

Strong duality for the GROW criterion

Ashwin Ram^{*} Martin Larsson[†] Johannes Ruf[‡] Aaditya Ramdas[§]

July 19, 2026

Abstract

This paper presents general strong duality results when testing hypotheses by betting against them. A bet is an e-variable for a composite null hypothesis $\mathcal{P}$: a nonnegative random variable X whose expected value is at most one under every $P \in \mathcal{P}$. Following Kelly, Breiman, Cover, Shafer, and Grünwald et al. [2024a], we study a natural minimax *log-optimality* criterion: given a composite alternative $\mathcal{Q}$, we characterize the “GROW value” $\sup_X \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X]$. This paper generalizes the results of Larsson et al. [2025] from (arbitrary $\mathcal{P}$ and) simple $\mathcal{Q}$ to arbitrary $\mathcal{Q}$. We prove that there always exists a minimizing information-projection pair between the weak-* closures of the convex hulls of arbitrary $\mathcal{P}$ and $\mathcal{Q}$, and show that the GROW value for *bounded* e-variables always equals their relative entropy. We also prove a similarly general strong duality for the REGROW criterion with bounded e-variables and arbitrary bounded offsets. Under various assumptions our results extend to unbounded e-variables, and examples show that without any assumptions such extensions fail. Our results are analogous to those in Larsson et al. [2026b], swapping tests for bounded e-variables, minimax risk for the GROW criterion, and total variation for relative entropy.

1 Introduction

Given a measurable space $(\Omega, \mathcal{F})$, we write $\mathcal{M}$ for the set of finite signed (countably additive) measures on $\mathcal{F}$, and $\mathcal{M}_+$ and $\mathcal{M}_1$ for the subsets of nonnegative measures and probability measures. Let $\mathcal{P} \subseteq \mathcal{M}_1$ be an arbitrary composite null hypothesis, and $\mathcal{Q} \subseteq \mathcal{M}_1$ an arbitrary composite alternative, both of which we always assume to be nonempty. An e-variable for $\mathcal{P}$ is a $[0, \infty]$ -valued random variable X such that $\mathbb{E}_P[X] \leq 1$ for all $P \in \mathcal{P}$. The set of all e-variables for $\mathcal{P}$ is denoted $\mathcal{E}$, and the subset of e-variables bounded from above is denoted $\mathcal{E}_b$.

The influential work of Grünwald et al. [2024a] proposed studying the *GROW (Growth Rate that is Optimal in the worst-case) value*,

$$G = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X], \tag{1}$$

and in particular, establishing duality results that equate G to the infimum relative entropy between (appropriate extensions of) $\mathcal{Q}$ and $\mathcal{P}$. For example, Larsson et al. [2025] point out that for simple $\mathcal{Q} = \{Q\}$, G equals a particular infimum relative entropy between Q and the *effective null hypothesis* corresponding to $\mathcal{P}$. Similarly, for simple $\mathcal{P} = \{P\}$ and convex $\mathcal{Q}$, Theorem 4.2

^{*}Carnegie Mellon University, aram2@andrew.cmu.edu

[†]Department of Mathematical Sciences, Carnegie Mellon University, larsson@cmu.edu.

[‡]Department of Mathematics, London School of Economics, j.ruf@lse.ac.uk

[§]Departments of Statistics and ML, Carnegie Mellon University, aramdas@cmu.edu

yields, when (for example) $H(Q | P) < \infty$ for each $Q \in \mathcal{Q}$, that G equals the infimum relative entropy between $\mathcal{Q}$ and $\mathcal{P}$.

When $\mathcal{P}$ and $\mathcal{Q}$ are both composite, such a “universal” (i.e., without any additional restrictions) strong duality result does not hold without additional restrictions. For example, the results of Huber and Strassen [1973] imply that if $\mathcal{P}$ and $\mathcal{Q}$ have a *least favorable distribution* pair (a strong assumption!), then

$$G = \inf_{P \in \mathcal{P}} \inf_{Q \in \mathcal{Q}} H(Q | P),$$

where H is the Kullback–Leibler divergence, or relative entropy; see Saha and Ramdas [2025] for a self-contained proof. Grünwald et al. [2024a] provide a different (also strong) sufficient condition under which the same statement holds. Our paper introduces several other conditions under which such characterizations hold; for example, under suitable compactness or domination assumptions on $\mathcal{P}$ and $\mathcal{Q}$, many of which are weaker or incomparable to existing ones.

However, our investigations of the above quantity led us to a related notion that remarkably satisfies a strong duality without any restrictions on $\mathcal{P}$ or $\mathcal{Q}$. In particular, we obtain a complete characterization of the *bounded* GROW value, where the supremum is taken over the set $\mathcal{E}_b$ of all *bounded* e-variables,

$$G_b = \sup_{X \in \mathcal{E}_b} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X]. \quad (2)$$

There are at least three reasons for considering bounded e-variables. First, Larsson et al. [2026b] provide a complete characterization of testable hypotheses, that is, of when there exists a nontrivial test between $\mathcal{P}$ and $\mathcal{Q}$; they point out this is equivalent to the existence of nontrivial *bounded* e-variables. Second, as we formalize below, when $\mathcal{Q}$ is a singleton, we have $G = G_b$, so both (1) and (2) may equally be considered generalizations of the singleton case. Third, G_b permits a strong duality result without any restrictions, while this does not appear to be possible for G .

The term GROW was originally motivated by the following betting context. We imagine a forecaster claiming that the data are well described by some $P \in \mathcal{P}$. A skeptic, believing instead that some $Q \in \mathcal{Q}$ is a better description, would like to bet against the forecaster. The e-variable corresponds to an available bet or wealth multiplier, meaning that the skeptic puts forward (say) one dollar along with a particular e-variable before seeing the data, and the forecaster must pay back the realized value of the e-variable (in dollars) after the data is observed. The forecaster is willing to accept any e-variable as a bet, because under their own belief $\mathcal{P}$, the Skeptic will have nonpositive expected net payoff. So which bet should the Skeptic choose? A long tradition dating back to Kelly [1956], Breiman [1961], Cover [1987], Shafer [2021] suggests that the Skeptic should maximize the expected logarithm of the wealth under the alternative Q . Picking the worst $Q \in \mathcal{Q}$ results in the GROW criterion/value. Since $\mathbb{E}_Q[\log X]$ is also sometimes called the e-power of X under Q (see, for example, Ramdas and Wang [2025]), one can also refer to G as the minimax e-power, but we stick to the former terminology for simplicity.

As mentioned above, a universal strong duality result was recently established for singleton $\mathcal{Q} = \{Q\}$ and arbitrary $\mathcal{P}$ by Larsson et al. [2025], extending earlier work by Lardy et al. [2024] and Li [1999]. Defining the *effective null* generated by $\mathcal{P}$ as

$$\mathcal{P}_{\text{eff}} = \{P \in \mathcal{M}_+ : \mathbb{E}_P[X] \leq 1 \text{ for all } X \in \mathcal{E}\}, \quad (3)$$

that paper showed that under no restrictions whatsoever on $\mathcal{Q}$ or $\mathcal{P}$,

$$\sup_{X \in \mathcal{E}} \mathbb{E}_Q[\log X] = \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P), \quad (4)$$

and identified a (sub-probability) measure $P^* \in \mathcal{P}_{\text{eff}}$ called the *reverse information projection* such that the right-hand side above equals $H(Q | P^*)$. One can view our work as generalizing these results to the setting of composite $\mathcal{P}$ and $\mathcal{Q}$.

Summary of our main results

To set the stage, we argue in Lemma 2.6 that for a singleton $\mathcal{Q}$,

$$\sup_{X \in \mathcal{E}} \mathbb{E}_{\mathcal{Q}}[\log X] = \sup_{X \in \mathcal{E}_b} \mathbb{E}_{\mathcal{Q}}[\log X]. \quad (5)$$

So the left-hand side in (4) could also have used bounded e-variables. Further, by monotone convergence the right-hand side of (4), which depends on $\mathcal{E}$ through the definition of the effective null, could also have used bounded e-variables (see (8)). Unsurprisingly, the story is more complicated for general composite $\mathcal{Q}$.

Obviously, $G \geq G_b$, but the inequality can be strict — remarkably, even for a point null $\mathcal{P} = \{\mathbf{P}\}$ and composite $\mathcal{Q}$, both supported on the naturals $\mathbb{N}$ and thus having a common dominating measure; see Example 5.1. It turns out that G_b has a duality theorem without any restrictions.

1. Our main result, Theorem 3.1, shows that *under no assumptions on $\mathcal{P}$ or $\mathcal{Q}$* , there always exists a pair $(\mu^*, \nu^*) \in \overline{\text{co}}^*(\mathcal{P}) \times \overline{\text{co}}^*(\mathcal{Q})$ such that

$$G_b = \inf_{\mu \in \overline{\text{co}}^*(\mathcal{P})} \inf_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} H(\nu \mid \mu) = H(\nu^* \mid \mu^*), \quad (6)$$

where $\overline{\text{co}}^*(\mathcal{P})$ and $\overline{\text{co}}^*(\mathcal{Q})$ are the closed convex hull of $\mathcal{P}$ and $\mathcal{Q}$, respectively. Here we take the weak-* closure in ba , the space of bounded finitely additive measures. (We recall these concepts later.) One can also prove that for any $\mathbf{Q} \in \mathcal{M}_1$,

$$\min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\mathbf{Q} \mid \mu) = \min_{\mathbf{P} \in \mathcal{P}_{\text{eff}}} H(\mathbf{Q} \mid \mathbf{P}), \quad (7)$$

so that (6), in conjunction with (7) and (5), implies (4), justifying our claim that Theorem 3.1 generalizes the work of Larsson et al. [2025] to arbitrary $\mathcal{Q}$.

2. Following the suggestion of Grünwald et al. [2024a], we also study a generalization of GROW, called REGROW (relative GROW):

$$\sup_{X \in \mathcal{E}_b} \inf_{\mathbf{Q} \in \mathcal{Q}} (\mathbb{E}_{\mathbf{Q}}[\log X] - \xi(\mathbf{Q})),$$

for some bounded functional $\xi : \mathcal{Q} \rightarrow \mathbb{R}$, for example $\xi(\mathbf{Q}) = \sup_X \mathbb{E}_{\mathbf{Q}}[\log X]$ (the GROW criterion for a singleton $\mathbf{Q}$, called GRO), when this is finite and bounded over $\mathcal{Q}$. Theorem 3.4 establishes a strong duality result for arbitrary $\mathcal{P}, \mathcal{Q}, \xi$ that involves the concave biconjugate of ξ .

Next, we examine several interesting special cases where one can use *unbounded* e-variables.

3. First, if $\mathcal{P} = \{\mathbf{P}\}$ is a singleton, and $\mathcal{Q}$ is convex with $H(\mathbf{Q} \mid \mathbf{P}) < \infty$ for all $\mathbf{Q} \in \mathcal{Q}$, we prove a strong duality result using a generalized information projection of $\mathbf{P}$ onto $\mathcal{Q}$:

$$G = \inf_{\mathbf{Q} \in \mathcal{Q}} H(\mathbf{Q} \mid \mathbf{P}).$$

Example 5.1 shows that despite strong duality holding for both G and G_b , the two quantities are not equal in general. Example 5.2 shows that if $H(\mathbf{Q} \mid \mathbf{P}) < \infty$ does not hold for all $\mathbf{Q} \in \mathcal{Q}$, then strong duality can fail, even though $\mathbf{Q} \ll \mathbf{P}$ for all $\mathbf{Q} \in \mathcal{Q}$.

4. We generalize the above result to the case of arbitrary $\mathcal{P}$, assuming that a joint information projection (JIPr) exists. A JIPr is defined as a pair (P^*, Q^*) such that the sub-probability P^* is the reverse information projection of the probability Q^* onto $\mathcal{P}_{\text{eff}}$ and Q^* is the generalized information projection (in the sense of Csiszár [1984]) of P^* onto $\mathcal{Q}$. When a JIPr exists, and (say) $H(Q | P^*) < \infty$ for all $Q \in \mathcal{Q}$, we show that

$$G = \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P),$$

and the left-hand side is achieved by the e-variable dQ^*/dP^* .

5. Additionally, if $\mathcal{Q}$ is convex and compact in the setwise topology, then we show that

$$G = G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q | P).$$

In the above two cases, if we further assume that (on Polish sample spaces) $\mathcal{P}$ is convex and weakly compact, we show that $\mathcal{P}_{\text{eff}}$ can be replaced by $\mathcal{P}$.

6. When working on Polish sample spaces, we show that if $\mathcal{P}$ and $\mathcal{Q}$ are both convex and weakly compact, then again

$$G = G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}} H(Q | P),$$

but this time one can also further restrict to *bounded continuous* e-variables.

7. Finally, we show that if Ω is finite, and $\mathcal{P}, \mathcal{Q}$ are compact convex sets in the finite-dimensional probability simplex and if $\sup_{P \in \mathcal{P}} P_i > 0$ for each coordinate i , then

$$G = G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}} H(Q | P).$$

Further the GROW value is finite, attained by a bounded e-variable, which can be represented as the likelihood ratio of JIPr pairs that minimize the above right-hand side.

The first and sixth result above have an intriguing strong parallel to the main results of Larsson et al. [2026b], after swapping e-variables for tests, the GROW criterion for minimax risk, and the relative entropy for total variation distance.

We show a number of other results of independent interest, along with several examples and counterexamples that demonstrate various subtleties in the aforementioned conditions.

The paper is organized as follows. Section 2 recaps some basic definitions and properties, especially of relative entropy when operating on finitely additive measures. Section 3 presents a strong duality for bounded GROW (and REGROW) without restriction on the sets of measures. Section 4 provides strong duality results for the unbounded GROW criterion, under some relatively mild restrictions. Section 5 discusses some interesting and nontrivial examples that demonstrate certain subtleties or challenges in these strong duality statements. The appendix contains all proofs not presented in the paper.

2 Preliminaries

We attempt to maintain the convention that for countably additive measures, we use Roman letters P, Q, R, S , whereas for measures that could be finitely additive, we use Greek letters μ, ν . Recall that the set of e-variables for $\mathcal{P}$ is

$$\mathcal{E} = \left\{ \text{all measurable } X: \Omega \rightarrow [0, \infty] \text{ such that } \int X dP \leq 1 \text{ for all } P \in \mathcal{P} \right\},$$

and the set of all bounded e-variables is given by

$$\mathcal{E}_b = \left\{ X \in \mathcal{E} : \sup_{\omega \in \Omega} X(\omega) < \infty \right\} = \mathcal{E} \cap \mathcal{B}_b,$$

where $\mathcal{B}_b$ is the Banach space of all bounded measurable functions $f : \Omega \rightarrow \mathbb{R}$ with the supremum norm $\|f\|_\infty$.

Since $X = 1$ is an e-variable, all elements of $\mathcal{P}_{\text{eff}}$, defined in (3), are sub-probabilities. Any sub-probability which is setwise dominated by some element of $\mathcal{P}_{\text{eff}}$ also belongs to $\mathcal{P}_{\text{eff}}$. Moreover, $\mathcal{P}$ and $\mathcal{P}_{\text{eff}}$ have the same nullsets A since if $P(A) = 0$ for all $P \in \mathcal{P}$, then $\infty \mathbf{1}_A \in \mathcal{E}$, and hence $P(A) = 0$ for all $P \in \mathcal{P}_{\text{eff}}$. Finally, in the definition of $\mathcal{P}_{\text{eff}}$ it is enough, thanks to the monotone convergence theorem, to let X range over the set $\mathcal{E}_b$. Hence we have

$$\mathcal{P}_{\text{eff}} = \left\{ P \in \mathcal{M}_+ : \int X dP \leq 1 \text{ for all } X \in \mathcal{E} \right\} = \left\{ P \in \mathcal{M}_+ : \int X dP \leq 1 \text{ for all } X \in \mathcal{E}_b \right\}. \quad (8)$$

We will also write $\text{ba} = \text{ba}(\Omega, \mathcal{F})$ for the Banach space of all bounded finitely additive signed measures on $(\Omega, \mathcal{F})$, endowed with the total variation norm. Its positive cone is denoted by ba_+ . Moreover, $\text{ba}_1 = \{\mu \in \text{ba}_+ : \mu(\Omega) = 1\}$ denotes the set of all finitely additive probability measures on $(\Omega, \mathcal{F})$. Each $\mu \in \text{ba}_1$ acts on $\mathcal{B}_b$ through finitely additive integration; for $f \in \mathcal{B}_b$, we write $\mathbb{E}_\mu[f] = \int f d\mu$. Thus, ba_1 may equivalently be viewed as the positive normalized part of the dual of $(\mathcal{B}_b, \|\cdot\|_\infty)$. For $\mu \in \text{ba}_1$, we will always let $\mu = \mu_c + \mu_p$ denote its Yosida-Hewitt decomposition, where $\mu_c \in \mathcal{M}_+$ is countably additive and μ_p is purely finitely additive. Moreover, $\overline{\text{co}}^*$ denotes the weak-* closed convex hull, that is, for a set $\mathcal{A} \subseteq \text{ba}$, $\overline{\text{co}}^*(\mathcal{A})$ denotes the closure of the convex hull of $\mathcal{A}$, taken in the space ba with the topology $\sigma(\text{ba}, \mathcal{B}_b)$.

Remark 2.1. Unless specified otherwise, no topology is imposed on Ω ; we work on a general measurable space $(\Omega, \mathcal{F})$. If Ω is a topological space and $\mathcal{F}$ is its Borel sigma algebra, then the usual weak topology on $\mathcal{M}$ is $\sigma(\mathcal{M}, C_b)$, where C_b denotes the family of bounded continuous functions. This topology is coarser than the setwise topology $\sigma(\mathcal{M}, \mathcal{B}_b)$ since $C_b \subseteq \mathcal{B}_b$. Thus setwise convergence implies weak convergence, but not conversely; for example, on $\mathbb{R}$, $\delta_{1/n}$ converges weakly to δ_0 , but not setwise. On the other hand, the total variation topology is finer than $\sigma(\mathcal{M}, \mathcal{B}_b)$. Indeed,

$$\|R\|_{\text{TV}} = \sup_{\|f\|_\infty \leq 1} \left| \int f dR \right|,$$

so total variation convergence is uniform convergence over the unit ball of $(\mathcal{B}_b, \|\cdot\|_\infty)$, whereas convergence in $\sigma(\mathcal{M}, \mathcal{B}_b)$ is only pointwise convergence against each fixed $f \in \mathcal{B}_b$.

For a nonnegative measurable function $f : \Omega \rightarrow [0, \infty]$, and $\mu \in \text{ba}_1$, we define the extended finitely additive integral as $\mathbb{E}_\mu[f] = \sup_{n \in \mathbb{N}} \mathbb{E}_\mu[f \wedge n] \in [0, \infty]$; equivalently, $\mathbb{E}_\mu[f] = \sup\{\mathbb{E}_\mu[g] : g \in \mathcal{B}_b, 0 \leq g \leq f\}$. Clearly, this definition coincides with the familiar Lebesgue integral whenever $\mu \in \mathcal{M}_1 \subseteq \text{ba}_1$. Let us adopt the conventions $\log 0 = -\infty$ and $\log \infty = \infty$. Whenever $\mathbb{E}_\mu[(\log f)^-] < \infty$, we set $\mathbb{E}_\mu[\log f] = \mathbb{E}_\mu[(\log f)^+] - \mathbb{E}_\mu[(\log f)^-] \in (-\infty, \infty]$. Whenever $\mathbb{E}_\mu[(\log f)^-] = \infty$, we define $\mathbb{E}_\mu[\log f] = -\infty$. Thus, by convention, this quantity equals $-\infty$ whenever the negative logarithmic part has infinite finitely additive integral.

The following convention is used when we evaluate the expression $q \log(q/p)$ for $q \in [0, 1]$ and $p \geq 0$:

$$q \log \frac{q}{p} = \begin{cases} 0, & q = 0, \\ \infty, & p = 0 \text{ and } q > 0. \end{cases} \quad (9)$$

This ensures that $q \log(q/p)$ is jointly convex and lower semicontinuous in (p, q) .

Definition 2.2 (Relative entropy over ba_1). We write Π for the set of all finite measurable partitions of Ω . For any $\nu \in \text{ba}_1$, $\mu \in \text{ba}_+$, and finite measurable partition $\pi = \{A_1, \dots, A_n\}$ we define, using the convention (9),

$$H_\pi(\nu \mid \mu) = \sum_{i=1}^n \nu(A_i) \log \frac{\nu(A_i)}{\mu(A_i)}.$$

Their Kullback–Leibler divergence or relative entropy is then given by

$$H(\nu \mid \mu) = \sup_{\pi \in \Pi} H_\pi(\nu \mid \mu).$$

For $Q \in \mathcal{M}_1$ and $P \in \mathcal{M}_+$, it can be shown that Definition 2.2 reduces to

$$H(Q \mid P) = \begin{cases} \int \log \left(\frac{dQ}{dP} \right) dQ, & Q \ll P, \\ \infty, & Q \not\ll P, \end{cases}$$

which we note can take negative values when $P(\Omega) > 1$.

Even for finitely additive probability measures, zero relative entropy implies equality.

Lemma 2.3. *For $\mu, \nu \in \text{ba}_1$, $H(\nu \mid \mu) \geq 0$ with equality if and only if $\mu = \nu$.*

All proofs for the statements in this section are deferred to Appendix A.

There is also a Donsker–Varadhan dual representation of relative entropy.

Lemma 2.4. *For $\nu \in \text{ba}_1$ and $\mu \in \text{ba}_+$, we have*

$$H(\nu \mid \mu) = \sup_{f \in \mathcal{B}_b} \left(\int f d\nu - \log \int e^f d\mu \right).$$

Moreover, for every measurable $f : \Omega \rightarrow [-\infty, \infty]$ such that $0 < \int e^f d\mu < \infty$, we have

$$\int f d\nu - \log \int e^f d\mu \leq H(\nu \mid \mu).$$

Proposition 2.5. *For $Q \in \mathcal{M}_1$ and $\mu \in \text{ba}_+$, we have $H(Q \mid \mu) = H(Q \mid \mu_c)$. For $\nu \in \text{ba}_1 \setminus \mathcal{M}_1$ and $P \in \mathcal{M}_+$, we have $H(\nu \mid P) = \infty$. Thus, for $\nu \in \text{ba}_1$ and $P \in \mathcal{M}_+$, $H(\nu \mid P) < \infty$ implies that $\nu \in \mathcal{M}_1$ and $\nu \ll P$.*

Last, we note that for simple alternatives, it is enough to optimize over *bounded* e-variables.

Lemma 2.6. *For any $Q \in \mathcal{M}_1$, $\sup_{X \in \mathcal{E}} \mathbb{E}_Q[\log X] = \sup_{X \in \mathcal{E}_b} \mathbb{E}_Q[\log X]$.*

Note however that the supremum on the left-hand side is achieved (by the numeraire), but on the right-hand side in general it is not.

3 Strong duality for bounded GROW and REGROW

This section will present and prove the main strong duality result. The reader may find the following weak duality useful for intuition for the results that follow:

$$\begin{aligned} \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] &= \sup_{X \in \mathcal{E}} \inf_{Q \in \text{co}(\mathcal{Q})} \mathbb{E}_Q[\log X] \leq \inf_{Q \in \text{co}(\mathcal{Q})} \sup_{X \in \mathcal{E}} \mathbb{E}_Q[\log X] \\ &= \inf_{Q \in \text{co}(\mathcal{Q})} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P) \leq \inf_{Q \in \text{co}(\mathcal{Q})} \inf_{P \in \text{co}(\mathcal{P})} H(Q \mid P), \end{aligned} \quad (10)$$

and note that $\mathcal{E}$ could have been replaced by $\mathcal{E}_b$ above. Our aim is to establish when *strong duality* holds, i.e., when we have equality in suitably adjusted versions of the above display.

Theorem 3.1. *The following strong duality holds:*

$$\sup_{X \in \mathcal{E}_b} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = \min_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \mu), \quad (11)$$

even if one side, and then also the other side, equals infinity. Further, there exists a pair $(\mu^*, \nu^*) \in \overline{\text{co}}^*(\mathcal{P}) \times \overline{\text{co}}^*(\mathcal{Q})$ that achieves the minimum on the right-hand side.

This theorem follows from Theorem 3.4 by setting $\xi = 0$ therein (in which case $\xi_Q^{**} = 0$ on $\overline{\text{co}}^*(\mathcal{Q})$).

One may ask whether considering the effective null or closures in ba is really necessary. Here we present an example where $\mathcal{Q}$ is a singleton, $\mathcal{P}$ is convex, $\inf_{P \in \mathcal{P}} H(Q \mid P)$ is achieved in $\mathcal{P}$ and yet this value is vastly different from $\inf_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(Q \mid \mu)$ and from $\inf_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P)$.

Example 3.2. Let $\Omega = [0, 1]$. Let $Q = U$ be the uniform distribution. Let $\mathcal{P}$ be the convex hull of P_0 and all $\{\delta_z : z \in [0, 1]\}$, where δ_z denotes a Dirac delta mass at z and $P_0 = \varepsilon U + (1 - \varepsilon)\delta_0$ for some small constant $\varepsilon > 0$. Clearly, P_0 achieves the infimum $\inf_{P \in \mathcal{P}} H(Q \mid P)$, which can be made arbitrarily large by letting ε tend to zero. However, $\mathcal{P}_{\text{eff}}$ contains all distributions on $[0, 1]$, and thus contains Q , causing $\inf_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P) = 0$. Larsson et al. [2026b, Theorem 3.3] shows that $\mathcal{P}_{\text{eff}} \cap \mathcal{M}_1 = \overline{\text{co}}^*(\mathcal{P}) \cap \mathcal{M}_1$ and so $Q \in \overline{\text{co}}^*(\mathcal{P})$, yielding $\inf_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(Q \mid \mu) = 0$.

The following example shows that every minimizing pair (μ^*, ν^*) on the right-hand side of (11) can in fact be purely finitely additive.

Example 3.3. Let $\Omega = \mathbb{N}$ be equipped with its power set sigma algebra. For each $n \in \mathbb{N}$, define

$$P_n = \frac{1}{n} \sum_{k=1}^n \delta_k, \quad Q_n = \frac{1}{n} \sum_{k=2}^{n+1} \delta_k$$

and set $\mathcal{P} = \{P_n : n \in \mathbb{N}\}$ and $\mathcal{Q} = \{Q_n : n \in \mathbb{N}\}$. Then the right-hand side of (11) is equal to zero. Moreover, every minimizing pair $(\mu^*, \nu^*) \in \overline{\text{co}}^*(\mathcal{P}) \times \overline{\text{co}}^*(\mathcal{Q})$ satisfies $\mu^* = \nu^* \in \overline{\text{co}}^*(\mathcal{P}) \cap \overline{\text{co}}^*(\mathcal{Q})$. Every such common minimizer is purely finitely additive. In particular, there is no minimizing pair with either $\mu^* \in \mathcal{M}_1$ or $\nu^* \in \mathcal{M}_1$. These claims are argued in Appendix B.

Grünwald et al. [2024a] point out that G can be too pessimistic as a criterion, as it is a worst case over all alternatives, and thus any (approximately) GROW e-variable may achieve poor e-power against an “easy” Q in order to preserve optimal e-power against a worst-case Q . To address this, they suggest the REGROW criterion, which normalizes the objective by the best achievable growth rate for that alternative. In this sense, a REGROW e-variable attempts to achieve (as much as possible) nearly optimal e-power for every $Q \in \mathcal{Q}$, and thus is a possibly more appropriate criterion to work with, motivating us to present the following theorem (for general bounded offsets ξ).

Theorem 3.4. Let $\xi : \mathcal{Q} \rightarrow \mathbb{R}$ be a bounded function, and define its concave biconjugate ξ_Q^{**} for $\nu \in \text{ba}_+$ by

$$\xi_Q^{**}(\nu) = \inf_{f \in \mathcal{B}_b} \left(\int f d\nu - \inf_{Q \in \mathcal{Q}} (\mathbb{E}_Q[f] - \xi(Q)) \right). \quad (12)$$

Then the following strong duality holds:

$$\sup_{X \in \mathcal{E}_b} \inf_{Q \in \mathcal{Q}} (\mathbb{E}_Q[\log X] - \xi(Q)) = \min_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} (H(\nu \mid \mu) - \xi_Q^{**}(\nu)), \quad (13)$$

even if one side, and then also the other side, equals infinity. Further, there exists a pair $(\mu^*, \nu^*) \in \overline{\text{co}}^*(\mathcal{P}) \times \overline{\text{co}}^*(\mathcal{Q})$ that achieves the minimum on the right-hand side.

Before proving this theorem we provide a remark and some corollaries and auxiliary lemmata.

Remark 3.5. Since the offset function ξ will typically be convex rather than concave, it cannot be used directly in the minimax argument. We therefore work with its concave biconjugate, given by (12). Let us write, for $Q \in \mathcal{Q}$,

$$G(Q) = G_b(Q) = \sup_{X \in \mathcal{E}_b} \int \log X dQ.$$

For $\mathcal{Q} = \{Q\}$, this agrees both with the GROW value G and the bounded GROW value G_b in (1) and (2), respectively. It is the maximal e-power against the simple alternative ν and coincides with the GRO value in the terminology of Grünwald et al. [2024a]. Let us now assume

$$\sup_{Q \in \mathcal{Q}} G(Q) < \infty$$

and set $\xi = G$. Then, the value in (13) presents REGROW (for bounded e-variables):

$$\sup_{X \in \mathcal{E}_b} \inf_{Q \in \mathcal{Q}} (\mathbb{E}_Q[\log X] - G(Q)) = \min_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} (H(\nu | \mu) - G_Q^{**}(\nu)) = \min_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} (G(\nu) - G_Q^{**}(\nu)),$$

where the final equality follows by Lemma 3.10, presented below.

Corollary 3.6. *For any $Q \in \mathcal{M}_1$,*

$$\min_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) = \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(Q | \mu).$$

Proof. We get the chain of equalities

$$\min_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) = \sup_{X \in \mathcal{E}} \mathbb{E}_Q[\log X] = \sup_{X \in \mathcal{E}_b} \mathbb{E}_Q[\log X] = \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(Q | \mu),$$

where the first equality follows from Larsson et al. [2025], the second from Lemma 2.6, and the last by taking $\mathcal{Q} = \{Q\}$ in (11). $\square$

There is also an interesting corollary regarding the existence of a *nontrivial* test, which we define as any test whose worst-case power exceeds its worst-case level. To set the stage, a test is a measurable function ϕ whose range is $[0, 1]$, its worst-case type-I error is $\sup_{P \in \mathcal{P}} \mathbb{E}_P[\phi]$ and its worst-case power is $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\phi]$. A classical theorem by Kraft [1955] (credited also to Le Cam) asserts that *if $\mathcal{P} \cup \mathcal{Q}$ are dominated by a common reference measure*, then a nontrivial test exists if and only if $\text{co}(\mathcal{P})$ and $\text{co}(\mathcal{Q})$ are separated in the total variation distance. Recently, Larsson et al. [2026b] proved that the aforementioned reference measure assumption can be dropped if $\text{co}(\mathcal{P})$ and $\text{co}(\mathcal{Q})$ are replaced by $\overline{\text{co}}^*(\mathcal{P})$ and $\overline{\text{co}}^*(\mathcal{Q})$, respectively. Our theorem above delivers the same corollary.

Corollary 3.7. *A nontrivial test for $\mathcal{P}$ against $\mathcal{Q}$ exists if and only if $d_{\text{TV}}(\overline{\text{co}}^*(\mathcal{P}), \overline{\text{co}}^*(\mathcal{Q})) > 0$.*

Proof. The sets $\overline{\text{co}}^*(\mathcal{P})$ and $\overline{\text{co}}^*(\mathcal{Q})$ are weak-* compact, and the map $(\mu, \nu) \mapsto \|\mu - \nu\|_{\text{TV}}$ is weak-* lower semicontinuous. Hence the total variation distance between the two sets is attained. Therefore this distance is positive if and only if $\overline{\text{co}}^*(\mathcal{P})$ and $\overline{\text{co}}^*(\mathcal{Q})$ are disjoint.

By Ramdas and Wang [2025], a nontrivial test exists if and only if there exists $X \in \mathcal{E}_b$ such that $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] > 0$. Equivalently, the left-hand side of (11) is strictly positive, and thus so is the right-hand side. Since relative entropy between two measures in ba_1 equals zero if and only if they are equal by Lemma 2.3, this is equivalent to $\overline{\text{co}}^*(\mathcal{P})$ and $\overline{\text{co}}^*(\mathcal{Q})$ being disjoint, and this concludes the proof. $\square$

Before proving Theorem 3.4, we establish a few facts concerning the concave biconjugate.

Lemma 3.8. *For $\mathcal{A} \subseteq \text{ba}_+$ nonempty, suppose $\xi : \mathcal{A} \rightarrow \mathbb{R}$ is bounded and define $\xi_{\mathcal{A}}^{**}$ as in (12). Then $\xi_{\mathcal{A}}^{**}$ is weak-* upper semicontinuous and concave on $\overline{\text{co}}^*(\mathcal{A})$, and for every $g \in \mathcal{B}_b$,*

$$\inf_{\nu \in \mathcal{A}} \left(\int g d\nu - \xi(\nu) \right) = \inf_{\nu \in \overline{\text{co}}^*(\mathcal{A})} \left(\int g d\nu - \xi_{\mathcal{A}}^{**}(\nu) \right). \quad (14)$$

Proof. For $g \in \mathcal{B}_b$, define the concave conjugate of ξ by

$$a_{\xi}(g) = \inf_{\nu \in \mathcal{A}} \left(\int g d\nu - \xi(\nu) \right)$$

and set

$$\psi_g(\nu) = \int g d\nu - a_{\xi}(g), \quad \nu \in \overline{\text{co}}^*(\mathcal{A}).$$

Then ψ_g is affine and weak-* continuous on $\overline{\text{co}}^*(\mathcal{A})$ and we have $\xi_{\mathcal{A}}^{**} = \inf_{g \in \mathcal{B}_b} \psi_g$. The pointwise infimum of continuous affine functions is upper semicontinuous and concave. Hence $\xi_{\mathcal{A}}^{**}$ is weak-* upper semicontinuous and concave on $\overline{\text{co}}^*(\mathcal{A})$.

It remains to prove (14). Fix $g \in \mathcal{B}_b$. From the definition of $\xi_{\mathcal{A}}^{**}$,

$$a_{\xi}(g) \leq \inf_{\nu \in \overline{\text{co}}^*(\mathcal{A})} \left(\int g d\nu - \xi_{\mathcal{A}}^{**}(\nu) \right). \quad (15)$$

For the reverse inequality, first note that for every $\nu_0 \in \mathcal{A}$ and every $h \in \mathcal{B}_b$,

$$a_{\xi}(h) = \inf_{\nu \in \mathcal{A}} \left(\int h d\nu - \xi(\nu) \right) \leq \int h d\nu_0 - \xi(\nu_0).$$

Thus

$$\int h d\nu_0 - a_{\xi}(h) \geq \xi(\nu_0).$$

Taking the infimum over $h \in \mathcal{B}_b$ gives $\xi_{\mathcal{A}}^{**}(\nu_0) \geq \xi(\nu_0)$. Consequently,

$$\inf_{\nu \in \overline{\text{co}}^*(\mathcal{A})} \left(\int g d\nu - \xi_{\mathcal{A}}^{**}(\nu) \right) \leq \inf_{\nu \in \mathcal{A}} \left(\int g d\nu - \xi_{\mathcal{A}}^{**}(\nu) \right) \leq \inf_{\nu \in \mathcal{A}} \left(\int g d\nu - \xi(\nu) \right) = a_{\xi}(g).$$

Combining this with (15) proves (14). $\square$

The following short lemma will be useful below.

Lemma 3.9. *Let $\mathcal{A} \subseteq \text{ba}_1$ be nonempty. Let $\xi : \mathcal{A} \rightarrow \mathbb{R}$ be a bounded function, and define*

$$\Gamma = \left\{ f \in \mathcal{B}_b : \sup_{P \in \mathcal{P}} \mathbb{E}_P[e^f] \leq 1 \right\}. \quad (16)$$

Then Γ is convex and

$$\sup_{X \in \mathcal{E}_b} \inf_{\nu \in \mathcal{A}} (\mathbb{E}_{\nu}[\log X] - \xi(\nu)) = \sup_{f \in \Gamma} \inf_{\nu \in \mathcal{A}} (\mathbb{E}_{\nu}[f] - \xi(\nu)).$$

Proof. We first observe that Γ is convex due to Hölder's inequality. The inequality " $\geq$ " is clear. For the reverse inequality, fix $X \in \mathcal{E}_b$ and $\varepsilon \in (0, 1)$, and set $X_{\varepsilon} = \varepsilon + (1 - \varepsilon)X \in \mathcal{E}_b$. Then $f_{\varepsilon} = \log X_{\varepsilon} \in \Gamma$. Since $X_{\varepsilon} \geq (1 - \varepsilon)X$, we have $f_{\varepsilon} \geq \log(1 - \varepsilon) + \log X$. Therefore

$$\inf_{\nu \in \mathcal{A}} \left(\int f_{\varepsilon} d\nu - \xi(\nu) \right) \geq \inf_{\nu \in \mathcal{A}} \left(\int \log X d\nu - \xi(\nu) \right) + \log(1 - \varepsilon).$$

Sending ε to zero and taking the supremum over $X \in \mathcal{E}_b$ proves the reverse inequality. $\square$

In preparation for the proof of Theorem 3.4, we also provide a simple-alternative duality. The main differences between this result and (4) are that ν is allowed to be in ba_1 (instead of just $\mathcal{M}_1$), and the right-hand side minimizes over $\overline{\text{co}}^*(\mathcal{P})$ instead of $\mathcal{P}_{\text{eff}}$.

Lemma 3.10. *Let Γ be as in (16). For $\nu \in \text{ba}_1$,*

$$\sup_{X \in \mathcal{E}_b} \int \log X \, d\nu = \sup_{f \in \Gamma} \int f \, d\nu = \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \mu).$$

Proof. By Lemma 3.9, applied with $\xi \equiv 0$, it suffices to argue

$$\sup_{f \in \Gamma} \int f \, d\nu = \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \mu). \quad (17)$$

To make headway, for $\mu \in \overline{\text{co}}^*(\mathcal{P})$ and $f \in \mathcal{B}_b$, define

$$\Phi(f, \mu) = \int f \, d\nu - \log \int e^f \, d\mu.$$

For each fixed $\mu \in \overline{\text{co}}^*(\mathcal{P})$, the map $f \mapsto \Phi(f, \mu)$ is concave on $\mathcal{B}_b$. Indeed, $f \mapsto \int f \, d\nu$ is affine, while $f \mapsto \log \int e^f \, d\mu$ is convex by Hölder's inequality. Moreover, this map is continuous with respect to the supremum norm. For each fixed $f \in \mathcal{B}_b$, the map $\mu \mapsto \Phi(f, \mu)$ is convex and continuous on $\overline{\text{co}}^*(\mathcal{P})$, since $\mu \mapsto \int e^f \, d\mu$ is affine and continuous, takes values in $(0, \infty)$, and $x \mapsto -\log x$ is continuous and convex on $(0, \infty)$.

Since $\overline{\text{co}}^*(\mathcal{P})$ is compact by the Banach–Alaoglu theorem and convex, and $\mathcal{B}_b$ is convex, Sion's minimax theorem applies. Together with the variational formula for relative entropy (Lemma 2.4), this yields

$$\begin{aligned} \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \mu) &= \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} \sup_{f \in \mathcal{B}_b} \Phi(f, \mu) = \sup_{f \in \mathcal{B}_b} \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} \Phi(f, \mu) \\ &= \sup_{f \in \mathcal{B}_b} \left(\int f \, d\nu - \sup_{\mu \in \overline{\text{co}}^*(\mathcal{P})} \log \int e^f \, d\mu \right). \end{aligned} \quad (18)$$

For $f \in \mathcal{B}_b$ set $c(f) = \sup_{\mu \in \overline{\text{co}}^*(\mathcal{P})} \log \int e^f \, d\mu$. Then we have $f - c(f) \in \Gamma$, yielding

$$\int f \, d\nu - c(f) = \int (f - c(f)) \, d\nu \leq \sup_{g \in \Gamma} \int g \, d\nu.$$

From (18) we hence get

$$\min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \mu) \leq \sup_{g \in \Gamma} \int g \, d\nu. \quad (19)$$

Conversely, suppose that $g \in \Gamma$. We know that $c(g) \leq 0$. Hence (18) yields

$$\int g \, d\nu \leq \int g \, d\nu - c(g) \leq \sup_{f \in \mathcal{B}_b} \left(\int f \, d\nu - c(f) \right) = \min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \mu).$$

Together with (19), this yields (17), concluding the proof. $\square$

Proof of Theorem 3.4. Let Γ be as in (16). We first note

$$\sup_{X \in \mathcal{E}_b} \inf_{\mathbf{Q} \in \mathcal{Q}} (\mathbb{E}_{\mathbf{Q}}[\log X] - \xi(\mathbf{Q})) = \sup_{f \in \Gamma} \inf_{\mathbf{Q} \in \mathcal{Q}} (\mathbb{E}_{\mathbf{Q}}[f] - \xi(\mathbf{Q})) = \sup_{f \in \Gamma} \inf_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \left(\int f \, d\nu - \xi_{\mathcal{Q}}^{**}(\nu) \right)$$

by Lemmata 3.9 and 3.8. We now intend to apply Sion's minimax theorem. To this end, recall that Γ is convex by Lemma 3.9 and $\overline{\text{co}}^*(\mathcal{Q})$ is compact by the Banach–Alaoglu theorem, respectively. We next define

$$L(f, \nu) = \int f d\nu - \xi_{\mathcal{Q}}^{**}(\nu).$$

By Lemma 3.8, $\nu \mapsto L(f, \nu)$ is lower semicontinuous and convex on $\overline{\text{co}}^*(\mathcal{Q})$. Moreover, $f \mapsto L(f, \nu)$ is linear and continuous. Further, $\xi_{\mathcal{Q}}^{**}(\nu) \in [\inf_{\mathbf{Q} \in \mathcal{Q}} \xi(\mathbf{Q}), \sup_{\mathbf{Q} \in \mathcal{Q}} \xi(\mathbf{Q})]$, and thus $\xi_{\mathcal{Q}}^{**}$ is real-valued on $\overline{\text{co}}^*(\mathcal{Q})$. Therefore, Sion's minimax theorem applies and we obtain

$$\sup_{f \in \Gamma} \inf_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \left(\int f d\nu - \xi_{\mathcal{Q}}^{**}(\nu) \right) = \inf_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \sup_{f \in \Gamma} \left(\int f d\nu - \xi_{\mathcal{Q}}^{**}(\nu) \right).$$

By Lemma 3.10, we get (13). Finally, attainment follows from lower semicontinuity and compactness of $\overline{\text{co}}^*(\mathcal{P}) \times \overline{\text{co}}^*(\mathcal{Q})$. $\square$

To provide further intuition for the concave biconjugate, we discuss the simpler case of a finite full simplex.

Example 3.11. Suppose that $\Omega = \{1, \dots, n\}$, and let $\mathbf{P} \in \Delta_n$ have full support, that is, $p_i = \mathbf{P}(\{i\}) > 0$ for every $i = 1, \dots, n$. Let $\mathcal{P} = \{\mathbf{P}\}$ and $\mathcal{Q} = \Delta_n$. For $\mathbf{Q} \in \mathcal{Q}$, we set again

$$\mathbf{G}(\mathbf{Q}) = H(\mathbf{Q} \mid \mathbf{P}).$$

We first compute the concave biconjugate of $\mathbf{G}$ on Δ_n . Write $q_i = \mathbf{Q}(\{i\})$. Then

$$\mathbf{G}(\mathbf{Q}) = \sum_{i=1}^n q_i \log \frac{q_i}{p_i}.$$

For $g = (g_1, \dots, g_n) \in \mathbb{R}^n$, the concave conjugate of $\mathbf{G}$ is

$$a_{\mathbf{G}}(g) = \inf_{q \in \Delta_n} \left(\sum_{i=1}^n q_i (g_i + \log p_i) - \sum_{i=1}^n q_i \log q_i \right).$$

Since $-\sum_i q_i \log q_i \geq 0$, this expression is bounded below by $\min_i (g_i + \log p_i)$, and this lower bound is attained by choosing q to be a Dirac mass at a minimizer of $g_i + \log p_i$. Hence

$$a_{\mathbf{G}}(g) = \min_{1 \leq i \leq n} (g_i + \log p_i).$$

Therefore

$$\mathbf{G}_{\Delta_n}^{**}(\mathbf{Q}) = \inf_{g \in \mathbb{R}^n} \left(\sum_{i=1}^n q_i g_i - \min_{1 \leq i \leq n} (g_i + \log p_i) \right) = - \sum_{i=1}^n q_i \log p_i + \inf_{h \in \mathbb{R}^n} \left(\sum_{i=1}^n q_i h_i - \min_i h_i \right),$$

where $h_i = g_i + \log p_i$. The last infimum is zero, since $\sum_i q_i h_i \geq \min_i h_i$ and equality is attained by taking $h_1 = \dots = h_n$. Thus

$$\mathbf{G}_{\Delta_n}^{**}(\mathbf{Q}) = - \sum_{i=1}^n \mathbf{Q}(\{i\}) \log p_i, \quad \mathbf{Q} \in \Delta_n.$$

Consequently,

$$\mathbf{G}(\mathbf{Q}) - \mathbf{G}_{\Delta_n}^{**}(\mathbf{Q}) = \sum_{i=1}^n q_i \log q_i.$$

This quantity is minimized over Δ_n at the uniform distribution $\mathbf{U}$, and the minimum value is $-\log n$. Thus the offset relative entropy side of (13) has value $-\log n$.

On the e -variable side, let $X = dU/dP$. Then X is an e -variable for P , and $X(i) = 1/(np_i)$. For $Q \in \Delta_n$,

$$\mathbb{E}_Q[\log X] - G(Q) = \sum_{i=1}^n q_i \log \frac{1}{np_i} - \sum_{i=1}^n q_i \log \frac{q_i}{p_i} = -\log n - \sum_{i=1}^n q_i \log q_i.$$

Taking the infimum over $Q \in \Delta_n$ gives $-\log n$, attained at any vertex of Δ_n . Hence the REGROW value is $-\log n$, and the REGROW e -variable is dU/dP . Notice that the minimizing Q on the e -variable side may be any Dirac mass, whereas the minimizing Q on the offset relative entropy side is the uniform distribution U .

4 Strong duality for unbounded GROW

We have been unable to derive an “assumption-free” theorem for unbounded GROW. The absence of such a theorem forces us to look for special cases where strong duality still holds with unbounded e -variables. While there is no hope of being complete, we present five important settings where strong duality holds. We present these below in the following order, one case per subsection. First, we handle the case of a singleton $\mathcal{P} = \{P\}$ together with convex $\mathcal{Q}$ such that $H(Q | P) < \infty$ for all $Q \in \mathcal{Q}$. We then extend this to the case where a joint information projection (defined later) exists. Third, we handle the case of convex and setwise compact $\mathcal{Q}$. Fourth, we handle the case of weakly compact $\mathcal{P}$ and $\mathcal{Q}$ on a Polish space Ω . Finally, we handle the case of finite Ω . In all cases, the final conclusions avoid finitely additive measures. In all relevant cases, we show using counterexamples that the conclusions of the theorems do not hold without restrictions, justifying our inability to find an assumption-free statement.

4.1 When $\mathcal{P} = \{P\}$ is a singleton

Fix a convex family $\mathcal{Q}$ of probability measures and a simple null $\mathcal{P} = \{P\}$ such that

$$\inf_{Q \in \mathcal{Q}} H(Q | P) < \infty.$$

Recall that then Csiszár’s I-projection (IPr), provided it exists, is the probability measure $Q^{\text{IPr}} \in \mathcal{Q}$ that achieves $\inf_{Q \in \mathcal{Q}} H(Q | P)$ [Csiszár, 1975], so that

$$H(Q^{\text{IPr}} | P) = \inf_{Q \in \mathcal{Q}} H(Q | P).$$

Csiszár proved that the I-projection of P onto $\mathcal{Q}$ exists if $\mathcal{Q}$ is convex and TV-closed and $\inf_{Q \in \mathcal{Q}} H(Q | P) < \infty$, in which case clearly $H(Q^{\text{IPr}} | P) < \infty$. Since the IPr does not always exist, the following generalization was conceived.

Following Topsøe [1979, Theorem 8] and Csiszár [1984], the generalized I-projection (GIPr) Q^{GIPr} of $P \in \mathcal{M}_1$ onto $\mathcal{Q} \subseteq \mathcal{M}_1$ is the unique probability measure that satisfies the following Pythagorean inequality for any $Q \in \mathcal{Q}$:

$$H(Q | Q^{\text{GIPr}}) + \inf_{R \in \mathcal{Q}} H(R | P) \leq H(Q | P). \quad (20)$$

The GIPr Q^{GIPr} was shown by the above authors to always exist for any convex $\mathcal{Q}$ with $\inf_{Q \in \mathcal{Q}} H(Q | P) < \infty$. However, $Q^{\text{GIPr}} \notin \mathcal{Q}$ in general, but (20) implies that it does lie in the I-closure of $\mathcal{Q}$, given by

$$\overline{\mathcal{Q}}^I = \left\{ R \in \mathcal{M}_1 : \inf_{Q \in \mathcal{Q}} H(Q | R) = 0 \right\}.$$

Further, one has

$$H(Q^{\text{GIPr}} \mid P) \leq \inf_{Q \in \mathcal{Q}} H(Q \mid P), \quad (21)$$

where strict inequality can occur as shown by [Csiszár, 1984, Example 3.2]. He also shows that one cannot obtain the generalized I-projection as a minimizer over some closure of $\mathcal{Q}$. Thanks to (21), and our standing assumption that $\inf_{Q \in \mathcal{Q}} H(Q \mid P) < \infty$, we have $H(Q^{\text{GIPr}} \mid P) < \infty$, and thus $Q^{\text{GIPr}} \ll P$, meaning that dQ^{GIPr}/dP is well defined, and it is clearly an e-variable for P . We now show that under some conditions, dQ^{GIPr}/dP is the log-optimal (unbounded) e-variable.

Assumption 4.1. $\inf_{R \in \mathcal{Q}} H(R \mid P) < \infty$ and there exists a random variable X^* that is a P -version of dQ^{GIPr}/dP , such that $\mathbb{E}_Q[\log X^*] \geq \inf_{R \in \mathcal{Q}} H(R \mid P)$ for each $Q \in \mathcal{Q}$.

The above assumption can sometimes be checked by verifying the stronger condition that for every $Q \in \mathcal{Q}$, $H(Q \mid Q^{\text{GIPr}}) < \infty$ holds, which in turn is implied by $H(Q \mid P) < \infty$. (Indeed, one can rearrange the Pythagorean inequality (20) even when $H(Q \mid P) = \infty$, and note that these stronger conditions imply $Q \ll Q^{\text{GIPr}} \ll P$.)

Theorem 4.2 (Singleton null strong duality). *For singleton $\mathcal{P} = \{P\}$ and convex $\mathcal{Q}$ such that Assumption 4.1 holds, the corresponding e-variable X^* achieves the supremum on the left-hand side of the following strong duality:*

$$G = \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^*] = \inf_{Q \in \mathcal{Q}} H(Q \mid P).$$

Further, if the GIPr is an IPr (i.e., it lies in $\mathcal{Q}$), then X^* is the P -a.s. unique GROW e-variable.

Proof. We have

$$\inf_{Q \in \mathcal{Q}} H(Q \mid P) \leq \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^*] \leq \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \leq \inf_{Q \in \mathcal{Q}} H(Q \mid P) < \infty,$$

where the first inequality follows from Assumption 4.1, the second one holds by noting $X^* \in \mathcal{E}$, the third by weak duality (10), and the last by assumption. This yields the strong duality. The uniqueness claim follows by Proposition 4.12. $\square$

An interesting application of this theorem, where $\mathcal{Q}$ is defined by a finite number of moment inequalities, is presented in Section 5.5.

The above result also corrects some inaccuracies in the statement and proof of Proposition 1 in Grünwald et al. [2024b]. In particular, the inequality in their Equation (1.11) would only follow if the GIPr belonged to $\mathcal{Q}$, which need not hold in general. This leads them to conclude in their Equation (1.10) that G equals $H(Q^{\text{GIPr}} \mid P)$, which is not true in general. It is worth noting that their assumption is *weaker* than the usual Pythagorean inequality, because their Equation (1.6) is equivalent to assuming our (20) but with the term $\inf_{R \in \mathcal{Q}} H(R \mid P)$ being replaced by the quantity $H(Q^{\text{GIPr}} \mid P)$, which (21) implies can be smaller. Thus the error leads to a weaker assumption but a stronger conclusion. Example 5.3 provides an explicit counterexample: it satisfies their assumptions but not their conclusion.

It is natural to ask whether we can relax the assumption to only needing $\inf_{Q \in \mathcal{Q}} H(Q \mid P) < \infty$ (which was anyway required for the GIPr to exist). The following example shows that this condition does not suffice, and we can still have

$$\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] < \inf_{Q \in \mathcal{Q}} H(Q \mid P). \quad (22)$$

Example 4.3. Let $\Omega = [0, 1]$, equipped with the Borel sigma algebra, and consider $P = U$, the uniform. Define Q_0 by $dQ_0/dP(\omega) = 2\omega$ so that Q_0 is a probability measure, and set

$$\mathcal{Q} = \text{co}(\{Q_0\} \cup \{\delta_\omega : \omega \in [0, 1]\}).$$

Let X be any e-variable with respect to P . Then necessarily $\inf_{\omega \in [0, 1]} X(\omega) \leq 1$; otherwise $X > 1$ everywhere and hence $\mathbb{E}_P[X] > 1$. Therefore, the left-hand side of (22) equals zero, achieved by the constant e-variable 1. On the other hand, among the elements of $\mathcal{Q}$ the only one absolutely continuous with respect to P is Q_0 , while $H(Q_0 | P) = \log 2 - 1/2 > 0$. Hence the right-hand side of (22) is finite and strictly positive, being attained by Q_0 .

The above example does not have $Q \ll P$ for every $Q \in \mathcal{Q}$. So it leaves open the possibility that assuming $\inf_{Q \in \mathcal{Q}} H(Q | P) < \infty$ along with $Q \ll P$ for every $Q \in \mathcal{Q}$ may possibly suffice. However, such hopes are quickly dashed by an extension of the above example, which due to its length is provided later as Example 5.2.

We end by noting that Assumption 4.1 is not necessary for strong duality to hold. Rather, it additionally ensures that the supremum over e-variables is attained. The following example emphasizes this point. (See also Example 5.5 for a more statistically relevant instance.)

Example 4.4. Let $\Omega = \mathbb{N} \cup \{0\}$, equipped with its power set sigma algebra, and let $\mathcal{P} = \{P\}$ with $P(n) = 2^{-(n+1)}$ for all $n \in \mathbb{N}_0$. Define two probability measures $Q_0 = \delta_0$ and $Q_\infty(n) = 1/(n(n+1))$ for $n \in \mathbb{N}$, and set

$$\mathcal{Q} = \text{co}(Q_0, Q_\infty) = \{tQ_0 + (1-t)Q_\infty : t \in [0, 1]\}.$$

We have $H(Q_0 | P) = \log 2$, but $H(Q_\infty | P) = \infty$, so for any $t \in [0, 1]$,

$$H(tQ_0 + (1-t)Q_\infty | P) = \infty.$$

Consequently, Q_0 is the IPr of P , with relative entropy $\log 2$. The likelihood ratio $X^* = dQ_0/dP$ is uniquely determined since P has full support. It satisfies $X^*(0) = 2$ and $X^*(n) = 0$ for all $n \in \mathbb{N}$. Thus, $\mathbb{E}_{Q_\infty}[\log X^*] = -\infty$, in particular $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^*] = -\infty$. To see that strong duality holds, for a small $\varepsilon > 0$, define the e-variable X^ε by $X^\varepsilon(0) = 2(1 - \varepsilon)$ and $X^\varepsilon(n) = \varepsilon Q_\infty(n)/P(n)$. Further, $\mathbb{E}_{Q_\infty}[\log X^\varepsilon] = \infty$, but $\mathbb{E}_{Q_0}[\log X^\varepsilon] = \log(2(1 - \varepsilon))$. Thus, $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^\varepsilon] = \log(2(1 - \varepsilon))$ which approaches $\log 2$ as $\varepsilon \downarrow 0$, meaning that strong duality holds. (In fact, Proposition 4.12, proved later, implies that no e-variable achieves strong duality.)

Remark 4.5. We explain the key idea behind the above example. Let Q^* denote the generalized I-projection of P onto some convex $\mathcal{Q}$ with $I = \inf_{Q \in \mathcal{Q}} H(Q | P) < \infty$. Let X^* be any fixed version of dQ^*/dP ; clearly $X^* \in \mathcal{E}$. Call a $Q \in \mathcal{Q}$ “bad” if $\mathbb{E}_Q[\log X^*] < I$, and “good” otherwise. Suppose there exists an e-variable Y which satisfies $\mathbb{E}_Q[\log Y] = \infty$ for every bad $Q \in \mathcal{Q}$. Then for $X_\delta = (1 - \delta)X^* + \delta Y \in \mathcal{E}$, the mixture with Y provides protection against the bad Q . Since $X_\delta \geq \delta Y$, we have $\mathbb{E}_Q[\log X_\delta] \geq \mathbb{E}_Q[\log Y] + \log \delta = \infty$ for bad Q . And for the “good” Q , $X_\delta \geq (1 - \delta)X^*$ gives us $\mathbb{E}_Q[\log X_\delta] \geq I + \log(1 - \delta)$. Put together, we get that $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X_\delta] \geq I + \log(1 - \delta)$, and letting $\delta \downarrow 0$ yields strong duality. We formalize these ideas further in Theorem 5.4.

4.2 When a joint information projection exists

Larsson et al. [2025] showed that for an arbitrary composite null $\mathcal{P}$ and singleton alternative Q , a log-optimal e-variable X^{num} , called the numeraire, always exists and is Q -almost-surely unique and positive. Then, they define the reverse information projection (RIPr) of Q onto $\mathcal{P}_{\text{eff}}$ as the sub-probability measure defined by $dP^*/dQ = 1/X^{\text{num}}$, and show that $P^* \in \mathcal{P}_{\text{eff}}$.

We define the GIPr (or IPr, if it exists) of a nonzero sub-probability $P \in \mathcal{M}_+$ onto a convex set $\mathcal{Q}$, as the GIPr (or IPr) of $\tilde{P} = P/P(\Omega)$ onto $\mathcal{Q}$ (assuming, as before, that $\inf_{Q \in \mathcal{Q}} H(Q | \tilde{P}) < \infty$). Since the Pythagorean inequality (20) contains $\tilde{P}$ on both sides, one finds that the normalization constant cancels out. Hence the inequality holds for P , despite P being a sub-probability, a fact that we record as a lemma for easier reference.

Lemma 4.6. *Let $\mathcal{Q}$ be convex. For any nonzero sub-probability measure $P \in \mathcal{M}_+$ with $\inf_{Q \in \mathcal{Q}} H(Q | P) < \infty$, its GIPr satisfies (20) verbatim, i.e.,*

$$H(Q | Q^{\text{GIPr}}) + \inf_{R \in \mathcal{Q}} H(R | P) \leq H(Q | P).$$

Now, we can define the joint information projection (JIPr).

Definition 4.7. Given arbitrary $\mathcal{P}$ and convex $\mathcal{Q}$, we will say that a pair $(P^*, Q^*) \in \mathcal{M}_+ \times \mathcal{M}_1$ is a joint information projection (JIPr) if

- (i) $P^* \in \mathcal{P}_{\text{eff}}$ is the RIPr of Q^* onto $\mathcal{P}_{\text{eff}}$;
- (ii) $\inf_{Q \in \mathcal{Q}} H(Q | P^*) < \infty$ and $Q^* \in \overline{\mathcal{Q}}^I$ is the GIPr of P^* onto $\mathcal{Q}$.

Also, assuming (i,ii) is weaker than assuming

- (iii) $(P^*, Q^*) \in \mathcal{P}_{\text{eff}} \times \mathcal{Q}$ achieves $\inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P)$, this value is finite, and $P^* \ll Q^*$.

One has the implication (iii) $\implies$ (i,ii), but in general the reverse implication may fail. We provide an example to illustrate that (i,ii) does not imply (iii), even when the GIPr is an IPr.

Example 4.8. Let $\Omega = \{0, 1, 2\}$ and set $P^* = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2$ and $Q^* = \delta_1$. Define the convex sets

$$\mathcal{P} = \{(1-s)P^* + s\delta_0 : s \in [0, 1]\}; \quad \mathcal{Q} = \{(1-s)Q^* + s\delta_0 : s \in [0, 1]\}.$$

Then the absolutely continuous part of P^* with respect to Q^* is $P_a^* = \frac{1}{2}\delta_1$. We will show that the pair (P_a^*, Q^*) satisfies both (i) and (ii). Since the e-variable $X^* = 2\mathbf{1}_{\{1\}}$ is the reciprocal of dP_a^*/dQ^* , P_a^* is the RIPr of Q^* onto $\mathcal{P}_{\text{eff}}$ [Larsson et al., 2025, Theorem 4.1], so condition (i) holds. Next, Q^* is the IPr of P_a^* onto $\mathcal{Q}$. Indeed, $H(Q^* | P_a^*) = \log 2 < \infty$, whereas every element $Q_s = (1-s)Q^* + s\delta_0$ with $s > 0$ assigns positive mass to $\{0\}$, so $H(Q_s | P_a^*) = \infty$. Hence condition (ii) holds. However, condition (iii) fails. Indeed, $\delta_0 \in \mathcal{P} \cap \mathcal{Q}$, and therefore $\inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) = 0$. On the other hand, $H(Q^* | P_a^*) = \log 2 > 0$. Thus (P_a^*, Q^*) does not attain (iii), even though conditions (i) and (ii) hold.

Another illustration for (i) and (ii) holding but (iii) failing is given in Example 5.3. This example is slightly longer, but also satisfies Assumption 4.9 below, hence Theorem 4.10 applies for that example.

Condition (i) implies that dQ^*/dP^* is an e-variable for $\mathcal{P}$ (it is the numeraire for $\mathcal{P}$ versus Q^*). However, it may not achieve the desired strong duality; recall Example 4.3, where (P, Q_0) was a JIPr pair satisfying condition (iii). Thus, for strong duality to hold, we need an additional condition, analogous to Assumption 4.1.

Assumption 4.9. A JIPr pair (P^*, Q^*) exists and there exists a random variable X^* that is a P^* -version of dQ^*/dP^* , such that $\mathbb{E}_Q[\log X^*] \geq \inf_{R \in \mathcal{Q}} H(R | P^*)$ for each $Q \in \mathcal{Q}$.

As before, one can sometimes check the above by verifying that either $H(Q | Q^*) < \infty$ holds, or the even stronger condition that $H(Q | P^*) < \infty$ holds, for each $Q \in \mathcal{Q}$. (To see that these imply the assumption, simply rearrange (20), which is allowed even when $H(Q | P^*) = \infty$).

Theorem 4.10 (Strong duality for composite null and alternative). *For arbitrary $\mathcal{P}$ and convex $\mathcal{Q}$, suppose there exists a joint information projection (P^*, Q^*) such that Assumption 4.9 holds. Then the corresponding e -variable X^* is a GROW e -variable, which satisfies the strong duality*

$$G = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^*] = \inf_{Q \in \mathcal{Q}} H(Q | P^*) = \inf_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q | P).$$

Further, if Q^* is an IPr (i.e., it lies in $\mathcal{Q}$), then X^* is the P^* -a.s. unique GROW e -variable.

The above theorem eliminates certain conditions in Grünwald et al. [2024a, Theorem 1, First Generalisation], such as assuming that $H(Q | Q') < \infty$ for all $Q, Q' \in \mathcal{Q}$, and also that these have full support.

We note above that if the GIPr actually was an IPr (meaning that the infimum relative entropy was achieved by $Q^* \in \mathcal{Q}$), then we can add to the strong duality a further equality to $H(Q^* | P^*)$.

Proof of Theorem 4.10. We begin with the weak duality

$$\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \leq \inf_{Q \in \mathcal{Q}} \sup_{X \in \mathcal{E}} \mathbb{E}_Q[\log X] = \inf_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) \leq \inf_{Q \in \mathcal{Q}} H(Q | P^*),$$

where the sole equality follows from (4), and the two inequalities are immediate.

In the opposite direction, we have

$$\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \geq \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^*] \geq \inf_{Q \in \mathcal{Q}} H(Q | P^*),$$

where the first inequality is immediate, and the second follows by taking infimum over $\mathcal{Q}$ in Assumption 4.9. Combining the two displays completes the proof of strong duality. The uniqueness claim follows by Proposition 4.12. $\square$

Example 4.11. Let $\mathcal{P}$ be $\{N(m, 1) : m \leq 0\}$ and $\mathcal{Q}$ be (the convex hull of) $\{N(m, 1) : m \geq 1\}$. Then the JIPr pair is given by $(N(0, 1), N(1, 1))$ and the GIPr is actually an IPr. In this case, the JIPr pair is also a least favorable distribution pair in the sense of Huber and Strassen [1973]; indeed Saha and Ramdas [2025] already applies to yield the GROW e -variable, which is the likelihood ratio of the JIPr pair.

Now take $\mathcal{P}$ to be set of all 1-sub-Gaussian distributions with nonpositive mean, and $\mathcal{Q}$ to be (the convex hull of) all 1-sub-Gaussian distributions with mean at least 1 and finite relative entropy to the standard Gaussian. For this nonparametric class, a least favorable distribution pair in the sense of Huber–Strassen does not exist, and it is not true that $H(Q | Q') < \infty$ for all $Q, Q' \in \mathcal{Q}$ (take, for example, $Q = U[1, 2]$, $Q' = U[3, 4]$), so the results in Grünwald et al. [2024a] do not apply. Nevertheless, the JIPr pair is still $(N(0, 1), N(1, 1))$, and the conditions of Theorem 4.10 are fulfilled, so the GROW e -variable is still their likelihood ratio.

Proposition 4.12. Assume that the following strong duality holds:

$$G = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) < \infty,$$

and that there exists $(P^*, Q^*) \in \mathcal{P}_{\text{eff}} \times \mathcal{Q}$ such that $H(Q^* | P^*) = \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P)$. If any e -variable X^* achieves the supremum above, then we must have $X^* = dQ^*/dP^*$, P^* -a.s.

Proof. Since $\mathbb{E}_{Q^*}[\log X^*] \geq G$ because $X^* \in \mathcal{E}$ witnesses strong duality, applying the Donsker–Varadhan inequality (Lemma 2.4) yields

$$G \leq \mathbb{E}_{Q^*}[\log X^*] \leq H(Q^* | P^*) + \log \mathbb{E}_{P^*}[X^*] \leq H(Q^* | P^*) = G.$$

It follows that equality holds throughout. In particular, $\log \mathbb{E}_{P^*}[X^*] = 0$, hence $\mathbb{E}_{P^*}[X^*] = 1$, and $\mathbb{E}_{Q^*}[\log X^*] = H(Q^* | P^*)$.

Define now a probability measure Q' by $\frac{dQ'}{dP^*} = X^*$. Since $H(Q^* | P^*) < \infty$, we have $Q^* \ll P^*$. Therefore, using the chain rule for relative entropy,

$$H(Q^* | Q') = H(Q^* | P^*) - \mathbb{E}_{Q^*}[\log X^*] = 0.$$

Thus $Q^* = Q'$, proving the claim. $\square$

Remark 4.13. If Q^* in Proposition 4.12 is not in $\mathcal{Q}$ but in the I-closure $\overline{\mathcal{Q}}^I$ only, then the conclusion of the statement still holds provided one assumes in addition $\mathbb{E}_{Q^*}[\log X^*] \geq G$.

We recall Example 4.4, which points out that Assumption 4.9 is not necessary. Indeed, this example shows that the JIPr can exist (satisfying the stronger condition (iii)) — it uniquely equals $(\delta_0/2, Q_0)$ — and strong duality can hold, but the likelihood ratio of the JIPr pair does not achieve the strong duality.

4.3 When $\mathcal{Q}$ is setwise compact

Definition 4.14. The *setwise topology* on $\mathcal{M}$ is the initial topology $\sigma(\mathcal{M}, \mathcal{B}_b)$ induced by the maps

$$R \mapsto \int f dR, \quad R \in \mathcal{M}, f \in \mathcal{B}_b.$$

A set $K \subseteq \mathcal{M}$ is called *setwise closed/compact* if it is closed/compact in $(\mathcal{M}, \sigma(\mathcal{M}, \mathcal{B}_b))$.

The setwise topology is the subspace topology induced by the weak-* topology in ba :

$$\sigma(\mathcal{M}, \mathcal{B}_b) = \sigma(\text{ba}, \mathcal{B}_b)|_{\mathcal{M}}.$$

Indeed, $\sigma(\text{ba}, \mathcal{B}_b)$ is the coarsest topology on ba making all maps $\nu \mapsto \int f d\nu, f \in \mathcal{B}_b$, continuous. Restricting these maps to $\mathcal{M} \subseteq \text{ba}$ gives exactly the maps defining $\sigma(\mathcal{M}, \mathcal{B}_b)$.

If we equip $\mathcal{M}$ with the total variation (TV) norm $\|\cdot\|_{\text{TV}}$ and its norm topology, TV convergence implies setwise convergence, because $\|R_n - R\|_{\text{TV}} \rightarrow 0$ implies $R_n \rightarrow R$ setwise. In addition, if we further suppose that Ω is Polish and $\mathcal{F}$ is the Borel sigma algebra, then setwise convergence implies weak convergence: clearly, $R_n \rightarrow R$ in $\sigma(\mathcal{M}, \mathcal{B}_b)$ implies that $R_n \rightarrow R$ in $\sigma(\mathcal{M}, C_b)$. All three topologies are Hausdorff.

Recall that $\mathcal{Q} \subseteq \mathcal{M}_1$ is said to be *uniformly absolutely continuous* with respect to a finite $R \in \mathcal{M}_+$ if $Q \ll R$ for all $Q \in \mathcal{Q}$, and for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $A \in \mathcal{F}$, $R(A) < \delta$ implies $\sup_{Q \in \mathcal{Q}} Q(A) \leq \varepsilon$. It follows directly from [Bogachev, 2007, Theorem 4.7.25] that $\text{cl}^{\sigma(\mathcal{M}, \mathcal{B}_b)}(\text{co}(\mathcal{Q}))$ is setwise compact if and only if there exists a finite measure $R \in \mathcal{M}_+$ such that $\mathcal{Q}$ is uniformly absolutely continuous with respect to R . From this we see that assuming $\mathcal{Q}$ to be convex and setwise compact is a rather strong assumption.

Two examples of convex and setwise compact $\mathcal{Q}$ are (i) the convex hull of Gaussians with means in $\{-1, 1\}$ and unit variance and (ii) all Bernoulli distributions with parameters in $[0.2, 0.8]$.

We shall use the following two lemmata.

Lemma 4.15. *If $\mathcal{Q}$ is convex and setwise compact, then $\overline{\text{co}}^*(\mathcal{Q}) = \mathcal{Q}$.*

Proof. Since the setwise topology on $\mathcal{M}_1$ is the subspace topology induced by $\sigma(\text{ba}, \mathcal{B}_b)$, the set $\mathcal{Q}$ is compact also as a subset of $(\text{ba}, \sigma(\text{ba}, \mathcal{B}_b))$. The latter space is Hausdorff, so compact subsets are closed. Hence $\mathcal{Q}$ is $\sigma(\text{ba}, \mathcal{B}_b)$ -closed, yielding the statement. $\square$

Lemma 4.16 (Distribution-uniform monotone convergence). *If $\mathcal{Q}$ is setwise compact then*

$$f_n \geq 0 \text{ and } f_n \uparrow f \implies \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[f_n] \uparrow \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[f]. \quad (23)$$

Further, if (23) holds, then $\mathcal{Q}$ is relatively setwise compact and $\overline{\text{co}}^(\mathcal{Q}) \subseteq \mathcal{M}_1$.*

Proof. Set

$$a_n = \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[f_n], \quad a = \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[f].$$

Then a_n is increasing and $a_n \leq a$. Let $L = \lim_n a_n$. Suppose, to the contrary, that $L < a$. For each n , choose $Q_n \in \mathcal{Q}$ such that

$$\mathbb{E}_{Q_n}[f_n] \leq a_n + \frac{1}{n}.$$

By setwise compactness, there is a subnet Q_{n_α} converging setwise to some $Q^* \in \mathcal{Q}$. Since n_α is a subnet of the sequence of indices, we have $n_\alpha \uparrow \infty$.

Fix now $k, m \in \mathbb{N}$. For all sufficiently large α , $n_\alpha \geq k$, and hence

$$\mathbb{E}_{Q_{n_\alpha}}[f_k \wedge m] \leq \mathbb{E}_{Q_{n_\alpha}}[f_{n_\alpha}] \leq a_{n_\alpha} + \frac{1}{n_\alpha}.$$

Since $f_k \wedge m \in \mathcal{B}_b$, setwise convergence yields $\mathbb{E}_{Q^*}[f_k \wedge m] \leq L$. By monotone convergence we get $\mathbb{E}_{Q^*}[f_k] \leq L$. Letting now $k \uparrow \infty$ and using monotone convergence once more gives $\mathbb{E}_{Q^*}[f] \leq L$. Therefore

$$a = \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[f] \leq \mathbb{E}_{Q^*}[f] \leq L,$$

contradicting $L < a$. Hence $L = a$, proving the claim.

Let $A_n \downarrow \emptyset$. Applying (23) to $f_n = \mathbf{1}_{A_n^c} \uparrow 1$ gives

$$1 - \sup_{Q \in \mathcal{Q}} Q(A_n) = \inf_{Q \in \mathcal{Q}} Q(A_n^c) \uparrow \inf_{Q \in \mathcal{Q}} Q(\Omega) = 1.$$

Hence $\sup_{Q \in \mathcal{Q}} Q(A_n) \downarrow 0$. Now let $\mu \in \overline{\text{co}}^*(\mathcal{Q})$. Thus there is a net Q_α in $\text{co}(\mathcal{Q})$ that weak-* converges to μ . For each fixed n , since $\mathbf{1}_{A_n} \in \mathcal{B}_b$,

$$\mu(A_n) = \lim_\alpha Q_\alpha(A_n) \leq \sup_{Q \in \mathcal{Q}} Q(A_n) \downarrow 0 \quad \text{whenever } A_n \downarrow \emptyset.$$

Thus μ is continuous from above, hence $\mu \in \mathcal{M}_1$, yielding $\overline{\text{co}}^*(\mathcal{Q}) \subseteq \mathcal{M}_1$. Since ba_1 is weak-* compact, this closure is weak-* compact. The weak-* topology restricted to $\mathcal{M}_1$ is exactly the setwise topology, and hence the setwise closure of $\mathcal{Q}$ is compact. Thus $\mathcal{Q}$ is relatively setwise compact. $\square$

With these two lemmata in place we can now prove the following.

Theorem 4.17. *If $\mathcal{P}$ is arbitrary and $\mathcal{Q}$ is convex and setwise compact, then*

$$G = G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P).$$

The proof below clarifies that distribution-uniform monotone convergence (23) is the only property of $\mathcal{Q}$ that we need in order for $G = G_b = \min_{Q \in \overline{\text{co}}^*(\mathcal{Q}) \cap \mathcal{M}_1} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P)$ to hold. (Indeed, Lemma 4.16 shows that under (23), the weak-* closure of $\mathcal{Q}$ lies in $\mathcal{M}_1$, and then Corollary 3.6 allows to replace $\overline{\text{co}}^*(\mathcal{P})$ with $\mathcal{P}_{\text{eff}}$.) Setwise compactness only helps to further restrict to just $\mathcal{Q}$ on the right hand side.

Proof of Theorem 4.17. We first argue that $G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P)$. Lemma 4.15 allows us to replace $\overline{\text{co}}^*(\mathcal{Q})$ in Theorem 3.1 with just $\mathcal{Q}$. Corollary 3.6 shows that for $Q \in \mathcal{M}_1$, $\min_{\mu \in \overline{\text{co}}^*(\mathcal{P})} H(Q \mid \mu) = \min_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P)$. This proves $G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q \mid P)$.

Clearly $G_b \leq G$, so we focus on the reverse inequality. Fix an arbitrary e-variable X . For $\varepsilon \in (0, 1)$ and $n \in \mathbb{N}$, define

$$X_{\varepsilon, n} = \varepsilon + (1 - \varepsilon)(X \wedge n), \quad X_\varepsilon = \varepsilon + (1 - \varepsilon)X.$$

Then $X_{\varepsilon, n}$ is strictly positive and $X_{\varepsilon, n} \in \mathcal{E}_b$. For fixed ε , we have $\log X_{\varepsilon, n} \uparrow \log X_\varepsilon$ and $\log X_{\varepsilon, n} \geq \log \varepsilon$. Hence Lemma 4.16 (applied to the nonnegative functions $\log X_{\varepsilon, n} - \log \varepsilon$) yields

$$G_b \geq \sup_n \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X_{\varepsilon, n}] = \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X_\varepsilon] \geq \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] + \log(1 - \varepsilon).$$

Letting $\varepsilon \downarrow 0$, we get $G_b \geq \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X]$. Since $X \in \mathcal{E}$ was arbitrary, we conclude that $G_b \geq G$. $\square$

We end this section with a slightly strengthened form of Ram and Ramdas [2026, Proposition 4.5]. In preparation, define the solid hull of $\mathcal{P}$ as

$$\text{sol}(\mathcal{P}) = \{S \in \mathcal{M}_+ : S \leq P \text{ for some } P \in \mathcal{P}\} \subseteq \mathcal{P}_{\text{eff}}.$$

Proposition 4.18. *Assume $(\Omega, \mathcal{F})$ is a Polish space with its Borel sigma algebra. If $\mathcal{P}$ is convex and weakly compact, then $\mathcal{P}_{\text{eff}}$ equals the solid hull of $\mathcal{P}$, and thus for any $Q \in \mathcal{M}_1$,*

$$\min_{S \in \mathcal{P}_{\text{eff}}} H(Q \mid S) = \min_{P \in \mathcal{P}} H(Q \mid P).$$

Thus, in Theorems 4.10 and 4.17, if $\mathcal{P}$ is convex and weakly compact, we can replace $\mathcal{P}_{\text{eff}}$ by $\mathcal{P}$ on the right-hand side.

Proof. To begin with, on a Polish space, the weak topology on probability measures is metrizable and thus the topology on uniformly bounded subprobabilities is metrizable also. Thus, we may work with sequences instead of nets. Weak compactness of $\mathcal{P}$ implies that $\text{sol}(\mathcal{P})$ is weakly closed. Indeed, suppose $S_n \in \text{sol}(\mathcal{P})$ with $S_n \leq P_n$ for some $P_n \in \mathcal{P}$ and suppose $S_n \rightarrow S$. By weak compactness of $\mathcal{P}$, we may pass to a subsequence such that $P_n \rightarrow P \in \mathcal{P}$. Then, $P_n - S_n \rightarrow P - S$ and since each $P_n - S_n$ is positive, so is $P - S$. This implies that $S \leq P$ and thus $S \in \text{sol}(\mathcal{P})$. This shows that $\text{sol}(\mathcal{P})$ is weakly closed.

We next prove $\text{sol}(\mathcal{P}) \supseteq \mathcal{P}_{\text{eff}}$. To this end, suppose there exists $R \in \mathcal{P}_{\text{eff}}$ with $R \notin \text{sol}(\mathcal{P})$. Now, $\text{sol}(\mathcal{P})$ being weakly closed and convex, there exists an $f \in C_b$ such that

$$\int f^+ dR \geq \int f dR > \sup_{S \in \text{sol}(\mathcal{P})} \int f dS.$$

Solidity yields

$$\sup_{S \in \text{sol}(\mathcal{P})} \int f dS = \sup_{P \in \mathcal{P}} \int f^+ dP.$$

Indeed, this is true since for each $P \in \mathcal{P}$ the supremum of $\int f dS$ over $S \leq P$ is achieved by $S = \mathbf{1}_{f > 0} P$. Hence, we have the existence of $f \in C_b$ such that

$$\int f^+ dR > \sup_{P \in \mathcal{P}} \int f^+ dP.$$

Now set $M = \sup_{P \in \mathcal{P}} \int f^+ dP$. If $M > 0$ then $X = f^+/M \in \mathcal{E}_b$ with $\mathbb{E}_R[X] > 1$. If $M = 0$, then $c = \int f^+ dR > 0$ but $X = 2f^+/c \in \mathcal{E}_b$ with $\mathbb{E}_R[X] = 2 > 1$. This proves that $R \notin \mathcal{P}_{\text{eff}}$. We conclude that $\mathcal{P}_{\text{eff}} = \text{sol}(\mathcal{P})$.

Now, if $S \leq P$, then $H(Q \mid S) \geq H(Q \mid P)$. Moreover, since $\mathcal{P} \subseteq \mathcal{P}_{\text{eff}}$, it follows that $\inf_{S \in \mathcal{P}_{\text{eff}}} H(Q \mid S) = \inf_{P \in \mathcal{P}} H(Q \mid P)$. Finally, since the map $P \mapsto H(Q \mid P)$ is weakly lower semicontinuous and $\mathcal{P}$ is weakly compact, we have the last infimum is attained. $\square$

4.4 When $\mathcal{P}$ and $\mathcal{Q}$ are convex and weakly compact

In this section, we will prove a countably additive version of our main result by imposing a Polish sample space Ω and a weak compactness assumption on $\mathcal{P}$ and $\mathcal{Q}$, which allows one to work with continuous e-variables and avoid assuming a common dominating reference measure.

To this end, in this section alone we will assume that Ω is a Polish space, equipped with the Borel sigma algebra. We will equip $\mathcal{M}_1$ with the weak topology $\sigma(\mathcal{M}, C_b)$ meaning that a sequence P_n converges weakly to P if and only if for every $h \in C_b$, $\int h dP_n \rightarrow \int h dP$. Also, we will write

$$\mathcal{E}_{bc} = \mathcal{E}_{bb} \cap C_b,$$

where $\mathcal{E}_{bb}$ is the subset of e-variables that are bounded from above and away from zero. We denote the corresponding GROW values as G_{bb} and G_{bc} .

Theorem 4.19. *Suppose that Ω is Polish, equipped with the Borel sigma algebra, and assume $\mathcal{P}$ and $\mathcal{Q}$ are both convex and weakly compact. Then,*

$$G = G_b = G_{bb} = G_{bc} = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}} H(Q | P).$$

By Theorem 3.1, G_b (and hence G, G_{bc}, G_{bb}) also equals $\min_{\nu \in \overline{co}^*(\mathcal{Q})} \min_{\mu \in \overline{co}^*(\mathcal{P})} H(\nu | \mu)$. The proof uses the Donsker–Varadhan variational representation on Polish spaces [Dembo and Zeitouni, 2009, Lemma 6.2.13],

$$H(Q | P) = \sup_{g \in C_b} (\mathbb{E}_Q[g] - \log \mathbb{E}_P[e^g]). \quad (24)$$

We first record a robust version over compact convex null classes, obtained by an application of Sion’s minimax theorem.

Lemma 4.20 (Robust Donsker–Varadhan formula). *Suppose that Ω is Polish, equipped with its Borel sigma algebra, and let $\mathcal{P}$ be convex and weakly compact. Then, for every $Q \in \mathcal{M}_1$,*

$$\sup_{g \in C_b} \inf_{P \in \mathcal{P}} (\mathbb{E}_Q[g] - \log \mathbb{E}_P[e^g]) = \min_{P \in \mathcal{P}} H(Q | P).$$

Proof. Fix $Q \in \mathcal{M}_1$ and define, for $g \in C_b$ and $P \in \mathcal{P}$, $\Phi(g, P) = \mathbb{E}_Q[g] - \log \mathbb{E}_P[e^g]$. For fixed $g \in C_b$, the map $P \mapsto \Phi(g, P)$ is weakly continuous, because $e^g \in C_b$. It is also convex in P , since $P \mapsto \mathbb{E}_P[e^g]$ is affine and $x \mapsto -\log x$ is convex on $(0, \infty)$. For fixed $P \in \mathcal{P}$, the map $g \mapsto \Phi(g, P)$ is continuous (with respect to the supremum norm) and concave on C_b . Hence Sion’s minimax theorem and (24) yield

$$\sup_{g \in C_b} \inf_{P \in \mathcal{P}} \Phi(g, P) = \inf_{P \in \mathcal{P}} \sup_{g \in C_b} \Phi(g, P) = \inf_{P \in \mathcal{P}} H(Q | P).$$

Finally, the map $P \mapsto H(Q | P)$ is weakly lower semicontinuous, again by (24), since it is the supremum over $g \in C_b$ of weakly continuous functions of P . As $\mathcal{P}$ is weakly compact, the infimum is attained. $\square$

Proof of Theorem 4.19. The inclusions $\mathcal{E}_{bc} \subseteq \mathcal{E}_{bb} \subseteq \mathcal{E}_b \subseteq \mathcal{E}$, together with (10), imply

$$G_{bc} \leq G_{bb} \leq G_b \leq G \leq \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}} H(Q | P).$$

It is therefore enough to prove the reverse inequality

$$G_{bc} \geq \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}} H(Q | P).$$

Indeed, this will imply equality throughout. The attainment of the two infima will follow from weak compactness and weak lower semicontinuity of relative entropy.

To make headway, for $g \in C_b$, define

$$c(g) = \log \sup_{P \in \mathcal{P}} \mathbb{E}_P[e^g], \quad \Psi(Q, g) = \mathbb{E}_Q[g] - c(g), \quad Q \in \mathcal{Q}.$$

The function c is finite-valued and continuous. It is also convex since for each $P \in \mathcal{P}$ the map $g \mapsto \log \mathbb{E}_P[e^g]$ is convex and c is the pointwise supremum of these convex functions. Consequently, for each $Q \in \mathcal{Q}$, the map $g \mapsto \Psi(Q, g)$ is concave and continuous on C_b . For each $g \in C_b$, the map $Q \mapsto \Psi(Q, g)$ is affine and weakly continuous on $\mathcal{Q}$.

We now claim that

$$G_{bc} = \sup_{g \in C_b} \inf_{Q \in \mathcal{Q}} \Psi(Q, g). \quad (25)$$

For $g \in C_b$, define $X_g = \exp(g - c(g))$. Then X_g is continuous and bounded from above and below. Moreover,

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P[X_g] = e^{-c(g)} \sup_{P \in \mathcal{P}} \mathbb{E}_P[e^g] = 1,$$

so $X_g \in \mathcal{E}_{bc}$. Since $\mathbb{E}_Q[\log X_g] = \Psi(Q, g)$, we obtain $G_{bc} \geq \sup_{g \in C_b} \inf_{Q \in \mathcal{Q}} \Psi(Q, g)$. Conversely, let $X \in \mathcal{E}_{bc}$ and set $g = \log X$. Then $g \in C_b$ and $c(g) = \log \sup_{P \in \mathcal{P}} \mathbb{E}_P[X] \leq 0$. Hence, for every $Q \in \mathcal{Q}$,

$$\mathbb{E}_Q[\log X] = \mathbb{E}_Q[g] \leq \mathbb{E}_Q[g] - c(g) = \Psi(Q, g).$$

It follows that $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \leq \sup_{h \in C_b} \inf_{Q \in \mathcal{Q}} \Psi(Q, h)$. Taking the supremum over $X \in \mathcal{E}_{bc}$ gives the reverse inequality in (25).

For fixed $g \in C_b$, the map $Q \mapsto \Psi(Q, g)$ is affine and weakly continuous. For fixed $Q \in \mathcal{Q}$, the map $g \mapsto \Psi(Q, g)$ is concave and continuous. Since $\mathcal{Q}$ is convex and weakly compact, Sion's theorem gives

$$\sup_{g \in C_b} \inf_{Q \in \mathcal{Q}} \Psi(Q, g) = \inf_{Q \in \mathcal{Q}} \sup_{g \in C_b} \Psi(Q, g).$$

Combining this with (25), we obtain

$$G_{bc} = \inf_{Q \in \mathcal{Q}} \sup_{g \in C_b} \left(\mathbb{E}_Q[g] - \log \sup_{P \in \mathcal{P}} \mathbb{E}_P[e^g] \right) = \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}} H(Q \mid P),$$

where the last equality follows from Lemma 4.20. Noting that relative entropy is jointly weakly lower semicontinuous and $\mathcal{Q} \times \mathcal{P}$ is weakly compact, we conclude that the infimum is attained. $\square$

The following corollary gives a simple criterion for when an e-variable witnesses strong duality. Its proof is immediate and hence omitted.

Corollary 4.21. *Assume either the setup of Theorem 4.17 or of Theorem 4.19, and let (P^*, Q^*) denote any pair that achieves the right-hand side of the strong duality. If $G < \infty$, then (P_a^*, Q^*) is a joint information projection. (Here, P_a^* is the absolutely continuous part of P^* with respect to Q^* .) If we further assume that Assumption 4.9 holds for the pair (P_a^*, Q^*) , then Theorem 4.10 implies that the e-variable $X^* = dQ^*/dP_a^*$ achieves the supremum in G .*

We also recall Example 4.4 to remind us that even with a simple null and setwise and weakly compact and convex $\mathcal{Q}$, strong duality may hold, but the likelihood ratio of the JIPr pair can have an arbitrarily bad worst-case e-power, and no e-variable witnesses the strong duality; see also Example 5.5. We conclude with an example frequently encountered in the recent e-variable literature.

Example 4.22 (Bounded means). Let $\Omega = [0, 1]$ and let $Z(\omega) = \omega$ denote the coordinate variable on $[0, 1]$. Set $\mathcal{P} = \{P : \mathbb{E}_P[Z] \leq a\}$ and $\mathcal{Q} = \{Q : \mathbb{E}_Q[Z] \geq b\}$, for some $0 < a < b < 1$. Larsson et al. [2026a] show that every e -variable is bounded by one of the form $1 + \lambda(Z - a)$ for $\lambda \in [0, 1/a]$, and is bounded by $1/a$. Therefore, $G = G_b < \infty$ is not surprising. Since $\mathcal{P}$ and $\mathcal{Q}$ are both convex and weakly compact, Corollary 4.21 applies to conclude that a JIPr pair (P^*, Q^*) exists (in this case, uniquely): it is in fact just the pair of Bernoulli distributions with parameters a, b . The continuous e -variable obtained by picking $\lambda^* = (b - a)/(a(1 - a))$ is easily checked to be a P^* -version of their likelihood ratio.

Assumption 4.9 can be verified as follows. Since $P^* = (1 - a)\delta_0 + a\delta_1$, $\inf_{R \in \mathcal{Q}} H(R | P^*)$ is achieved by Q^* . (Indeed, any R not supported on $\{0, 1\}$ has infinite relative entropy, so we only check amongst Bernoullis.) So we need to check that for every $Q \in \mathcal{Q}$,

$$\mathbb{E}_Q[\log X^*] \geq H(Q^* | P^*) = b \log \frac{b}{a} + (1 - b) \log \frac{1 - b}{1 - a}.$$

Denote $\phi(\omega) = \log X^*(\omega) = \log(1 + \lambda^*(\omega - a))$. Since ϕ is concave on $[0, 1]$, it lies above its chord between 0 and 1: $\phi(\omega) \geq (1 - \omega)\phi(0) + \omega\phi(1)$. So for any $Q \in \mathcal{Q}$, $\mathbb{E}_Q[\phi(Z)] \geq \mathbb{E}_Q[1 - Z]\phi(0) + \mathbb{E}_Q[Z]\phi(1)$. Since $\mathbb{E}_Q[Z] \geq b$ and $\phi(1) > \phi(0)$, we get $\mathbb{E}_Q[\log X^*] \geq (1 - b)\phi(0) + b\phi(1) = H(Q^* | P^*)$, as required.

The GROW optimality of $1 + \lambda^*(Z - a)$ thus follows from Corollary 4.21. Arnold and Clerico [2026] proved the GROW optimality of this e -variable as well, but it is reassuring that we can derive the same result as a special case of our general theory. To emphasize this point, we present similar calculations for a slightly more complicated problem in Appendix C, where unbounded data have bounded second moment, and the null and alternative classes differ in their means.

Other natural statistical examples that are convex and weakly compact include relative entropy balls $\mathcal{P} = \{P : H(P | P_0) \leq c_0\}$ and $\mathcal{Q} = \{Q : H(Q | Q_0) \leq c_1\}$, for some $P_0, Q_0 \in \mathcal{M}_1$ and $c_0, c_1 > 0$, or Hellinger balls on compact Polish Ω .

4.5 When the sample space is finite

Below, we consider the case when $\Omega = \{1, \dots, d\}$, and identify probability measures on Ω with vectors in the simplex

$$\Delta_d = \left\{ p \in [0, 1]^d : \sum_{i=1}^d p_i = 1 \right\}.$$

Theorem 4.23 (Finite sample spaces with no null coordinates). Let $\mathcal{P}, \mathcal{Q} \subseteq \Delta_d$ be convex and compact sets. Assume that $\mathcal{P}$ is supported on all coordinates, meaning that $\sup_{P \in \mathcal{P}} P_i > 0$ for all $i \in \{1, \dots, d\}$. Then $\mathcal{E} = \mathcal{E}_b$ and

$$G = G_b = \max_{X \in \mathcal{E}_b} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}} H(Q | P) < \infty,$$

meaning that the GROW value is finite and is attained by some bounded e -variable $X^* \in \mathcal{E}_b$. Moreover, if $(P^*, Q^*) \in \mathcal{P} \times \mathcal{Q}$ minimizes the right-hand side, then $X^* = dQ^*/dP^*$, P^* -a.s.

Proof. Any $X \in \mathcal{E}$ takes only finitely many values, all finite due to the support assumption on $\mathcal{P}$. This shows $\mathcal{E} = \mathcal{E}_b$, so $G = G_b$. Next, since Ω is finite, weak convergence on Δ_d is the same as coordinatewise, hence Euclidean, convergence. Thus the assumed compactness of $\mathcal{P}$ and $\mathcal{Q}$ implies weak compactness. Moreover, weak-* closure agrees with Euclidean closure in finite dimension, and therefore

$$\overline{\text{co}}^*(\mathcal{P}) = \mathcal{P}, \quad \overline{\text{co}}^*(\mathcal{Q}) = \mathcal{Q}.$$

Theorem 3.1 then gives

$$G = G_b = \min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}} H(Q | P). \quad (26)$$

To show that the right-hand side above is finite, for each i choose $P^{(i)} \in \mathcal{P}$ with $P_i^{(i)} > 0$, and define $\bar{P} = \frac{1}{d} \sum_{i=1}^d P^{(i)} \in \mathcal{P}$, by convexity. Then every $Q \in \mathcal{Q}$ satisfies $Q \ll \bar{P}$, so $H(Q | \bar{P}) < \infty$.

Finally, we prove that the supremum over e -variables is attained. Note that

$$\mathcal{E} \subseteq \prod_{i=1}^d \left[0, \frac{1}{P_i^{(i)}} \right].$$

Moreover, $\mathcal{E}$ is closed, since it is defined by the closed linear inequalities

$$\sum_{i=1}^d P_i X_i \leq 1, \quad P \in \mathcal{P}.$$

Thus $\mathcal{E}$ is compact. For fixed $Q \in \mathcal{Q}$, the map $X \mapsto \mathbb{E}_Q[\log X]$ is upper semicontinuous, hence so is $X \mapsto \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X]$. Therefore this objective attains its maximum on the compact set $\mathcal{E}$.

Let $(P^*, Q^*) \in \mathcal{P} \times \mathcal{Q}$ now minimize the right-hand side of (26), and let $X^* \in \mathcal{E}$ be any GROW optimizer. By Proposition 4.18 we have

$$\min_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) = H(Q^* | P^*).$$

Applying Proposition 4.12 then yields the final claim. $\square$

5 Additional examples

5.1 $G > G_b$ with singleton $\mathcal{P} = \{P\}$ and countable $\mathcal{Q}$ on $\mathbb{N}$

While our main Theorem 3.1 establishes strong duality for $\mathcal{E}_b$ without imposing any assumptions on $\mathcal{P}$ or $\mathcal{Q}$, the strong duality results for $\mathcal{E}$ in Section 4 require additional restrictions on these families. A natural question is whether such restrictions are necessary. We answer this in part by presenting a small modification of Example 3.15 in Ramdas and Wang [2025], which shows that $G > G_b$ can occur. Notably, the example only involves a singleton null $\mathcal{P} = \{P\}$ (which is hence compact in any reasonable topology), and a countable alternative $\mathcal{Q} = \{Q_n : n \in \mathbb{N}\}$ on $\mathbb{N}$. In particular, these measures admit a common dominating measure. Thus, even in a very regular setting, the equality $G = G_b$ may fail.

Example 5.1. Let $\Omega = \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, let $P = (1/2)\delta_0 + (1/2)\delta_1$, where δ_x denotes the point mass at x , and let $\mathcal{Q} = \{Q_n : n \in \mathbb{N}\}$, where

$$Q_n = \frac{1}{n}\delta_n + \left(1 - \frac{1}{n}\right)\delta_0.$$

Clearly, the family admits a common dominating reference measure. Moreover,

$$\inf_{Q \in \mathcal{Q}} H(Q | P) = H(Q_1 | P) = \log 2 < \infty,$$

since $Q_1 = \delta_1$, whereas $H(Q_n | P) = \infty$ for all $n \geq 2$. The same remains true after passing to $\text{co}(\mathcal{Q})$, and hence the (generalized) information projection of P onto $\text{co}(\mathcal{Q})$ equals Q_1 . We further have the following observations.

(i) For $\varepsilon \in (0, 1)$, define

$$X^\varepsilon(n) = \exp(n \log(2 - \varepsilon) - (n - 1) \log \varepsilon), \quad n \in \mathbb{N}_0.$$

Then $X^\varepsilon(0) = \varepsilon$ and $X^\varepsilon(1) = 2 - \varepsilon$, hence $\mathbb{E}_P[X^\varepsilon] = 1$ and X^ε is an e -variable for P . Moreover, for every $n \in \mathbb{N}$,

$$\mathbb{E}_{Q_n}[\log X^\varepsilon] = \frac{1}{n} (n \log(2 - \varepsilon) - (n - 1) \log \varepsilon) + \left(1 - \frac{1}{n}\right) \log \varepsilon = \log(2 - \varepsilon).$$

Therefore $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X^\varepsilon] = \log(2 - \varepsilon)$, which tends to $\log 2$ as $\varepsilon \downarrow 0$. By weak duality (recall (10)), this is the largest possible value. Hence $G = \log 2$, and the conclusion of Theorem 4.2 holds, although its assumption fails.

(ii) There does not exist any bounded e -variable X for which $\mathbb{E}_Q[X] > 1$ for all $Q \in \mathcal{Q}$. Indeed, meeting this requirement for Q_1 alone would imply $X(1) > 1$. To meet this requirement for Q_n , we would need $(1/n)X(n) + (1 - 1/n)X(0) > 1$ for all $n \in \mathbb{N}$. Since X is bounded, this then implies $X(0) \geq 1$. Together with $X(1) > 1$, we have $\mathbb{E}_P[X] > 1$, conflicting the fact that X is an e -variable for P .

(iii) Point (ii) above and Jensen's inequality then imply that there is no bounded e -variable X for which $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] > 0$. This means that $G_b = 0$.

(iv) Note that here P does not lie in $\text{co}(\mathcal{Q})$, but it does lie within its TV-closed convex hull (hence also within its weak-* closure). To see this, note that by sending $n \uparrow \infty$, Q_n converges to δ_0 in total variation. Averaging δ_0 with Q_1 yields P . Thus the information projection of P onto $\text{co}(\mathcal{Q})$, which equals Q_1 , differs from its information projection onto the TV-closure of $\text{co}(\mathcal{Q})$, which equals P itself. This reinforces a point made by [Csiszár, 1984, Example 3.1] that the generalized I -projection of P onto $\text{co}(\mathcal{Q})$ cannot in general be seen as the I -projection onto some closure of $\text{co}(\mathcal{Q})$.

The above example shows that strong duality alone does not imply $G = G_b$. Indeed, strong duality holds for G by (i), while Theorem 3.1 gives the corresponding strong-duality representation for G_b ; nevertheless, in this example $G > G_b$. This failure occurs despite substantial structure: the null class is a singleton, the families admit a common dominating reference measure, and the relevant information projection is explicit.

5.2 Extension of Example 4.3 that maintains $Q \ll P$ for all $Q \in \mathcal{Q}$

We now provide the promised extension of Example 4.3, which additionally maintains $Q \ll P$ for all $Q \in \mathcal{Q}$. Here, an IPr exists, the supremum in G is achieved, but strong duality fails.

Example 5.2. As in Example 4.3, we consider $\Omega = [0, 1]$, equipped with the Borel sigma algebra, $P = U$, and the probability measure Q_0 given by $dQ_0/dP(\omega) = 2\omega$. Let

$$\mathcal{I}_\infty = \{Q \ll P : H(Q | P) = \infty\},$$

and define

$$\mathcal{Q} = \{Q_0\} \cup \{(1 - t)Q_0 + tR : t \in (0, 1], R \in \mathcal{I}_\infty\}.$$

Note that $\mathcal{I}_\infty$ is convex, since positive mixtures of elements of $\mathcal{I}_\infty$ remain in $\mathcal{I}_\infty$: if $H(Q | P) = \infty$, then $H((1 - t)R + tQ | P) = \infty$ for every $R \ll P$ and $t \in (0, 1]$. Thus $\mathcal{Q}$ is also convex. By the same token, the only finite-entropy element of $\mathcal{Q}$ is Q_0 , so

$$\inf_{Q \in \mathcal{Q}} H(Q | P) = H(Q_0 | P) = \log 2 - \frac{1}{2}. \quad (27)$$

Moreover, $Q^{\text{GIPr}} = Q_0$. Indeed, the Pythagorean inequality (20) holds with equality at Q_0 , while for every other element of $\mathcal{Q}$ the left-hand side is infinite.

Now, let X be any e -variable for P . Define

$$A = \{\omega : X(\omega) \leq 1\}.$$

As in Example 4.3 we have $P(A) > 0$. Since P is nonatomic, there exists a probability measure $Q_A \ll P$ supported on A such that $H(Q_A | P) = \infty$. Consequently $Q_A \in \mathcal{Q}$. Since $X \leq 1$ Q_A -almost surely, $\mathbb{E}_{Q_A}[\log X] \leq 0$, and therefore

$$\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \leq \mathbb{E}_{Q_A}[\log X] \leq 0.$$

Since this holds for every e -variable X , $\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \leq 0$. On the other hand, choosing $X \equiv 1$ yields $\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = 0$. Hence

$$\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = 0.$$

Combining this with (27),

$$\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = 0 < \log 2 - \frac{1}{2} = \inf_{Q \in \mathcal{Q}} H(Q | P).$$

The above example violates strong duality in Theorem 4.2, but (as expected) the strong duality in Theorem 3.1 still holds. To see this, first note that $\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = 0$ implies immediately that $\sup_{X \in \mathcal{E}_b} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = 0$, both achieved by $X \equiv 1$. But the weak-* closure $\overline{\text{co}}^*(\mathcal{Q})$ contains P , so the right-hand side of Theorem 3.1 also equals zero.

5.3 Strong duality for singleton $\mathcal{P} = \{P\}$ holds, but $G > H(Q^* | P)$ for GIPr Q^*

We now revisit Example 3.2 from Csiszár [1984], which makes the point that $\inf_{Q \in \mathcal{Q}} H(Q | P) > H(Q^* | P)$ for the GIPr Q^* . We now complete the duality picture for this example. We show that strong duality for G holds, but despite having a singleton $\mathcal{P} = \{P\}$, G may not equal $H(Q | P)$ for any $Q \in \mathcal{Q}$, nor equal $H(Q^* | P)$.

Example 4.8 already shows that it is possible to satisfy conditions (i) and (ii) in the JIPr Definition 4.7 without satisfying condition (iii). The following example goes one step further. It shows that even if Assumption 4.1 is satisfied, it is possible to satisfy conditions (i) and (ii) in the JIPr definition 4.7 without satisfying condition (iii).

Example 5.3. Let $\Omega = [0, \infty)$, equipped with its Borel sigma algebra. Let

$$\mathcal{Q}_a = \{Q \in \mathcal{M}_1 : \int s dQ(s) \geq a\}$$

for some $a > 2/3$, which is convex. Let P be defined by the density $p(s) = \exp(-s)(s+4)/(s+1)^4$ for $s \geq 0$, which has mean about $0.298 < 2/3$, and thus lies outside $\mathcal{Q}_a$. Csiszár proves that the GIPr of P onto $\mathcal{Q}_a$ is a distribution Q^* with density $q^*(s) = \frac{2}{3}(s+4)/(s+1)^4$ for $s \geq 0$, which remarkably does not depend on a , and its mean is $2/3$, so $Q^* \notin \mathcal{Q}_a$. We have the following facts.

1. For any $Q \in \mathcal{Q}_a$ with $Q \ll Q^*$, the chain rule gives

$$H(Q | P) = H(Q | Q^*) + \mathbb{E}_Q \left[\log \frac{dQ^*}{dP} \right] = H(Q | Q^*) + \int s dQ(s) + \log \frac{2}{3} \geq a + \log \frac{2}{3}.$$

If $Q \not\ll Q^*$, since $Q^* \sim P$, we have $H(Q | P) = \infty$, and the same lower bound is trivial. Put together, $\inf_{Q \in \mathcal{Q}_a} H(Q | P) \geq a + \log \frac{2}{3}$. In fact, equality holds, as justified in Appendix D.

2. Since

$$\frac{q^*(s)}{p(s)} = \frac{\frac{2}{3}(s+4)/(s+1)^4}{e^{-s}(s+4)/(s+1)^4} = \frac{2}{3}e^s, \quad s \geq 0,$$

we choose the Borel version of the Radon–Nikodym derivative dQ^*/dP given by $X^*(s) = \frac{2}{3}e^s$. Then $X^* = \frac{dQ^*}{dP}$, P -a.s., hence $X^* \in \mathcal{E}$. For every $Q \in \mathcal{Q}_a$,

$$\mathbb{E}_Q[\log X^*] = \int s dQ(s) + \log \frac{2}{3} \geq a + \log \frac{2}{3} = \inf_{R \in \mathcal{Q}_a} H(R | P).$$

Thus, Assumption 4.1 is satisfied, and, by Theorem 4.2, strong duality holds and X^* is GROW optimal.

3. $H(Q^* | P) = 2/3 - \log(3/2) < a + \log(2/3)$.

To summarize, we have

$$a + \log \frac{2}{3} = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}_a} \mathbb{E}_Q[\log X] = \inf_{Q \in \mathcal{Q}_a} H(Q | P) > H(Q^* | P),$$

meaning that strong duality holds, and the left-hand side supremum is achieved by the e -variable $X^* = dQ^*/dP$. Furthermore, $\inf_{Q \in \mathcal{Q}_a} \mathbb{E}_Q[\log X^*]$ is not achieved by Q^* , but it is achieved by any $Q \in \mathcal{Q}_a$ with mean equal to a . Finally, (P, Q^*) also forms a JIPr pair satisfying conditions (i),(ii) but not condition (iii) in Definition 4.7.

5.4 Gaussian mean shift with heavy-tailed outliers

We now formalize and generalize Remark 4.5, which allows us to extend Example 4.4 to a more statistically meaningful context.

Theorem 5.4 (Shielding theorem). *Suppose that $\mathcal{Q}$ is partitioned as $\mathcal{Q} = \mathcal{Q}_I \cup \mathcal{Q}_\infty$, where $\mathcal{Q}_I$ is nonempty. Assume that strong duality holds over $\mathcal{Q}_I$ with a finite value, i.e.,*

$$I = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}_I} \mathbb{E}_Q[\log X] = \inf_{Q \in \mathcal{Q}_I} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) < \infty.$$

Assume further that there exists $Y \in \mathcal{E}$ such that

$$\mathbb{E}_Q[\log Y] = \infty \quad \text{for every } Q \in \mathcal{Q}_\infty. \quad (28)$$

Then strong duality holds over the full alternative $\mathcal{Q}$, i.e.,

$$I = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = \inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P). \quad (29)$$

Appendix D provides the proof of Theorem 5.4.

Example 5.5 (Gaussian mean shift with heavy-tailed outliers). *Let $\Omega = \mathbb{R}$ and let $\mathcal{P} = \{P\}$ with $P = N(0, 1)$. Fix $\theta > 0$ and let $Q_0 = N(\theta, 1)$. The likelihood ratio of Q_0 with respect to P is $L(\omega) = \exp\left(\theta\omega - \frac{\theta^2}{2}\right)$, which is an e -variable with $\mathbb{E}_{Q_0}[\log L] = H(Q_0 | P) = \frac{\theta^2}{2}$. Thus strong duality holds for the singleton alternative $\mathcal{Q}_I = \{Q_0\}$, with value $I = \frac{\theta^2}{2}$. Now let R be a centered Student- t_κ distribution with $1 < \kappa < 2$. For $t \in [0, 1]$, define $Q_t = (1-t)Q_0 + tR$, and set*

$$\mathcal{Q} = \{Q_t : 0 \leq t \leq 1\}, \quad \mathcal{Q}_\infty = \{Q_t : 0 < t \leq 1\}.$$

This gives the partition $\mathcal{Q} = \mathcal{Q}_I \cup \mathcal{Q}_\infty$, and we note that $\mathcal{Q}$ is both setwise and weakly compact, so Theorem 4.17 already gives strong duality, but we show below that it is not attained by any e -variable, and shielding helps verify this strong duality.

Choose now $a \in (0, 1/2)$ and define the random variable

$$Y_a(\omega) = \sqrt{1 - 2a} \exp(a\omega^2).$$

Since $\mathbb{E}_{\mathbf{P}}[Y_a] = 1$, we get $Y_a \in \mathcal{E}$. Moreover, $\log Y_a(\omega) = \frac{1}{2} \log(1 - 2a) + a\omega^2$. Therefore, for every $t > 0$,

$$\mathbb{E}_{\mathbf{Q}_t}[\log Y_a] = (1 - t)\mathbb{E}_{\mathbf{Q}_0}[\log Y_a] + t\mathbb{E}_{\mathbf{R}}[\log Y_a] = \infty,$$

because $\mathbf{R}$ has infinite second moment. Thus Y_a is a shielding e -variable for $\mathcal{Q}_\infty$. By Theorem 5.4, strong duality holds:

$$\frac{\theta^2}{2} = \sup_{X \in \mathcal{E}} \inf_{\mathbf{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbf{Q}}[\log X] = \inf_{\mathbf{Q} \in \mathcal{Q}} H(\mathbf{Q} \mid \mathbf{P}).$$

Proposition 4.12 implies that this strong duality is not attained by any e -variable. Indeed, if it was obtained, such e -variable would have to equal $d\mathbf{Q}_0/d\mathbf{P}$, $\mathbf{P}$ -a.s., and hence also $\mathbf{R}$ -a.s., but $\mathbb{E}_{\mathbf{R}}[\log d\mathbf{Q}_0/d\mathbf{P}] = \mathbb{E}_{\mathbf{R}}[\theta Z - \theta^2/2] = -\theta^2/2 < I$, where $Z(\omega) = \omega$.

5.5 Finite bounded moment inequalities

We next present an example of a natural class that was recently studied by Larsson et al. [2026a]; there it served the role of $\mathcal{P}$, and below it serves the role of $\mathcal{Q}$ in Theorem 4.2.

Example 5.6. Let $\mathbf{P} \in \mathcal{M}_1$, let $f_1, \dots, f_K \in \mathcal{B}_b$, and define

$$\mathcal{Q} = \{\mathbf{Q} \in \mathcal{M}_1 : \mathbb{E}_{\mathbf{Q}}[f_k] \leq 0, \quad k = 1, \dots, K\}.$$

I -projections under finitely many linear inequality constraints were studied by Dykstra and Wollan [1987]; see also Bhattacharya and Dykstra [1995]. Assume the strict-feasibility condition that there exists $\mathbf{R} \in \mathcal{Q}$ such that $\mathbf{R} \ll \mathbf{P}$ and

$$H(\mathbf{R} \mid \mathbf{P}) < \infty, \quad \mathbb{E}_{\mathbf{R}}[f_k] < 0, \quad k = 1, \dots, K.$$

For example, this holds if there exist $A \in \mathcal{F}$ and $\eta > 0$ such that

$$\mathbf{P}(A) > 0, \quad f_k \leq -\eta \text{ on } A, \quad k = 1, \dots, K,$$

by taking $\mathbf{R} = \mathbf{P}(\cdot \mid A)$. Define

$$\psi(\lambda) = \log \mathbb{E}_{\mathbf{P}} \left[\exp \left(- \sum_{k=1}^K \lambda_k f_k \right) \right], \quad \lambda \in [0, \infty)^K.$$

Since the functions f_k are bounded, ψ is finite, convex, and continuously differentiable. Moreover, since

$$\eta = \min_{1 \leq k \leq K} (-\mathbb{E}_{\mathbf{R}}[f_k]) > 0,$$

the Donsker–Varadhan inequality (Lemma 2.4) gives

$$\psi(\lambda) \geq -H(\mathbf{R} \mid \mathbf{P}) - \sum_{k=1}^K \lambda_k \mathbb{E}_{\mathbf{R}}[f_k] \geq -H(\mathbf{R} \mid \mathbf{P}) + \eta \|\lambda\|_1.$$

Thus ψ attains its minimum at some $\lambda^* \in [0, \infty)^K$. Set

$$X^* = \exp \left(- \sum_{k=1}^K \lambda_k^* f_k - \psi(\lambda^*) \right)$$

Then $\mathbb{E}_{\mathbf{P}}[X^*] = 1$. Define next $\mathbf{Q}^*$ by $\frac{d\mathbf{Q}^*}{d\mathbf{P}} = X^*$. The first-order optimality conditions at λ^* give

$$\mathbb{E}_{\mathbf{Q}^*}[f_k] = -\frac{\partial\psi}{\partial\lambda_k}(\lambda^*) \leq 0, \quad \lambda_k^* \mathbb{E}_{\mathbf{Q}^*}[f_k] = -\lambda_k^* \frac{\partial\psi}{\partial\lambda_k}(\lambda^*) = 0, \quad k = 1, \dots, K.$$

Hence $\mathbf{Q}^* \in \mathcal{Q}$ and

$$H(\mathbf{Q}^* | \mathbf{P}) = -\sum_{k=1}^K \lambda_k^* \mathbb{E}_{\mathbf{Q}^*}[f_k] - \psi(\lambda^*) = -\psi(\lambda^*).$$

For every $\mathbf{Q} \in \mathcal{Q}$,

$$\mathbb{E}_{\mathbf{Q}}[\log X^*] = -\sum_{k=1}^K \lambda_k^* \mathbb{E}_{\mathbf{Q}}[f_k] - \psi(\lambda^*) \geq -\psi(\lambda^*) = H(\mathbf{Q}^* | \mathbf{P}).$$

Combining the above with weak duality, we get

$$H(\mathbf{Q} | \mathbf{P}) \geq \mathbb{E}_{\mathbf{Q}}[\log X^*] \geq H(\mathbf{Q}^* | \mathbf{P}), \quad \mathbf{Q} \in \mathcal{Q}.$$

Thus $\mathbf{Q}^*$ is the IPr of $\mathbf{P}$ onto $\mathcal{Q}$, and the preceding display verifies Assumption 4.1. Consequently,

$$\mathbf{G} = \inf_{\mathbf{Q} \in \mathcal{Q}} H(\mathbf{Q} | \mathbf{P}) = H(\mathbf{Q}^* | \mathbf{P}) = -\psi(\lambda^*),$$

and $X^* = d\mathbf{Q}^*/d\mathbf{P}$ is the GROW e -variable.

5.6 $\mathbf{G}_b > \mathbf{G}_{bc}$ can hold in general

In the context of Subsection 4.4, $\mathbf{G}_b \geq \mathbf{G}_{bc}$ holds always. We provide an example where $\mathbf{G}_b > \mathbf{G}_{bc}$ in order to show that some assumptions on $\mathcal{Q}$ are needed for $\mathbf{G}_b = \mathbf{G}_{bc}$ to hold.

Example 5.7. Let $\Omega = [0, 1]$ with its Borel sigma algebra, set $D = \mathbb{Q} \cap [0, 1]$, and let $\mathcal{P} = \{\mathbf{U}\}$ and $\mathcal{Q} = \{\delta_q : q \in D\}$. For $N \in \mathbb{N}$, define $X_N = N\mathbf{1}_D + \mathbf{1}_{D^c} \in \mathcal{E}_b$. Then

$$\sup_{X \in \mathcal{E}_b} \inf_{\mathbf{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbf{Q}}[\log X] = \infty.$$

On the other hand, let $X \in \mathcal{E}_{bc}$. Then X is continuous, nonnegative, and satisfies $\mathbb{E}_{\mathbf{U}}[X] \leq 1$. Therefore $\inf_{\omega \in [0, 1]} X(\omega) \leq 1$. Since D is dense in $[0, 1]$ and X is continuous,

$$\inf_{q \in D} X(q) = \inf_{\omega \in [0, 1]} X(\omega) \leq 1.$$

Consequently,

$$\inf_{\mathbf{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbf{Q}}[\log X] = \inf_{q \in D} \log X(q) \leq 0.$$

Since the constant function $X \equiv 1$ belongs to $\mathcal{E}_{bc}$, we obtain

$$\sup_{X \in \mathcal{E}_{bc}} \inf_{\mathbf{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbf{Q}}[\log X] = 0.$$

Thus $\mathbf{G}_b > \mathbf{G}_{bc}$ in general.

6 Summary

Theorem 3.1 establishes strong duality for the minimax e-power of bounded e-variables, equivalently for the *bounded* GROW value G_b . It shows that G_b equals the minimum relative entropy between the weak-* closed convex hulls of $\mathcal{P}$ and $\mathcal{Q}$, without imposing any restrictions on either family. We also extend this result to the REGROW criterion for arbitrary bounded offsets. We then prove several strong duality results for the *unbounded* GROW value G . Theorem 4.2 treats the case of a singleton null $\mathcal{P}$, improving on previous work. Theorem 4.10 extends this result to the case that a joint information projection exists. We then prove strong duality results for unbounded GROW under assumptions such as setwise or weak compactness, that do not require taking weak-* closures. It remains open to what extent such restrictions are necessary. While our examples show that G can be strictly larger than G_b , we do not know whether a universal strong duality theorem for G holds for arbitrary $\mathcal{P}$ and $\mathcal{Q}$. Indeed, Examples 4.3 and 5.2 demonstrate that a naive entropy duality does not hold without assumptions; it remains open whether some alternative universal dual representation for unbounded GROW exists.

Acknowledgments and Disclosures

AR acknowledges support from the National Science Foundation under grant DMS-2310718. ML acknowledges support from the National Science Foundation under grant NSF DMS-2510965.

References

- Sebastian Arnold and Eugenio Clerico. Optimal e-values for testing the mean of a bounded random variable against a composite alternative. *arXiv preprint arXiv:2601.11347*, 2026.
- Bhaskar Bhattacharya and Richard L. Dykstra. A general duality approach to I-projections. *Journal of Statistical Planning and Inference*, 47(3):203–216, 1995. doi: 10.1016/0378-3758(94)00142-I.
- Vladimir I. Bogachev. *Measure Theory*, volume 1 and 2. Springer-Verlag, Berlin, Heidelberg, 2007. ISBN 978-3-540-34513-8. doi: 10.1007/978-3-540-34514-5.
- Leo Breiman. Optimal gambling systems for favorable games. *Berkeley Symp. on Math. Statist. and Prob*, 1961.
- Thomas M. Cover. Log optimal portfolios. In *Chapter in “Gambling Research: Gambling and Risk Taking,” Seventh International Conference*, volume 4, 1987.
- Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1):146–158, 1975.
- Imre Csiszár. Sanov property, generalized I -projection and a conditional limit theorem. *The Annals of Probability*, 12(3):768–793, 08 1984. doi: 10.1214/aop/1176993227.
- Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38. Springer Science & Business Media, 2009.
- Richard L. Dykstra and Peter C. Wollan. Finding I-projections subject to a finite set of linear inequality constraints. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 36(3):377–383, 1987. doi: 10.2307/2347801.

- Peter Grünwald, Rianne de Heide, and Wouter Koolen. Safe testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(5):1091–1128, 11 2024a. doi: 10.1093/jrsssb/qkae011.
- Peter Grünwald, Yunda Hao, and Akshay Balsubramani. Growth-optimal e-variables and an extension to the multivariate Csiszar-Sanov-Chernoff theorem. *arXiv preprint arXiv:2412.17554*, 2024b.
- Peter J. Huber and Volker Strassen. Minimax tests and the Neyman-Pearson lemma for capacities. *Ann. Statist.*, 1:251–263, 1973. ISSN 0090-5364,2168-8966.
- John L. Kelly. A new interpretation of information rate. *The Bell System Technical Journal*, 35(4):917–926, 1956.
- Charles Kraft. Some conditions for consistency and uniform consistency of statistical procedures. *University of California Publication in Statistics*, 2:125–141, 1955.
- Tyron Lardy, Peter Grünwald, and Peter Harremoës. Reverse information projections and optimal e-statistics. *IEEE Transactions on Information Theory*, 2024.
- Martin Larsson, Aaditya Ramdas, and Johannes Ruf. The numeraire e-variable and reverse information projection. *The Annals of Statistics*, 53(3):1015–1043, 2025. doi: 10.1214/24-AOS2487.
- Martin Larsson, Aaditya Ramdas, and Johannes Ruf. Testing hypotheses generated by constraints. *Mathematics of Operations Research*, 2026a.
- Martin Larsson, Johannes Ruf, and Aaditya Ramdas. A complete characterization of testable hypotheses. *arXiv preprint arXiv:2601.05217*, 2026b.
- Qiang Jonathan Li. *Estimation of Mixture Models*. PhD thesis, Yale University, 1999.
- Ashwin Ram and Aaditya Ramdas. The optimal betting wealth growth rate. *arXiv preprint arXiv:2604.25280*, 2026.
- Aaditya Ramdas and Ruodu Wang. Hypothesis testing with e-values. *Foundations and Trends® in Statistics*, 1(1-2):1–390, 2025.
- Aytijhya Saha and Aaditya Ramdas. Huber-robust likelihood ratio tests for composite nulls and alternatives. *IEEE Transactions on Information Theory*, 2025.
- Glenn Shafer. Testing by betting: A strategy for statistical and scientific communication. *Journal of the Royal Statistical Society Series A: Statistics in Society, with discussion*, 184(2):407–431, 2021.
- Flemming Topsøe. Information-theoretical optimization techniques. *Kybernetika*, 15(1):8–27, 1979.

A Proofs for Section 2

Proof of Lemma 2.3. The implication $\mu = \nu \Rightarrow H(\nu \mid \mu) = 0$ is immediate from the definition. Conversely, suppose that $H(\nu \mid \mu) = 0$ and consider some $A \in \mathcal{F}$, along with the partition $\pi = \{A, A^c\}$. Then the finite-dimensional entropy

$$H_\pi(\nu \mid \mu) = \sum_{A \in \pi} \nu(A) \log \frac{\nu(A)}{\mu(A)}$$

is nonnegative, since it is the relative entropy between the two probability vectors $(\nu(A), \nu(A^c))$ and $(\mu(A), \mu(A^c))$. Hence, we must have $H_\pi(\nu \mid \mu) = 0$. The finite-dimensional relative entropy of two probability vectors is zero if and only if the two vectors are equal. Therefore

$$\nu(A) = \mu(A) \quad \text{and} \quad \nu(A^c) = \mu(A^c).$$

Since $A \in \mathcal{F}$ was arbitrary, it follows that $\nu = \mu$. $\square$

Proof of Lemma 2.4. Write

$$F(f) = \int f \, d\nu - \log \int e^f \, d\mu, \quad f \in \mathcal{B}_b.$$

We first show that $F(f) \leq H(\nu \mid \mu)$ for every $f \in \mathcal{B}_b$. Suppose first that f is simple, say $f = \sum_{i=1}^n c_i 1_{A_i}$ for some finite measurable partition $\pi = \{A_1, \dots, A_n\}$. Set

$$q_i = \nu(A_i), \quad p_i = \mu(A_i), \quad \kappa = \sum_{i=1}^n p_i e^{c_i}.$$

If $q_i > 0$ and $p_i = 0$ for some i , then $H_\pi(\nu \mid \mu) = \infty$, so there is nothing to prove. Otherwise $\kappa > 0$ and we define $r_i = p_i e^{c_i} / \kappa$. Then (r_i) is a probability vector, and recalling convention (9),

$$F(f) = \sum_{i=1}^n q_i (c_i - \log \kappa) = \sum_{i=1}^n q_i \log \frac{r_i}{p_i} = \sum_{i=1}^n q_i \log \frac{q_i}{p_i} - \sum_{i=1}^n q_i \log \frac{q_i}{r_i}.$$

Since $\sum_{i=1}^n q_i \log(q_i/r_i) \geq 0$, it follows that

$$F(f) \leq \sum_{i=1}^n q_i \log \frac{q_i}{p_i} = H_\pi(\nu \mid \mu) \leq H(\nu \mid \mu).$$

Now let $f \in \mathcal{B}_b$ be arbitrary. Choose simple functions $f_m \in \mathcal{B}_b$ such that $\|f_m - f\|_\infty \rightarrow 0$. Then

$$\left| \int f_m \, d\nu - \int f \, d\nu \right| \leq \|f_m - f\|_\infty \rightarrow 0.$$

Also, if $\varepsilon_m = \|f_m - f\|_\infty$, then

$$e^{-\varepsilon_m} \int e^{f_m} \, d\mu \leq \int e^f \, d\mu \leq e^{\varepsilon_m} \int e^{f_m} \, d\mu,$$

so

$$\left| \log \int e^f \, d\mu - \log \int e^{f_m} \, d\mu \right| \leq \varepsilon_m \rightarrow 0.$$

Therefore $F(f_m) \rightarrow F(f)$. Since $F(f_m) \leq H(\nu \mid \mu)$ for every m , we conclude that

$$\sup_{f \in \mathcal{B}_b} F(f) \leq H(\nu \mid \mu).$$

For the reverse inequality, fix a finite measurable partition $\pi = \{A_1, \dots, A_n\}$, and again write $q_i = \nu(A_i)$ and $p_i = \mu(A_i)$. If $q_i > 0$ and $p_i = 0$ for some i , then $H_\pi(\nu \mid \mu) = \infty$. In that case, letting $f_m = m 1_{A_i}$, we obtain

$$F(f_m) \geq m q_i - \log \int e^{f_m} \, d\mu = m q_i - \log \mu(A_i^c) \uparrow \infty.$$

Hence $\sup_{f \in \mathcal{B}_b} F(f) = \infty = H_\pi(\nu \mid \mu)$. Now assume instead that $q_i = 0$ whenever $p_i = 0$. For $m \in \mathbb{N}$, define

$$c_i^{(m)} = \begin{cases} \log(q_i/p_i), & q_i > 0, \\ -m, & q_i = 0, \end{cases} \quad f_m = \sum_{i=1}^n c_i^{(m)} \mathbf{1}_{A_i} \in \mathcal{B}_b.$$

Then

$$\int f_m d\nu = \sum_{q_i > 0} q_i \log \frac{q_i}{p_i},$$

while

$$\int e^{f_m} d\mu = \sum_{q_i > 0} p_i \frac{q_i}{p_i} + \sum_{q_i = 0} p_i e^{-m} = 1 + e^{-m} \sum_{q_i = 0} p_i.$$

Therefore

$$F(f_m) = \sum_{q_i > 0} q_i \log \frac{q_i}{p_i} - \log \left(1 + e^{-m} \sum_{q_i = 0} p_i \right) \rightarrow \sum_{i=1}^n q_i \log \frac{q_i}{p_i} = H_\pi(\nu \mid \mu).$$

Hence $H_\pi(\nu \mid \mu) \leq \sup_{f \in \mathcal{B}_b} F(f)$. Since π was arbitrary, taking the supremum over all finite measurable partitions gives $H(\nu \mid \mu) \leq \sup_{f \in \mathcal{B}_b} F(f)$, concluding the proof of the first part.

For the second part, if $\int f^- d\nu = \infty$, the claim follows from our convention for extended integrals. Otherwise, set $f_n = (f \vee (-n)) \wedge n \in \mathcal{B}_b$. From the first part of the lemma, we then have

$$\int f_n d\nu - \log \int e^{f_n} d\mu \leq H(\nu \mid \mu).$$

By monotone convergence applied separately to the positive and negative parts, $\int f_n d\nu \rightarrow \int f d\nu$. Moreover,

$$e^f \wedge e^n \leq e^{f_n} \leq e^f \wedge e^n + e^{-n},$$

and hence $\int e^{f_n} d\mu \rightarrow \int e^f d\mu$. Letting $n \uparrow \infty$ proves the claim. $\square$

Proof of Proposition 2.5. Fix $\mu \in \text{ba}_+$ and let $\mu = \mu_c + \mu_p$ denote its Yosida–Hewitt decomposition. Since $\mu_c \leq \mu$, we immediately have $H(\mathbf{Q} \mid \mu) \leq H(\mathbf{Q} \mid \mu_c)$.

We next prove the reverse inequality. Let $\mathbf{R} = \mathbf{Q} + \mu_c$. Since $\mathbf{R}$ is countably additive and μ_p is purely finitely additive, there is no nonzero positive finitely additive measure dominated by both μ_p and $\mathbf{R}$. Equivalently, we may choose a sequence $(B_n) \subseteq \mathcal{F}$ such that

$$\mu_p(B_n) + \mathbf{R}(B_n^c) \rightarrow 0.$$

Fix now a finite measurable partition $\pi = \{A_1, \dots, A_m\}$. For each n , consider the finite measurable partition $\pi_n = \{A_1 \cap B_n, \dots, A_m \cap B_n, B_n^c\}$. Then $H(\mathbf{Q} \mid \mu) \geq H_{\pi_n}(\mathbf{Q} \mid \mu)$. For each $i = 1, \dots, m$, we have $\mathbf{Q}(A_i \cap B_n) \rightarrow \mathbf{Q}(A_i)$, and

$$\mu(A_i \cap B_n) = \mu_c(A_i \cap B_n) + \mu_p(A_i \cap B_n) \rightarrow \mu_c(A_i).$$

Hence, using (9), which implies the lower semicontinuity of $q \log(q/p)$, yields

$$\liminf_{n \uparrow \infty} \sum_{i=1}^m \mathbf{Q}(A_i \cap B_n) \log \frac{\mathbf{Q}(A_i \cap B_n)}{\mu(A_i \cap B_n)} \geq \sum_{i=1}^m \mathbf{Q}(A_i) \log \frac{\mathbf{Q}(A_i)}{\mu_c(A_i)} = H_\pi(\mathbf{Q} \mid \mu_c).$$

The remaining atom B_n^c does not affect the lower bound, since $\mathbf{Q}(B_n^c) \rightarrow 0$ and $\mathbf{Q}(B_n^c) \log(\mathbf{Q}(B_n^c)/\mu(B_n^c))$ has nonnegative limit inferior. Therefore,

$$H(\mathbf{Q} \mid \mu) \geq \liminf_{n \uparrow \infty} H_{\pi_n}(\mathbf{Q} \mid \mu) \geq H_\pi(\mathbf{Q} \mid \mu_c).$$

Taking the supremum over all finite measurable partitions π yields $H(Q \mid \mu) \geq H(Q \mid \mu_c)$, hence the first claim is established.

Now consider $\nu \in \text{ba}_1 \setminus \mathcal{M}_1$. Then continuity from above fails for ν . Hence there exists a nonincreasing sequence $(A_n) \subseteq \mathcal{F}$ such that $A_n \downarrow \emptyset$ but there exists some $\delta > 0$ such that $\nu(A_n) \geq \delta$ for all n . On the other hand, since $P \in \mathcal{M}_+$ is countably additive and finite, we have $P(A_n) \downarrow 0$. Consider the two-point partitions $\pi_n = \{A_n, A_n^c\}$. Then

$$H(\nu \mid P) \geq H_{\pi_n}(\nu \mid P) \geq \nu(A_n) \log \frac{\nu(A_n)}{P(A_n)} + \nu(A_n^c) \log \frac{\nu(A_n^c)}{P(\Omega)}.$$

The first term now tends to ∞ , whereas the second term is bounded from below. This yields the claim. $\square$

Proof of Lemma 2.6. Since $\mathcal{E}_b \subseteq \mathcal{E}$, we only have to argue $\sup_{Y \in \mathcal{E}_b} \mathbb{E}_Q[\log Y] \geq \mathbb{E}_Q[\log X]$ for all $X \in \mathcal{E}$ such that $\mathbb{E}_Q[(\log X)^-] < \infty$. We fix such an X and define $X_m = X \wedge m \in \mathcal{E}_b$ for $m \in \mathbb{N}$. Then by monotone convergence, $\mathbb{E}_Q[\log X_m] \uparrow \mathbb{E}_Q[\log X]$. This yields the statement. $\square$

B Details of Example 3.3

We first show that the weak-* closed convex hulls of $\mathcal{P}$ and $\mathcal{Q}$ have a common finitely additive element. Since ba_1 is weak-* compact, the sequence P_n has a weak-* convergent subnet P_{n_α} , say with limit $\mu \in \text{ba}_1$. Since n_α is a subnet of the sequence of integers, we have $n_\alpha \uparrow \infty$. For any $f \in \mathcal{B}_b$,

$$\int f dQ_{n_\alpha} - \int f dP_{n_\alpha} = \frac{1}{n_\alpha} \sum_{k=2}^{n_\alpha+1} f(k) - \frac{1}{n_\alpha} \sum_{k=1}^{n_\alpha} f(k) = \frac{f(n_\alpha+1) - f(1)}{n_\alpha}.$$

Thus

$$\left| \int f dQ_{n_\alpha} - \int f dP_{n_\alpha} \right| \leq \frac{2\|f\|_\infty}{n_\alpha} \downarrow 0.$$

Since $P_{n_\alpha} \rightarrow \mu$ in $\sigma(\text{ba}, \mathcal{B}_b)$, it follows that also $Q_{n_\alpha} \rightarrow \mu$ in $\sigma(\text{ba}, \mathcal{B}_b)$. Hence $\mu \in \overline{\text{co}}^*(\mathcal{P}) \cap \overline{\text{co}}^*(\mathcal{Q})$.

We next show that μ is not countably additive. Fix $m \in \mathbb{N}$. Since $n_\alpha \uparrow \infty$, we have $n_\alpha \geq m$ eventually, and hence $P_{n_\alpha}(\{m\}) = \frac{1}{n_\alpha}$ eventually. Therefore $\mu(\{m\}) = \lim_\alpha P_{n_\alpha}(\{m\}) = 0$. If μ were countably additive, then

$$1 = \mu(\mathbb{N}) = \sum_{m=1}^{\infty} \mu(\{m\}) = 0,$$

a contradiction. Thus $\mu \notin \mathcal{M}_1$.

Lemma 2.3 yields $H(\mu \mid \mu) = 0$. Since $\mu \in \overline{\text{co}}^*(\mathcal{P}) \cap \overline{\text{co}}^*(\mathcal{Q})$ and since relative entropy is nonnegative, we conclude that

$$\min_{\nu \in \overline{\text{co}}^*(\mathcal{Q})} \min_{\lambda \in \overline{\text{co}}^*(\mathcal{P})} H(\nu \mid \lambda) = 0. \quad (30)$$

Let (ν^*, λ^*) be now any minimizing pair in (30). By Lemma 2.3, this implies $\nu^* = \lambda^*$. Hence every minimizing pair is of the form (λ^*, λ^*) , with $\lambda^* \in \overline{\text{co}}^*(\mathcal{P}) \cap \overline{\text{co}}^*(\mathcal{Q})$. It remains to show that this intersection contains only purely finitely additive probability measures.

Indeed, consider any $\lambda \in \overline{\text{co}}^*(\mathcal{P}) \cap \overline{\text{co}}^*(\mathcal{Q})$. For every $n, k \in \mathbb{N}$, we have $P_n(\{k\}) \geq P_n(\{k+1\})$. This inequality is preserved under finite convex combinations and under weak-* limits because evaluation at $\{k\}$ and $\{k+1\}$ is weak-* continuous. Hence, for every $k \in \mathbb{N}$, $\lambda(\{k\}) \geq \lambda(\{k+1\})$.

On the other hand, every Q_n satisfies $Q_n(\{1\}) = 0$, and this property is also preserved under finite convex combinations and weak-* limits. Since $\lambda \in \overline{\text{co}}^*(\mathcal{Q})$, it follows that $\lambda(\{1\}) = 0$. Therefore,

$$0 \leq \lambda(\{k\}) \leq \lambda(\{1\}) = 0, \quad k \in \mathbb{N},$$

and hence $\lambda(\{k\}) = 0$ for every k . Let now $\lambda = \lambda_c + \lambda_p$ be the Yosida–Hewitt decomposition of λ into its countably additive and purely finitely additive parts. Since $\lambda_c \leq \lambda$, we have $\lambda_c(\{k\}) = 0$ for every k . Countable additivity of λ_c then yields

$$\lambda_c(\mathbb{N}) = \sum_{k=1}^{\infty} \lambda_c(\{k\}) = 0.$$

Hence $\lambda_c = 0$, and therefore λ is purely finitely additive.

C Details from Section 4

Example 4.22 presented a simple example of $\mathcal{P}, \mathcal{Q}$ that are convex and weakly compact, and showed how to apply this paper’s techniques to derive strong duality, along with the GROW e-variable. We now do the same for a slightly more sophisticated example, which involves mean shifts of unbounded random variables with bounded second moment.

Example C.1 (Mean shift with second-moment constraints and no domination). *Let $\Omega = \mathbb{R}$ and $Z(\omega) = \omega$, and define*

$$\mathcal{P} = \left\{ P \in \mathcal{M}_1 : \mathbb{E}_P[Z] \leq 0, \mathbb{E}_P[Z^2] \leq 1 \right\},$$

and

$$\mathcal{Q} = \left\{ Q \in \mathcal{M}_1 : \mathbb{E}_Q[Z] \geq 1, \mathbb{E}_Q[Z^2] \leq 2 \right\}.$$

Both classes are convex and weakly compact. Indeed, the second-moment bounds imply tightness, while the moment constraints are closed under weak convergence because the second moments are uniformly bounded and because of the Portmanteau theorem. The classes are not dominated by any common finite reference measure, since $\mathcal{P}$ contains δ_ω for every $\omega \in [-1, 0]$ and $\mathcal{Q}$ contains δ_ω for every $\omega \in [1, \sqrt{2}]$.

Let

$$\varphi = \frac{1 + \sqrt{5}}{2}, \quad \omega_- = -\frac{1}{\varphi}, \quad \omega_+ = \varphi.$$

Define

$$P^* = \frac{\varphi^2}{\varphi^2 + 1} \delta_{\omega_-} + \frac{1}{\varphi^2 + 1} \delta_{\omega_+},$$

and

$$Q^* = \frac{1}{\varphi^2 + 1} \delta_{\omega_-} + \frac{\varphi^2}{\varphi^2 + 1} \delta_{\omega_+}.$$

Then $\mathbb{E}_{P^*}[Z] = 0, \mathbb{E}_{P^*}[Z^2] = 1$, so $P^* \in \mathcal{P}$, while $\mathbb{E}_{Q^*}[Z] = 1, \mathbb{E}_{Q^*}[Z^2] = 2$, so $Q^* \in \mathcal{Q}$.

On the support $\{\omega_-, \omega_+\}$, the likelihood ratio is

$$\frac{dQ^*}{dP^*}(\omega_-) = \frac{1}{\varphi^2}, \quad \frac{dQ^*}{dP^*}(\omega_+) = \varphi^2,$$

yielding $H(Q^* | P^*) = \frac{2 \log \varphi}{\sqrt{5}}$. We next choose a useful pointwise version of the likelihood ratio. Larsson et al. [2026a] indicate that all e-variables can be dominated by one with a quadratic form, so below we choose the quadratic to match the value of $\frac{dQ^*}{dP^*}$ at ω_- and ω_+ . Let

$$A = \frac{2}{5} \left(1 + \frac{4 \log \varphi}{\sqrt{5}} \right),$$

and define

$$X^*(x) = A(1+x) + (1-A)x^2, \quad x \in \mathbb{R}.$$

Then $2/3 < A < 3/4$, and $X^*(x) > 0$ for all $x \in \mathbb{R}$. Moreover, since $1 + \varphi = \varphi^2$ and $1 - 1/\varphi = 1/\varphi^2$, we have

$$X^*(\omega_-) = \frac{1}{\varphi^2}, \quad X^*(\omega_+) = \varphi^2.$$

Thus X^* is a P^* -version of dQ^*/dP^* , and

$$\mathbb{E}_P[X^*] = A(1 + \mathbb{E}_P[Z]) + (1-A)\mathbb{E}_P[Z^2] \leq 1$$

yields $X^* \in \mathcal{E}$ and hence that P^* is the RIPr of Q^* onto $\mathcal{P}_{\text{eff}}$ (by Larsson et al. [2025]).

We now show that Assumption 4.9 holds. First note that

$$\inf_{R \in \mathcal{Q}} H(R \mid P^*) = H(Q^* \mid P^*).$$

Indeed, since P^* is supported on $\{\omega_-, \omega_+\}$, any R with $H(R \mid P^*) < \infty$ must also be supported on $\{\omega_-, \omega_+\}$. Writing

$$R = (1-q)\delta_{\omega_-} + q\delta_{\omega_+},$$

the constraints $\mathbb{E}_R[Z] \geq 1$ and $\mathbb{E}_R[Z^2] \leq 2$ force $q = \frac{\varphi^2}{\varphi^2+1}$, and hence $R = Q^*$.

It remains to check that $\mathbb{E}_Q[\log X^*] \geq H(Q^* \mid P^*)$ for every $Q \in \mathcal{Q}$. For the above choice of A , a direct one-dimensional calculus check shows that with

$$c = \frac{A-2}{4} < 0, \quad \lambda = 4A-2 > 0, \quad \beta = \frac{3}{2}A-1 > 0,$$

we have

$$h(\omega) = c + \lambda\omega - \beta\omega^2 \leq \log X^*(\omega) \quad \text{for all } \omega \in \mathbb{R},$$

with equality at ω_- and ω_+ . Therefore, for any $Q \in \mathcal{Q}$,

$$\mathbb{E}_Q[\log X^*] \geq \mathbb{E}_Q[h(Z)] = c + \lambda\mathbb{E}_Q[Z] - \beta\mathbb{E}_Q[Z^2] \geq c + \lambda - 2\beta.$$

Since Q^* has $\mathbb{E}_{Q^*}[Z] = 1$ and $\mathbb{E}_{Q^*}[Z^2] = 2$, and since $h(Z) = \log(X^*)$ on the support of Q^* ,

$$c + \lambda - 2\beta = \mathbb{E}_{Q^*}[h(Z)] = \mathbb{E}_{Q^*}[\log X^*] = H(Q^* \mid P^*).$$

Thus

$$\mathbb{E}_Q[\log X^*] \geq \inf_{R \in \mathcal{Q}} H(R \mid P^*) \quad \text{for every } Q \in \mathcal{Q}.$$

This is Assumption 4.9. Consequently, the likelihood-ratio version X^* witnesses strong duality:

$$G = \sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] = H(Q^* \mid P^*) = \frac{2 \log \varphi}{\sqrt{5}}.$$

Clearly X^* is also continuous, though not bounded, but Theorem 4.19 yields that $G = G_b = G_{bb} = G_{bc}$ as well.

D Details from Section 5

Derivation of $\inf_{Q \in \mathcal{Q}_a} H(Q | P)$ in Example 5.3. Let $A_m = [m, \infty)$, $\alpha_m = Q^*(A_m)$, and let R_m be the conditional distribution of Q^* on A_m , i.e., $R_m(\cdot) = Q^*(\cdot | A_m)$. A direct calculation gives $\alpha_m = \frac{1}{3(m+1)^2} + \frac{2}{3(m+1)^3}$. Moreover, $M_m = \int s dR_m(s) \uparrow \infty$. Indeed,

$$\frac{1}{\alpha_m} \int_m^\infty s q^*(s) ds = \frac{2}{3\alpha_m} \left(\frac{1}{m+1} + \frac{1}{(m+1)^2} - \frac{1}{(m+1)^3} \right) \geq 2m.$$

For all large m , define $\eta_m = \frac{a - \frac{2}{3}}{M_m - \frac{2}{3}} \in (0, 1)$ and set $Q_m = (1 - \eta_m)Q^* + \eta_m R_m$. Then

$$\int s dQ_m(s) = (1 - \eta_m) \frac{2}{3} + \eta_m M_m = a,$$

so $Q_m \in \mathcal{Q}_a$. By convexity of relative entropy we get,

$$H(Q_m | Q^*) \leq (1 - \eta_m)H(Q^* | Q^*) + \eta_m H(R_m | Q^*) = \eta_m \log \frac{1}{\alpha_m} \rightarrow 0.$$

Therefore,

$$H(Q_m | P) = H(Q_m | Q^*) + \mathbb{E}_{Q_m} \left[\log \frac{dQ^*}{dP} \right] = H(Q_m | Q^*) + a + \log \frac{2}{3} \rightarrow a + \log \frac{2}{3}.$$

Hence $\inf_{Q \in \mathcal{Q}_a} H(Q | P) = a + \log \frac{2}{3}$. □

Proof of Theorem 5.4. For every $Q \in \mathcal{Q}_\infty$, $\sup_{X \in \mathcal{E}} \mathbb{E}_Q[\log X] \geq \mathbb{E}_Q[\log Y] = \infty$, and thus from (4), $\inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) = \infty$. Therefore,

$$\inf_{Q \in \mathcal{Q}} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) = \min \left(\inf_{Q \in \mathcal{Q}_I} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P), \inf_{Q \in \mathcal{Q}_\infty} \inf_{P \in \mathcal{P}_{\text{eff}}} H(Q | P) \right) = I.$$

Thus, we only have to prove the first equality in (29), and indeed only $\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \geq I$, needs to be argued. To this end, fix $\varepsilon > 0$. By the definition of I , there exists $X_\varepsilon \in \mathcal{E}$ such that

$$I - \varepsilon \leq \inf_{Q \in \mathcal{Q}_I} \mathbb{E}_Q[\log X_\varepsilon].$$

For $\delta \in (0, 1)$ define now

$$X_{\varepsilon, \delta} = (1 - \delta)X_\varepsilon + \delta Y \in \mathcal{E}.$$

If $Q \in \mathcal{Q}_I$, then $X_{\varepsilon, \delta} \geq (1 - \delta)X_\varepsilon$, and therefore

$$\mathbb{E}_Q[\log X_{\varepsilon, \delta}] \geq \mathbb{E}_Q[\log X_\varepsilon] + \log(1 - \delta) \geq I - \varepsilon + \log(1 - \delta).$$

If $Q \in \mathcal{Q}_\infty$, then $X_{\varepsilon, \delta} \geq \delta Y$, and hence, by (28),

$$\mathbb{E}_Q[\log X_{\varepsilon, \delta}] \geq \log \delta + \mathbb{E}_Q[\log Y] = \infty.$$

The last two displays together yield

$$\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X_{\varepsilon, \delta}] \geq I - \varepsilon + \log(1 - \delta).$$

Letting $\delta \downarrow 0$ and $\varepsilon \downarrow 0$ gives $\sup_{X \in \mathcal{E}} \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[\log X] \geq I$, concluding the proof. □