

Article

Not peer-reviewed version

Minimal Entropy and Entropic Risk Measures: A Unified Framework via Relative Entropy

[Moritz Sohns](#) *

Posted Date: 3 March 2025

doi: 10.20944/preprints202503.0111.v1

Keywords: Entropic Risk Measure; Minimal Entropy Risk Measure; Relative Entropy; Risk-Neutral Measures; Dynamic Risk Measures



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Minimal Entropy and Entropic Risk Measures: A Unified Framework via Relative Entropy

Moritz Sohns^{1,2,3}

¹ Faculty of Economic Studies, University of Finance and Administration, Prague; sohns@mokel.science

² Mathematical Institute, University of Oxford

³ Department of Mathematical Sciences, University of South Africa

Abstract: We introduce a new coherent risk measure called the minimal entropy risk measure. This measure is based on the minimal entropy σ -martingale measure, which itself is inspired by the minimal entropy martingale measure well-known in option pricing. While the minimal entropy martingale measure is commonly used for pricing and hedging, the minimal entropy σ -martingale measure has not previously been studied, nor has it been analyzed as a traditional risk measure. We address this gap by clearly defining this new risk measure and examining its fundamental properties. In addition, we revisit the entropic risk measure, typically expressed through an exponential formula. We provide an alternative definition using a supremum over Kullback–Leibler divergences, making its connection to entropy clearer. We verify important properties of both risk measures, such as convexity and coherence, and extend these concepts to dynamic situations. We also illustrate their behavior in scenarios involving optimal risk transfer. Our results link entropic concepts with incomplete-market pricing and demonstrate how both risk measures share a unified entropy-based foundation once market constraints are considered.

Keywords: entropic risk measure; minimal entropy risk measure; relative entropy; risk-neutral measures; dynamic risk measures

MSC: 91B70;60H30;91G80

1. Introduction

Risk measures play an essential role in both academic research and financial practice, as they provide a systematic way to assess the potential losses of a financial position. Their importance has been highlighted ever since the seminal work of [2], which introduced an axiomatic framework for coherent risk measures. Subsequent studies, such as ([23], Chapter 4) or [12], have refined and extended these ideas.

In the monetary risk-measure framework, one models a financial position by a real-valued random variable X on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The number $X(\omega)$ is the discounted net worth of the position if scenario ω occurs. A monetary risk measure ρ assigns a real number $\rho(X)$, which represents the minimal amount of capital needed to make X acceptable according to certain risk criteria. Desirable properties include Monotonicity (increasing payoffs lowers the risk), and Translation Invariance (adding a sure amount of cash decreases the risk by the same amount), and when ρ is further assumed to be convex or coherent, it reflects the benefits of diversification. We refer to [2,12,23] for standard references on these axioms.

A preferred way to price financial claims in financial mathematics (and hence risk measurement, when viewed from a market perspective) is to determine the price as expectation of the discounted payoff under an equivalent (local) martingale measure. However, in general markets and for certain price processes, a (local) martingale measure need not exist. Instead, from the more general perspective of no-arbitrage (or no free lunch with vanishing risk, to be more precise), one can only ensure the

existence of an equivalent σ -martingale measure (E σ MM). See, for instance, [15,33,48] for details on σ -martingale arguments. In addition, in incomplete markets, there may be multiple such measures, and a natural question is how to select one “best” or “preferred” measure among the many.

One popular selection criterion in incomplete markets is to pick the measure that is “closest” to the real-world probability measure \mathbf{P} in terms of the relative entropy (also known as the Kullback–Leibler divergence); see [13,24,27,40]. Minimizing relative entropy leads to the well-known minimal entropy martingale measure, a construction which has proven valuable in option pricing and hedging and which is closely connected to maximizing the investor’s utility (see Section 4). However, most of the literature focuses on the local martingale setting. The corresponding minimal entropy σ -martingale measure has not been studied, even though, in some markets, one must work with E σ MMs. We close this gap by introducing and studying the minimal entropy σ -martingale measure:

$$\mathbf{Q}^E = \operatorname{argmin}_{\mathbf{Q} \in \mathcal{M}_\sigma} H(\mathbf{Q} \mid \mathbf{P}),$$

where \mathcal{M}_σ denotes the set of E σ MMs and $H(\cdot \mid \mathbf{P})$ is the relative entropy with respect to \mathbf{P} . The associated minimal entropy risk measure is then

$$\rho^E(X) = \mathbf{E}_{\mathbf{Q}^E}[-X].$$

This measure is an extension of the classical minimal entropy martingale measure (since an EMM is a special case of an E σ MM) but is strictly more general whenever no local martingale measure exists. Notably, while the minimal entropy martingale measure has been studied for pricing and hedging, it has not been viewed or analyzed as a traditional risk measure in the classical sense. Nor has the σ -martingale version been examined at all. In this paper, we fill this gap by proving that the minimal entropy σ -martingale measure leads to a coherent risk measure with desirable properties.

One of the most popular risk measures is the entropic risk measure. At first glance, one might suspect connections between entropic risk measures and minimal entropy methods because of the shared “entropy” label. Nevertheless, the entropic risk measure is typically introduced via the exponential formula

$$e_\gamma(X) = \frac{1}{\gamma} \log(\mathbf{E}[e^{-\gamma X}]),$$

which does not reveal any connection to entropy. Its name stems from its robust representation

$$e_\gamma(X) = \sup_{\mathbf{Q} \ll \mathbf{P}} \left\{ \mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} \mid \mathbf{P}) \right\},$$

showing that it penalizes deviations from \mathbf{P} in proportion to the relative entropy. Although this measure is well understood as a convex, time-consistent risk measure, the label “entropic” might not be fully transparent when beginning from an exponential definition. Here, we show that one can equivalently start with a relative-entropy-based formulation, making the “entropic” nature manifest. We show that the measure can be equivalently defined, and one arrives at the same results and conclusions quite easily.

This measure is strictly convex (rather than linear) in the payoff X . In contrast, the minimal entropy risk measure only picks out a single measure – the one that eliminates arbitrage and is of the least relative entropy. As a result, it turns out to be coherent. Despite these structural differences, both measures share a fundamental entropy-based underpinning.

In summary, the main contributions of this paper are:

- We introduce the minimal entropy σ -martingale measure for general semimartingale models (with dividends), which is new to the literature.
- We prove that the induced minimal entropy risk measure is coherent and extends the classical minimal entropy martingale measure results.

- We define the entropic risk measure via its robust representation and, therefore, provide an alternative approach which highlights its connection to entropy.
- We demonstrate key properties including convexity, coherence, dynamic consistency, and optimal risk transfer for both measures, thereby revealing that minimal-entropy techniques are not only pricing tools but also valid risk measures in their own right.
- We provide some estimates, comparing the two risk measures to their real-world expectations.

The paper is organized as follows. Section 2 contains the precise definitions of both the minimal entropy σ -martingale measure (by minimizing relative entropy under the no-arbitrage condition) and the entropic risk measure (via a relative-entropy supremum), along with existence criteria. We establish their existence and compare their properties, highlighting the new σ -martingale generalization. Section 3 focuses on the definition of monetary risk measures, convexity and coherence, proving that the minimal entropy risk measure is coherent while the entropic measure is convex. In Section 4, we explore duality and highlight the deeper relationship between entropy-based valuations and risk measures and, in particular, show that our definition of the entropic risk measure is equivalent to the more common definition in the literature. Section 5 provides dynamic versions, establishing time consistency. Finally, Section 6 discusses optimal risk transfer and how each of these risk measures behaves in that context.

2. Definition and Existence

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, L^∞ the space of bounded random variables and \mathcal{P}_L the set of probability measures on $L^\infty(\mathcal{F})$. Furthermore, let $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ be a filtration with $\mathcal{F}_T = \mathcal{F}$, and let S be a (potentially multi-dimensional) stochastic process adapted to \mathbb{F} . We assume $\hat{S} = (\hat{S}_t)_{0 \leq t \leq T}$ is the discounted price process (for details on discounting, see [47]).

In most models, there exists at least one probability measure such that the discounted process \hat{S} is a local martingale under this probability measure. However, [15] showed that in an arbitrage-free market (more precise a market that satisfies no risk with vanishing risk), you can only assume that a probability measure exists such that \hat{S} is a σ -martingale under this probability measure. Therefore, it makes sense to study σ -martingales, equivalent σ -martingale measures and derived risk measures.

First, let us recall:

Definition 1. A one-dimensional semimartingale S is called a σ -martingale if there exists a sequence of predictable sets D_n such that:

- (i) $D_n \subset D_{n+1}$ for all n ;
- (ii) $\bigcup_{n=1}^\infty D_n = \Omega \times \mathbb{R}_+$;
- (iii) for each $n \geq 1$, the process $\mathbb{1}_{D_n} \cdot S$ is a uniformly integrable martingale.

Such a sequence $(D_n)_{n \in \mathbb{N}}$ is called a σ -localizing sequence. A d -dimensional semimartingale is called a σ -martingale if each of its components is a one-dimensional σ -martingale.

These notions, introduced by Chou [7], further explored by Émery [51], and subsequently discussed by, for example, Jacod and Shiryaev [31] and Kallsen [33], allow certain processes to fail to be true local martingales yet still satisfy a limiting or piecewise martingale property.

Definition 2. An equivalent σ -martingale measure (E σ MM) is a probability measure \mathbf{Q} with $\mathbf{Q} \sim \mathbf{P}$ such that \hat{S} is a σ -martingale under \mathbf{Q} . The set of all E σ MM is denoted by \mathcal{M}_σ .

We also define a broader sets of absolutely continuous measures ($\mathbf{Q} \ll \mathbf{P}$) under which \hat{S} is a σ -martingale:

$$\mathcal{M}_\sigma^{ac} := \{\mathbf{Q}; \mathbf{Q} \ll \mathbf{P}, \hat{S} \text{ is a } \sigma\text{-martingale under } \mathbf{Q}\},$$

For more details on σ -martingales, see, for example, [28,33,48].

Definition 3. The relative entropy $H(\mathbf{Q} \mid \mathbf{P})$ of a probability measure \mathbf{Q} with respect to \mathbf{P} is defined as

$$H(\mathbf{Q} \mid \mathbf{P}) = \begin{cases} \int_{\Omega} \log\left(\frac{d\mathbf{Q}}{d\mathbf{P}}\right) d\mathbf{Q} = \mathbf{E}_{\mathbf{Q}}\left(\log\left(\frac{d\mathbf{Q}}{d\mathbf{P}}\right)\right) & \text{for } \mathbf{Q} \ll \mathbf{P}, \\ \infty & \text{otherwise.} \end{cases}$$

The relative entropy provides a notion of distance between a probability measure \mathbf{Q} and a reference probability measure \mathbf{P} . Even though it can be interpreted as a distance, it is not a metric since neither the symmetry property nor the triangle inequality holds.

Beyond financial mathematics, relative entropy is a crucial concept in statistical physics and information theory, where it is referred to as the Kullback–Leibler divergence [46]. It also plays a fundamental role in large deviations theory, where it underpins results such as Sanov’s theorem [8]. A comprehensive summary of its applications can be found in [6].

Further illustrations of this definition are available in [11], including a demonstration of how relative entropy can be understood as a distance measure. In statistics, this interpretation is reinforced by results such as Stein’s Lemma, which can be found in [30, Satz 10.4].

Definition 4.

- (a) An equivalent σ -martingale measure $\mathbf{Q}^E \in \mathcal{M}_{\sigma}$ is called the minimal entropy σ -martingale measure if it minimizes the relative entropy (or Kullback–Leibler divergence) $H(\cdot \mid \mathbf{P})$ among all $\mathbf{Q} \in \mathcal{M}_{\sigma}$, i.e.,

$$H(\mathbf{Q}^E \mid \mathbf{P}) = \inf_{\mathbf{Q} \in \mathcal{M}_{\sigma}} H(\mathbf{Q} \mid \mathbf{P}).$$

The corresponding minimal entropy risk measure ρ^E is defined by

$$\rho^E(X) := \mathbf{E}_{\mathbf{Q}^E}[-X].$$

- (b) For $X \in L^{\infty}$ and $\gamma > 0$, the entropic risk e_{γ} is defined as

$$e_{\gamma}(X) := \sup_{\mathbf{Q} \in \mathcal{P}_L} \left(\mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} \mid \mathbf{P}) \right). \quad (1)$$

Remark 1. In the above definition of e_{γ} , the parameter $\gamma > 0$ can be viewed as a risk-aversion level or scaling factor, much like in exponential-utility frameworks where larger γ corresponds to greater risk aversion. Thus, e_{γ} can be seen as an “entropic” or “exponential-penalized” valuation: the higher the γ , the stronger the penalty on deviating from the reference measure \mathbf{P} .

Note that the existence of these risk measures is not immediately clear. In particular, the existence of a minimal entropy σ -martingale measure can still depend on the properties of S . For instance, if the market is not arbitrage-free (e.g., the No Free Lunch With Vanishing Risk condition fails) and S represents the price process of a tradable asset, then no equivalent σ -martingale measure can exist. Even if the market does satisfy NFLVR, one may still require additional conditions on the price processes (e.g., local boundedness or boundedness from below) to ensure that there is some equivalent σ -martingale measure carrying finite relative entropy. For further details, see [14], [15], and [16].

As opposed to the minimal entropy σ -martingale measure, the entropic risk always exists.

Theorem 1. The entropic risk always exists.

Proof. It suffices to show that e_{γ} is indeed a finite number.

Let $X \in L^{\infty}(\mathcal{F})$. It is straightforward to verify:

$$\lim_{\gamma \searrow 0} |e_{\gamma}(X)| = |\mathbf{E}_{\mathbf{P}}[-X]| < \infty$$

and

$$\lim_{\gamma \rightarrow \infty} |e_\gamma(X)| = |\text{ess sup}(-X)| < \infty$$

because X is almost surely bounded.

Since $e_\gamma(X)$ is increasing in γ , we conclude for a fixed $\tilde{\gamma} > 0$:

$$-\infty < \lim_{\gamma \searrow 0} e_\gamma(X) \leq e_{\tilde{\gamma}}(X) \leq \lim_{\gamma \rightarrow \infty} e_\gamma(X) < \infty. \quad \square$$

While the minimal entropy σ -martingale measure is new in the literature, the notion of the minimal entropy martingale measure has been extensively studied in the literature. The definition of the latter goes analogue to ours, but instead of equivalent σ -martingale measures, one minimizes the entropy among all equivalent martingale measures (see, for example, [29,38,39,43,45]). There are numerous criteria in the literature for the existence of the minimal entropy martingale measure. Most of these criteria impose strong conditions on the stochastic process \hat{S} . A prominent example is when \hat{S} is a Lévy process. For this case, Theorem 3.1 in [26] provides a proof of existence, albeit with a highly technical approach. Another simple result is the existence of the minimal entropy martingale measure if \hat{S} is bounded as shown in [24, Theorem 2.1]. Other results depend on the specific form of the Radon-Nikodym derivative, e.g., [29] or are applicable only in discrete time settings, e.g., [23, Corollary 3.27].

For the minimal entropy σ -martingale measure, things are a bit easier. We now state and prove a theorem showing that, under a mild condition – namely, the existence of at least one σ -martingale measure \mathbf{Q} with finite $H(\mathbf{Q} | \mathbf{P})$ – one obtains the existence (and uniqueness) of a minimal entropy σ -martingale measure. This holds even if \hat{S} is not locally bounded and even if no equivalent local martingale measures exists.

For the proof, we need some lemmas.

Lemma 1. *Let \mathcal{P} be the set of probability measures on (Ω, \mathcal{A}) . We have:*

- (a) $H(\mathbf{Q} | \mathbf{P}) \geq 0$. Furthermore, $H(\mathbf{Q} | \mathbf{P}) = 0$ if and only if $\mathbf{Q} = \mathbf{P}$.
- (b) The mapping

$$: \begin{cases} \mathcal{P} & \rightarrow \mathbb{R} \\ \mathbf{Q} & \mapsto H(\mathbf{Q} | \mathbf{P}) \end{cases}$$

is convex and strictly convex on the set of probability measures which are absolutely continuous with respect to \mathbf{P} .

Proof. We have:

$$H(\mathbf{Q} | \mathbf{P}) = \begin{cases} \mathbb{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} \log \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right) \right] & \text{for } \mathbf{Q} \ll \mathbf{P}, \\ \infty & \text{else.} \end{cases}$$

hence, it is reasonable to take a closer look at the function ϕ , which we define as a continuous extension of the function $x \mapsto x \log x$ to $[0, \infty)$. Because $\phi(x)'' = \frac{1}{x}$ and Theorem A1, we have $\phi''(x) > 0$ for $x > 0$, which proves the strict convexity of ϕ .

It is easy to see that the function $x \mapsto x \log x - x$ takes its minimum at $x = 1$. Hence, we have $x \log x - x \geq -1$ and thus

$$\phi(x) \geq x - 1 \quad \text{for all } x \geq 0. \quad (2)$$

(a) According to Equation (2), we have for $\mathbf{Q} \ll \mathbf{P}$:

$$\begin{aligned} H(\mathbf{Q} | \mathbf{P}) &= \mathbf{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} \log \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right) \right] \\ &\geq \mathbf{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} - 1 \right] = \mathbf{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} \right] - 1 \\ &= \mathbf{E}_{\mathbf{Q}}[1] - 1 = 0. \end{aligned}$$

Now let $H(\mathbf{Q} | \mathbf{P}) = 0$. We must show that the equality $\mathbf{Q} = \mathbf{P}$ follows. We have:

$$\begin{aligned} 0 &= H(\mathbf{Q} | \mathbf{P}) = \mathbf{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} \log \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right) \right] \\ &= \mathbf{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} \log \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right) \right] - \underbrace{\mathbf{E}_{\mathbf{P}} \left[\frac{d\mathbf{Q}}{d\mathbf{P}} - 1 \right]}_{=0} \\ &= \mathbf{E}_{\mathbf{P}} \left[\underbrace{\frac{d\mathbf{Q}}{d\mathbf{P}} \log \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right) - \frac{d\mathbf{Q}}{d\mathbf{P}} + 1}_{\substack{\text{Equation (2)} \\ \geq 0}} \right]. \end{aligned}$$

Hence \mathbf{P} -almost everywhere we have $\frac{d\mathbf{Q}}{d\mathbf{P}} \log \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right) = \frac{d\mathbf{Q}}{d\mathbf{P}} - 1$. Now it follows that $\frac{d\mathbf{Q}}{d\mathbf{P}} = 1$ \mathbf{P} -almost surely. If \mathbf{Q} is not absolutely continuous with respect to \mathbf{P} , the statement is obvious.

(b) For probability measures that are not absolutely continuous with respect to \mathbf{P} , the statement is obvious. So we focus on the case where $\mathbf{Q}_1, \mathbf{Q}_2 \ll \mathbf{P}$. For $\lambda_1, \lambda_2 \geq 0$ with $\lambda_1 + \lambda_2 = 1$, we have:

$$\begin{aligned} H(\lambda_1 \mathbf{Q}_1 + \lambda_2 \mathbf{Q}_2 | \mathbf{P}) &= \mathbf{E}_{\mathbf{P}} \left[\phi \left(\lambda_1 \frac{d\mathbf{Q}_1}{d\mathbf{P}} + \lambda_2 \frac{d\mathbf{Q}_2}{d\mathbf{P}} \right) \right] \\ &\leq \mathbf{E}_{\mathbf{P}} \left[\lambda_1 \phi \left(\frac{d\mathbf{Q}_1}{d\mathbf{P}} \right) + \lambda_2 \phi \left(\frac{d\mathbf{Q}_2}{d\mathbf{P}} \right) \right] \\ &= \lambda_1 \mathbf{E}_{\mathbf{P}} \left[\phi \left(\frac{d\mathbf{Q}_1}{d\mathbf{P}} \right) \right] + \lambda_2 \mathbf{E}_{\mathbf{P}} \left[\phi \left(\frac{d\mathbf{Q}_2}{d\mathbf{P}} \right) \right] \\ &= \lambda_1 H(\mathbf{Q}_1 | \mathbf{P}) + \lambda_2 H(\mathbf{Q}_2 | \mathbf{P}). \end{aligned}$$

Thus, strict convexity follows. \square

Lemma 2. Suppose there exists a measure $\mathbf{Q}_0 \in \mathcal{M}_{\sigma}^{ac}$ with

$$H(\mathbf{Q}_0 | \mathbf{P}) < \infty \quad \text{and} \quad H(\mathbf{Q}_0 | \mathbf{P}) \leq H(\mathbf{Q} | \mathbf{P}) \quad \text{for all } \mathbf{Q} \in \mathcal{M}_{\sigma}^{ac}.$$

Then \mathbf{Q}_0 is the unique Minimal Entropy σ -Martingale Measure.

Proof. Define the set

$$\Gamma := \{ \mathbf{Q} \in \mathcal{M}_{\sigma}^{ac} ; H(\mathbf{Q} | \mathbf{P}) < \infty \}.$$

By assumption, Γ is nonempty because $\mathbf{Q}_0 \in \Gamma$. Thus, we have

$$I(\Gamma, \mathbf{P}) := \inf_{\mathbf{Q} \in \Gamma} H(\mathbf{Q} | \mathbf{P}) < \infty.$$

Next, we show that Γ is convex and closed in total variation:

For any $\mathbf{Q}_1, \mathbf{Q}_2 \in \Gamma$, their convex combination $\alpha \mathbf{Q}_1 + (1 - \alpha) \mathbf{Q}_2$, $\alpha \in [0, 1]$, is also in Γ . Indeed, convex combinations preserve both the finite entropy condition (due to convexity of relative entropy) and the σ -martingale property as σ -martingales form a vector space.

The set \mathcal{M}_σ^{ac} can be expressed through linear constraints as

$$\mathcal{M}_\sigma^{ac} = \{\mathbf{Q} \ll \mathbf{P} : \mathbf{E}_{\mathbf{Q}}[f_i] = 0 \text{ for a suitable family of bounded measurable functions } f_i\}.$$

Sets defined via linear constraints of the form $\{\mathbf{Q} \ll \mathbf{P} : \mathbf{E}_{\mathbf{Q}}[f_i] = 0\}$ are closed in total variation topology. Intersecting with the closed set of measures having finite entropy preserves this closedness.

By applying Csiszár's Theorem A6, we conclude that the set Γ has a unique I -projection $\tilde{\mathbf{Q}}$ of \mathbf{P} , which satisfies:

$$H(\tilde{\mathbf{Q}} | \mathbf{P}) = \inf_{\mathbf{Q} \in \Gamma} H(\mathbf{Q} | \mathbf{P}).$$

By hypothesis, \mathbf{Q}_0 already satisfies this minimality, so it must coincide with $\tilde{\mathbf{Q}}$. Thus, uniqueness follows directly from the strict convexity of relative entropy.

Finally, we verify that \mathbf{Q}_0 is equivalent to \mathbf{P} . Suppose this is not the case. Then there exists $\tilde{\mathbf{Q}} \in \Gamma$, which is equivalent to \mathbf{P} , as given by hypothesis. If \mathbf{Q}_0 is not equivalent to \mathbf{P} , its support is strictly smaller than that of $\tilde{\mathbf{Q}}$, which contradicts minimality via Csiszár–Gibbs inequality (Theorem A7). Hence, $\mathbf{Q}_0 \sim \mathbf{P}$.

Thus, $\mathbf{Q}_0 \in \mathcal{M}_\sigma^{ac}$, is equivalent to \mathbf{P} , and minimizes entropy. Therefore, it is the unique minimal entropy σ -martingale measure. \square

Lemma 3. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $\tilde{\mathbf{Q}}, \mathbf{Q}^1, \mathbf{Q}^2, \dots$ be probability measures on (Ω, \mathcal{F}) with $\tilde{\mathbf{Q}}, \mathbf{Q}^n \sim \mathbf{P}$ for each n and

$$\frac{d\mathbf{Q}^n}{d\mathbf{P}} \xrightarrow[n \rightarrow \infty]{L^1(\mathbf{P})} \frac{d\tilde{\mathbf{Q}}}{d\mathbf{P}}.$$

Suppose X is a semimartingale that is a σ -martingale under each measure \mathbf{Q}^n . Then X is also a σ -martingale under $\tilde{\mathbf{Q}}$.

Proof. Since X is a σ -martingale under each measure \mathbf{Q}^n , there exists, for every n , a suitable family of predictable sets making each localized piece of X a uniformly integrable martingale under \mathbf{Q}^n . By reindexing or combining these families appropriately, we can select a single sequence $(D_k)_{k \in \mathbb{N}}$ of predictable sets for which $\mathbb{1}_{D_k} \cdot X$ is a uniformly integrable martingale simultaneously under every measure \mathbf{Q}^n .

To show that $\mathbb{1}_{D_k} \cdot X$ remains a uniformly integrable martingale under $\tilde{\mathbf{Q}}$, fix arbitrary times $s < t$ and an arbitrary set $A \in \mathcal{F}_s$. Define the bounded random variable

$$Z = \mathbb{1}_A((\mathbb{1}_{D_k} \cdot X)_t - (\mathbb{1}_{D_k} \cdot X)_s).$$

Since $\mathbb{1}_{D_k} \cdot X$ is a martingale under each \mathbf{Q}^n , we immediately have $\mathbf{E}_{\mathbf{Q}^n}[Z] = 0$ for every n . Using the convergence of the Radon–Nikodým derivatives in $L^1(\mathbf{P})$ and boundedness of Z , we get:

$$\mathbf{E}_{\tilde{\mathbf{Q}}}[Z] = \mathbf{E}_{\mathbf{P}}\left[Z \frac{d\tilde{\mathbf{Q}}}{d\mathbf{P}}\right] = \mathbf{E}_{\mathbf{P}}\left[\lim_{n \rightarrow \infty} Z \frac{d\mathbf{Q}^n}{d\mathbf{P}}\right] = \lim_{n \rightarrow \infty} \mathbf{E}_{\mathbf{P}}\left[Z \frac{d\mathbf{Q}^n}{d\mathbf{P}}\right] = \lim_{n \rightarrow \infty} \mathbf{E}_{\mathbf{Q}^n}[Z] = 0.$$

This equality implies that the conditional expectation of $(\mathbb{1}_{D_k} \cdot X)_t$ given \mathcal{F}_s under $\tilde{\mathbf{Q}}$ equals $(\mathbb{1}_{D_k} \cdot X)_s$, establishing the martingale property for each localized piece. Since this argument applies to each predictable set D_k , it follows directly that X is a σ -martingale under $\tilde{\mathbf{Q}}$. \square

Theorem 2. Suppose there exists at least one measure $\mathbf{Q}_1 \in \mathcal{M}_\sigma$ with finite entropy $H(\mathbf{Q}_1 | \mathbf{P}) < \infty$. Then there exists a unique minimal entropy σ -martingale measure $\mathbf{Q}^E \in \mathcal{M}_\sigma$, characterized by

$$H(\mathbf{Q}^E | \mathbf{P}) = \inf_{\mathbf{Q} \in \mathcal{M}_\sigma} H(\mathbf{Q} | \mathbf{P}) < \infty.$$

Proof. We define

$$\mathcal{M}^0 = \{\mathbf{Q} \in \mathcal{M}_\sigma^{ac} : H(\mathbf{Q} | \mathbf{P}) < \infty\}.$$

By assumption, there exists a sequence $\{\mathbf{Q}_n\}_{n \geq 1} \in \mathcal{M}^0$ such that $(H(\mathbf{Q}_n | \mathbf{P}))_{n \in \mathbb{N}}$ is decreasing and satisfies

$$H(\mathbf{Q}_n | \mathbf{P}) \searrow \inf_{\mathbf{Q} \in \mathcal{M}_\sigma^{ac}} H(\mathbf{Q} | \mathbf{P}).$$

Because the sequence of the relative entropies is monotone, we obtain for the convex function $\phi(x) = x \log x$

$$\sup_{n \geq 1} \mathbf{E}_\mathbf{P} \left(\phi \left(\frac{d\mathbf{Q}_n}{d\mathbf{P}} \right) \right) \leq H(\mathbf{Q}_1 | \mathbf{P}),$$

and therefore the sequence $\left(\frac{d\mathbf{Q}_n}{d\mathbf{P}} \right)$ is uniformly integrable according to Theorem A3. By Theorem A4, there exists a subsequence $\left(\frac{d\mathbf{Q}_{n_k}}{d\mathbf{P}} \right)_{k \geq 1}$ which converges weakly in L^1 . However, by Mazur's Lemma (Theorem A5) there exists a sequence $\left(\widetilde{\frac{d\mathbf{Q}_n}{d\mathbf{P}}} \right)_{n \geq 1}$ consisting of convex combinations

$$\frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}} = \sum_{k=n}^{N_n} \lambda_{n_k} \frac{d\mathbf{Q}_{n_k}}{d\mathbf{P}} \quad \left(\lambda_{n_k} \geq 0, \sum_k \lambda_{n_k} = 1 \right), \quad (3)$$

which converges in $L^1(\mathbf{P})$ to some $\frac{d\mathbf{Q}_0}{d\mathbf{P}}$.

The measure $\tilde{\mathbf{Q}}_n$ is uniquely defined if we interpret $\frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}}$ as a Radon–Nikodym derivative. One can see directly that $\tilde{\mathbf{Q}}_n \ll \mathbf{P}$. By assumption, \hat{S} is a σ -martingale with respect to all \mathbf{Q}_n (and hence also w.r.t. all \mathbf{Q}_{n_k}). Now, we want to show it is also a σ -martingale under $\tilde{\mathbf{Q}}_n$. To prove this, let $(D_i)_i$ be a sequence of predictable sets such that $\mathbb{1}_{D_i} \cdot \hat{S}$ is a martingale under a fixed \mathbf{Q}_{n_k} . For each $A \in \mathcal{F}_s$,

$$\begin{aligned} \int_A (\mathbb{1}_{D_i} \cdot \hat{S}_t - \mathbb{1}_{D_i} \cdot \hat{S}_s) d\tilde{\mathbf{Q}}_n &= \int_A (\mathbb{1}_{D_i} \cdot \hat{S}_t - \mathbb{1}_{D_i} \cdot \hat{S}_s) \frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}} d\mathbf{P} \\ &= \int_A (\mathbb{1}_{D_i} \cdot \hat{S}_t - \mathbb{1}_{D_i} \cdot \hat{S}_s) \left(\sum_{k=n}^{N_n} \lambda_{n_k} \frac{d\mathbf{Q}_{n_k}}{d\mathbf{P}} \right) d\mathbf{P} \\ &= \sum_{k=n}^{N_n} \lambda_{n_k} \int_A (\mathbb{1}_{D_i} \cdot \hat{S}_t - \mathbb{1}_{D_i} \cdot \hat{S}_s) \frac{d\mathbf{Q}_{n_k}}{d\mathbf{P}} d\mathbf{P} \\ &= \sum_{k=n}^{N_n} \lambda_{n_k} \int_A (\mathbb{1}_{D_i} \cdot \hat{S}_t - \mathbb{1}_{D_i} \cdot \hat{S}_s) d\mathbf{Q}_{n_k} = 0, \end{aligned}$$

because $\mathbf{Q}_{n_k} \in \mathcal{M}_\sigma^{ac}$. Thus $\mathbb{1}_{D_i} \cdot \hat{S}$ is also a martingale under $\tilde{\mathbf{Q}}_n$, hence \hat{S} is a σ martingale under $\tilde{\mathbf{Q}}_n$.

Moreover, $\frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}}$ converges in L^1 to $\frac{d\mathbf{Q}_0}{d\mathbf{P}}$, and so, by Lemma 3 we have $\mathbf{Q}_0 \in \mathcal{M}_\sigma^{ac}$.

Because of the convexity of the relative entropy (Lemma 1), Equation (3), and since $(H(\mathbf{Q}_n | \mathbf{P}))_{n \in \mathbb{N}}$ is a decreasing sequence, we have

$$\begin{aligned} H(\tilde{\mathbf{Q}}_n | \mathbf{P}) &= H \left(\sum_{k=n}^{N_n} \lambda_k \mathbf{Q}_{n_k} \mid \mathbf{P} \right) \\ &\leq \sum_{k=n}^{N_n} \lambda_k H(\mathbf{Q}_n | \mathbf{P}) \leq \max_{k \in \{n, \dots, N_n\}} H(\mathbf{Q}_{n_k} | \mathbf{P}) = H(\mathbf{Q}_n | \mathbf{P}). \end{aligned}$$

Thus, by applying Fatou's Lemma, we obtain

$$\begin{aligned} H(\mathbf{Q}_0 | \mathbf{P}) &= \mathbf{E}_\mathbf{P} \left[\liminf_{n \rightarrow \infty} \frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}} \log \left(\frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}} \right) \right] \leq \liminf_{n \rightarrow \infty} \mathbf{E}_\mathbf{P} \left[\frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}} \log \left(\frac{d\widetilde{\mathbf{Q}_n}}{d\mathbf{P}} \right) \right] \\ &= \liminf_{n \rightarrow \infty} H(\tilde{\mathbf{Q}}_n | \mathbf{P}) \leq \liminf_{n \rightarrow \infty} H(\mathbf{Q}_n | \mathbf{P}). \end{aligned}$$

Since $H(\mathbf{Q}_n | \mathbf{P})$ is decreasing, it follows that $H(\mathbf{Q}_0 | \mathbf{P})$ is already minimal. Now, all conditions of Lemma 2 are satisfied, so we conclude that \mathbf{Q}_0 is equivalent to \mathbf{P} , proving the existence of the minimal entropy martingale measure.

Uniqueness follows from Lemma 1 and the strict convexity of $H(\cdot | \mathbf{P})$. \square

Remark 2. Theorem 2 shows that the “entropy-minimizing” approach extends smoothly to σ -martingale models: as soon as there is one finite-entropy σ -martingale measure, there must be a unique one of minimal entropy, even if \hat{S} is unbounded or not locally bounded. This measure can be used for entropy-based valuation, risk measurement, or other applications. The local martingale approach, in contrast, might be impossible if $\mathcal{M}_{\text{loc}} = \emptyset$.

3. Convex and Coherent Risk Measures

In this section, we define the notions of monetary, convex, and coherent risk measures. We also highlight the relationships among translation invariance, monotonicity, convexity, positive homogeneity, and subadditivity, emphasizing how these conditions characterize different levels of risk-measure structure.

Definition 5. Let \mathcal{X} be the space of linear bounded, real-valued random variables. and

(a) A mapping $\rho : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ is called a monetary risk measure, if it satisfies for all $X, Y \in \mathcal{X}$

- Monotonicity: If $X \leq Y$ almost surely, then $\rho(X) \geq \rho(Y)$.
- Translation Invariance: For all $m \in \mathbb{R}$, $\rho(X + m) = \rho(X) - m$.

(b) The monetary risk measure ρ is called convex if it also satisfies

- Convexity: For all $X, Y \in \mathcal{X}$ and $\lambda \in [0, 1]$,

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y).$$

(c) A convex risk measure ρ is called coherent if, in addition to monotonicity and translation invariance, it satisfies

- Positive Homogeneity: For all $\alpha > 0$,

$$\rho(\alpha X) = \alpha \rho(X).$$

- Subadditivity: For all $X, Y \in \mathcal{X}$,

$$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

Theorem 3. Let ρ be a monetary risk measure on a linear space of random variables \mathcal{X} .

(a) We say that ρ is normalized if $\rho(0) = 0$. In particular, if ρ is positively homogeneous, then it follows that $\rho(0) = 0$ automatically.

(b) Suppose ρ satisfies translation invariance and monotonicity (as in Definition 5). Then, any two of the following three properties imply the remaining third:

- Convexity,
- Positive Homogeneity,
- Subadditivity.

For a detailed discussion and proof, see [23, Chapter 4].

Theorem 4. The minimal entropy martingale measure and the entropic risk are both convex risk measures. The minimal entropy martingale measure is also a coherent risk measure.

Proof. We start with the minimal entropy martingale measure. It suffices to prove the coherence. The convexity follows from the fact that positive homogeneity and subadditivity imply convexity.

- Monotonicity. Let $X, Y \in L^\infty(\mathcal{F})$ with $\mathbf{P}[X \geq Y] = 1$. We have

$$\rho^E(X) = \mathbf{E}_{\mathbf{Q}^E}[-X] \leq \mathbf{E}_{\mathbf{Q}^E}[-Y] = \rho^E(Y).$$

- Cash Translability. This simply follows from

$$\rho^E(X + c) = \rho^E(-X - c) = \rho^E(-X) - c = \rho^E(X) - c.$$

- Positive Homogeneity and Subadditivity. These follow directly from the linearity of the expectation.

We now address the entropic risk.

- Monotonicity. From $X \geq Y$ almost surely, $\mathbf{E}_{\mathbf{Q}}[-Y] \geq \mathbf{E}_{\mathbf{Q}}[-X]$ follows for all $\mathbf{Q} \in \mathcal{P}_L$, and hence

$$\mathbf{E}[-Y] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \geq \mathbf{E}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P})$$

for all \mathbf{Q} . Thus $e_\gamma(X) \leq e_\gamma(Y)$.

- Cash Translability. With

$$\mathbf{E}[-(X + c)] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) = \mathbf{E}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) - c,$$

we obtain $e_\gamma(X + c) = e_\gamma(X) - c$.

- Convexity. We have

$$\begin{aligned} & \mathbf{E}[-(\lambda X + (1 - \lambda)Y)] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \\ &= \lambda \mathbf{E}[-X] + (1 - \lambda) \mathbf{E}[-Y] - \lambda \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) - (1 - \lambda) \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \\ &= \lambda \left(\mathbf{E}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \right) + (1 - \lambda) \left(\mathbf{E}[-Y] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \right). \end{aligned}$$

And since for general functions f, g and $c_1, c_2 \in \mathbb{R}$, we have

$$\sup_x (c_1 f(x) + c_2 g(x)) \leq c_1 \sup_x f(x) + c_2 \sup_x g(x),$$

the result follows. \square

In general, the entropic risk is not a coherent risk measure, even though it is additive for independent positions. By considering Lemma 1, one can easily see from the definition that in order to be a coherent risk measure, the entropic risk and the risk measure ρ defined as $\rho(X) = \mathbf{E}_{\mathbf{P}}[-X]$ have to coincide.

There is also a coherent version of the entropic risk measure described in [21], which is quite similar to the two risk measures we are examining here.

Finally, to better understand these risk measures, we explore their relationship with the real-world probability measure. The following lemma will be helpful.

Lemma 4. Let $\mathbf{Q}, \mathbf{R} \in \mathcal{M}_\sigma^{ac}$ be absolutely continuous σ -martingale measures, i.e. each of $\frac{d\mathbf{Q}}{d\mathbf{P}}, \frac{d\mathbf{R}}{d\mathbf{P}}$ is in $L^1(\mathbf{P})$, and assume $H(\mathbf{Q} | \mathbf{P})$ and $H(\mathbf{R} | \mathbf{P})$ are both finite. Then, for any bounded random variable $X \in L^\infty(\mathcal{F})$, we have

$$|\mathbf{E}_{\mathbf{Q}}[X] - \mathbf{E}_{\mathbf{R}}[X]| \leq \|X\|_\infty \left(\sqrt{2H(\mathbf{Q} | \mathbf{P})} + \sqrt{2H(\mathbf{R} | \mathbf{P})} \right). \quad (4)$$

Proof. First note that for any probability measures \mathbf{Q}, \mathbf{R} (absolutely continuous w.r.t. \mathbf{P}),

$$\begin{aligned} |\mathbf{E}_{\mathbf{Q}}[X] - \mathbf{E}_{\mathbf{R}}[X]| &= |\mathbf{E}_{\mathbf{P}}[X \frac{d\mathbf{Q}}{d\mathbf{P}}] - \mathbf{E}_{\mathbf{P}}[X \frac{d\mathbf{R}}{d\mathbf{P}}]| = |\mathbf{E}_{\mathbf{P}}[X(\frac{d\mathbf{Q}}{d\mathbf{P}} - \frac{d\mathbf{R}}{d\mathbf{P}})]| \\ &\leq \|X\|_{\infty} \mathbf{E}_{\mathbf{P}}\left[\left|\frac{d\mathbf{Q}}{d\mathbf{P}} - \frac{d\mathbf{R}}{d\mathbf{P}}\right|\right] = \|X\|_{\infty} \|\mathbf{Q} - \mathbf{R}\|_{\text{TV}}, \end{aligned}$$

where $\|\mathbf{Q} - \mathbf{R}\|_{\text{TV}}$ denotes the total variation distance:

$$\|\mathbf{Q} - \mathbf{R}\|_{\text{TV}} = \mathbf{E}_{\mathbf{P}}\left[\left|\frac{d\mathbf{Q}}{d\mathbf{P}} - \frac{d\mathbf{R}}{d\mathbf{P}}\right|\right].$$

Next, by the triangle inequality in total variation,

$$\|\mathbf{Q} - \mathbf{R}\|_{\text{TV}} \leq \|\mathbf{Q} - \mathbf{P}\|_{\text{TV}} + \|\mathbf{P} - \mathbf{R}\|_{\text{TV}}.$$

Finally, we invoke the Pinsker inequality Theorem A8 and obtain

$$\begin{aligned} \|\mathbf{Q} - \mathbf{P}\|_{\text{TV}}^2 &\leq 2H(\mathbf{Q} | \mathbf{P}), \\ \|\mathbf{R} - \mathbf{P}\|_{\text{TV}}^2 &\leq 2H(\mathbf{R} | \mathbf{P}), \end{aligned}$$

whenever $H(\mathbf{Q} | \mathbf{P})$ (resp. $H(\mathbf{R} | \mathbf{P})$) is finite. Hence

$$\|\mathbf{Q} - \mathbf{P}\|_{\text{TV}} \leq \sqrt{2H(\mathbf{Q} | \mathbf{P})}, \quad \|\mathbf{R} - \mathbf{P}\|_{\text{TV}} \leq \sqrt{2H(\mathbf{R} | \mathbf{P})}.$$

Putting these pieces together yields

$$\|\mathbf{Q} - \mathbf{R}\|_{\text{TV}} \leq \|\mathbf{Q} - \mathbf{P}\|_{\text{TV}} + \|\mathbf{P} - \mathbf{R}\|_{\text{TV}} \leq \sqrt{2H(\mathbf{Q} | \mathbf{P})} + \sqrt{2H(\mathbf{R} | \mathbf{P})}.$$

Therefore,

$$|\mathbf{E}_{\mathbf{Q}}[X] - \mathbf{E}_{\mathbf{R}}[X]| \leq \|X\|_{\infty} \|\mathbf{Q} - \mathbf{R}\|_{\text{TV}} \leq \|X\|_{\infty} \left(\sqrt{2H(\mathbf{Q} | \mathbf{P})} + \sqrt{2H(\mathbf{R} | \mathbf{P})} \right).$$

That completes the proof. \square

Remark 3. Inequality (4) shows that two absolutely continuous σ -martingale measures \mathbf{Q}, \mathbf{R} which both have small entropy $H(\cdot | \mathbf{P})$ must produce close expectations of any bounded payoff X . In particular, one recovers the fact that a measure with finite entropy is “close” to \mathbf{P} in the sense of total variation if the entropy is small. Hence, if two measures are entirely determined by their Radon–Nikodým derivatives in $L^1(\mathbf{P})$, a bound in terms of $\sqrt{H(\cdot | \mathbf{P})}$ is typically the best general estimate on how far their integrals can differ.

Theorem 5. Let $X \in L^{\infty}(\mathcal{F})$ be bounded and $\gamma > 0$. Then, the entropic risk

$$e_{\gamma}(X) = \sup_{\mathbf{Q} \in \mathcal{P}_L} \left(\mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \right)$$

satisfies the following two-sided bound:

$$\mathbf{E}_{\mathbf{P}}[-X] \leq e_{\gamma}(X) \leq \mathbf{E}_{\mathbf{P}}[-X] + \frac{\gamma \|X\|_{\infty}^2}{2}. \quad (5)$$

Proof. By taking the specific choice $\mathbf{Q} = \mathbf{P}$ in the definition of $e_{\gamma}(X)$, we see

$$e_{\gamma}(X) \geq \mathbf{E}_{\mathbf{P}}[-X] - \frac{1}{\gamma} H(\mathbf{P} | \mathbf{P}) = \mathbf{E}_{\mathbf{P}}[-X],$$

since $H(\mathbf{P} | \mathbf{P}) = 0$. This proves the left inequality in (5).

For any probability measure $\mathbf{Q} \ll \mathbf{P}$ with finite entropy $H(\mathbf{Q} | \mathbf{P})$, we use the triangle bound from total variation Lemma 4 to obtain

$$|\mathbf{E}_{\mathbf{Q}}[-X] - \mathbf{E}_{\mathbf{P}}[-X]| \leq \|X\|_{\infty} \sqrt{2H(\mathbf{Q} | \mathbf{P})}. \quad (6)$$

Hence,

$$\mathbf{E}_{\mathbf{Q}}[-X] \leq \mathbf{E}_{\mathbf{P}}[-X] + \|X\|_{\infty} \sqrt{2H(\mathbf{Q} | \mathbf{P})}.$$

Therefore, for any such \mathbf{Q} ,

$$\mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \leq \mathbf{E}_{\mathbf{P}}[-X] + \|X\|_{\infty} \sqrt{2H(\mathbf{Q} | \mathbf{P})} - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}).$$

Define $h := H(\mathbf{Q} | \mathbf{P}) \geq 0$. Then, we want to maximize $\|X\|_{\infty} \sqrt{2h} - (1/\gamma)h$ over $h \geq 0$. A short calculus argument shows that the maximum occurs at

$$h^* = \left(\frac{\|X\|_{\infty} \sqrt{2}}{2/\gamma} \right)^2 = \frac{\gamma^2 \|X\|_{\infty}^2}{2},$$

and that the maximum value is $\frac{\gamma \|X\|_{\infty}^2}{2}$. Consequently,

$$\sup_{h \geq 0} \left[\|X\|_{\infty} \sqrt{2h} - \frac{h}{\gamma} \right] = \frac{\gamma \|X\|_{\infty}^2}{2}.$$

Hence, for all $\mathbf{Q} \in \mathcal{P}_L$ with $H(\mathbf{Q} | \mathbf{P}) < \infty$,

$$\mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \leq \mathbf{E}_{\mathbf{P}}[-X] + \frac{\gamma \|X\|_{\infty}^2}{2}.$$

Taking the supremum over all $\mathbf{Q} \in \mathcal{P}_L$ leads to

$$e_{\gamma}(X) = \sup_{\mathbf{Q}} \left(\mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \right) \leq \mathbf{E}_{\mathbf{P}}[-X] + \frac{\gamma \|X\|_{\infty}^2}{2}.$$

This completes the proof of (5). \square

Corollary 1. Let \mathbf{Q}^E be the minimal entropy σ -martingale measure for a (bounded) price process \hat{S} . Then,

$$e_{\gamma}(X) \geq \mathbf{E}_{\mathbf{Q}^E}[-X] - \frac{1}{\gamma} H(\mathbf{Q}^E | \mathbf{P}) \quad \text{for all bounded } X. \quad (7)$$

Furthermore, combining this estimate with the two-sided bound of Theorem 5, we have:

$$\max \left\{ \mathbf{E}_{\mathbf{P}}[-X], \rho^E(X) - \frac{1}{\gamma} H(\mathbf{Q}^E | \mathbf{P}) \right\} \leq e_{\gamma}(X) \leq \mathbf{E}_{\mathbf{P}}[-X] + \frac{\gamma \|X\|_{\infty}^2}{2}. \quad (8)$$

Proof. Since $e_{\gamma}(X)$ is defined as the supremum $\sup_{\mathbf{Q} \in \mathcal{P}_L} \{ \mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \}$, it clearly holds that for any \mathbf{Q} in that admissible set,

$$e_{\gamma}(X) \geq \mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}).$$

If \mathbf{Q}^E (the minimal entropy σ -martingale measure) is in \mathcal{P}_L (i.e. $\mathbf{Q}^E \ll \mathbf{P}$ and $H(\mathbf{Q}^E | \mathbf{P}) < \infty$), then simply take $\mathbf{Q} = \mathbf{Q}^E$ in the supremum. This yields the lower bound (7).

Finally, to deduce (8), observe that $e_{\gamma}(X) \geq \mathbf{E}_{\mathbf{P}}[-X]$ also holds by choosing $\mathbf{Q} = \mathbf{P}$. Then Theorem 5 provides the upper bound, and so taking the maximum of these lower estimates proves (8). \square

Remark 4. The assumption that $\mathbf{Q}^E \in \mathcal{P}_L$ requires that the minimal entropy measure \mathbf{Q}^E be absolutely continuous w.r.t. \mathbf{P} with finite relative entropy. If, for instance, the market admits no arbitrage of the first kind (NA1) and there exists a finite-entropy $E\sigma\text{MM}$, one can show $\mathbf{Q}^E \in \mathcal{P}_L$. In that case, (7) provides a non-trivial lower estimate on the entropic risk in terms of the minimal entropy measure.

Remark 5. When the risk-aversion parameter γ tends to zero, the penalty factor $\frac{1}{\gamma}$ in the entropic risk measure

$$e_\gamma(X) = \sup_{\mathbf{Q} \ll \mathbf{P}} \left\{ \mathbf{E}_{\mathbf{Q}}[-X] - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \right\}$$

becomes very large. As a result, any measure $\mathbf{Q} \neq \mathbf{P}$ with $H(\mathbf{Q} | \mathbf{P}) > 0$ gets heavily penalized in the objective, so that the supremum is forced increasingly toward $\mathbf{Q} = \mathbf{P}$.

Formally, it can be shown that

$$\lim_{\gamma \rightarrow 0} e_\gamma(X) = \mathbf{E}_{\mathbf{P}}[-X].$$

Economically, this reflects a situation of ultra-strong aversion to deviating from the reference measure \mathbf{P} . In the limit $\gamma \rightarrow 0$, the only measure not incurring an infinite penalty is $\mathbf{Q} = \mathbf{P}$. Thus, the entropic risk measure collapses to $\mathbf{E}_{\mathbf{P}}[-X]$, the plain (negative) expectation under the real-world probability \mathbf{P} .

4. Duality

Any convex or coherent risk measure admits a so-called robust representation. This has been discussed in several works, including [2], [23], and [25].

The usual approach to defining the entropic risk measure is:

$$\frac{1}{\gamma} \log \mathbf{E} \left[e^{-\gamma X} \right] \quad (9)$$

and its robust representation, (1), is typically derived using dual representation theorems.

In this paper, we defined the entropic risk using its robust representation and now demonstrate that this definition is equivalent to the traditional one found in literature. Since the convexity of (9) has not been explicitly shown here, we present an alternative proof.

Theorem 6. For $\gamma > 0$ and $X \in L^\infty(\mathcal{F})$, we have:

$$\sup_{\mathbf{Q} \in \mathcal{P}_L} \left(\mathbf{E}_{\mathbf{Q}}(-X) - \frac{1}{\gamma} H(\mathbf{Q} | \mathbf{P}) \right) = \frac{1}{\gamma} \log \mathbf{E}_{\mathbf{P}} \left[e^{-\gamma X} \right].$$

Proof. Define a new probability measure \mathbf{P}_X by:

$$\frac{d\mathbf{P}_X}{d\mathbf{P}} = \frac{e^{-\gamma X}}{\mathbf{E}_{\mathbf{P}} e^{-\gamma X}}.$$

Then:

$$\begin{aligned} H(\mathbf{Q} | \mathbf{P}) &= \mathbf{E}_{\mathbf{Q}} \left[\log \frac{d\mathbf{Q}}{d\mathbf{P}} \right] \\ &= \mathbf{E}_{\mathbf{Q}} \log \left[\frac{d\mathbf{Q}}{d\mathbf{P}_X} \frac{d\mathbf{P}_X}{d\mathbf{P}} \right] \\ &= \mathbf{E}_{\mathbf{Q}} \left[\log \left(\frac{d\mathbf{Q}}{d\mathbf{P}_X} \right) + \log \left(e^{-\gamma X} \right) - \log \mathbf{E}_{\mathbf{P}} e^{-\gamma X} \right] \\ &= \mathbf{E}_{\mathbf{Q}} \left[\log \frac{d\mathbf{Q}}{d\mathbf{P}_X} \right] + \mathbf{E}_{\mathbf{Q}}[-\gamma X] - \log \mathbf{E}_{\mathbf{P}}[e^{-\gamma X}]. \end{aligned}$$

Substituting this into the optimization problem:

$$\begin{aligned}
 & \exp \sup_{\mathbf{Q} \in \mathcal{P}_L} (\mathbf{E}_{\mathbf{Q}}[-\gamma X] - H(\mathbf{Q} | \mathbf{P})) \\
 &= \exp \left(\sup_{\mathbf{Q} \in \mathcal{P}_L} \left(\mathbf{E}_{\mathbf{Q}}[-\gamma X] - \mathbf{E}_{\mathbf{Q}} \left[\log \frac{d\mathbf{Q}}{d\mathbf{P}_X} \right] - \mathbf{E}_{\mathbf{Q}}[-\gamma X] + \log(\mathbf{E}_{\mathbf{P}}[e^{-\gamma X}]) \right) \right) \\
 &= \exp \left(\sup_{\mathbf{Q} \in \mathcal{P}_L} \left(-\mathbf{E}_{\mathbf{Q}} \left[\log \frac{d\mathbf{Q}}{d\mathbf{P}_X} \right] + \log(\mathbf{E}_{\mathbf{P}}[e^{-\gamma X}]) \right) \right) \\
 &= \mathbf{E}_{\mathbf{P}}[e^{-\gamma X}] \exp \left(- \inf_{\mathbf{Q} \in \mathcal{P}_L} \mathbf{E}_{\mathbf{Q}} \left[\log \frac{d\mathbf{Q}}{d\mathbf{P}_X} \right] \right) \\
 &= \mathbf{E}_{\mathbf{P}}(e^{-\gamma X}),
 \end{aligned}$$

where the last equality follows from $\mathbf{P}_X \in \mathcal{P}_L$ and Lemma 1.

Taking the logarithm and dividing by γ yields the result. \square

A detailed analysis of the dual problem for the minimal entropy martingale measure is more complex but provides valuable insights into its economic interpretation as an equivalent local martingale measure. The dual problem, in essence, connects minimizing relative entropy to maximizing a specific utility function. Below, we elaborate briefly on this connection.

Let U be a utility function of the form:

$$U(x) = -\exp(-x).$$

We aim to maximize the expected utility of the terminal value of a wealth process by selecting an optimal strategy ϕ from a set of admissible strategies Φ . If the terminal value is represented as $\phi \cdot \hat{S}_T := \phi_0 \hat{S}_0 + \int_0^T \phi_t d\hat{S}_t$, then the objective becomes:

$$\sup_{\phi \in \Phi} \mathbf{E}[U(\phi \cdot \hat{S}_T)] = \sup_{\phi \in \Phi} \mathbf{E}[-\exp(-\phi \cdot \hat{S}_T)].$$

Under weak conditions, [13] show:

$$\sup_{\phi \in \Phi} \mathbf{E}[-\exp(-\phi \cdot \hat{S}_T)] = -\exp \left(\inf_{\mathbf{Q} \in \mathcal{M}_{\text{loc}}} H(\mathbf{Q} | \mathbf{P}) \right) \quad (10)$$

as well as:

$$\frac{d\mathbf{Q}^E}{d\mathbf{P}} = \exp \left(-H(\mathbf{Q}^E | \mathbf{P}) + \phi^* \cdot \hat{S}_T \right),$$

where ϕ^* is the strategy that maximizes the utility in Equation (10).

Further details on this duality can be found in [5], [13], and [32].

5. Dynamic Consistency

In this section, we present dynamic versions of the entropic risk measure and the risk measure associated with the minimal entropy martingale measure, proving that they are time-consistent. Dynamic consistency is particularly important when dealing with practical financial applications. For instance, dynamic risk measures based on entropy have been successfully applied to energy markets and stochastic volatility models, as discussed in [49].

We start by introducing the relevant definitions for this section.

Let $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ be a filtration, L_t^∞ the set of bounded \mathbb{F}_t -measurable functions, \mathcal{Q}_t the set of probability measures on \mathcal{F}_t which are equivalent to \mathbf{P} on \mathcal{F}_t , and let S be an adapted stochastic process.

Definition 6.

- (a) A map $\rho_t : L^\infty \rightarrow L_t^\infty$ is called a dynamic risk measure.
 (b) A dynamic risk measure is called time consistent if for $s \leq t$,

$$\mathbf{P}[\rho_t(X) \geq \rho_t(Y)] = 1 \implies \mathbf{P}[\rho_s(X) \geq \rho_s(Y)] = 1.$$

- (c) The dynamic entropic risk measure $e_{\gamma,t}$ is defined as

$$e_{\gamma,t}(X) := \frac{1}{\gamma} \log(\mathbf{E}[\exp(-\gamma X) | \mathcal{F}_t]).$$

- (d) The dynamic minimal entropy risk measure ρ_t^E is defined as

$$\rho_t^E(X) := \mathbf{E}_{\mathbf{Q}^E}[-X | \mathcal{F}_t],$$

where \mathbf{Q}^E is the minimal entropy martingale measure.

It is possible to allow γ to be an adapted process instead of a constant, resulting in a slightly more general definition of the dynamic entropic risk measure. However, in such cases, time consistency may not be guaranteed.

If for $\mathbf{Q} \in \mathcal{Q}_t$ we define the conditional relative entropy $H_t(\mathbf{Q} | \mathbf{P})$ as

$$H_t(\mathbf{Q} | \mathbf{P}) = \mathbf{E}_{\mathbf{Q}} \left[\log \frac{d\mathbf{Q}}{d\mathbf{P}} | \mathcal{F}_t \right],$$

the dynamic entropic risk measure can also be expressed in a robust representation involving the relative entropy:

$$e_{\gamma,t}(X) = \sup_{\mathbf{Q} \in \mathcal{Q}_t} \left(\mathbf{E}_{\mathbf{Q}}[-X | \mathcal{F}_t] - \frac{1}{\gamma} H_t(\mathbf{Q} | \mathbf{P}) \right).$$

Further details on this duality and on dynamic risk measures in general (in both continuous and discrete time) can be found in [18], [22], [1], and [41].

It is straightforward to verify that the two risk measures defined above satisfy the properties of dynamic risk measures.

Theorem 7. *The dynamic entropic risk measure and the dynamic minimal entropy risk measure are time-consistent.*

Proof. We start with the entropic risk measure. We aim to show that $e_{\gamma,s}(X) = e_{\gamma,s}(-e_{\gamma,t}(X))$ for $0 \leq s \leq t$. Time consistency then follows from the intertemporal monotonicity theorem [22, Proposition 4.2].

Using the tower property of conditional expectation, we have:

$$\begin{aligned} e_{\gamma,s}(X) &= \frac{1}{\gamma} \log \left(\mathbf{E} \left[e^{-\gamma X} | \mathcal{F}_s \right] \right) \\ &= \frac{1}{\gamma} \log \left(\mathbf{E} \left[\mathbf{E} \left[e^{-\gamma X} | \mathcal{F}_t \right] | \mathcal{F}_s \right] \right) \\ &= \frac{1}{\gamma} \log \left(\mathbf{E} \left[\exp \left(-\gamma \underbrace{\left(-\frac{1}{\gamma} \log \left(\mathbf{E} \left[e^{-\gamma X} | \mathcal{F}_t \right] \right)}_{=-e_{\gamma,t}(X)} \right) | \mathcal{F}_s \right] \right) \right) \\ &= e_{\gamma,s}(-e_{\gamma,t}(X)). \end{aligned}$$

For the minimal entropy martingale measure, we proceed similarly:

$$\begin{aligned}\rho_s^E(X) &= \mathbf{E}_{\mathbf{Q}^E}[-X \mid \mathcal{F}_s] \\ &= \mathbf{E}_{\mathbf{Q}^E}[\mathbf{E}_{\mathbf{Q}^E}[-X \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\ &= \rho_s^E(-\rho_t^E(X)).\end{aligned}$$

This proves time consistency for both measures. \square

6. Optimal Risk Transfer

Optimal risk transfer is a classical topic in the study of risk measures and has been studied extensively in, e.g., [3]. In this section, we present a somewhat more general and extended view, highlighting in particular the property of γ -dilated families of risk measures, the associated inf-convolution approach, and how these results can be restricted to the risk-neutral setting. We then provide a brief discussion of why no non-trivial γ -dilation family can be constructed from the minimal entropy coherent risk measure, whereas the standard entropic risk measure naturally fits into this framework. Finally, we show how restricting to (local) risk-neutral measures force the minimal entropy measure to emerge as the unique solution to the risk-transfer problem.

Let ρ be any (convex or coherent) risk measure on $L^\infty(\mathcal{F})$. The fundamental question in optimal risk sharing between two entities with risk measures ρ_1 and ρ_2 is to split a position X into two parts $X - F$ and F . The goal is to minimize the combined risk $\rho_1(X - F) + \rho_2(F)$. This optimization problem is well expressed via the inf-convolution operator:

$$(\rho_1 \square \rho_2)(X) := \inf_F \{\rho_1(X - F) + \rho_2(F)\}. \quad (11)$$

The optimal risk transfer is solved when the infimum in (11) is achieved by some F^* . Well-known results, see [3], show that certain risk measures, particularly those arising from γ -dilated families, enjoy a simple solution: the optimal allocation is often a linear split of X .

Definition 7 (γ -Dilated Family). *A family $\{\rho_\gamma\}_{\gamma \in I}$ of risk measures on $L^\infty(\mathcal{F})$ is called γ -dilated if there exists a base risk measure ρ such that*

$$\rho_\gamma(X) = \frac{1}{\gamma} \rho(\gamma X) \quad \text{for each } \gamma \in I,$$

where $I \subseteq (0, \infty)$ is some index set.

An immediate (but important) example is the entropic risk measure, which is γ -dilated by construction:

$$e_\gamma(X) = \frac{1}{\gamma} \log(\mathbf{E}[e^{-\gamma X}]) = \frac{1}{\gamma} \rho(\gamma X)$$

if one sets $\rho(Z) := \log(\mathbf{E}[e^{-Z}])$. By contrast, a coherent risk measure with positive homogeneity cannot usually form a non-trivial γ -dilated family unless it is purely linear in γ and thus collapses to the \mathbf{E} -type functional (more details below).

We restate here a known result, cf. [3], slightly paraphrased to fit our notation.

Theorem 8. *Let $\{\rho_\gamma\}_{\gamma > 0}$ be a γ -dilated family of convex risk measures on $L^\infty(\mathcal{F})$. Then the following statements hold:*

(a) *For any $\gamma_1, \gamma_2 > 0$, we have the inf-convolution identity:*

$$\rho_{\gamma_1 + \gamma_2}(X) = (\rho_{\gamma_1} \square \rho_{\gamma_2})(X). \quad (12)$$

(b) The optimal allocation F^* solving $\inf_F \{\rho_{\gamma_1}(X - F) + \rho_{\gamma_2}(F)\}$ is linear in X , specifically

$$F^* = \frac{\gamma_2}{\gamma_1 + \gamma_2} X. \quad (13)$$

(c) Consequently, we also have

$$\rho_{\gamma_1 + \gamma_2}(X) = \rho_{\gamma_1}(X - F^*) + \rho_{\gamma_2}(F^*).$$

(d) Moreover, if ρ^1 and ρ^2 are two convex risk measures that each can be embedded in a single γ -dilated family (i.e., they are both instances of ρ_γ for some base ρ), then for any $\gamma_1, \gamma_2 > 0$,

$$\rho_{\gamma_1} \square \rho_{\gamma_2} = (\rho \square \rho)_{\gamma_1 + \gamma_2}$$

under a suitable identification. Hence, the γ -dilation structure is preserved under inf-convolution.

Sketch of Proof. The key insight is that for a γ -dilated family, scaling X by γ and adjusting the risk measure by $\frac{1}{\gamma}$ yields a consistent re-parametrization of the same base functional ρ . Once this is established, standard inf-convolution arguments for the exponentially (or more generally, ρ -) tilted function imply the linear sharing rule (13). For full details, see [3] or references therein. \square

We now remark why a coherent risk measure – in particular, the minimal entropy risk measure ρ^E – cannot generally be part of a non-trivial γ -dilated family. Recall that a coherent risk measure ρ is positively homogeneous, i.e.

$$\rho(\lambda X) = \lambda \rho(X) \quad \text{for all } \lambda \geq 0.$$

Combining this with the γ -dilation requirement $\rho_\gamma(X) = \frac{1}{\gamma} \rho(\gamma X)$ quickly forces $\rho_\gamma(X) \equiv \rho(X)$ for all γ , unless $\rho \equiv 0$ or $\rho \equiv$ some fixed linear functional. Indeed, for coherent ρ , we would have

$$\rho(\gamma X) = \gamma \rho(X), \quad \text{so} \quad \frac{1}{\gamma} \rho(\gamma X) = \rho(X).$$

Hence, ρ_γ cannot genuinely “depend” on γ . In short, non-trivial γ -dilated families (like the entropic class) are not coherent. This explains why the minimal entropy risk measure, $\rho^E(X) = \mathbf{E}_{\mathbf{Q}^E}[-X]$, cannot appear in a standard γ -dilated framework.

Appendix A. Some well-known theorems

The following theorem is from [37, page 146].

Theorem A1. Let f be a function which is continuous on $[a, b]$ and two times differentiable (a, b) . Then we have

- (a) f is convex in $[a, b]$ if and only if $f'' \geq 0$,
- (b) f is strictly convex if $f'' > 0$.

A thorough presentation of this and closely related inequalities can be found, for instance, in [44, Section 4] or in standard real analysis texts.

Theorem A2 (Classical Young–Gibbs Inequality). Let $x > 0$ and $y \in \mathbb{R}$. Then

$$x y - \ln(x) \leq y, \quad \text{with equality if and only if } x = 1.$$

The following theorem is a special case of *de La Vallée Poussin-criterion* and is proved for example in ([17], T22).

Theorem A3. Let $(X_\lambda)_{\lambda \in \Lambda}$ be a family of integrable random variables and $\phi : \mathbb{R} \rightarrow \mathbb{R}_+$ a measurable function with $\lim_{t \rightarrow \infty} \frac{\phi(t)}{t} = \infty$. If we have

$$K := \sup_{\lambda \in \Lambda} \mathbb{E}[\phi(|X_\lambda|)] < \infty,$$

then $(X_\lambda)_{\lambda \in \Lambda}$ is uniformly integrable.

A proof of the following theorem can be found in [19, page 294].

Theorem A4 (Dunford-Pettis criterion). Let $\mathcal{K} \subseteq L^1$. Then the following are equivalent:

- (i) \mathcal{K} is uniformly integrable.
- (ii) \mathcal{K} is relatively compact in L^1 , if we consider the weak convergence in L^1 as corresponding convergence.
- (iii) For each sequence in \mathcal{K} , there exists a subsequence which converges weakly in L^1 .

A proof of the next statement can be found in [20, page 6] or [50, Theorem III.2.5].

Theorem A5 (Mazur's lemma). Let $(X, \|\cdot\|)$ be a Banach space and let $(u_n)_{n \in \mathbb{N}}$ be a sequence in X that converges weakly to some u_0 in X :

$$u_n \rightarrow u_0 \text{ as } n \rightarrow \infty.$$

That is, for every continuous linear functional $f \in X^*$, the continuous dual space of X ,

$$f(u_n) \rightarrow f(u_0) \text{ as } n \rightarrow \infty.$$

Then, there exists a function $N : \mathbb{N} \rightarrow \mathbb{N}$ and a sequence of sets of real numbers

$$\{\alpha(n)_k | k = n, \dots, N(n)\}$$

such that $\alpha(n)_k \geq 0$ and

$$\sum_{k=n}^{N(n)} \alpha(n)_k = 1$$

such that the sequence $(v_n)_{n \in \mathbb{N}}$ defined by the convex combination

$$v_n = \sum_{k=n}^{N(n)} \alpha(n)_k u_k$$

converges strongly in X to u_0 , i.e.

$$\|v_n - u_0\| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

We denote with \mathcal{L} the set of probability measures defined on (Ω, \mathcal{F}) .

For $\Gamma \subseteq \mathcal{L}$ we set:

$$I(\Gamma, \mathbf{P}) = \inf\{I(\mathbf{Q}, \mathbf{P}) : \mathbf{Q} \in \Gamma\}.$$

A probability $\mathbf{Q} \in \Gamma$ satisfying

$$I(\mathbf{Q}, \mathbf{P}) = I(\Gamma, \mathbf{P})$$

is called the I -projection of \mathbf{P} on Γ .

Theorem A6 (Th. 2.1 [9]). Let $\Gamma \subseteq \mathcal{L}$ be a convex set such that $I(\Gamma, \mathbf{P}) < \infty$. If Γ is closed in variation then \mathbf{P} has a (unique) I -projection on Γ .

Theorem A7 (Th. 2.2 [9]). Let $\Gamma \subseteq \mathcal{L}$ be a convex set. A $\mathbf{Q} \in \Gamma$ such that $I(\mathbf{Q}, \mathbf{P}) < \infty$ is the I -projection of \mathbf{P} on Γ if and only if:

$$I(\mathbf{R}, \mathbf{P}) \geq I(\mathbf{R}, \mathbf{Q}) + I(\mathbf{Q}, \mathbf{P}) \quad \forall \mathbf{R} \in \Gamma.$$

The following theorem is known as the Pinsker inequality (also sometimes called Kullback–Csiszár–Kemperman inequality), linking total variation norm and relative entropy (Kullback–Leibler divergence) between two probability measures.

Theorem A8 (Pinsker inequality). *Let \mathbf{P} and \mathbf{Q} be two probability measures on the same measurable space. Then*

$$\|\mathbf{Q} - \mathbf{P}\|_{\text{TV}}^2 \leq 2 H(\mathbf{Q} | \mathbf{P}), \quad (\text{A1})$$

where $H(\mathbf{Q} | \mathbf{P})$ is the relative entropy or Kullback–Leibler divergence defined by

$$H(\mathbf{Q} | \mathbf{P}) = \int \log\left(\frac{d\mathbf{Q}}{d\mathbf{P}}\right) d\mathbf{Q},$$

provided $\mathbf{Q} \ll \mathbf{P}$ and otherwise $H(\mathbf{Q} | \mathbf{P}) = +\infty$.

This inequality can be found explicitly, for example, in [4, Section 2]. An earlier result with weaker constants appears in [42]. The final and more general form was established independently by [35], [10], and [34].

References

1. Acciaio, Beatrice and Irina Penner. 2011. Dynamic risk measures. In *Advanced Mathematical Methods for Finance*, pp. 1–34. Springer. https://doi.org/10.1007/978-3-642-18412-3_1.
2. Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical finance* 9(3), 203–228. <https://doi.org/10.1111/1467-9965.00068>.
3. Barrieu, Pauline and Nicole El Karoui. 2005. Inf-convolution of risk measures and optimal risk transfer. *Finance and stochastics* 9(2), 269–298. <https://doi.org/10.1007/s00780-005-0152-0>.
4. Barron, Andrew R. 1986. Entropy and the central limit theorem. *The Annals of Probability* 14(1), 336–342.
5. Bellini, Fabio and Marco Frittelli. 2002. On the existence of minimax martingale measures. *Mathematical Finance* 12(1), 1–21. <https://doi.org/10.1111/1467-9965.00001>.
6. Cherny, Alexander S and Victor P Maslov. 2004. On minimization and maximization of entropy in various disciplines. *Theory of Probability & Its Applications* 48(3), 447–464. <https://doi.org/10.1137/s0040585x9798052x>.
7. Chou, Ching-Sung. 1979. Caractérisation d’une classe de semimartingales. In *Séminaire de Probabilités XIII*, pp. 250–252. Springer.
8. Cover, Thomas M and Joy A Thomas. 2012. *Elements of information theory*. John Wiley & Sons.
9. Csiszár, Imre. 1975. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability* 3(1), 146–158. <https://doi.org/10.1214/aop/1176996454>.
10. Csiszár, Imre. 1967. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica* 2, 299–318.
11. Dacunha-Castelle, Didier and Marie Duflo. 1986. *Probability and Statistics* (1 ed.). Springer.
12. Delbaen, Freddy. 2000. Coherent risk measures. *Blätter der DGVFM* 24(4), 733–739. <https://doi.org/10.1007/bf02809088>.
13. Delbaen, Freddy, Peter Grandits, Thorsten Rheinländer, Davide Samperi, Martin Schweizer, and Christophe Stricker. 2002. Exponential hedging and entropic penalties. *Mathematical Finance* 12(2), 99–123. <https://doi.org/10.1111/1467-9965.02001>.
14. Delbaen, Freddy and Walter Schachermayer. 1994. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* 300(1), 463–520. <https://doi.org/10.1007/BF01450498>.
15. Delbaen, Freddy and Walter Schachermayer. 1998. The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen* 312(2), 215–250. <https://doi.org/10.1007/s002080050220>.
16. Delbaen, Freddy and Walter Schachermayer. 2006. *The Mathematics of Arbitrage*. Springer Finance. Springer.
17. Dellacherie, Claude and Paul-André Meyer. 1982. *Probabilities and Potential, B: Theory of Martingales*. North-Holland Mathematics Studies. Elsevier Science.
18. Detlefsen, Kai and Giacomo Scandolo. 2005. Conditional and dynamic convex risk measures. *Finance and stochastics* 9(4), 539–561. <https://doi.org/10.1007/s00780-005-0159-6>.

19. Dunford, Nelson and Jacob Theodore Schwartz. 1958. *Linear Operators. Part 1. General Theory*. Pure and Applied Mathematics. Interscience Publishers.
20. Ekeland, Ivar and Roger Temam. 1976. *Convex Analysis and Variational Problems*. Studies in Mathematics and its Applications. Elsevier Science.
21. Föllmer, Hans and Thomas Knispel. 2011. Entropic risk measures: Coherence vs. convexity, model ambiguity and robust large deviations. *Stochastics and Dynamics* 11(02n03), 333–351. <https://doi.org/10.1142/s0219493711003334>.
22. Föllmer, Hans and Irina Penner. 2006. Convex risk measures and the dynamics of their penalty functions. *Statistics & Decisions* 24(1/2006), 61–96. <https://doi.org/10.1524/std.2006.24.1.61>.
23. Föllmer, Hans and Alexander Schied. 2011. *Stochastic Finance: An Introduction in Discrete Time* (2nd ed.), Volume 27 of *De Gruyter Studies in Mathematics*. Berlin: Walter de Gruyter. Updated and expanded edition, introducing discrete-time methods for stochastic finance., <https://doi.org/10.1515/9783110218053>.
24. Frittelli, Marco. 2000. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance* 10(1), 39–52. <https://doi.org/10.1111/1467-9965.00079>.
25. Frittelli, Marco and Emanuela Rosazza Gianin. 2002. Putting order in risk measures. *Journal of Banking & Finance* 26(7), 1473–1486. [https://doi.org/10.1016/s0378-4266\(02\)00270-4](https://doi.org/10.1016/s0378-4266(02)00270-4).
26. Fujiwara, Tsukasa. 2004. From the minimal entropy martingale measures to the optimal strategies for the exponential utility maximization: the case of geometric lévy processes. *Asia-Pacific Financial Markets* 11(4), 367–391.
27. Fujiwara, Tsukasa and Yoshio Miyahara. 2003. The minimal entropy martingale measures for geometric lévy processes. *Finance and Stochastics* 7(4), 509–531.
28. Goll, Thomas, Jan Kallsen, et al. 2003. A complete explicit solution to the log-optimal portfolio problem. *The Annals of Applied Probability* 13(2), 774–799. <https://doi.org/10.1214/aoap/1050689603>.
29. Grandits, Peter and Thorsten Rheinländer. 2002. On the minimal entropy martingale measure. *The Annals of Probability* 30(3), 1003–1038. <https://doi.org/10.1214/aop/1029867119>.
30. Hesse, Christian. 2003. *Angewandte Wahrscheinlichkeitstheorie*. Vieweg.
31. Jacod, Jean and Albert N. Shiryaev. 2003. *Limit theorems for stochastic processes* (Second ed.), Volume 288 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. <https://doi.org/10.1007/978-3-662-05265-5>.
32. Kabanov, Yuri M. and Christophe Stricker. 2002. On the optimal portfolio for the exponential utility maximization: remarks to the six-author paper. *Mathematical Finance* 12(2), 125–134.
33. Kallsen, Jan. 2004. σ -localization and σ -martingales. *Theory of Probability & Its Applications* 48(1), 152–163.
34. Kemperman, Johannes H. B. 1969. On the optimum rate of transmitting information. *Annals of Mathematical Statistics* 40, 2156–2177.
35. Kullback, Solomon. 1967. A lower bound for discrimination information in terms of variation. *IEEE Transactions on Information Theory* 13(1), 126–127.
36. Kupper, Michael and Walter Schachermayer. 2009. Representation results for law invariant time consistent functions. *Mathematics and Financial Economics* 2(3), 189–210. <https://doi.org/10.1007/s11579-009-0019-9>.
37. Königsberger, Konrad. 2003. *Analysis 1*. Springer-Lehrbuch. Springer Berlin Heidelberg.
38. Lee, Young and Thorsten Rheinländer. 2013. The minimal entropy martingale measure for exponential markov chains. *Journal of Applied Probability* 50(2), 344–358. <https://doi.org/10.1239/jap/1371648945>.
39. Miyahara, Yoshio. 1999. Minimal entropy martingale measures of jump type price processes in incomplete assets markets. *Asia-Pacific Financial Markets* 6(2), 97–113. <https://doi.org/10.1023/A:1010062625672>.
40. Miyahara, Yoshio. 2004. A note on esscher transformed martingale measures for geometric lévy processes. *Discussion Papers in Economics, Nagoya City University* 379, 1–14.
41. Penner, Irina. 2007. *Dynamic convex risk measures: time consistency, prudence, and sustainability*. Ph. D. thesis, Humboldt-Universität zu Berlin.
42. Pinsker, Mark S. 1964. *Information and Information Stability of Random Variables and Processes*. Holden-Day.
43. Rheinländer, Thorsten and Gallus Steiger. 2006. The minimal entropy martingale measure for general barndorff-nielsen/shephard models. *The Annals of Applied Probability* 16(3), 1319–1351. <https://doi.org/10.1214/105051606000000240>.
44. Rockafellar, R. Tyrrell. 1970. *Convex Analysis*. Princeton, NJ: Princeton University Press.
45. Schweizer, Martin. 2010. Minimal entropy martingale measure. In *Encyclopedia of Quantitative Finance*. Wiley. <https://doi.org/10.1002/9780470061602.eqf04014>.
46. Shunsuke, Ihara. 1993. *Information theory for continuous systems*, Volume 2. World Scientific.

47. Sohns, Moritz. 2023. The general semimartingale model with dividends. In O. Roubai (Ed.), *Proceedings of the 10th Ročník Konference Doktorandů Na Vysoké škole Finanční A Správní*. <https://doi.org/10.37355/KD-2023-11>.
48. Sohns, Moritz. 2025. σ -martingales: Foundations, properties, and a new proof of the Ansel–Stricker lemma. *Mathematics* 13(4), 682. <https://doi.org/10.3390/math13040682>.
49. Swishchuk, Anatoliy. 2007. Change of time method in mathematical finance. *Canad. Appl. Math. Quart* 15(3), 299–336.
50. Werner, Dirk. 2008. *Funktionalanalysis (Springer-Lehrbuch) (German Edition)* (6., korrig. ed.). Springer.
51. Émery, Michel. 1980. Compensation de processus vf non localement intégrables. In *Séminaire de Probabilités XIV 1978/79*, pp. 152–160. Springer.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.