A Differential Equation Approach to the Most-Informative Boolean Function Conjecture

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

Abstract

We study the most-informative Boolean function conjecture using a differential equation approach. This leads to a formulation of a functional inequality on finite-dimensional random variables. We also develop a similar inequality in the case of the Hellinger conjecture. Finally, we conjecture a specific finite dimensional inequality that, if proved, will lead to a proof of the Boolean function conjecture in the balanced case.

I. Introduction

Let $\mathbb{H}^n = \{-1,1\}^n$ denote the n-dimensional Boolean Hypercube centered at $\mathbf{0}$. The most informative Boolean function conjecture states the following:

Conjecture 1 ([1])**.** *Let $\mathbf{X} \sim \text{Unif}(\mathbb{H}^n)$ be a random variable distributed uniformly on the Boolean Hypercube and $\mathbf{Y}$ be received after passing each bit of $\mathbf{X}$ through a BSC channel with cross-over probability $p = \frac{1-\rho}{2} \in [0,1]$. Let $f : \mathbb{H}^n \mapsto \mathbb{H}$ be a Boolean function that maps a Boolean sequence to a binary value. Then the following inequality holds:*

$$I(f(\mathbf{X});\mathbf{Y}) \leq 1 - H_2(p),$$

where $H_2(p) = -p\log_2(p) - (1-p)\log_2(1-p)$ is the binary entropy function.

A weaker form of the conjecture was studied in [2]. Samorodnitsky [3] demonstrated the existence of a positive constant ρ_0 such that the conjecture is true for balanced Boolean functions when $|\rho| \leq \rho_0$. This result was subsequently strengthened by Lei Yu [4], who confirmed that the conjecture holds for balanced Boolean functions when $|\rho| \leq 0.44$. Li and Médard [5] investigated the Boolean function that maximizes $\mathbb{E}[|T_\rho f|^\alpha]$ for a fixed mean, where $\alpha \in [1,2]$. They conjectured that the dictator function is optimal. Subsequently, Barnes and Ozgur in [6] showed that the conjecture is related to solving the α-Non-Interactive Correlation Distillation (NICD) problem raised by Li and Médard in [5]. All of the previous approaches view the problem as a given instance, with parameters given by the channel transition probability and the mean of the Boolean function. In this work, we traverse a path in the space of channels and analyze our objective function along this trajectory.

It is also known that Conjecture 1 is implied by the following stronger conjecture, which we call the "Hellinger Conjecture":

Conjecture 2 ([7])**.** *Under the same assumptions as in Conjecture 1, the following holds:*

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

where $T_\rho f(\boldsymbol{y}) = \mathbb{E}[f(\mathbf{X})|\mathbf{Y} = \boldsymbol{y}]$.

Recently, [8] found a parametric class of conjectures that interpolates between Conjecture 1 and Conjecture 2. Conjecture 2 implies a conjecture about the sensitivity of Boolean functions. For a Boolean function f, the sensitivity at a point $\boldsymbol{x}$, represented as $s_f(\boldsymbol{x})$, is defined as the count of $\boldsymbol{x}$'s neighbors for which the function produces a value contrary to $f(\boldsymbol{x})$. Isoperimetric inequalities can help set limits on this sensitivity. For example, the Talagrand isoperimetric inequality demonstrates that for any balanced function, the following holds:

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

Conjecture 2 implies a strengthened version of this inequality (still unproved) as follows: for any balanced functions, we have

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

The sensitivity of Boolean functions has received attention in the math community; see [9]–[12] for some related results on the sensitivity of Boolean functions. Building on the work in [11], it was shown very recently that for all balanced functions

$$\mathbb{E}\left[s_f^\beta(\mathbf{X})\right] \geq 1,$$

for all $\beta \geq 0.50057$.

In this paper, we introduce a differential equation approach to investigate the two conjectures mentioned earlier. We consider a path comprising Binary Symmetric Channels (BSC) with a crossover probability that evolves over time, resulting in a channel output denoted as $\mathbf{Y}_t$. We compute the derivative of our objective function, which represents the mutual information between $f(\mathbf{X})$ and the output $\mathbf{Y}_t$ along the path. By establishing bounds on the derivative, we derive new constraints on the endpoints of this path. Observe that this approach is akin to the auxiliary receiver approach in [13] but uses a continuum of auxiliary receivers instead of just one auxiliary receiver, i.e., BSC channels whose crossover probability is smaller than that of the given channel.

Assume that $\mathbf{Y}_t$ and $\mathbf{Y}_{t+\epsilon}$ are two channel outputs at times t and $t+\epsilon$. Instead of single-letterizing the difference between mutual information terms $I(f(\mathbf{X});\mathbf{Y}_t) - I(f(\mathbf{X});\mathbf{Y}_{t+\epsilon})$ using the past/future of $\mathbf{Y}_t$ and $\mathbf{Y}_{t+\epsilon}$, we take the derivative of $I(f(\mathbf{X});\mathbf{Y}_t)$ with respect to t and then combine the derivative with a natural induction technique on the dimension of the hypercube. This process leads to a functional inequality (on finite dimensions) whose solutions provide requisite lower bounds to the quantity of interest. In particular, backed by numerical simulations, we also conjecture that a particular function satisfies the functional inequality induced from Conjecture 1, which, if true, would establish it for balanced functions (and perhaps more).

This paper is organized as follows: Section II describes the differential equation framework through the example of the most informative Boolean function conjecture. Section III applies the framework to the Hellinger conjecture. All proofs are given in Section V.

Notation: We use the following notation in this paper. We use the bold letter $\mathbf{X}$ to denote the vector of random variables $\mathbf{X} = (X_1, X_2, \cdots, X_n)$. We use uppercase letter to denote random variables while their values are depicted in lowercase letters. Let

$$H_2(x) = x\log_2\frac{1}{x} + (1-x)\log_2\frac{1}{1-x}$$

be the binary entropy function,

$$J(x) = \log_2\frac{1-x}{x}$$

be the derivative of $H_2(x)$. The inverse function H_2^{-1} will map $[0,1]$ to $[0,\frac{1}{2}]$. Let $D_2(x\|y) = x\log_2\frac{x}{y} + (1-x)\log_2\frac{1-x}{1-y}$ be the binary KL-divergence between the distributions $(x, 1-x)$ and $(y, 1-y)$.

II. The general framework of the differential equation approach

Consider the setting in Conjecture 1, and let

$$p_t = \frac{1-e^{-2t}}{2}$$

be the crossover probability for some $t \in [0,\infty)$, i.e., $\rho_t = e^{-2t}$. Let

$$\mathbf{Y}_t = (Y_{t,1}, Y_{t,2}, \cdots, Y_{t,n})$$

be the output of the BSC channel with crossover probability p_t. Take some arbitrary $P_{F|\mathbf{X}}$ where $F \in \{-1,1\}$ is a binary random variable such that $P_{F|\mathbf{X}}(F=-1|\mathbf{X}=\boldsymbol{x}) \in (0,1)$ for all $\boldsymbol{x}$. Note that we are excluding the case of F being a function of $\mathbf{X}$, but the function case can be considered to be a limiting case when the probabilities tend to 0 and 1. Let $P_{F,\mathbf{X},\mathbf{Y}_t} = P_{F|\mathbf{X}}P_{\mathbf{X},\mathbf{Y}_t}$. Without loss of generality, we can assume that $F \to \mathbf{X} \to \mathbf{Y}_{t_1} \to \mathbf{Y}_{t_2}$ holds for any $t_1 < t_2$. Define

$$\gamma(t) = H(F|\mathbf{Y}_t).$$

Note that $\gamma(0) = H(F|\mathbf{X})$, $\gamma(\infty) = H(F)$, $I(F;\mathbf{X}) = \gamma(\infty) - \gamma(0)$ and $I(F;\mathbf{Y}_t) = \gamma(\infty) - \gamma(t)$. The definition shows that $\gamma(t) = H(F) - I(F;\mathbf{Y}_t)$ is an increasing function in t because of the data processing inequality. The following lemma (whose proof is given in Section V) provides the exact derivative.

Lemma 1.

$$\frac{d\gamma(t)}{dt} = \frac{1}{2^n}\sum_{\boldsymbol{x}\sim\boldsymbol{y}}(v_{\boldsymbol{x}}(t) - v_{\boldsymbol{y}}(t))(J(v_{\boldsymbol{y}}(t)) - J(v_{\boldsymbol{x}}(t))),$$

where $v_{\boldsymbol{x}}(t) = \Pr(F=-1|\mathbf{Y}_t = \boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{H}^n$. Here $\boldsymbol{x} \sim \boldsymbol{y}$ stands for the Hamming distance $d_H(\boldsymbol{x},\boldsymbol{y}) = 1$ and the tuple $(\boldsymbol{x},\boldsymbol{y}), (\boldsymbol{y},\boldsymbol{x})$ are only counted once in the summation.

Proof. We have

$$\gamma(t) = H(F|\mathbf{Y}_t) = \mathbb{E}\left[H(F|\mathbf{Y}_t = \boldsymbol{y})\right] = \mathbb{E}\left[H_2(v_{\boldsymbol{y}}(t))\right] = \frac{1}{2^n}\sum_{\boldsymbol{y}} H_2(v_{\boldsymbol{y}}(t)).$$

This implies that

$$\begin{aligned}
\frac{d\gamma(t)}{dt} &= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\frac{dv_{\boldsymbol{y}}(t)}{dt}\\
&= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\lim_{\varepsilon\downarrow 0}\frac{v_{\boldsymbol{y}}(t+\varepsilon)-v_{\boldsymbol{y}}(t)}{\varepsilon}\\
&= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\lim_{\varepsilon\downarrow 0}\frac{\sum_{\boldsymbol{x}} P(F=-1,\mathbf{Y}_t=\boldsymbol{x}|\mathbf{Y}_{t+\varepsilon}=\boldsymbol{y})-v_{\boldsymbol{y}}(t)}{\varepsilon}\\
&= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\lim_{\varepsilon\downarrow 0}\frac{\sum_{\boldsymbol{x}} v_{\boldsymbol{x}}(t)P(\mathbf{Y}_t=\boldsymbol{x}|\mathbf{Y}_{t+\varepsilon}=\boldsymbol{y})-v_{\boldsymbol{y}}(t)}{\varepsilon}\\
&= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\lim_{\varepsilon\downarrow 0}\frac{\sum_{\boldsymbol{x}} v_{\boldsymbol{x}}(t)p_\varepsilon^{d_H(\boldsymbol{x},\boldsymbol{y})}\left(1-p_\varepsilon\right)^{n-d_H(\boldsymbol{x},\boldsymbol{y})}-v_{\boldsymbol{y}}(t)}{\varepsilon}\\
&= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\lim_{\varepsilon\downarrow 0}\frac{\sum\limits_{\boldsymbol{x}\sim\boldsymbol{y}}(v_{\boldsymbol{x}}(t)-v_{\boldsymbol{y}}(t))p_\varepsilon\left(1-p_\varepsilon\right)^{n-1}}{\varepsilon}+o(\varepsilon)\\
&= \frac{1}{2^n}\sum_{\boldsymbol{y}} J(v_{\boldsymbol{y}}(t))\sum_{\boldsymbol{x}\sim\boldsymbol{y}}(v_{\boldsymbol{x}}(t)-v_{\boldsymbol{y}}(t))\\
&= \frac{1}{2^n}\sum_{\boldsymbol{x}\sim\boldsymbol{y}}(J(v_{\boldsymbol{y}}(t))-J(v_{\boldsymbol{x}}(t)))(v_{\boldsymbol{x}}(t)-v_{\boldsymbol{y}}(t)).
\end{aligned}$$

□

Remark 1. *Note that*

$$(u-w)(J(w)-J(u)) = D_2(u\|w)+D_2(w\|u)$$

where D_2 is the binary KL divergence.

Definition 1. *Let Ψ be the class of all non-negative functions $\psi(a,b):(0,1)^2\mapsto\mathbb{R}$ that satisfy the following two conditions:*

- *$\psi(a,b)=0$ when $H_2(a)\le b$.*
- *$\psi(1-a,b)=\psi(a,b)$.*
- *Let P_X be an arbitrary distribution on $\mathcal{X}=\{1,2,3,4,5\}$, and $(u_x,w_x)\in(0,1)^2$ for $x\in\mathcal{X}$ be arbitrary. Then, the following holds:*

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}_X[(u_X-w_X)(J(w_X)-J(u_X))]\\
&\quad\ge\psi\left(\frac{\mathbb{E}\left[u_X+w_X\right]}{2},\frac{\mathbb{E}\left[H_2(u_X)+H_2(w_X)\right]}{2}\right)-\frac{\psi(\mathbb{E}\left[u_X\right],\mathbb{E}\left[H_2(u_X)\right])+\psi(\mathbb{E}\left[w_X\right],\mathbb{E}\left[H_2(w_X)\right])}{2}.
\end{aligned}\tag{2}$$

The following remark follows from Caratheodery's theorem. See Section V for details.

Remark 2. *Instead of imposing the third condition in Definition 1, i.e.* (2)*, for $\mathcal{X}$ of cardinality size* 5*, we can equivalently require it for any P_X on a set $\mathcal{X}$ of arbitrary size.*

Theorem 2. *For every $\psi\in\Psi$ we have*

$$\frac{d\gamma(t)}{dt}\ge\psi(\Pr[F=-1],\gamma(t)).\tag{3}$$

Consequently, if we let

$$g(x)=\int_{\gamma(0)}^{x}\frac{du}{\psi(\Pr[F=-1],u)},\qquad\forall x\ge\gamma(0).$$

we obtain

$$\gamma(t)\ge g^{-1}(t).$$

Proof. Take some arbitrary function $\psi(a,b)$. We would like to show that

$$\frac{d\gamma(t)}{dt}\ge\psi(\Pr(F=-1),\gamma(t)).$$

We prove the statement by induction on n. For the base case of $n=1$, we have

$$\begin{aligned}\gamma(t) &= \frac{1}{2}H(v_1(t)) + \frac{1}{2}H(v_{-1}(t)) \\ \Pr(F=-1) &= \frac{1}{2}v_1(t) + \frac{1}{2}v_{-1}(t), \\ \frac{d\gamma(t)}{dt} &= \frac{1}{2}(J(v_1(t)) - J(v_{-1}(t)))(v_{-1}(t) - v_1(t)).\end{aligned}$$

Thus, we need to show that

$$\frac{1}{2}(J(v_1(t)) - J(v_{-1}(t)))(v_{-1}(t) - v_1(t)) \geq \psi\left(\frac{1}{2}v_1(t) + \frac{1}{2}v_{-1}(t), \frac{1}{2}H(v_1(t)) + \frac{1}{2}H(v_{-1}(t))\right).$$

By the first property of ψ, we have

$$\psi\left(v_i(t), H(v_i(t))\right) = 0, \qquad i \in \{1,-1\}$$

and we can rewrite the above inequality as

$$\begin{aligned}&\frac{1}{2}(J(v_1(t)) - J(v_{-1}(t)))(v_{-1}(t) - v_1(t)) \\ &\quad \geq \psi\left(\frac{1}{2}v_1(t) + \frac{1}{2}v_{-1}(t), \frac{1}{2}H(v_1(t)) + \frac{1}{2}H(v_{-1}(t))\right) - \frac{1}{2}\psi\left(v_1(t), H(v_1(t))\right) - \frac{1}{2}\psi\left(v_{-1}(t), H(v_{-1}(t))\right).\end{aligned}$$

The induction basis is established.

Next, assume that the desired inequality holds for $n-1$. We show it for n. Define $\tilde{\mathbf{Y}}_t = (Y_{t,1}, Y_{t,2}, \cdots, Y_{t,n-1})$ to be the subsequence of the first $n-1$ random variables in $\mathbf{Y}_t$. Let

$$\begin{aligned}v^+_{\tilde{\boldsymbol{y}}}(t) = v_{(\tilde{\boldsymbol{y}},1)}(t) = \Pr(F=-1|\tilde{\mathbf{Y}}_t = \tilde{\boldsymbol{y}}, Y_{t,n} = 1) \\ v^-_{\tilde{\boldsymbol{y}}}(t) = v_{(\tilde{\boldsymbol{y}},-1)}(t) = \Pr(F=-1|\tilde{\mathbf{Y}}_t = \tilde{\boldsymbol{y}}, Y_{t,n} = -1)\end{aligned}$$

Let F_+, F_- be two binary random variables, jointly distributed with with $\tilde{\mathbf{Y}}_t$ according to conditional laws

$$\Pr(F_+ = -1|\tilde{\mathbf{Y}}_t = \tilde{\boldsymbol{y}}) = v^+_{\tilde{\boldsymbol{y}}}(t)$$

and

$$\Pr(F_- = -1|\tilde{\mathbf{Y}}_t = \tilde{\boldsymbol{y}}) = v^-_{\tilde{\boldsymbol{y}}}(t)$$

respectively. Letting $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \in \mathbb{H}^{n-1}$ be the first $n-1$ digits of $\boldsymbol{x}, \boldsymbol{y}$ on the $(n-1)$-dimensional subcube, we can write

$$\begin{aligned}\frac{d\gamma(t)}{dt} &= \frac{1}{2^n}\sum_{\boldsymbol{x}\sim\boldsymbol{y}}(J(v_{\boldsymbol{y}}(t)) - J(v_{\boldsymbol{x}}(t)))(v_{\boldsymbol{x}}(t) - v_{\boldsymbol{y}}(t)) \\ &= \frac{1}{2^n}\sum_{\tilde{\boldsymbol{x}}}(J(v^+_{\tilde{\boldsymbol{x}}}(t)) - J(v^-_{\tilde{\boldsymbol{x}}}(t)))(v^-_{\tilde{\boldsymbol{x}}}(t) - v^+_{\tilde{\boldsymbol{x}}}(t)) \\ &\quad + \frac{1}{2^n}\sum_{\tilde{\boldsymbol{x}}\sim\tilde{\boldsymbol{y}}}(J(v^-_{\tilde{\boldsymbol{y}}}(t)) - J(v^-_{\tilde{\boldsymbol{x}}}(t)))(v^-_{\tilde{\boldsymbol{x}}}(t) - v^-_{\tilde{\boldsymbol{y}}}(t)) + \frac{1}{2^n}\sum_{\tilde{\boldsymbol{x}}\sim\tilde{\boldsymbol{y}}}(J(v^+_{\tilde{\boldsymbol{y}}}(t)) - J(v^+_{\tilde{\boldsymbol{x}}}(t)))(v^+_{\tilde{\boldsymbol{x}}}(t) - v^+_{\tilde{\boldsymbol{y}}}(t)) \\ &\geq \frac{1}{2^n}\sum_{\tilde{\boldsymbol{x}}}(J(v^+_{\tilde{\boldsymbol{x}}}(t)) - J(v^-_{\tilde{\boldsymbol{x}}}(t)))(v^-_{\tilde{\boldsymbol{x}}}(t) - v^+_{\tilde{\boldsymbol{x}}}(t)) + \frac{1}{2}\psi\left(\mathbb{E}\left[v^+_{\tilde{\mathbf{Y}}}\right], \mathbb{E}\left[H_2\left(v^+_{\tilde{\mathbf{Y}}}\right)\right]\right) + \frac{1}{2}\psi\left(\mathbb{E}\left[v^-_{\tilde{\mathbf{Y}}}\right], \mathbb{E}\left[H_2\left(v^-_{\tilde{\mathbf{Y}}}\right)\right]\right)\end{aligned} \tag{4}$$

where the last step follows from the induction hypothesis. Next, from the last property of ψ for the choice of $X = \tilde{\mathbf{Y}}$, $u_X = v^+_{\tilde{\mathbf{Y}}}$ and $w_X = v^-_{\tilde{\mathbf{Y}}}$ we obtain

$$\begin{aligned}&\frac{1}{2^n}\sum_{\tilde{\boldsymbol{x}}}(J(v^+_{\tilde{\boldsymbol{x}}}(t)) - J(v^-_{\tilde{\boldsymbol{x}}}(t)))(v^-_{\tilde{\boldsymbol{x}}}(t) - v^+_{\tilde{\boldsymbol{x}}}(t)) \\ &\quad \geq \psi\left(\frac{\mathbb{E}\left[v^+_{\tilde{\mathbf{Y}}} + v^-_{\tilde{\mathbf{Y}}}\right]}{2}, \frac{\mathbb{E}\left[H_2\left(v^-_{\tilde{\mathbf{Y}}}\right) + H_2\left(v^+_{\tilde{\mathbf{Y}}}\right)\right]}{2}\right) - \frac{1}{2}\psi\left(\mathbb{E}\left[v^+_{\tilde{\mathbf{Y}}}\right], \mathbb{E}\left[H_2\left(v^+_{\tilde{\mathbf{Y}}}\right)\right]\right) - \frac{1}{2}\psi\left(\mathbb{E}\left[v^-_{\tilde{\mathbf{Y}}}\right], \mathbb{E}\left[H_2\left(v^-_{\tilde{\mathbf{Y}}}\right)\right]\right).\end{aligned}$$

This equation, along with (4) shows that

$$\frac{d\gamma(t)}{dt} = \frac{1}{2^n}\sum_{\boldsymbol{x}\sim\boldsymbol{y}}(J(v_{\boldsymbol{y}}(t)) - J(v_{\boldsymbol{x}}(t)))(v_{\boldsymbol{x}}(t) - v_{\boldsymbol{y}}(t))$$

$$\geq \psi\left(\frac{\mathbb{E}\left[v^{+}_{\tilde{\mathbf{Y}}}+v^{-}_{\tilde{\mathbf{Y}}}\right]}{2}, \frac{\mathbb{E}\left[H_2\left(v^{-}_{\tilde{\mathbf{Y}}}\right)+H_2\left(v^{+}_{\tilde{\mathbf{Y}}}\right)\right]}{2}\right) \tag{5}$$

$$= \psi(\Pr(F=-1), \gamma(t)). \tag{6}$$

The proof is complete. □

Lemma 3. *Ψ is a non-empty closed convex set, which is also closed under pointwise maximum, i.e., if $\psi_i \in \Psi$ for $i \in \{1,2\}$, then $\psi(a,b) = \max(\psi_1(a,b), \psi_2(a,b)) \in \Psi$. Consequently, the class Ψ has maximal element $\psi^*(a,b)$ that pointwise dominates all the other members of Ψ.*

Proof. The proof of this Lemma can be found in Section V. □

A. A conjecture

Given a function ψ, observe that verifying whether ψ belongs to Ψ requires verifying an inequality with 14 free variables. While this is a finite-dimensional optimization problem, the space of variables is large. However, we can also think of ψ as follows: given a four tuple (m_u, m_w, e_u, e_w) satisfying $H_2(m_u) \geq e_u$ and $H_2(m_w) \geq e_w$, let

$$\zeta(m_u, m_w, e_u, e_w) := \inf_{(U,W)\in\mathcal{S}} \frac{1}{2}\mathbb{E}\left[(U-W)(J(W)-J(U))\right],$$

where the set $\mathcal{S}$ is the set of all pairs of random variables $(U,W) \in (0,1)^2$ such that

$$\mathbb{E}\left[U\right] = m_u, \mathbb{E}\left[W\right] = m_w, \mathbb{E}\left[H_2(U)\right] = e_u, \mathbb{E}\left[H_2(W)\right] = e_w.$$

It follows from the definition of ζ that it is a jointly convex function on four variables, and equation (2) states a lower bound on ζ in terms of ψ:

$$\zeta(m_u, m_w, e_u, e_w) \geq \psi\left(\frac{m_u+m_w}{2}, \frac{e_u+e_w}{2}\right) - \frac{1}{2}\psi(m_u, e_u) - \frac{1}{2}\psi(m_w, e_w).$$

In particular, when $m_u = m$, $m_w = 1-m$ and $e_u = e_w = e$, we obtain

$$\zeta(m, 1-m, e, e) \geq \psi\left(\frac{1}{2}, e\right) - \psi(m, e). \tag{7}$$

Theorem 4. *We have*

$$\zeta(1-m, m, e, e) = \phi\left(\frac{1}{2}, e\right) - \phi(m, e)$$

where $\phi(m,e)$ is defined as follows: let

$$\phi(x,y) = \begin{cases} \Phi(y) - \frac{y}{r}\Phi(r) & H_2(x) > y \\ 0 & H_2(x) \leq y \end{cases}$$

where

$$\Phi(x) = (1-2H_2^{-1}(x)) \cdot J(H_2^{-1}(x)), \qquad \forall x \in (0,1]$$

and $r \in (0,1]$ is defined as follows: if $x = 1/2$, we set $r = 1$; else if $x \neq \frac{1}{2}$, $r \in (0,1)$ is the unique solution of the following equation:

$$\frac{r}{1-2H_2^{-1}(r)} = \frac{y}{|1-2x|}.$$

Proof. The proof of this Theorem can be found in Section V. □

Based on the above partial characterization of the function ζ, numerical simulations, and some lower bounds on ζ, we make the following conjecture:

Conjecture 3. *The function ϕ belongs to the class Ψ.*

Next, we show that the conjecture, if true, implies the most informative Boolean function conjecture in the balanced case:

Lemma 5. *If Conjecture 3 holds, then Conjecture 1 holds whenever F is balanced, i.e.* $\Pr(F=-1) = \Pr(F=1) = \frac{1}{2}$.

Proof. First, observe that $\phi(x,y) = \phi(1-x,y)$ and

$$\phi\left(\frac{1}{2}, y\right) = \Phi(y)$$

Then,

$$\begin{aligned} g(x) &= \int_{\gamma(0)}^{x} \frac{du}{\phi(\frac{1}{2},u)} \\ &= \int_{\gamma(0)}^{x} \frac{du}{(1-2H_2^{-1}(u))\cdot J(H_2^{-1}(u))} \\ &= \int_{H_2^{-1}(\gamma(0))}^{H_2^{-1}(x)} \frac{dt}{1-2t} \qquad (8) \\ &= \frac{1}{2}\log\frac{1-2H_2^{-1}(\gamma(0))}{1-2H_2^{-1}(x)} \end{aligned}$$

where in (8), we apply the change to the variables $u = H_2(t)$. Thus, we get

$$g(\gamma(t)) \geq t$$

or

$$e^{-2t}\left[1-2H_2^{-1}(\gamma(0))\right] \geq 1-2H_2^{-1}(\gamma(t)). \tag{9}$$

or

$$\gamma(t) \geq H_2\left[\frac{1}{2}\left\{1-e^{-2t}\left[1-2H_2^{-1}(\gamma(0))\right]\right\}\right]. \tag{10}$$

Since $\gamma(t) = 1 - I(F;\mathbf{Y}_t)$ and $p_t = \frac{1-e^{-2t}}{2}$, we can rewrite the above as

$$1 - I(F;\mathbf{Y}_t) \geq H_2(p_t * H_2^{-1}(1 - I(F;\mathbf{X})))$$

where $a * b = a\cdot(1-b)+(1-a)\cdot b$ is the binary convolution. Given a balanced Boolean function B, define $P(F_\epsilon = -1|\mathbf{X}=\mathbf{x}) = 1-\epsilon$, if $B(\mathbf{x}) = -1$ and $P(F_\epsilon = 1|\mathbf{X}=\mathbf{x}) = 1-\epsilon$, if $B(\mathbf{x}) = 1$. As $\epsilon \to 0$, we have $I(F_\epsilon;\mathbf{X}) \to 1$, and $I(F_\epsilon;\mathbf{Y}_t) \to I(B(\mathbf{X});Y_t)$, establishing the desired inequality. □

Remark 3. *It is possible that a similar statement can be made about unbalanced functions as well. However, in this case, the integral of the reciprocal of ϕ does not seem to have a closed form.*

Remark 4. *The following natural guess for $\psi(x,y)$ does not belong to Ψ:*

$$\psi(x,y) = \Phi(1 - H_2(x) + y).$$

It has the following counterexample: let $X \in \{1,2,3\}$ be a ternary random variable with the probability distribution $P_X = (0.1, 0.45, 0.45)$. Let $(u_1,u_2,u_3) = (0.99, 0.9999, 0.0001), (w_1,w_2,w_3) = (0.01, 0.9999, 0.0001)$. One can verify that the above example does not satisfy the inequality (2).

III. The Hellinger Conjecture

The Hellinger Conjecture states that

$$\sqrt{1-\mathbb{E}\left[f(\mathbf{X})\right]^2} - \mathbb{E}\left[\sqrt{1-((T_\rho f)(\mathbf{Y}))^2}\right] \leq 1-\sqrt{1-\rho^2}, \tag{11}$$

where $T_\rho f(\boldsymbol{y}) = \mathbb{E}\left[f(\mathbf{X})|\mathbf{Y}=\boldsymbol{y}\right]$. Again, take some arbitrary $P_{F|\mathbf{X}}$ where $F \in \{-1,1\}$ is a binary random variable such that $P_{F|\mathbf{X}}(F=-1|\mathbf{X}=\boldsymbol{x}) \in (0,1)$ for all $\boldsymbol{x}$. Let $\rho_t = e^{-2t}$ and $v_{\boldsymbol{x}}(t) = Pr(F=-1|\mathbf{Y}_t = \boldsymbol{x})$. Let

$$d_{\boldsymbol{x}}(t) \triangleq \mathbb{E}\left[F|\mathbf{Y}_t = \boldsymbol{x}\right] = 1 - 2v_{\boldsymbol{x}}(t)$$

Define

$$r(t) = \mathbb{E}_{\mathbf{X}}\sqrt{1-\left(\mathbb{E}\left[F|\mathbf{Y}_t=\mathbf{X}\right]\right)^2} = \mathbb{E}\left[\sqrt{1-d_{\boldsymbol{X}}(t)^2}\right] = \frac{1}{2^n}\sum_{\boldsymbol{x}}\sqrt{1-d_{\boldsymbol{x}}(t)^2}.$$

Observe that

$$r(\infty) = \sqrt{1-\mathbb{E}\left[F\right]^2}.$$

The function $r(t)$ is increasing and direct calculation shows that its derivative can be calculated as follows:

Lemma 6. *We have*

$$r'(t) = \frac{1}{2^n}\sum_{(\boldsymbol{x},\boldsymbol{y}):\boldsymbol{x}\sim\boldsymbol{y}}(d_{\boldsymbol{x}}-d_{\boldsymbol{y}})\left(\frac{d_{\boldsymbol{x}}}{\sqrt{1-d_{\boldsymbol{x}}^2}}-\frac{d_{\boldsymbol{y}}}{\sqrt{1-d_{\boldsymbol{y}}^2}}\right).$$

where $\boldsymbol{x} \sim \boldsymbol{y}$ stands for the Hamming distance $d_H(\boldsymbol{x},\boldsymbol{y}) = 1$ and the tuple $(\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{y},\boldsymbol{x})$ are only counted once in the summation.

Let us define the corresponding set for the Hellinger conjecture:

Definition 2. *Let Ψ_H be the class of all non-negative functions $\psi(a,b) : (-1,1) \times [0,1) \mapsto \mathbb{R}$ that satisfy the following two conditions:*

- *$\psi(a,b) = 0$ when $\sqrt{1-a^2} \leq b$.*
- *$\psi(a,b) = \psi(-a,b)$.*
- *Let P_X be an arbitrary distribution on $\mathcal{X} = \{1,2,3,4,5\}$, and $(u_x,w_x) \in (-1,1)^2$ for $x \in \mathcal{X}$ be arbitrary. Then, the following holds:*

$$\frac{1}{2}\mathbb{E}_X\left[(u_X - w_X)\left(\frac{u_X}{\sqrt{1-u_X^2}} - \frac{w_X}{\sqrt{1-w_X^2}}\right)\right] \tag{12}$$

$$\geq \psi\left(\frac{\mathbb{E}\left[u_X + w_X\right]}{2}, \frac{\mathbb{E}\left[\sqrt{1-u_X^2}+\sqrt{1-w_X^2}\right]}{2}\right) - \frac{\psi\left(\mathbb{E}\left[u_X\right], \mathbb{E}\left[\sqrt{1-u_X^2}\right]\right) + \psi\left(\mathbb{E}\left[w_X\right], \mathbb{E}\left[\sqrt{1-w_X^2}\right]\right)}{2}. \tag{13}$$

Remark 5. *The following observations from the previous section carry over almost verbatim: Ψ_H is a non-empty closed convex set, which is also closed under pointwise maximum, i.e., if $\psi_i \in \Psi_H$ for $i \in \{1,2\}$, then $\psi(a,b) = \max(\psi_1(a,b),\psi_2(a,b)) \in \Psi_H$. Consequently, the class Ψ_H has maximal element $\psi_H^*(a,b)$ that pointwise dominates all the other members of Ψ_H.*

The following lemma follows:

Theorem 7. *For every $\psi \in \Psi_H$ we have*

$$\frac{dr(t)}{dt} \geq \psi\left(\mathbb{E}[F], r(t)\right).$$

Consequently, if we let

$$g(x) = \int_{r(0)}^{x} \frac{du}{\psi(\mathbb{E}[F],u)}, \qquad \forall x \geq r(0),$$

we obtain

$$r(t) \geq g^{-1}(t).$$

Proof. The proof of this theorem mimics that of Theorem 2 and is omitted. □

Lemma 8. *The following hold:*

1) For $(u,w) \in (-1,1)$,

$$\frac{1}{2}(u-w)\left(\frac{u}{\sqrt{1-u^2}} - \frac{w}{\sqrt{1-w^2}}\right) = \left(\left(\frac{u-w}{2}\right)^2 + \left(\frac{\sqrt{1-u^2}-\sqrt{1-w^2}}{2}\right)^2\right)\left(\frac{1}{\sqrt{1-u^2}} + \frac{1}{\sqrt{1-w^2}}\right).$$

2) For $(u_x,w_x) \in (-1,1)^2$

$$\frac{1}{2}\mathbb{E}_X\left[(u_X - w_X)\left(\frac{u_X}{\sqrt{1-u_X^2}} - \frac{w_X}{\sqrt{1-w_X^2}}\right)\right]$$

$$\geq \left(\left(\frac{\mathbb{E}\left[u_X - w_X\right]}{2}\right)^2 + \left(\frac{\mathbb{E}\left[\sqrt{1-u_X^2}-\sqrt{1-w_X^2}\right]}{2}\right)^2\right)\left(\frac{1}{\mathbb{E}\left[\sqrt{1-u_X^2}\right]} + \frac{1}{\mathbb{E}\left[\sqrt{1-w_X^2}\right]}\right).$$

Proof. The proof of this theorem can be found in Section V. □

Definition 3. *Let $\hat{\Psi}_H$ be the class of all non-negative functions $\psi(a,b) : (-1,1) \times [0,1) \mapsto \mathbb{R}$ that satisfy the following two conditions:*

- *$\psi(a,b) = 0$ when $\sqrt{1-a^2} \leq b$.*
- *$\psi(a,b) = \psi(-a,b)$*
- *For $(a_1,a_2) \in (-1,1)^2$ and $(b_1,b_2) \in [0,1)^2$, we have*

$$\left(\left(\frac{a_1-a_2}{2}\right)^2 + \left(\frac{b_1-b_2}{2}\right)^2\right)\left(\frac{1}{b_1}+\frac{1}{b_2}\right) \geq \psi\left(\frac{a_1+a_2}{2}, \frac{b_1+b_2}{2}\right) - \frac{1}{2}\psi(a_1,b_1) - \frac{1}{2}\psi(a_2,b_2).$$

Remark 6. *The following points are worth noting:*

1) *From Lemma 8, it is immediate that* $\hat{\Psi}_H \subseteq \Psi_H$.
2) *We do not have a conjecture for the maximal element of* Ψ_H*, but we have verified that the natural choice*

$$\psi(a,b) = \frac{1-a^2-b^2}{b}$$

has counterexamples. Let $X \in \{1,2,3\}$ *be a ternary random variable and consider the following setting of parameters:*

$$\begin{aligned} P_X &= (0.9118, 0.0760, 0.0122), \\ (u_1, u_2, u_3) &= (0.9996, 0.7316, 0.2996), \\ (w_1, w_2, w_3) &= (0.9992, 0.1416, 0.5866). \end{aligned}$$

One can show that this choice violates the constraint in (12).

IV. Discussion and Conclusion

The most-informative Boolean function conjecture and Talagrand's isoperimetric inequality are well-studied questions in information theory and combinatorics, respectively. The latter is a limiting case of the Hellinger conjecture. Despite a naturally degraded structure with respect to the noise operator, previous studies have not explicitly used this observation. In this work, we study both conjectures via the lens of the degraded channels and work on the derivative of the quantity of interest. This approach yields a finite-dimensional functional inequality whose solutions provide a dimension-independent lower bound to the original problem.

Acknowledgements

The authors would like to thank Lau Chin Wa for discussions at the preliminary stages of the work, as well as Venkat Anantharam with whom they had shared some initial forms of the conjecture.

References

[1] G. R. Kumar and T. A. Courtade, "Which Boolean functions are most informative?" in *2013 IEEE International Symposium on Information Theory*. IEEE, 2013, pp. 226–230.

[2] V. Anantharam, A. Gohari, S. Kamath, and C. Nair, "On hypercontractivity and the mutual information between Boolean functions," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2013, pp. 13–19.

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

[4] L. Yu, "On the Φ-stability and related conjectures," *Probability Theory and Related Fields*, vol. 186, no. 3, pp. 1045–1080, Aug 2023. [Online]. Available: https://doi.org/10.1007/s00440-023-01209-5

[5] 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.

[6] L. P. Barnes and A. Ozgur, "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)*, 2020, pp. 2205–2209.

[7] 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.

[8] Z. Chen and C. Nair, "On the optimality of dictator functions and isoperimetric inequalities on Boolean hypercubes," in *2024 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2024, pp. 3380–3385.

[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, no. 1, pp. 206–214, 01 1997. [Online]. Available: http://dx.doi.org/10.1214/aop/1024404285

[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," *Proceedings of the American Mathematical Society*, vol. 148, no. 10, pp. 4213–4224, 2020.

[12] P. Durcik, P. Ivanisvili, and J. Roos, "Sharp isoperimetric inequalities on the Hamming cube near the critical exponent," 2024. [Online]. Available: https://arxiv.org/abs/2407.12674

[13] A. Gohari and C. Nair, "Outer bounds for multiuser settings: the auxiliary receiver approach," *IEEE Transactions on Information Theory*, vol. 68, no. 2, pp. 701–736, 2021.

V. Proofs

A. Proof of Remark 2

This remark follows from Caratheodery's theorem because, given P_X one can find a distribution Q_X with support of size at most five such that

$$\begin{aligned}\mathbb{E}_{P_X}[(u_X - w_X)(J(w_X) - J(u_X))] \geq \mathbb{E}_{Q_X}[(u_X - w_X)(J(w_X) - J(u_X))],\\ \mathbb{E}_{P_X}(w_X) = \mathbb{E}_{Q_X}(w_X),\\ \mathbb{E}_{P_X}(u_X) = \mathbb{E}_{Q_X}(u_X),\\ \mathbb{E}_{P_X}(H_2(w_X)) = \mathbb{E}_{Q_X}(H_2(w_X)),\\ \mathbb{E}_{P_X}(H_2(u_X)) = \mathbb{E}_{Q_X}(H_2(u_X)).\end{aligned}$$

B. Proof of Lemma 3

Clearly $\psi(a,b) = 0$ for all a, b belongs to Ψ. Therefore, Ψ is non-empty. If $\psi_1, \psi_2 \in \Psi$, it is immediate at $\alpha\psi_1 + (1-\alpha)\psi_2 \in \Psi$ for $\alpha \in [0,1]$, and hence Ψ is convex.

Let $\psi_n \in \Psi, n \in N$ be a sequence of functions that converge (pointwise) to ψ_∞ on $(0,1)^2$. We have the following pointwise inequalities

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\left[(u_X - w_X)(J(w_X) - J(u_X))\right]\\ &\quad\geq \psi_n\left(\frac{\mathbb{E}\left[u_X + w_X\right]}{2}, \frac{\mathbb{E}\left[H_2(u_X) + H_2(w_X)\right]}{2}\right) - \frac{\psi_n(\mathbb{E}\left[u_X\right], \mathbb{E}\left[H_2(u_X)\right]) + \psi_n(\mathbb{E}\left[w_X\right], \mathbb{E}\left[H_2(w_X)\right])}{2},\end{aligned}$$

by taking limits on both side with respect to n, we have that

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\left[(u_X - w_X)(J(w_X) - J(u_X))\right]\\ &\quad\geq \psi_\infty\left(\frac{\mathbb{E}\left[u_X + w_X\right]}{2}, \frac{\mathbb{E}\left[H_2(u_X) + H_2(w_X)\right]}{2}\right) - \frac{\psi_\infty(\mathbb{E}\left[u_X\right], \mathbb{E}\left[H_2(u_X)\right]) + \psi_\infty(\mathbb{E}\left[w_X\right], \mathbb{E}\left[H_2(w_X)\right])}{2},\end{aligned}$$

Therefore, Ψ is closed.

On the other hand, Ψ is closed under pointwise maximum. Let $\psi_1, \psi_2 \in \Psi$, and set $\psi(a,b) = \max(\psi_1(a,b), \psi_2(a,b))$. We have that for any $i \in \{1,2\}$:

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\left[(u_X - w_X)(J(w_X) - J(u_X))\right]\\ &\quad\geq \psi_i\left(\frac{\mathbb{E}\left[u_X + w_X\right]}{2}, \frac{\mathbb{E}\left[H_2(u_X) + H_2(w_X)\right]}{2}\right) - \frac{\psi_i(\mathbb{E}\left[u_X\right], \mathbb{E}\left[H_2(u_X)\right]) + \psi_i(\mathbb{E}\left[w_X\right], \mathbb{E}\left[H_2(w_X)\right])}{2}\\ &\quad\geq \psi_i\left(\frac{\mathbb{E}\left[u_X + w_X\right]}{2}, \frac{\mathbb{E}\left[H_2(u_X) + H_2(w_X)\right]}{2}\right) - \frac{\psi(\mathbb{E}\left[u_X\right], \mathbb{E}\left[H_2(u_X)\right]) + \psi(\mathbb{E}\left[w_X\right], \mathbb{E}\left[H_2(w_X)\right])}{2}.\end{aligned}$$

By taking the maximum of the right-hand side over $i \in \{1,2\}$, we get that $\psi \in \Psi$.

C. Proof of Theorem 4

1) Step 1: Existence of minimizer in the expanded space: We need the following lemma:

Lemma 9. *The infimum*

$$\zeta(m_u, m_w, e_u, e_w) = \inf_{p(x), \{u_x, w_x\}} \frac{1}{2}\mathbb{E}\left[(u_X - w_X)(J(w_X) - J(u_X))\right]$$

subject to $(u_x, w_x) \in (0,1)^2$

$$\begin{aligned}\mathbb{E}\left[u_X\right] = m_u,\\ \mathbb{E}\left[w_X\right] = m_w,\\ \mathbb{E}\left[H_2(u_X)\right] = e_u,\\ \mathbb{E}\left[H_2(w_X)\right] = e_w,\end{aligned}$$

can be written as a minimum if we expand the domain of (u_x, w_x) *to* $(u_x, w_x) \in (0,1)^2 \cup \{(0,0),(1,1)\}$ *and define* $(J(w) - J(u))(u - w) = 0$ *when* $(u, w) \in \{(0,0),(1,1)\}$*. Moreover, any minimizer must satisfy the following property: if* $p(x_1), p(x_2) > 0$ *then one cannot simultaneously have* $u_{x_1} > u_{x_2}$ *and* $w_{x_1} < w_{x_2}$*.*

Proof. Fix some alphabet $\mathcal{X}$ and consider a sequence $\{(p_i(x), u_{ix}, w_{ix})\}$ for $i = 1, 2, \cdots$ converging to the infimum defining $\zeta(m_u, m_w, e_u, e_w)$. Since $\{(p_i(x), u_{ix}, w_{ix})\} \in [0,1]^{3|\mathcal{X}|}$ lies in a compact source, without loss of generality, we can assume that $\{(p_i(x), u_{ix}, w_{ix})\}$ converges to some sequence $\{(p(x), u_x, w_x)\} \in [0,1]^{3|\mathcal{X}|}$, where $p(x)$ is a probability distribution on $\mathcal{X}$. Let $\mathcal{S} = \{x \in \mathcal{X} : p(x) > 0\}$ be the support set of X.

Observe that the limiting u_x and w_x might be equal to 0 or 1. Note that $(u-w)(J(w)-J(u))$ converges to infinity if one of u and v converges to 0 and the other one converges to 1. Therefore, the only possible case is that the limiting (u_x, w_x) might be equal to $(0,0)$ or $(1,1)$ for some x. Note that $(J(w)-J(u))(u-w) \geq 0$ for every $u, w \in (0,1)$. Consequently, if $(u(x), w(x)) \in \{(0,0),(1,1)\}$ for some $x \in \mathcal{S}$, changing the value of w_{ix} to be equal to u_{ix} for such x would decrease $(J(w_{xi}) - J(u_{xi}))(u_{xi} - w_{xi})$ to 0, and cannot affect the limit of the other terms. Therefore, if $(u(x), w(x)) \in \{(0,0),(1,1)\}$ for some $x \in \mathcal{S}$, without loss of generality we can assume that $u_{ix} = w_{ix}$ for every i. Let

$$\begin{aligned} \mathcal{S}_0 &= \{x \in \mathcal{S} : u_x = w_x = 0\}, \\ \mathcal{S}_1 &= \{x \in \mathcal{S} : u_x = w_x = 1\}, \\ \mathcal{S}_2 &= \{x \in \mathcal{S} : u_x, w_x \in (0,1)\}. \end{aligned}$$

In this case, the infimum of the expression will be equal to

$$\frac{1}{2}\sum_{x \in \mathcal{S}_2} p(x)(J(w_x) - J(u_x))(u_x - w_x).$$

This shows that if we extend the domain of (u,w) from $(0,1)^2$ to $(0,1)^2 \cup \{(0,0),(1,1)\}$ and define $(J(w)-J(u))(u-w) = 0$ when $(u,w) \in \{(0,0),(1,1)\}$, the infimum of the expression will also be a minimum and will be attained at some $p(x), u_x, w_x$ where $(u_x, w_x) \in (0,1)^2 \cup \{(0,0),(1,1)\}$ for every x.

Next, let us consider a minimizer $p(x), u_x, w_x$ for $x \in \mathcal{X}$. Assume that $u_{x_1} > u_{x_2}$ and $w_{x_1} < w_{x_2}$. Then, using the fact that $J(x)$ is a decreasing function, we have

$$(J(w_{x_1}) - J(w_{x_2}))(u_{x_1} - u_{x_2}) + (J(u_{x_1}) - J(u_{x_2}))(w_{x_1} - w_{x_2}) > 0.$$

Take x_1 and x_2 where $p(x_1) \geq p(x_2) > 0$. Take a symbol $x_3 \notin \mathcal{X}$ and expand the alphabet of X as $\mathcal{X}' = \mathcal{X} \cup \{x_3\}$. Let $p(X' = x) = p(X = x)$ if $x \notin \{x_1, x_3\}$, $p(X' = x_1) = p(x_1) - p(x_2)$ and $p(X' = x_3) = p(x_2)$. Let $(\hat{u}_{x'}, \hat{w}_{x'}) = (u_{x'}, w_{x'})$ if $x' \notin \{x_2, x_3\}$, and $(\hat{u}_{x_2}, \hat{w}_{x_2}) = (u_{x_2}, w_{x_1})$ and $(\hat{u}_{x_3}, \hat{w}_{x_3}) = (u_{x_1}, w_{x_2})$. Observe that

$$\begin{aligned} &(u_{x_1} - w_{x_1})(J(w_{x_1}) - J(u_{x_1})) + (u_{x_2} - w_{x_2})(J(w_{x_2}) - J(u_{x_2})) \\ &\quad > (u_{x_1} - w_{x_2})(J(w_{x_2}) - J(u_{x_1})) + (u_{x_2} - w_{x_1})(J(w_{x_1}) - J(u_{x_2})) \end{aligned}$$

because the difference between the left-hand side and the right-hand side equals

$$(J(w_{x_1}) - J(w_{x_2}))(u_{x_1} - u_{x_2}) + (J(u_{x_1}) - J(u_{x_2}))(w_{x_1} - w_{x_2}).$$

One can then deduce that X' along with $u_{X'}$ and $w_{X'}$ will yield a smaller value of $\frac{1}{2}\mathbb{E}\left[(u_{X'} - w_{X'})(J(w_{X'}) - J(u_{X'}))\right]$ than X, u_X and w_X. Moreover, $\mathbb{E}\hat{u}_{X'} = \mathbb{E}u_X$, $\mathbb{E}\hat{w}_{X'} = \mathbb{E}w_X$, $\mathbb{E}H(\hat{u}_{X'}) = \mathbb{E}H(u_X)$, $\mathbb{E}H(\hat{w}_{X'}) = \mathbb{E}H(w_X)$. This contradicts the assumption of minimality of X, u_X and w_X. □

Let us return to the proof of Theorem 4. Suppose $p(x), u_x, w_x$ is a minimizer (which exists due to the above lemma when we allow $(0,0)$ and $(1,1)$ in the domain).

2) Step 2: finding a minimizer satisfying $u_x + w_x = 1$ for all x where $(u_x, w_x) \in (0,1)^2$: We first show that one can find another minimizer satisfying $u_x + w_x = 1$ for all x where $(u_x, w_x) \in (0,1)^2$. To construct the new minimizer, we first apply a symmetrization argument. Let $\mathcal{X}' = \mathcal{X} \times \{1,2\}$. For any $x' = (x, j)$ where $x \in \mathcal{X}$ and $j \in \{1,2\}$ define

$$p(x') = \frac{1}{2}p(x)$$

and define $u_{x'}$ and $v_{x'}$ as follows:

$$u_{(x,1)} = u_x, \qquad w_{(x,1)} = w_x$$

$$u_{(x,2)} = 1 - w_x, \qquad w_{(x,2)} = 1 - u_x$$

One can verify that

$$\begin{aligned} \mathbb{E}\left[u_{X'}\right] &= m, \\ \mathbb{E}\left[w_{X'}\right] &= 1 - m, \\ \mathbb{E}\left[H_2(u_{X'})\right] &= e, \\ \mathbb{E}\left[H_2(w_{X'})\right] &= e. \end{aligned}$$

Moreover,

$$\mathbb{E}\left[(u_X - w_X)(J(w_X) - J(u_X))\right] = \mathbb{E}\left[(u_{X'} - w_{X'})(J(w_{X'}) - J(u_{X'}))\right]$$

Thus, $p(x'), u_{x'}, w_{x'}$ is also a minimizer.

Next, for every x where $(u_x, w_x) \in (0,1)^2$, let v_x be the unique solution of

$$\frac{H_2(v_x)}{2v_x - 1} = \frac{H_2(u_x) + H_2(w_x)}{2(u_x - w_x)}.$$

Here, we set $v_x = 1/2$ if $u_x = w_x$. Next, let

$$r_x = \frac{H_2(u_x) + H_2(w_x)}{2H_2(v_x)}.$$

Lemma 10. *We have that* $r_x \in [0,1]$. *Moreover,*

$$r_x(1 - 2v_x)J(v_x) \leq \frac{1}{2}(u_x - w_x)(J(w_x) - J(u_x)). \tag{14}$$

The proof of the lemma is given below. We use this lemma as follows: we are choosing the pair

$$u_{(x,1)} = u_x, \qquad w_{(x,1)} = w_x$$

with probability $p(x)/2$ and the pair

$$u_{(x,2)} = 1 - w_x, \qquad w_{(x,2)} = 1 - u_x$$

with probability $p(x)/2$. Instead, let us choose the pair $(0,0)$ with probability $(1-r_x)p(x)/2$, the pair $(1,1)$ with probability $(1-r_x)p(x)/2$ and finally the pair $(v_x, 1-v_x)$ with probability $r_x p(x)/2$. We apply this replacement for every x where $(u_x, w_x) \in (0,1)^2$. One can verify that this transformation preserves $\mathbb{E}[u_{X'}]$, $\mathbb{E}[w_{X'}]$ as well as $\mathbb{E}[H_2(u_{X'})]$, $\mathbb{E}[H_2(w_{X'})]$. Moreover, (14) implies that the value of the objective function after the transformation is less than or equal to the original value. Observe that after the transformation, we are using either the pairs $(0,0)$, $(1,1)$ or $(v_x, 1-v_x)$. This establishes the statement desired in Step 2.

It remains to prove Proof of Lemma 10.

Proof of Lemma 10. Assume that $u_x \geq w_x$ (the proof for the other case is similar). In this case, $v_x \geq 1/2$. We have

$$\frac{H_2(u_x) + H_2(w_x)}{2} = \frac{H_2(1-u_x) + H_2(w_x)}{2} \leq H_2\left(\frac{1 - u_x + w_x}{2}\right).$$

Thus,

$$H_2^{-1}\left(\frac{H_2(u_x) + H_2(w_x)}{2}\right) \leq \frac{1 - u_x + w_x}{2}$$

and we obtain

$$0 \leq u_x - w_x \leq 1 - 2H_2^{-1}\left(\frac{H_2(u_x) + H_2(w_x)}{2}\right)$$

and we can write

$$\frac{H_2(1-v_x)}{1 - 2(1-v_x)} = \frac{H_2(u_x) + H_2(w_x)}{2(u_x - w_x)} \geq \frac{\frac{H_2(u_x)+H_2(w_x)}{2}}{1 - 2H_2^{-1}(\frac{H_2(u_x)+H_2(w_x)}{2})} = \frac{H_2(H_2^{-1}(\frac{H_2(u_x)+H_2(w_x)}{2}))}{1 - 2H_2^{-1}(\frac{H_2(u_x)+H_2(w_x)}{2})}.$$

Since $H_2(x)/(1-2x)$ is increasing in $x \in [0, 0.5]$, we get

$$1 - v_x \geq H_2^{-1}(\frac{H_2(u_x) + H_2(w_x)}{2})$$

Thus,

$$H_2(v_x) \geq \frac{H_2(u_x) + H_2(w_x)}{2}$$

This completes the proof for $r_x \in [0,1]$.

We now prove that

$$r_x(1 - 2v_x)J(v_x) \leq \frac{1}{2}(u_x - w_x)(J(w_x) - J(u_x)). \tag{15}$$

From the definition of r_x and v_x we have

$$r_x(2v_x - 1) = (u_x - w_x).$$

We show the inequality when $u_x \geq w_x$ (which implies $v_x \geq 1/2$). The proof for the other case is similar. Observe that (15) is equivalent to

$$-J(v_x) \leq \frac{J(w_x) - J(u_x)}{2}.$$

Let

$$z=\frac{\sqrt{w_x(1-u_x)}}{\sqrt{w_x(1-u_x)}+\sqrt{u_x(1-w_x)}}\in(0,0.5]\,. \tag{16}$$

Then, one can verify that

$$J(z)=\frac{J(w_x)-J(u_x)}{2}\geq 0. \tag{17}$$

We wish to show that $-J(v_x)=J(1-v_x)\leq J(z)$. Since J is a decreasing function, this is equivalent to $1-v_x\geq z$. Since $H_2(x)/(1-2x)$ is increasing in $x\in[0,0.5]$, this inequality is equivalent to

$$\frac{H_2(z)}{1-2z}\leq\frac{H_2(1-v_x)}{1-2(1-v_x)}=\frac{H_2(u_x)+H_2(w_x)}{2(u_x-w_x)}.$$

To sum this up, for z defined by (16), we need to show the above inequality. Equivalently,

$$\begin{aligned}&H_2(u_x)+H_2(w_x)\\&\geq 2\left(\sqrt{w_x(1-u_x)}+\sqrt{u_x(1-w_x)}\right)^2 H_2\left(\frac{\sqrt{w_x(1-u_x)}}{\sqrt{w_x(1-u_x)}+\sqrt{u_x(1-w_x)}}\right)\\&=-w_x(1-u_x)\log_2(w_x(1-u_x))-u_x(1-w_x)\log_2(u_x(1-w_x))\\&\quad-2\sqrt{w_x(1-u_x)u_x(1-w_x)}\log_2\sqrt{w_x(1-u_x)u_x(1-w_x)}\\&\quad+\left(\sqrt{u_x(1-w_x)}+\sqrt{w_x(1-u_x)}\right)^2\log_2\left(\sqrt{u_x(1-w_x)}+\sqrt{w_x(1-u_x)}\right)^2.\end{aligned}$$

By expanding the left-hand side, the inequality is equivalent to proving

$$\begin{aligned}&2\sqrt{w_x(1-u_x)u_x(1-w_x)}\log_2\sqrt{w_x(1-u_x)u_x(1-w_x)}\\&\quad\geq w_xu_x\log_2(w_xu_x)+(1-w_x)(1-u_x)\log_2((1-w_x)(1-u_x))\\&\qquad+\left(\sqrt{u_x(1-w_x)}+\sqrt{w_x(1-u_x)}\right)^2\log_2\left(\sqrt{u_x(1-w_x)}+\sqrt{w_x(1-u_x)}\right)^2.\end{aligned} \tag{18}$$

Let

$$a=\frac{\frac{u_x}{1-u_x}+\frac{w_x}{1-w_x}}{2},$$

and

$$b=\sqrt{\frac{u_x}{1-u_x}\frac{w_x}{1-w_x}}.$$

Observe that we have $a\geq b\geq 0$. If we divide both sides of (18) by $(1-u_x)(1-w_x)$ and apply a change of variable, the inequality will be equivalent to

$$(1+2a+b^2)\log_2(1+2a+b^2)-2(a+b)\log_2(2a+2b)\geq 2b^2\log_2 b-2b\log_2(b).$$

Taking the derivative in a, we obtain $2\log(1+2a+b^2)-2\log(2a+2b)\geq 0$. Therefore, it suffices to prove the above inequality for $a=b$. In other words, we need to show that

$$2b\log(b)+(1+2b+b^2)\log(1+2b+b^2)-2b^2\log b-4b\log 4b\geq 0.$$

Equivalently, we need to show that

$$f(b):=(1+b)^2\log(1+b)-b^2\log b-b\log(b)-2b\log 4\geq 0.$$

Observe that $b^2f\left(\frac{1}{b}\right)=f(b)$. Therefore, it suffices to consider $b\in(0,1]$. □

Note that

$$f'(b)=2(1+b)\log(1+b)-2b\log b-\log b-2\log 4.$$

One can verify that there is exactly one solution for $f''(b)=0$ in $(0,1)$. Further, one can observe that $f(b)$ initially increases and is concave, attains a maximum, then decreases, and then turns convex in the interval $(0,1)$. Note that $f'(1)=0$ and hence decreasing as $b\uparrow 1$. As $f(0)=0$ and $f(1)=0$, and $f'(0)=\infty$, if $f(b)$ had a local minimum with a negative value, it must have had two local maximums in $(0,1)$. This contradicts that there is exactly one solution for $f''(b)=0$ in $(0,1)$. Therefore, $f(b)\geq 0$ for $b\in(0,1)$, and equality only holds when $b\in\{0,1\}$.

3) Step 3: Simplifying the minimizer: Assume that we have a minimizer satisfying $u_x + w_x = 1$ for all x where $(u_x, w_x) \in (0,1)^2$. Take x_1 and x_2 such that $p(x_1), p(x_2) > 0$ and $u_{x_1} + w_{x_1} = 1$ and $u_{x_2} + w_{x_2} = 1$. We claim that $u_{x_1} = u_{x_2}$ and $w_{x_1} = w_{x_2}$. Assume that $u_{x_1} > u_{x_2}$. Then, by Lemma 9, we get $w_{x_1} \geq w_{x_2}$. This would imply $1 - u_{x_1} \geq 1 - u_{x_2}$ or $u_{x_1} \leq u_{x_2}$ which is a contradiction. Thus, $u_{x_1} = u_{x_2}$ and $w_{x_1} = w_{x_2}$. This shows that we may put weights on at most a single term $(v, 1-v)$. Thus, we can take the pairs $(0,0)$ with some probability p_0, $(1,1)$ with some probability p_1 and $(v, 1-v)$ for some v with some probability p_2. Since $\mathbb{E}[u_X] = m$ and $\mathbb{E}[w_X] = 1 - m$, we get

$$vp_2 + p_1 = m \tag{19}$$

$$(1-v)p_2 + p_1 = 1 - m \tag{20}$$

Summing up, we get, $p_2 + 2p_1 = 1$ which implies $p_0 = p_1 = (1-p_2)/2$ and we get

$$vp_2 + (1-p_2)/2 = m. \tag{21}$$

Using

$$e = p_2 H_2(v)$$

the objective function reduces to the expression given in the statement of the theorem.

D. Proof of Lemma 8

We wish to show that

$$\frac{1}{2}(u-w)\left(\frac{u}{\sqrt{1-u^2}} - \frac{w}{\sqrt{1-w^2}}\right) = \left(\left(\frac{u-w}{2}\right)^2 + \left(\frac{\sqrt{1-u^2}-\sqrt{1-w^2}}{2}\right)^2\right)\left(\frac{1}{\sqrt{1-u^2}} + \frac{1}{\sqrt{1-w^2}}\right).$$

Let $\sqrt{1-u^2} = x$ and $\sqrt{1-w^2} = y$. Then, the desired equality can be written as, requiring to show that,

$$\frac{1}{2}(u-w)\frac{(uy-wx)}{xy} = \left(\left(\frac{u-w}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2\right)\frac{(x+y)}{xy}.$$

Using $u^2 + x^2 = 1$ and $w^2 + y^2 = 1$, the desired equality can be written as, requiring to show that,

$$(u-w)(uy-wx) = (1 - uw - xy)(x+y).$$

which is immediate by expansion and using $u^2 + x^2 = 1$ and $w^2 + y^2 = 1$.

Cauchy-Schwarz implies that if $Y > 0$, then $\mathbb{E}\left[\frac{X^2}{Y}\right] \geq \frac{\mathbb{E}[X]^2}{\mathbb{E}[Y]}$. Now, the second part is immediate, by applying the earlier part and expanding the expression, naturally, into four terms.