 1$, the monotonicity and convexity follow immediately from the corresponding properties of the binary Shannon entropy function.

For the proof of convexity, if $\alpha > 1$, we have that

$$\begin{aligned}\Phi''_\alpha(p) = &\frac{\alpha^3 \log_2 \alpha}{(1-\alpha)^2}(p^\alpha + (1-p)^\alpha)^{\frac{\log_2 \alpha}{\alpha-1}-2} \\ &\times \left(\frac{\log_2 \alpha}{\alpha - 1} - 1\right)(p^{\alpha-1} - (1-p)^{\alpha-1})^2 \\ &+ \frac{\alpha^2 \log_2 \alpha}{(\alpha-1)}(p^\alpha + (1-p)^\alpha)^{\frac{\log_2 \alpha}{\alpha-1}-1} \\ &\times (p^{\alpha-2} + (1-p)^{\alpha-2}).\end{aligned}$$

Notice that all the terms here are non-negative for $p \in [0, 1]$, hence $\Phi_\alpha(p)$ is convex on $[0, 1]$.

When $\alpha \in [\frac{1}{2}, 1)$, it suffices to show that

$$\begin{aligned}&\left(\frac{1}{\alpha} - 1\right)(p^\alpha + (1-p)^\alpha)(p^{\alpha-2} + (1-p)^{\alpha-2}) \\ &\geq \left(\frac{\log_2 \alpha}{\alpha - 1} - 1\right)(p^{\alpha-1} - (1-p)^{\alpha-1})^2.\end{aligned}$$

This holds when $p = \frac{1}{2}$. We divide both sides by $p^{2\alpha-2}$ and let $t = \frac{1-p}{p}, t \neq 1$. By symmetry (function on the left-hand-side remains unchanged when t is replaced by $\frac{1}{t}$), it suffices to show that for $t > 1$

$$\frac{(1+t^\alpha)(1+t^{\alpha-2})}{(1-t^{\alpha-1})^2} \geq \frac{\alpha(\alpha - 1 - \log_2 \alpha)}{(1-\alpha)^2}.$$

It is immediate to verify that $\alpha(\alpha - 1 - \log_2 \alpha)$ is monotonically decreasing with respect to α when $\alpha \in [\frac{1}{2}, 1]$. The derivative in the interval has a maximum value $1 - \frac{1}{\ln 2} < 0$. Therefore, the right-hand-side is upper bounded by $\frac{1}{4(1-\alpha)^2}$.

Hence, it suffices to show that

$$(1+t^\alpha)(1+t^{\alpha-2}) \geq \frac{(1-t^{\alpha-1})^2}{4(1-\alpha)^2}.$$

By Remark 7, it suffices to show

$$4(1+t^\alpha)(1+t^{\alpha-2}) \geq \left[\left(\frac{1+t^{\alpha-1}}{2}\right)\ln t\right]^2.$$

This is equivalent to

$$16 + \frac{16t^{\alpha-2}(t-1)^2}{(1+t^{\alpha-1})^2} \geq \ln^2 t.$$

When $\ln t \leq 4$, the inequality holds trivially.

When $\ln t > 4$, since $t^{\alpha-2} \geq t^{-\frac{3}{2}}$ and $(1+t^{\alpha-1})^2 \leq 2$, it suffices to show that

$$16 + \frac{4(t-1)^2}{t^{\frac{3}{2}}} \geq \ln^2 t.$$

Let $s = \frac{1}{2}\ln t \in [2, \infty)$, it suffices to show that

$$4 + e^s - 2e^{-s} + e^{-3s} \geq s^2.$$

Since $e^{-s} < 1$, it suffices to verify that

$$2 + e^s \geq s^2,$$

which is trivially true when $s \geq 2$.

Combining the above results, for $\forall \alpha \in [\frac{1}{2}, 2]$, we have that $\Phi_\alpha(p)$ is convex with respect to p for $\forall p \in [0,1]$. □

Proof of generalized version of Lemma 1. We first prove the non-negativity and the monotonicity of the function.

When $\alpha < 1$, to show the non-negativity only suffices to show that

$$n^{1-\alpha} \geq \sum_{i=1}^{n} p_i^\alpha.$$

This is equivalent to

$$1 \geq \sum_{i=1}^{n} \frac{(np_i)^\alpha}{n}.$$

Since $\alpha < 1$, by Power Mean Inequality, we have that

$$\begin{aligned}\left(\sum_{i=1}^{n} \frac{(np_i)^\alpha}{n}\right)^{\frac{1}{\alpha}} &\leq \sum_{i=1}^{n} \frac{np_i}{n} \\ &= 1.\end{aligned}$$

To show the monotonicity, it suffices to show that $\Gamma(t) = \sum_{i=1}^{n} \left(t\left(p_i - \frac{1}{n}\right) + \frac{1}{n}\right)^\alpha$ is monotonically decreasing.

Notice that

$$\begin{aligned}\Gamma'(t) &= \alpha \sum_{i=1}^{n} \left(t\left(p_i - \frac{1}{n}\right) + \frac{1}{n}\right)^{\alpha-1} \left(p_i - \frac{1}{n}\right) \\ \Gamma''(t) &= \alpha(\alpha-1) \sum_{i=1}^{n} \left(t\left(p_i - \frac{1}{n}\right) + \frac{1}{n}\right)^{\alpha-2} \left(p_i - \frac{1}{n}\right)^2.\end{aligned}$$

We have that $\Gamma''(t) \leq 0, \forall t$. Therefore, $\Gamma'(t)$ is monotonically decreasing. Since we also have that $\Gamma'(0) = 0$, we conclude that $\Gamma'(t) \leq 0$ and this implies that $\Gamma(t)$ is monotonically decreasing. Combining all the results, we conclude that for $\alpha < 1$, $\Phi_\alpha(t\mathbf{p} + (1-t)\text{Unif}(n))$ is non-negative and monotonically increasing with respect to $t \in [0,1]$.

When $\alpha > 1$, to show the non-negativity only suffices to show that

$$n^{1-\alpha} \leq \sum_{i=1}^{n} p_i^\alpha.$$

Since $\alpha > 1$, by Power Mean Inequality, we have that

$$\begin{aligned}\left(\sum_{i=1}^{n} \frac{(np_i)^\alpha}{n}\right)^{\frac{1}{\alpha}} &\geq \sum_{i=1}^{n} \frac{np_i}{n} \\ &= 1.\end{aligned}$$

Hence, the non-negativity is true.

To show the monotonicity, it suffices to show that $\Gamma(t) = \sum_{i=1}^{n} \left(t\left(p_i - \frac{1}{n}\right) + \frac{1}{n}\right)^\alpha$ is monotonically increasing.

Notice that

$$\Gamma'(t) = \alpha \sum_{i=1}^{n} \left(t \left(p_i - \frac{1}{n} \right) + \frac{1}{n} \right)^{\alpha-1} \left(p_i - \frac{1}{n} \right)$$
$$\Gamma''(t) = \alpha(\alpha-1) \sum_{i=1}^{n} \left(t \left(p_i - \frac{1}{n} \right) + \frac{1}{n} \right)^{\alpha-2} \left(p_i - \frac{1}{n} \right)^2 .$$

We have that $\Gamma''(t) \geq 0, \forall t$. Therefore, $\Gamma'(t)$ is monotonically increasing. Since we also have that $\Gamma'(0) = 0$, we conclude that $\Gamma'(t) \geq 0$ and this implies that $\Gamma(t)$ is monotonically increasing. Combining all the results, we conclude that for $\alpha > 1$, $\Phi_\alpha(t\mathbf{p} + (1-t)\text{Unif}(n))$ is non-negative and monotonically increasing with respect to $t \in [0, 1]$.

For $1 < \alpha \leq 2$, to show that $\Phi_\alpha(\mathbf{p})$ is jointly convex, it suffices to show $\zeta(\mathbf{p}) = (\sum_{i=1}^n p_i^\alpha)^{\frac{\log_2 \alpha}{\alpha-1}}$ is jointly convex.

$$\frac{\partial^2 \zeta(\mathbf{p})}{\partial p_k \partial p_j} = \alpha^2 \frac{\log_2 \alpha}{\alpha-1} \left(\frac{\log_2 \alpha}{\alpha-1} - 1 \right) \left(\sum_{i=1}^{n} p_i^\alpha \right)^{\frac{\log_2 \alpha}{\alpha-1}-2} p_j^{\alpha-1} p_k^{\alpha-1}, \forall j \neq k$$
$$\frac{\partial^2 \zeta(\mathbf{p})}{\partial p_j^2} = \left(\alpha^2 \frac{\log_2 \alpha}{\alpha-1} \left(\frac{\log_2 \alpha}{\alpha-1} - 1 \right) \left(\sum_{i=1}^{n} p_i^\alpha \right)^{\frac{\log_2 \alpha}{\alpha-1}-2} p_j^{2\alpha-2} + \alpha(\alpha-1) \frac{\log_2 \alpha}{\alpha-1} \left(\sum_{i=1}^{n} p_i^\alpha \right)^{\frac{\log_2 \alpha}{\alpha-1}-1} p_j^{\alpha-2} \right)$$

Notice that this is the sum of a PSD diagonal matrix and an all-one matrix which multiplies a diagonal matrix on the left and right. Therefore, this is PSD and $\Phi_\alpha(\mathbf{p})$ is jointly convex.

When $0 < \alpha < 1$, we require the summation constraint and $p_k \geq 0 \forall k$.

Let $D = \sum_{i=1}^n p_i^\alpha, \phi(\alpha) = \frac{\log_2 \alpha}{\alpha-1}$ and we have that $p_n = 1 - \sum_{i=1}^{n-1} p_i$.

Therefore, we have that

$$\frac{\partial D}{\partial p_k} = \alpha p_k^{\alpha-1} - \alpha p_n^{\alpha-1}, \forall k \neq n,$$
$$\frac{\partial^2 D}{\partial p_k \partial p_l} = \alpha(\alpha-1) p_n^{\alpha-2}, \forall l \neq k, l \neq n, k \neq n,$$
$$\frac{\partial^2 D}{\partial p_k^2} = \alpha(\alpha-1) p_k^{\alpha-2} + \alpha(\alpha-1) p_n^{\alpha-2}, k \neq n,$$
$$\frac{\partial D^{\phi(\alpha)}}{\partial p_k} = \phi(\alpha) D^{\phi(\alpha)-1} \frac{\partial D}{\partial p_k}, \forall k \neq n,$$
$$\frac{\partial^2 D^{\phi(\alpha)}}{\partial p_k \partial p_l} = \phi(\alpha)(\phi(\alpha)-1) D^{\phi(\alpha)-2} \frac{\partial D}{\partial p_k} \frac{\partial D}{\partial p_l} + \phi(\alpha) D^{\phi(\alpha)-1} \frac{\partial^2 D}{\partial p_k \partial p_l}, \forall k \neq n, l \neq n.$$

It suffices to show that

$$\sum_{k,l} \left\{ (\phi(\alpha)-1)\alpha(p_k^{\alpha-1} - p_n^{\alpha-1})(p_l^{\alpha-1} - p_n^{\alpha-1}) + D(\alpha-1)p_n^{\alpha-2} \right\} x_k x_l + \sum_k D(\alpha-1) p_k^{\alpha-2} x_k^2 \leq 0, \forall x_k, x_l.$$

We first verify that the diagonal term is negative.

This is equivalent to

$$D(1-\alpha)(p_k^{\alpha-2} + p_n^{\alpha-2}) \geq (\phi(\alpha)-1)\alpha(p_k^{\alpha-1} - p_n^{\alpha-1})^2.$$

When $p_k = p_n$, the above inequality is true, we verify for $p_k \neq p_n$. It suffices to show that

$$\frac{(p_k^\alpha + p_n^\alpha)(p_k^{\alpha-2} + p_n^{\alpha-2})}{(p_k^{\alpha-1} - p_n^{\alpha-1})^2} \geq \frac{\alpha(\phi(\alpha)-1)}{1-\alpha}.$$

Let $t = \frac{p_k}{p_n}$, the left-hand side is equivalent to $1 + \frac{t^{\alpha-2}(t+1)^2}{(t^{\alpha-1}-1)^2}$ and the derivative with respect to α is $-\frac{(t+1)^2 t^\alpha (t^\alpha + t) \log t}{(t^\alpha - t)^3}$. Notice that this equation is symmetric with respect to p_k and p_n, therefore we can assume that $t < 1$. Let $P(\alpha) = 1 + \frac{t^{\alpha-2}(t+1)^2}{(t^{\alpha-1}-1)^2} - \frac{\alpha(\phi(\alpha)-1)}{1-\alpha}$, we have that

$$P'(\alpha) = -\frac{(t+1)^2 t^\alpha (t^\alpha + t) \log t}{(t^\alpha - t)^3} - \frac{(\alpha+1)\log_2 \alpha + (1-\alpha)(1 + \frac{1}{\ln 2})}{(\alpha-1)^3}$$
$$= -\frac{1}{(t^\alpha - t)^3 (\alpha-1)^3} \left\{ (\alpha-1)^3 (t+1)^2 t^\alpha (t^\alpha + t) \log t + (t^\alpha - t)^3 \left((\alpha+1)\log_2 \alpha + (1-\alpha)(\frac{1}{\ln 2} + 1) \right) \right\}.$$

When $\frac{1}{2} \le \alpha < 1$, we show that $P(\alpha)$ is monotonically increasing. To show this, it suffices to show that

$$-\frac{1}{(t^\alpha - t)^3(\alpha-1)^3} \ge 0$$
$$(t^\alpha - t)^3(1-\alpha)(\frac{1}{\ln 2}+1) \ge 0$$
$$(\alpha-1)^3(t+1)^2 t^\alpha (t^\alpha + t)\log t + (t^\alpha - t)^3(\alpha+1)\log_2 \alpha \ge 0.$$

The first two inequalities are obvious, we show the third inequality.

Notice by 7, we have that

$$\begin{aligned}\frac{(t^\alpha - t)^3}{(\alpha-1)^3 \ln^3 t} &< \left(\frac{t^\alpha + t}{2}\right)^3 \\ &\le t^{2\alpha}\left(\frac{t^\alpha + t}{2}\right) \\ &\le t\left(\frac{t^\alpha + t}{2}\right)\end{aligned}$$

It suffices to show that

$$2(t+1)^2 t^\alpha \ge t\log^2 t(\alpha+1)\log_2 \frac{1}{\alpha}$$

Notice that

$$\begin{aligned}t\log^2 t(\alpha+1)\log_2\frac{1}{\alpha} &\le 2t\log^2 t \\ &\le \frac{8}{e^2} \\ &< 2.\end{aligned}$$

Hence, we showed that the function is monotonically increasing with respect to α. Since when $\alpha = \frac{1}{2}$, we have that $\frac{\alpha(\phi(\alpha)-1)}{1-\alpha} = 1$, we imply that $P(\alpha) > 0$. Therefore, when $\frac{1}{2} \le \alpha < 1$, we have that the diagonal entries are negative.

When $0 < \alpha < \frac{1}{2}$, notice that $-\frac{(t+1)^2 t^\alpha (t^\alpha + t)\log t}{(t^\alpha - t)^3}$ is always positive, which means that $1 + \frac{t^{\alpha-2}(t+1)^2}{(t^{\alpha-1}-1)^2}$ is monotonically increasing with respect to α. Hence, it suffices to show that

$$1 + \frac{(t+1)^2}{(t-1)^2} \ge \frac{\alpha(\phi(\alpha)-1)}{1-\alpha}.$$

It suffices to show that

$$2 \ge \frac{\alpha(\phi(\alpha)-1)}{1-\alpha}.$$

Notice that the right-hand side is monotonically increasing with respect to α.

Therefore, it suffices to show that

$$\begin{aligned}2 &\ge \frac{\frac{1}{2}(\phi(\frac{1}{2})-1)}{1-\frac{1}{2}} \\ &= 1.\end{aligned}$$

The inequality is trivially true.

Combining the results above, we showed that for $0 < \alpha < 1$, all diagonal entries are negative.

Toward the proof of the convexity, we use arguments from linear algebra. We denote the following symbols

$$\begin{aligned}\mathbf{A} &= D(1-\alpha)\mathrm{Diag}(p_k^{\alpha-2}), \\ \mathbf{B} &= D(1-\alpha)p_n^{\alpha-2}\mathbf{1}_{(n-1)\times(n-1)}, \\ \mathbf{v} &= \sqrt{(\phi(\alpha)-1)\alpha}\left(p_1^{\alpha-1} - p_n^{\alpha-1}, \cdots, p_{n-1}^{\alpha-1} - p_n^{\alpha-1}\right)^T.\end{aligned}$$

Hence, it suffices to show that

$$\mathbf{A} + \mathbf{B} - \mathbf{v}\mathbf{v}^T \succeq 0.$$

Here we proved the above results for $\forall k, p_k > 0$. If $p_k = 0$, the convexity goes through by continuity arguments of the probability distribution.

Notice that $\mathbf{A}+\mathbf{B}$ is a PSD matrix of size $(n-1)\times(n-1)$ and $\mathbf{v}\mathbf{v}^T$ is a rank-1 $(n-1)\times(n-1)$ matrix. Applying Weyl's inequality, we have that

$$\begin{aligned}
\lambda_{n-1}(\mathbf{A}+\mathbf{B}-\mathbf{v}\mathbf{v}^T) &\geq \lambda_{n-1}(\mathbf{A}+\mathbf{B})-\lambda_{n-1}(\mathbf{v}\mathbf{v}^T),\\
\lambda_{n-2}(\mathbf{A}+\mathbf{B}-\mathbf{v}\mathbf{v}^T) &\geq \lambda_{n-1}(\mathbf{A}+\mathbf{B})-\lambda_{n-2}(\mathbf{v}\mathbf{v}^T)\\
&= \lambda_{n-1}(\mathbf{A}+\mathbf{B})\\
&\geq \lambda_{n-1}(\mathbf{A})+\lambda_{n-1}(\mathbf{B})\\
&= \lambda_{n-1}(\mathbf{A})\\
&= p_{n-1}^{\alpha-2}\\
&> 0.
\end{aligned}$$

Therefore, the matrix has at most 1 negative eigenvalues and $n-2$ strictly positive eigenvalues, it suffices to show that

$$\det(\mathbf{A}+\mathbf{B}-\mathbf{v}\mathbf{v}^T)\geq 0.$$

Applying the matrix determinant lemma, we have that

$$\det(\mathbf{A}+\mathbf{B}-\mathbf{v}\mathbf{v}^T)=\det(\mathbf{A}+\mathbf{B})(1-\mathbf{v}^T(\mathbf{A}+\mathbf{B})^{-1}\mathbf{v}).$$

Since $\det(\mathbf{A}+\mathbf{B})\geq 0$, it suffices to show that

$$1-\mathbf{v}^T(\mathbf{A}+\mathbf{B})^{-1}\mathbf{v}\geq 0.$$

By Sherman-Morrison formula, we have that

$$(\mathbf{A}+\mathbf{B})^{-1}=\mathbf{A}^{-1}-\frac{\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}}{1+D(1-\alpha)p_n^{\alpha-2}\mathbf{1}_{1\times(n-1)}\mathbf{A}^{-1}\mathbf{1}_{(n-1)\times 1}}.$$

It suffices to show that

$$1-\mathbf{v}^T\mathbf{A}^{-1}\mathbf{v}+\frac{\mathbf{v}^T\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{v}}{1+D(1-\alpha)p_n^{\alpha-2}\mathbf{1}_{1\times(n-1)}\mathbf{A}^{-1}\mathbf{1}_{(n-1)\times 1}}\geq 0.$$

This is equivalent to

$$(1-\mathbf{v}^T\mathbf{A}^{-1}\mathbf{v})(1+D(1-\alpha)p_n^{\alpha-2}\mathbf{1}_{1\times(n-1)}\mathbf{A}^{-1}\mathbf{1}_{(n-1)\times 1})+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{v}\geq 0.$$

Notice that we have

$$\mathbf{A}^{-1}=\frac{\mathrm{Diag}(p_k^{2-\alpha})}{D(1-\alpha)}.$$

We can simplify the above equation to

$$\left(D(1-\alpha)-(\phi(\alpha)-1)\alpha\sum_{k=1}^{n-1}p_k^{2-\alpha}(p_k^{\alpha-1}-p_n^{\alpha-1})^2\right)\left(1+p_n^{\alpha-2}\sum_{k=1}^{n-1}p_k^{2-\alpha}\right)+(\phi(\alpha)-1)\alpha p_n^{\alpha-2}\left(\sum_{k=1}^{n-1}p_k^{2-\alpha}(p_k^{\alpha-1}-p_n^{\alpha-1})\right)^2\geq 0$$

We observe that the inequality can be written in the following form

$$\begin{aligned}
&D(1-\alpha)(1+p_n^{\alpha-2}\sum_{k=1}^{n-1}p_k^{2-\alpha})-(\phi(\alpha)-1)\alpha\sum_{k=1}^{n-1}p_k^{2-\alpha}(p_k^{\alpha-1}-p_n^{\alpha-1})^2\\
&\geq(\phi(\alpha)-1)\alpha p_n^{\alpha-2}\left(\sum_{k=1}^{n-1}p_k^{2-\alpha}\sum_{k=1}^{n-1}p_k^{2-\alpha}(p_k^{\alpha-1}-p_n^{\alpha-1})^2-\left(\sum_{k=1}^{n-1}p_k^{2-\alpha}(p_k^{\alpha-1}-p_n^{\alpha-1})\right)^2\right).
\end{aligned}$$

Observe that by Cauchy inequality, the right-hand side is positive, a necessary condition for the inequality to hold is the left-hand side is positive.

$$D(1-\alpha)(1+p_n^{\alpha-2}\sum_{k=1}^{n-1}p_k^{2-\alpha})-(\phi(\alpha)-1)\alpha\sum_{k=1}^{n-1}p_k^{2-\alpha}(p_k^{\alpha-1}-p_n^{\alpha-1})^2\geq 0.$$

□

B. Proof of Lemma 2

We start with the following lemma.

Lemma 6. *For $t > 1$ and $\alpha \in [\frac{1}{2}, 1)$ we have the following inequalities:*

(i) $\ln t \leq \frac{(t-1)(t^2+10t+1)}{6t(t+1)}$,

(ii) $1 \geq \left(-\frac{\alpha \ln \alpha}{\ln 2} - \alpha(1-\alpha)\right) + \left(\frac{\ln \alpha}{\ln 2} + \frac{1}{\ln 2} + 1 - 2\alpha\right)$,

(iii) $(1+t^\alpha)^2(1-5t^{-1}+5t^{1-\alpha}-t^{-\alpha})$
$\geq (\frac{1-t^{\alpha-1}}{1-\alpha})^2(t^{2-\alpha}-1+5t-5t^{1-\alpha})$,

(iv) $(1+t^\alpha)^2(1-5t^{-1}+5t^{1-\alpha}-t^{-\alpha})$
$\geq \frac{6(1-t^{\alpha-1})}{1-\alpha}(1+t^\alpha)(1+t^{1-\alpha})$.

(i) $\ln t \leq \frac{(t-1)(t^2+10t+1)}{6t(t+1)}$,

Proof. Set $t = \frac{1+z}{1-z}$. As $t > 1$, we have $z \in (0,1)$. Note that

$$\begin{aligned}\ln t &= \ln(1+z) - \ln(1-z) \\ &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1} z^n}{n} + \sum_{n=1}^{\infty} \frac{z^n}{n} \\ &= 2\sum_{k=1}^{\infty} \frac{z^{2k-1}}{2k-1} \\ &\leq 2z + \frac{2}{3} z \sum_{k=1}^{\infty} z^{2k} \\ &= 2z + \frac{2z^3}{3(1-z^2)}.\end{aligned}$$

We plugin $z = \frac{t-1}{t+1}$ and the result follows. □

(ii) $1 \geq \left(-\frac{\alpha \ln \alpha}{\ln 2} - \alpha(1-\alpha)\right) + \left(\frac{\ln \alpha}{\ln 2} + \frac{1}{\ln 2} + 1 - 2\alpha\right)$,

Proof. This is equivalent to

$$\alpha(1-\alpha) + 2\alpha + \frac{-(1-\alpha)\ln \alpha}{\ln 2} \geq \frac{1}{\ln 2}.$$

This can be rewritten as

$$\frac{9}{4} - \left(\frac{3}{2} - \alpha\right)^2 + \frac{-(1-\alpha)\ln \alpha}{\ln 2} \geq \frac{1}{\ln 2}.$$

Note that $\frac{9}{4} - \left(\frac{3}{2} - \alpha\right)^2$ is increasing in the interval $\alpha \in [\frac{1}{2}, 1]$, while $\frac{-(1-\alpha)\ln\alpha}{\ln 2}$ is decreasing and non-negative in the same interval.
For $\alpha \in [\frac{2}{3}, 1]$, the left-hand-side is at least as large as $\frac{9}{4} - \left(\frac{3}{2} - \frac{2}{3}\right)^2 = \frac{14}{9} \geq \frac{1}{\ln 2}$, and hence the desired inequality holds in this interval. For $\alpha \in [\frac{1}{2}, \frac{2}{3}]$, the left-hand-side is at least as large as $\frac{9}{4} - \left(\frac{3}{2} - \frac{1}{2}\right)^2 + \frac{\ln\left(\frac{3}{2}\right)}{3\ln 2} \geq \frac{1}{\ln 2}$. Therefore, the inequality also holds in the interval $\alpha \in [\frac{1}{2}, \frac{2}{3}]$. □

(iii) $(1+t^\alpha)^2(1 - 5t^{-1} + 5t^{1-\alpha} - t^{-\alpha})$
$\geq (\frac{1-t^{\alpha-1}}{1-\alpha})^2(t^{2-\alpha} - 1 + 5t - 5t^{1-\alpha})$,

Proof. This is equivalent to

$$\begin{aligned}&(1+t^\alpha)^2 t(t - 5 + 5t^{2-\alpha} - t^{1-\alpha}) \\ &\quad \geq (\frac{t-t^\alpha}{1-\alpha})^2(t^{2-\alpha} - 1 + 5t - 5t^{1-\alpha}).\end{aligned}$$

This can be rewritten as

$$\begin{aligned}&\left(5(1+t^\alpha)^2 t - (\frac{t-t^\alpha}{1-\alpha})^2\right)(t^{2-\alpha} - 1) \\ &\quad + \left((1+t^\alpha)^2 t - 5(\frac{t-t^\alpha}{1-\alpha})^2\right)(t - t^{1-\alpha}) \geq 0.\end{aligned}$$

Observe that $t^{1-\alpha}(t - t^{1-\alpha}) \leq t^{2-\alpha} - 1$. Divide both sides by $t - t^{1-\alpha}$, it suffices to show that

$$\begin{aligned}&\left(5(1+t^\alpha)^2 t - \left(\frac{t-t^\alpha}{1-\alpha}\right)^2\right) t^{1-\alpha} \\ &\quad + (1+t^\alpha)^2 t - 5\left(\frac{t-t^\alpha}{1-\alpha}\right)^2 \geq 0.\end{aligned}$$

Which is equivalent to showing that

$$(5t^{1-\alpha} + 1)(1+t^\alpha)^2 t \geq (t^{1-\alpha} + 5)\left(\frac{t-t^\alpha}{1-\alpha}\right)^2.$$

Dividing $(1+t^\alpha)^2(t^{1-\alpha}+5)$ on both sides, we need to show that

$$\left(5-\frac{24}{t^{1-\alpha}+5}\right)t \geq \left(\frac{t-t^\alpha}{(1-\alpha)(1+t^\alpha)}\right)^2.$$

By using Remark 7 with $a=1, b=\alpha$, we have that

$$\frac{t-t^\alpha}{(1-\alpha)(1+t^\alpha)} < \frac{\ln t}{2}.$$

Hence, it suffices to show that

$$\left(5-\frac{24}{t^{1-\alpha}+5}\right)t \geq \left(\frac{\ln t}{2}\right)^2.$$

Since $t^{1-\alpha} \geq 1$, it further suffices to show that

$$t \geq \left(\frac{\ln t}{2}\right)^2.$$

This is true as $2\sqrt{t}-\ln t$ increases in t when $t>1$ completing the proof. □

(iv) $(1+t^\alpha)^2(1-5t^{-1}+5t^{1-\alpha}-t^{-\alpha})$
$\geq \frac{6(1-t^{\alpha-1})}{1-\alpha}(1+t^\alpha)(1+t^{1-\alpha})$,

Proof. This is equivalent to showing that

$$\begin{aligned}(5-5\alpha)(t^2-1)+(1-\alpha)(t^{1+\alpha}-t^{1-\alpha})\\ -(5\alpha+1)(t^{2-\alpha}-t^\alpha)\geq 0.\end{aligned} \tag{8}$$

Denote $g(\mu)=\frac{t^{1+\mu}-t^{1-\mu}}{\mu}$, for $\mu \neq 0$, and $g(0)=\lim_{\mu\to 0} g(\mu) = 2t\ln t$. Note that $g(\mu)$ is non-negative, $g(-\mu)=g(\mu)$ and

$$g'(\mu)=\frac{t^{1-\mu}}{\mu^2}(-t^{2\mu}+\mu(t^{2\mu}+1)\ln t+1).$$

When $x \geq 1$, we have that $\ln x \geq \frac{x^2-1}{x^2+1}$. Therefore $g'(\mu)\geq 0$, when $\mu \geq 0$.
Note that inequality (8) is equivalent to

$$\begin{aligned}(5-5\alpha)g(1)+(1-\alpha)\alpha g(\alpha)\\ -(5\alpha+1)(1-\alpha)g(1-\alpha)\geq 0.\end{aligned}$$

Canceling out $(1-\alpha)$ (as $\alpha \leq 1$) we need to show that

$$5g(1)+\alpha g(\alpha)-(5\alpha+1)g(1-\alpha)\geq 0.$$

Since $g(x)$ is increasing when $x \geq 0$ and $\alpha \in [\frac{1}{2}, 1]$, we have $g(1)\geq g(\alpha)\geq g(1-\alpha)\geq 0$. Hence the inequality will follow if we show that

$$5+\alpha-(5\alpha+1)\geq 0.$$

However, this is immediate when $\alpha \leq 1$, and completes the proof. □

Proof of Lemma 2. By taking derivatives with respect to p, the above convexity is equivalent to showing that for $p \in (0, \frac{1}{2})$

$$\Psi''_\beta(p)\Psi'_\alpha(p)-\Psi'_\beta(p)\Psi''_\alpha(p)\leq 0. \tag{9}$$

This is equivalent to showing that $(\ln \Psi'_\alpha(p))'$ is monotonically increasing with respect to α. Let $U(\alpha) := \frac{\log_2 \alpha}{1-\alpha}$ and $t=\frac{1-p}{p}$. Note that $t>1$.
$\frac{d}{d\alpha}(\ln \Psi'_\alpha(p))' \geq 0$ is equivalent to verifying

$$\begin{aligned}&(U(\alpha)+\alpha U'(\alpha)-1)\frac{(p^{\alpha-1}-(1-p)^{\alpha-1})}{(p^\alpha+(1-p)^\alpha)}\\ &\quad+\frac{(p^{\alpha-2}+(1-p)^{\alpha-2})}{(p^{\alpha-1}-(1-p)^{\alpha-1})}\\ &\geq \Bigg((1-\alpha)\frac{p^{\alpha-2}(1-p)^{\alpha-2}}{(p^{\alpha-1}-(1-p)^{\alpha-1})^2}\\ &\quad+(\alpha U(\alpha)-\alpha)\frac{p^{\alpha-1}(1-p)^{\alpha-1}}{(p^\alpha+(1-p)^\alpha)^2}\Bigg)\ln\left(\frac{1-p}{p}\right).\end{aligned}$$

This can rewritten as requiring to verify

$$(U(\alpha)+\alpha U'(\alpha)-1)\frac{(1-t^{\alpha-1})}{(1+t^{\alpha})}+\frac{(1+t^{\alpha-2})}{(1-t^{\alpha-1})} \tag{10}$$
$$\geq (1+t)\ln(t)$$
$$\times\left((1-\alpha)\frac{t^{\alpha-2}}{(1-t^{\alpha-1})^2}+(\alpha U(\alpha)-\alpha)\frac{t^{\alpha-1}}{(1+t^{\alpha})^2}\right).$$

Part (i) of 6 yields that

$$\ln t^{1-\alpha}\leq\frac{(t^{1-\alpha}-1)(t^{2-2\alpha}+10t^{1-\alpha}+1)}{6t^{1-\alpha}(t^{1-\alpha}+1)}.$$

Therefore, it suffices to show that

$$(U(\alpha)+\alpha U'(\alpha)-1)(1-t^{\alpha-1})^3(1+t^{\alpha})$$
$$+(1-t^{\alpha-1})(1+t^{\alpha})^2(1+t^{\alpha-2})$$
$$\geq\left((1-\alpha)t^{\alpha-2}(1+t^{\alpha})^2+(\alpha U(\alpha)-\alpha)t^{\alpha-1}(1-t^{\alpha-1})^2\right)$$
$$\times(1+t)\frac{(t^{1-\alpha}-1)(t^{2-2\alpha}+10t^{1-\alpha}+1)}{6t^{1-\alpha}(t^{1-\alpha}+1)(1-\alpha)}.$$

Substituting for $U(\alpha)$ and rearranging, this is equivalent to showing that

$$(1+t^{\alpha})^2(1-5t^{-1}+5t^{1-\alpha}-t^{-\alpha})$$
$$\geq\left(-\frac{\alpha\ln\alpha}{\ln 2}-\alpha(1-\alpha)\right)\left(\frac{1-t^{\alpha-1}}{1-\alpha}\right)^2$$
$$\times(t^{2-\alpha}-1+5t-5t^{1-\alpha})$$
$$+\left(\frac{\ln\alpha}{\ln 2}+\frac{1}{\ln 2}+1-2\alpha\right)\frac{6(1-t^{\alpha-1})}{1-\alpha}$$
$$\times(1+t^{\alpha})(1+t^{1-\alpha}).$$

Note that $-\frac{\ln\alpha}{\ln 2}\geq(1-\alpha)$ for $\alpha\in[\frac{1}{2},1]$, and therefore it suffices to show that the following three inequalities hold:

$$1\geq\left(-\frac{\alpha\ln\alpha}{\ln 2}-\alpha(1-\alpha)\right)+\left(\frac{\ln\alpha}{\ln 2}+\frac{1}{\ln 2}+1-2\alpha\right),$$
$$(1+t^{\alpha})^2(1-5t^{-1}+5t^{1-\alpha}-t^{-\alpha})$$
$$\geq\left(\frac{1-t^{\alpha-1}}{1-\alpha}\right)^2(t^{2-\alpha}-1+5t-5t^{1-\alpha}),$$
$$(1+t^{\alpha})^2(1-5t^{-1}+5t^{1-\alpha}-t^{-\alpha})$$
$$\geq\frac{6(1-t^{\alpha-1})}{1-\alpha}(1+t^{\alpha})(1+t^{1-\alpha}).$$

These are established in Lemma 6 and this completes the proof. □

C. Proof of Lemma 4

As before we begin with a similar lemma lemma.

Lemma 7. *Let* $U(\alpha):=\frac{\log_2\alpha}{1-\alpha}$, $t>1$, *and* $\alpha\in(1,2]$. *The following inequalities hold:*

(i) $(1-U(\alpha)-\alpha U'(\alpha))(\alpha-1)+2-\alpha\geq(\alpha U(\alpha)-\alpha)$,

(ii) $t^{\frac{\alpha}{2}}(1+t^{\alpha})\geq t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2}$,

(iii) $\frac{(1+t^{\alpha})^2}{2}\geq t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2}$.

(i) $(1-U(\alpha)-\alpha U'(\alpha))(\alpha-1)+2-\alpha\geq(\alpha U(\alpha)-\alpha)$,

Proof. Expanding and simplifying this is equivalent to verifying that

$$1+\alpha+\alpha(1-\alpha)U'(\alpha)\geq U(\alpha)(2\alpha-1).$$

Since $(1-\alpha)U(\alpha)=\log_2\alpha$, differentiating both sides we obtain

$$(1-\alpha)U'(\alpha)-U(\alpha)=\frac{1}{\alpha\ln 2}.$$

Substituting this above, we need to show that

$$1+\alpha+\alpha U(\alpha)+\frac{1}{\ln 2}\geq U(\alpha)(2\alpha-1),$$

or equivalently

$$1+\alpha+\frac{1}{\ln 2}\geq U(\alpha)(\alpha-1).$$

Note that $U(\alpha)<0$ for $\alpha>1$, and the left-hand-side is positive; so we are done. □

(ii) $t^{\frac{\alpha}{2}}(1+t^{\alpha})\geq t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2}$,

Proof. This is equivalent to

$$2t^{\alpha}+2\geq t^{\alpha}+t+t^{\alpha-1}+1,$$

and can be simplified to

$$(t-1)(t^{\alpha-1}-1)\geq 0.$$

This holds when $t\geq 1$ and $\alpha\geq 1$, as is the case here. □

(iii) $\frac{(1+t^{\alpha})^2}{2}\geq t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2}$,

Proof. This is equivalent to

$$t^{2\alpha}+2t^{\alpha}+1-t^{\frac{\alpha}{2}}-t^{\frac{3\alpha}{2}-1}-t^{\frac{3\alpha}{2}}-t^{\frac{\alpha}{2}+1}\geq 0.$$

This can be rewritten as

$$(t^{\frac{\alpha}{2}+1}-1)(t^{\frac{3\alpha}{2}}-1)\geq(t^{\frac{3\alpha}{4}}-t^{\frac{\alpha}{4}})^2.$$

Since $t^{\frac{\alpha}{2}+1}\geq t^{\alpha}$ as $t>1$ and $\alpha\in(1,2]$, it suffices to show that

$$(t^{\alpha}-1)(t^{\frac{3\alpha}{2}}-1)\geq(t^{\frac{3\alpha}{4}}-t^{\frac{\alpha}{4}})^2.$$

Canceling $(t^{\frac{\alpha}{2}}-1)$ from both sides, this becomes equivalent to

$$(t^{\frac{\alpha}{2}}+1)(t^{\frac{3\alpha}{2}}-1)\geq t^{\frac{\alpha}{2}}(t^{\frac{\alpha}{2}}-1).$$

Expanding, we need to show that

$$t^{2\alpha}+t^{\frac{3\alpha}{2}}\geq t^{\alpha}+1.$$

However, this is immediate as $t^{2\alpha}\geq t^{\frac{3\alpha}{2}}\geq t^{\alpha}\geq 1$, as $t>1$ and $\alpha>1$. □

Proof of Lemma 4. As in the proof of Lemma 2, with $U(\alpha)$ and t as defined earlier, convexity will follow (see 10) if we show that the following inequality holds:

$$\begin{aligned}&(U(\alpha)+\alpha U'(\alpha)-1)\frac{(1-t^{\alpha-1})}{(1+t^{\alpha})}+\frac{(1+t^{\alpha-2})}{(1-t^{\alpha-1})}\\&\quad\geq(1+t)\ln t\\&\qquad\times\left((1-\alpha)\frac{t^{\alpha-2}}{(1-t^{\alpha-1})^2}+(\alpha U(\alpha)-\alpha)\frac{t^{\alpha-1}}{(1+t^{\alpha})^2}\right).\end{aligned}$$

This can be rewritten as required to show that

$$\begin{aligned}&(U(\alpha)+\alpha U'(\alpha)-1)(1-t^{\alpha-1})^3(1+t^{\alpha})\\&\quad+(\alpha-1)t^{\alpha-2}(1+t^{\alpha})^2(1+t)\ln t\\&\geq(t^{\alpha-1}-1)(1+t^{\alpha})^2(1+t^{\alpha-2})\\&\quad+(\alpha U(\alpha)-\alpha)t^{\alpha-1}(1-t^{\alpha-1})^2(1+t)\ln t.\end{aligned}$$

Applying Remark 7 with $a=\alpha-1$ and $b=0$ yields that $t^{\alpha-1}-1\leq\frac{(\alpha-1)(t^{\alpha-1}+1)\ln t}{2}$. Using this estimate for the first term on the right-hand-side, we see that it suffices to show that

$$\begin{aligned}&(U(\alpha)+\alpha U'(\alpha)-1)(1-t^{\alpha-1})^3(1+t^{\alpha})\\&\quad+(\alpha-1)t^{\alpha-2}(1+t^{\alpha})^2(1+t)\ln t\\&\geq\frac{(\alpha-1)(t^{\alpha-1}+1)\ln t}{2}(1+t^{\alpha})^2(1+t^{\alpha-2})\\&\quad+(\alpha U(\alpha)-\alpha)t^{\alpha-1}(1-t^{\alpha-1})^2(1+t)\ln t.\end{aligned}$$

This can be rewritten as required to show

$$\begin{aligned}&(1-U(\alpha)-\alpha U'(\alpha))(t^{\alpha-1}-1)^2(1+t^{\alpha})\\&\quad+\frac{(\alpha-1)\ln t(1+t^{\alpha})^2}{2}(1-t^{\alpha-2})\\&\qquad\geq(\alpha U(\alpha)-\alpha)t^{\alpha-1}(t^{\alpha-1}-1)(1+t)\ln t.\end{aligned}$$

Now, we apply Remark 7 to obtain

$$\begin{aligned}&(t^{\alpha-1}-1)^2>(\alpha-1)^2t^{\alpha-1}(\ln t)^2\\&\qquad(\text{set } a=\alpha-1, b=0),\\&1-t^{\alpha-2}>t^{\frac{\alpha-2}{2}}(2-\alpha)\ln t\\&\qquad(\text{set } a=0, b=\alpha-2),\\&t^{\alpha-1}-1<\frac{(\alpha-1)(t^{\alpha-1}+1)\ln t}{2}\\&\qquad(\text{set } a=\alpha-1, b=0).\end{aligned}$$

Using the above inequalities, it suffices to show that

$$\begin{aligned}&(1-U(\alpha)-\alpha U'(\alpha))(\alpha-1)^2t^{\alpha-1}(\ln t)^2(1+t^{\alpha})\\&\quad+\frac{(\alpha-1)(1+t^{\alpha})^2}{2}(2-\alpha)t^{\frac{\alpha-2}{2}}(\ln t)^2\\&\qquad\geq(\alpha U(\alpha)-\alpha)t^{\alpha-1}(1+t)\frac{(\alpha-1)(t^{\alpha-1}+1)(\ln t)^2}{2}.\end{aligned}$$

This is equivalent to showing that

$$\begin{aligned}&(1-U(\alpha)-\alpha U'(\alpha))(\alpha-1)t^{\frac{\alpha}{2}}(1+t^{\alpha})+\frac{(1+t^{\alpha})^2}{2}(2-\alpha)\\&\quad\geq(\alpha U(\alpha)-\alpha)t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2}.\end{aligned}$$

It is easy to verify that $1-U(\alpha)-\alpha U'(\alpha)\geq 0$ for $1<\alpha\leq 2$. Therefore the above inequality follows as long as the following three inequalities hold:

$$\begin{aligned}&(1-U(\alpha)-\alpha U'(\alpha))(\alpha-1)+2-\alpha\geq(\alpha U(\alpha)-\alpha),\\&t^{\frac{\alpha}{2}}(1+t^{\alpha})\geq t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2},\\&\frac{(1+t^{\alpha})^2}{2}\geq t^{\frac{\alpha}{2}}(1+t)\frac{(t^{\alpha-1}+1)}{2}.\end{aligned}$$

These are established in Lemma 7 and this completes the proof. □