

Risk-Sensitive Investment Management via Free Energy–Entropy Duality

Sébastien Lleo*

Wolfgang Runggaldier†

April 28, 2026

Abstract

We study a benchmarked risk-sensitive portfolio problem in a factor-based setting to bring together three strands of the literature: benchmarked risk-sensitive investment management, the Kuroda–Nagai change-of-measure method, and the free energy–entropy duality of Dai Pra et al. (1996). We show that the duality yields a direct solution of the benchmarked problem by reformulating it as a linear-quadratic-Gaussian stochastic differential game under a suitable equivalent probability measure, with an entropic regularization. The resulting value function is quadratic, the optimal controls are explicit affine feedback maps, and the optimal allocation admits two complementary interpretations: as a fractional Kelly strategy and as a Kelly portfolio adjusted via the entropic regularization. This formulation, therefore, contributes both a direct analytical route to the solution and a clearer interpretation of risk sensitivity, thereby embedding the classical Kuroda–Nagai change-of-measure approach within a more general framework. An added benefit of this formulation is that it is suitable for implementation via an RL algorithm. A simple implementation on U.S. equity data illustrates the tractability of the framework and numerically confirms the equivalence of the two approaches.

Keywords: risk-sensitive control, fractional Kelly strategies, free energy–entropy duality, stochastic games, portfolio optimization.

JEL Classification: C32; C44; C61; C73; G11; G12.

MSC classes: 91G10; 91G80; 93E20.

1 Introduction

This paper brings together three strands of literature in a single tractable framework: benchmarked risk-sensitive investment management, the Kuroda–Nagai change-of-measure method, and the free energy–relative entropy duality of Dai Pra et al. (1996).

Over the past quarter-century, risk-sensitive control has become a well-established approach to dynamic portfolio selection. One reason for its appeal in financial applications is that it admits several useful interpretations: it connects naturally to utility theory, provides a dynamic counterpart to mean–variance optimization, converges to the Kelly criterion in the limit as the

*inance Department and ‘AI, Data Science & Business’ AE, NEOMA Business School, France

†University of Padova, Italy, and Fellow, Institut Louis Bachelier, Paris, France

risk sensitivity parameter $\theta \rightarrow 0$, and, for $\theta > 0$, the associated objective relates closely to the entropic risk measure (Bielecki & Pliska, 2003; Lleo & MacLean, 2025). At the same time, risk-sensitive investment problems are not standard control problems. In the benchmarked setting, the state process X_s is uncontrolled, while the terminal reward depends on a controlled, stochastic rate-of-return process that contains an Itô integral. Because the criterion is exponential, this Itô integral does not disappear under expectation. For some time, this feature made such problems difficult to solve directly (see Bielecki & Pliska, 1999). Kuroda and Nagai later showed that a suitable change of measure recasts the risk-sensitive investment criterion into a standard linear exponential-of-quadratic Gaussian (LEQG) control problem (Kuroda & Nagai, 2002), which can then be solved explicitly; see also Bensoussan (1992). Together, the factor-based framework of Bielecki and Pliska and the Kuroda–Nagai change-of-measure method shaped a large subsequent literature, including models of factor-based asset management (Bielecki & Pliska, 1999, 2000; Hata, 2017), insider trading (Hata, 2018), benchmarked asset management (Davis & Lleo, 2008), asset–liability management (Davis & Lleo, 2014, Chapter 4), stochastic interest rates (Bielecki & Pliska, 2004; Hata, 2021), jump-diffusion dynamics (Davis & Lleo, 2011, 2013b,c, 2015), partial observation Nagai & Peng (2002); Hata & Iida (2006); Lleo & Runggaldier (2024), and, more recently, expert opinions (Davis & Lleo, 2013a, 2021, 2024). In parallel, risk-sensitive control also saw methodological developments, particularly in the context of model uncertainty Bielecki et al. (2022).

The benchmarked formulation we adopt here is economically natural in delegated portfolio management, where investors are often evaluated relative to a reference portfolio rather than in absolute terms. It also substantially broadens the scope of the framework: by varying the benchmark, one can represent a continuum of investment styles, ranging from strongly active mandates to tightly benchmarked or index-tracking mandates, while retaining the same underlying factor structure. At the same time, this generalization comes at essentially no additional mathematical cost, since it amounts to replacing log wealth by the log excess return relative to the benchmark. In particular, the traditional asset-only formulation is recovered as a special case by setting the benchmark to the investor’s initial wealth.

This paper makes two main contributions. First, it provides a direct solution to the benchmarked risk-sensitive portfolio problem by leveraging the free energy–entropy duality and an equivalent change of probability measure (see Dai Pra et al., 1996). Second, we rigorously demonstrate that this novel solution technique is fully equivalent to the classical Kuroda–Nagai approach. Overall, the advantages of the free energy–entropy duality formulation are twofold: on the one hand, it grounds the solution method in solid theoretical principles; on the other hand, it enhances the transparency of the resulting allocation structure. Notably, the optimal portfolio can now be interpreted as a fractional Kelly strategy, or equivalently, as a Kelly portfolio corrected by a term induced by the duality. Thus, our approach not only introduces a new analytical pathway to the solution but also sharpens the understanding of risk-sensitive portfolio choice as optimization under a penalized adverse probability tilt. In a companion paper, Lleo & Runggaldier (2026) relax the known-parameters assumption and demonstrate how this framework extends naturally to reinforcement learning settings.

Section 2 develops the analytical core of the paper. Using free energy–entropy duality, we

reformulate the original benchmarked problem as a linear-quadratic-Gaussian (LQG) stochastic differential game under a suitable equivalent probability measure, with an entropic regularization arising from the change of measure. This reformulation makes the problem tractable: the optimal controls are explicit, affine feedback rules, the value function is quadratic, and its coefficients solve a Riccati equation together with related linear equations. It also yields structural insight. The optimal strategy can be decomposed into a fractional Kelly allocation, a benchmark-tracking portfolio, and an intertemporal hedging portfolio, or equivalently represented as a Kelly allocation corrected by a term induced by the duality. In this way, the dual formulation does not merely solve the problem; it also interprets risk sensitivity through the lens of robustness to adverse probability tilts, with entropy regularization shaping the control.

Section 3 shows that free energy–entropy duality offers a direct route to the classical solution of the benchmarked risk-sensitive investment problem. First, it establishes equivalence between the duality-based approach and the Kuroda–Nagai change-of-measure method. Second, the section shows that the change of probability measure induced by duality decomposes into the Kuroda–Nagai transformation followed by an entropy-induced change. Third, this analysis not only recovers the standard optimal investment rule but also clarifies the underlying measure-theoretic and PDE mechanisms. Thus, Section 3 both validates the direct method and highlights the advantages of reframing risk-sensitive control problems in a free energy–entropy duality framework.

Section 4 provides a simple implementation on U.S. equity data. Its purpose is illustrative rather than empirical. Even so, the exercise strengthens the paper in three respects. First, it shows that the duality-based strategy can be computed and simulated in a realistic multi-factor setting. Second, it numerically confirms that the direct duality route and the Kuroda–Nagai route yield identical outcomes, as the theory predicts. Third, by comparing the performance of the resulting allocation with that of the benchmark and of the Kelly portfolio, the model’s economic content becomes more concrete and highlights the trade-off between growth and risk control. The implementation is therefore not intended as a forecasting exercise, but as a proof of concept for the mathematical framework developed in the paper.

2 Solving the Risk-Sensitive Benchmarked Investment Management Problem via the Free Energy–Entropy Duality

This paper connects stochastic control theory and stochastic differential games, with a view toward a formulation that can later be embedded in a reinforcement-learning framework based on policy-gradient methods. To keep the terminology consistent with that broader objective, we distinguish between *strategy*, *control* (or *action*), and *policy*. A *strategy* is a stochastic process $H = (h_s)_{s \in [0, T]}$ defined on the time interval $[0, T]$. The value $h_s \in \mathbb{R}^m$ denotes the *control*, or *action*, applied at time s . Informally, a *policy* specifies how the control is selected as a function of the current time and state. Formally, a policy is a measurable map

$$h : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^m,$$

and the associated control process is called *Markov* when it takes the form $h_s = h(s, X_s)$ —see Remark 3(i). This terminology is important for the rest of the paper: the control problem will later be recast as a game, the optimal policy will be shown to be Markov, and this same object is the natural one to parameterize in a policy-gradient framework.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{0 \leq s \leq T}, \mathbb{P})$ be a filtered complete probability space supporting a d -dimensional (\mathcal{F}_s) -Wiener process $W = (W_s)_{0 \leq s \leq T}$, with components W_s^i , $i = 1, \dots, d$. We take the Brownian motion to have dimension $d = n + m + 1$, where $n \geq 1$ is the number of factors, $m \geq 1$ is the number of financial assets, and the additional dimension captures the component of benchmark risk that is not spanned by the traded assets¹.

2.1 Model for the Financial Market

Using the money market account as numéraire, we model the discounted prices of the m risky assets by

$$\frac{dS_s^i}{S_s^i} = (a_s + A_s X_s)^i ds + \sum_{j=1}^d \sigma_s^{ij} dW_s^j, \quad i = 1, \dots, m, \quad (2.1)$$

with initial values S_0^i , $i = 1, \dots, m$. The deterministic coefficient functions $a : [0, T] \rightarrow \mathbb{R}^m$, $A : [0, T] \rightarrow \mathbb{R}^{m \times n}$, $\Sigma : [0, T] \rightarrow \mathbb{R}^{m \times d}$, where $\Sigma_s = (\sigma_s^{ij})_{1 \leq i \leq m, 1 \leq j \leq d}$, are assumed to be of class C^1 .

Remark 1 (Notation). Throughout the paper, we write a_s , A_s , Σ_s , and similarly for other time-dependent coefficients, in place of $a(s)$, $A(s)$, $\Sigma(s)$, and so forth.

We impose the following nondegeneracy condition on the risky assets.

Assumption 2.1. For every $s \in [0, T]$, the matrix $\Sigma_s \Sigma_s'$ is positive definite.

Remark 2. Under Assumption 2.1, the instantaneous covariance matrix of risky-asset returns is nonsingular, so the traded assets generate distinct directions of risk exposure.

The discounted asset drifts depend affinely on an n -dimensional factor process $X = (X_s)_{s \in [0, T]}$, modeled as a Gaussian diffusion with dynamics

$$dX_s = (b_s + B_s X_s) ds + \Lambda_s dW_s, \quad (2.2)$$

where $b : [0, T] \rightarrow \mathbb{R}^n$, $B : [0, T] \rightarrow \mathbb{R}^{n \times n}$, $\Lambda : [0, T] \rightarrow \mathbb{R}^{n \times d}$ are suitable C^1 functions. We refer to X as the *state process*. The initial condition is $X_0 = x$, where x may be either deterministic or random. In the latter case, we assume $X_0 \sim N(\mu, P_0)$, for known $\mu \in \mathbb{R}^n$ and covariance matrix $P_0 \in \mathbb{R}^{n \times n}$, and that X_0 is independent of the Wiener process $(W_s)_{s \in [0, T]}$.

The investor evaluates portfolio performance relative to a benchmark, such as a market index or a passive reference portfolio. We model the discounted benchmark level $L = (L_s)_{s \in [0, T]}$ by

$$\frac{dL_s}{L_s} = (c_s + C_s X_s) ds + \Xi_s' dW_s, \quad (2.3)$$

where $c : [0, T] \rightarrow \mathbb{R}$, $C : [0, T] \rightarrow \mathbb{R}^n$, $\Xi : [0, T] \rightarrow \mathbb{R}^d$ are C^1 functions.

¹If the benchmark is perfectly spanned by the traded assets, one may instead take $d = n + m$.

2.2 Risk-Sensitive Control Problem

Let $H = (h_s)_{s \in [0, T]}$ be a \mathcal{F}_s -adapted, self-financing investment strategy that represents the vector of portfolio weights in the risky assets. Short selling and leverage are allowed, so H takes its values in \mathbb{R}^m . The corresponding discounted wealth process $V = (V_s)_{s \in [0, T]}$ then satisfies

$$\frac{dV_s}{V_s} = h'_s(a_s + A_s X_s) ds + h'_s \Sigma_s dW_s. \quad (2.4)$$

To measure cumulative performance relative to the benchmark, we introduce the log price relative

$$R_s := \ln \frac{V_s}{L_s}, \quad s \in [0, T].$$

Applying Itô's formula to (2.4) and (2.3) yields

$$dR_s = \ell(s, X_s, h_s) ds + (h'_s \Sigma_s - \Xi'_s) dW_s, \quad (2.5)$$

where

$$\ell(s, x, h) := -\frac{1}{2} h' \Sigma_s \Sigma'_s h + h' a_s + \frac{1}{2} \Xi'_s \Xi_s - c_s + (h' A_s - C_s) x.$$

The cumulative log excess return of the portfolio relative to the benchmark over $[0, T]$ is therefore $R_T - R_0$.

The investor seeks to maximize the risk-sensitive benchmarked performance criterion

$$\begin{aligned} J(H, \theta) &:= -\frac{1}{\theta} \ln \mathbf{E} \left[e^{-\theta(R_T - R_0)} \right] \\ &= -\frac{1}{\theta} \ln \mathbf{E} \left[\exp \left\{ -\theta \int_0^T \ell(s, X_s, h_s) ds - \theta \int_0^T (h'_s \Sigma_s - \Xi'_s) dW_s \right\} \right], \end{aligned} \quad (2.6)$$

where $\theta \in (-1, 0) \cup (0, \infty)$ is the risk-sensitivity parameter and $T < \infty$ is the investment horizon. Equivalently, one may minimize the exponentially transformed criterion

$$I(H, \theta) := e^{-\theta J(H, \theta)} = \mathbf{E} \left[e^{-\theta(R_T - R_0)} \right]. \quad (2.7)$$

In most asset-management applications, the case $\theta > 0$ is the primary one of interest. In the benchmarked asset-management interpretation, $\theta > 0$ corresponds to more conservative behavior than the Kelly/log-utility benchmark. By contrast, $\theta \in (-1, 0)$ leads to overbetting strategies that leverage the Kelly portfolio and therefore trade expected return for additional volatility (Davis & Lleo, 2021).

We restrict attention to the following admissible class, which is sufficient for the well-posedness of the wealth equation and for the duality arguments developed later.

Definition 2.2. A strategy $H = (h_s)_{s \in [0, T]}$ belongs to the class \mathcal{A}^H if:

- (i) h_s is a Markov control;
- (ii) $\mathbb{P} \left(\int_0^T |h_s|^2 ds < \infty \right) = 1$.

Remark 3.

- (i) A Markov control is a process of the form $h_s = f(s, X_s)$ for some measurable map $f : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^m$. We refer the reader to p. 159 in Fleming & Soner (2006) for a formal definition. Markov controls are a natural admissibility class for the present problem. Moreover, Theorem 2.9 below shows that the optimal strategies derived are indeed generated by affine Markov controls.
- (ii) Condition (ii) imposes square integrability on the control process. This guarantees that the controlled wealth dynamics are well defined and supports the duality arguments developed later in the paper.

In this setup, the affine factor structure gives tractability while connecting to the empirical asset-pricing and factor-investing literatures. The benchmarked formulation focuses on relative performance and accommodates strategies ranging from active management to passive benchmark replication. Finally, the exponential performance criterion captures risk sensitivity in a form that is compatible with the free energy–entropy duality approach.

2.3 Free Energy–Entropy Duality Relation

We begin by recalling the free energy–entropy duality from Dai Pra et al. (1996), and then show how it applies to our risk-sensitive investment management problem.

2.3.1 General Definitions

The *free energy* of a random variable ψ with respect to the probability measure \mathbb{P} , whenever e^ψ is integrable, is defined by

$$\mathcal{E}^{\mathbb{P}}(\psi) := \ln \left(\int e^\psi d\mathbb{P} \right). \quad (2.8)$$

Let \mathbb{Q} be another probability measure. The *relative entropy* of \mathbb{Q} with respect to \mathbb{P} is

$$D_{\text{KL}}(\mathbb{Q} \parallel \mathbb{P}) := \int \ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{Q}. \quad (2.9)$$

The *free energy–entropy duality* relation, as given in Dai Pra et al. (1996, Proposition 2.3(ii)), states that

$$\mathcal{E}^{\mathbb{P}}(\psi) = \sup_{\mathbb{Q} \ll \mathbb{P}} \left\{ \int \psi d\mathbb{Q} - D_{\text{KL}}(\mathbb{Q} \parallel \mathbb{P}) \right\}. \quad (2.10)$$

2.3.2 Applying the Free Energy–Entropy Duality to the Risk-Sensitive Benchmarked Investment Problem

Letting $\psi = -\theta(R_T - R_0)$, the duality relation implies that, for every $H \in \mathcal{A}^H$,

$$\ln I(H, \theta) = \ln \mathbf{E} \left[e^{-\theta(R_T - R_0)} \right] = \sup_{\mathbb{P}^\Gamma \ll \mathbb{P}} \left\{ \mathbf{E}^{\mathbb{P}^\Gamma} [-\theta(R_T - R_0)] - D_{\text{KL}}(\mathbb{P}^\Gamma \parallel \mathbb{P}) \right\}, \quad (2.11)$$

where the supremum is taken over probability measures absolutely continuous with respect to \mathbb{P} ; see Dai Pra et al. (1996, Section 3.1.2). Here, $\mathbf{E}^{\mathbb{P}^\Gamma}[\cdot]$ denotes expectation under \mathbb{P}^Γ .

We now make the variational representation in (2.11) explicit by introducing a process Γ that generates the measure change. By Theorem 7.11 in Liptser & Shiryaev (2004), there exists an (\mathcal{F}_s) -progressively measurable process $\Gamma = (\gamma_s)_{s \in [0, T]}$ in \mathbb{R}^d such that $\int_0^T \|\gamma_s\|^2 ds < \infty$ almost surely and

$$W_s^\Gamma := W_s - \int_0^s \gamma_u du \quad (2.12)$$

is a \mathbb{P}^Γ -Wiener process. Define the density process

$$\chi_s^\Gamma := \exp \left\{ -\frac{1}{2} \int_0^s \|\gamma_u\|^2 du + \int_0^s \gamma_u' dW_u \right\}, \quad s \in [0, T]. \quad (2.13)$$

By Girsanov's theorem, $\frac{d\mathbb{P}^\Gamma}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = \chi_T^\Gamma$, where $\frac{d\mathbb{P}^\Gamma}{d\mathbb{P}} \Big|_{\mathcal{F}_T}$ is the restriction of the Radon-Nikodym derivative $\frac{d\mathbb{P}^\Gamma}{d\mathbb{P}}$ to the filtration \mathcal{F}_T .

Under \mathbb{P}^Γ , W^Γ is a Brownian motion, and the drift terms are shifted accordingly. Under \mathbb{P}^Γ , the state process satisfies

$$dX_s = (b_s + B_s X_s + \Lambda_s \gamma_s) ds + \Lambda_s dW_s^\Gamma. \quad (2.14)$$

The corresponding dynamics of the log price relative process become

$$\begin{aligned} dR_s = & \left[-\frac{1}{2} h_s' \Sigma_s \Sigma_s' h_s + h_s' a_s + \frac{1}{2} \Xi_s' \Xi_s - c_s + (h_s' \Sigma_s - \Xi_s') \gamma_s + (h_s' A_s - C_s) X_s \right] ds \\ & + (h_s' \Sigma_s - \Xi_s') dW_s^\Gamma. \end{aligned} \quad (2.15)$$

Moreover, standard entropy calculations yield

$$D_{\text{KL}}(\mathbb{P}^\Gamma \| \mathbb{P}) = \mathbf{E}^{\mathbb{P}^\Gamma} [\ln \chi_T^\Gamma] = \frac{1}{2} \mathbf{E}^{\mathbb{P}^\Gamma} \left[\int_0^T \|\gamma_s\|^2 ds \right]. \quad (2.16)$$

Therefore, we can express the *free energy-entropy duality* as

$$\ln I(H, \theta) = \sup_{\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[-\theta(R_T - R_0) - \frac{1}{2} \int_0^T \|\gamma_s\|^2 ds \right], \quad (2.17)$$

for all choices of $H \in \mathcal{A}^H$, and where the supremum is taken over progressively measurable processes $\Gamma = (\gamma_s)_{s \in [0, T]}$.

To prepare the dynamic programming analysis, we restrict attention to a Markovian admissible class for the maximizing player.

Definition 2.3. A strategy $\Gamma = (\gamma_s)_{s \in [0, T]} \in \mathbb{R}^d$ is in class \mathcal{A}^Γ if the control process γ_s satisfies the following conditions:

- (i) γ_s is a Markov control;
- (ii) $\mathbb{P} \left(\int_0^T |\gamma_s|^2 ds < +\infty \right) = 1$;
- (iii) $(\chi_s^\Gamma)_{s \in [0, T]}$, defined at (2.13) above, is an exponential martingale.

The Markovian structure of the problem suggests that an optimizer Γ^* should lie in \mathcal{A}^Γ . We therefore impose the following restriction, which is later validated by the verification theorem.

Assumption 2.4.

$$\sup_{\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[-\theta(R_T - R_0) - \frac{1}{2} \int_0^T \|\gamma_s\|^2 ds \right] = \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[-\theta(R_T - R_0) - \frac{1}{2} \int_0^T \|\gamma_s\|^2 ds \right]. \quad (2.18)$$

Expanding (2.17) using the dynamics of R under \mathbb{P}^Γ , we obtain

$$\begin{aligned} & \ln I(H, \theta) \\ &= \sup_{\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[-\theta \int_0^T \left\{ -\frac{1}{2} h'_s \Sigma_s \Sigma'_s h_s + h'_s a_s + \frac{1}{2} \Xi'_s \Xi_s - c_s + (h'_s \Sigma_s - \Xi'_s) \gamma_s + (h'_s A_s - C_s) X_s \right\} ds \right. \\ & \quad \left. - \theta \int_0^T (h'_s \Sigma_s - \Xi'_s) dW_s^\Gamma - \frac{1}{2} \int_0^T \|\gamma_s\|^2 ds \right] \\ & \stackrel{\text{Assumption 2.4}}{=} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[\theta \int_0^T g(s, X_s, h_s, \gamma_s; \theta) ds \right], \end{aligned} \quad (2.19)$$

where the running payoff g is defined by

$$g(s, x, h, \gamma; \theta) := \frac{1}{2} h'_s \Sigma_s \Sigma'_s h - h'_s a_s - \frac{1}{2} \Xi'_s \Xi_s + c_s - (h'_s \Sigma_s - \Xi'_s) \gamma - (h'_s A_s - C_s) x - \frac{1}{2\theta} \|\gamma\|^2. \quad (2.20)$$

The last equality follows because the stochastic integral, $\int_0^T (h'_s \Sigma_s - \Xi'_s) dW_s^\Gamma$, is a \mathbb{P}^Γ -martingale for $H \in \mathcal{A}^H$ and $\Gamma \in \mathcal{A}^\Gamma$.

In what follows, we focus on the practically most relevant case $\theta > 0$. The case $\theta \in (-1, 0)$ can be handled similarly², but it leads to overbetting relative to the Kelly benchmark and is therefore of limited practical interest. Since the logarithm is strictly increasing, minimizing $I(H, \theta)$ is equivalent to minimizing $\ln I(H, \theta)$. We make this equivalence explicit in the next display.

$$\begin{aligned} \inf_{H \in \mathcal{A}^H} \ln I(H, \theta) &= \inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[\theta \int_0^T g(s, X_s, h_s, \gamma_s; \theta) ds \right] \\ \Leftrightarrow \ln \inf_{H \in \mathcal{A}^H} I(H, \theta) &= \inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[\theta \int_0^T g(s, X_s, h_s, \gamma_s; \theta) ds \right] \\ \Leftrightarrow \inf_{H \in \mathcal{A}^H} I(H, \theta) &= \exp \left\{ \inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[\theta \int_0^T g(s, X_s, h_s, \gamma_s; \theta) ds \right] \right\}, \end{aligned} \quad (2.21)$$

where the first equivalence follows from Lemma A.1 in Appendix A.1.

The right-hand side of (2.21) admits a natural interpretation as a two-player stochastic differential game, which we now define.

²For the criterion J , instead of maximizing a concave function when $\theta > 0$, we minimize a convex function when $\theta < 0$.

Definition 2.5 (The implied stochastic differential game). For $H = (h_s)_{s \in [0, T]} \in \mathcal{A}^H$ and $\Gamma = (\gamma_s)_{s \in [0, T]} \in \mathcal{A}^\Gamma$, the implied stochastic differential game is defined by

$$\inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}^{\mathbb{P}^\Gamma} \left[\theta \int_0^T g(s, X_s, h_s, \gamma_s; \theta) ds \right], \quad (2.22)$$

where H is the minimizing player's strategy, Γ is the maximizing player's strategy, and the state dynamics are given by (2.14).

2.3.3 Solving the Stochastic Differential Game

This subsection studies the stochastic differential game introduced in Definition 2.5. The game is linear-quadratic-Gaussian. We show that the optimal strategies are generated by affine policies of the state variable X , and hence are Markov controls. Under Markov controls, the controlled state process remains Markov. Therefore, at any intermediate time $t \in [0, T]$, conditional expectations of functionals of $(X_s, h_s, \gamma_s)_{s \in [t, T]}$, given $X_t = x$, can be expressed as functions of (t, x) . We then show that the game's value function, defined in (2.27), is quadratic in the state variable x , with time-varying coefficients.

As in the previous subsection, we assume that $\theta > 0$ and consider a generic starting time $t \in [0, T]$. For convenience, we normalize the benchmark level at time t by setting $L_t = V_t$, so that the corresponding log price relative satisfies $R_t = 0$. The corresponding continuation criteria are then

$$J(t, x; H, \theta) := -\frac{1}{\theta} \ln \mathbf{E}_{t,x} \left[e^{-\theta R_T} \right] \quad (2.23)$$

$$I(t, x; H, \theta) := e^{-\theta J(t,x;H,\theta)} = \mathbf{E}_{t,x} \left[e^{-\theta R_T} \right]. \quad (2.24)$$

The continuation version of the last relation in (2.21) is

$$\inf_{H \in \mathcal{A}^H} I(t, x; H, \theta) = \exp \left\{ \inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\theta \int_t^T g(s, X_s, h_s, \gamma_s; \theta) ds \right] \right\}, \quad (2.25)$$

which yields the continuation version of the stochastic differential game:

$$\inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\theta \int_t^T g(s, X_s, h_s, \gamma_s; \theta) ds \right]. \quad (2.26)$$

We define the associated value function $u(t, x)$ by

$$u(t, x) := \inf_{H \in \mathcal{A}^H} \sup_{\Gamma \in \mathcal{A}^\Gamma} \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\theta \int_t^T g(s, X_s, h_s, \gamma_s; \theta) ds \right]. \quad (2.27)$$

Lemma 2.6. *The value function satisfies*

$$u(t, x) = \ln \inf_{H \in \mathcal{A}^H} I(t, x; H, \theta) = -\theta \sup_{H \in \mathcal{A}^H} J(t, x; H, \theta). \quad (2.28)$$

Proof. The first equality follows directly from the definition of u in (2.27) together with (2.25).

For the second equality, use (2.24) and the identity

$$\inf_{H \in \mathcal{A}^H} (-\theta J(t, x; H, \theta)) = -\theta \sup_{H \in \mathcal{A}^H} J(t, x; H, \theta), \quad \theta > 0.$$

□

Remark 4. The free energy–entropy duality rewrites the original risk-sensitive investment criterion under \mathbb{P} as the objective of an LQG game under \mathbb{P}^Γ . The running payoff g depends on both the investment strategy H and the change-of-measure process Γ . In particular, it contains the interaction term $h'_s \Sigma_s \gamma_s$. By contrast, the state dynamics in (2.14) are controlled only by the change-of-measure process Γ .

The function g is affine in the state variable x and quadratic in the control variables h and γ . Hence, the stochastic differential game associated with $u(t, x)$ is a linear-quadratic-Gaussian (LQG) game. The associated Bellman–Isaacs partial differential equation is

$$\frac{\partial u(s, x)}{\partial s} + \mathcal{H}^+(s, x, Du(s, x), D^2u(s, x)) = 0, \quad (2.29)$$

with terminal condition $u(T, x) = 0$. Here, $Du(s, x) = \left(\frac{\partial u}{\partial x_1}(s, x), \dots, \frac{\partial u}{\partial x_n}(s, x) \right)'$ and $D^2u(s, x) = \left[\frac{\partial^2 u}{\partial x_i \partial x_j}(s, x) \right]_{i, j=1}^n$.

The Hamiltonian is given by

$$\mathcal{H}^+(s, x, p, M) := \inf_{h \in \mathbb{R}^m} \sup_{\gamma \in \mathbb{R}^d} \left\{ [b_s + B_s x + \Lambda_s \gamma]' p + \frac{1}{2} \text{tr}(\Lambda_s \Lambda_s' M) + \theta g(s, x, h, \gamma; \theta) \right\}. \quad (2.30)$$

For completeness, we also define the Hamiltonian

$$\mathcal{H}^-(s, x, p, M) := \sup_{\gamma \in \mathbb{R}^d} \inf_{h \in \mathbb{R}^m} \left\{ [b_s + B_s x + \Lambda_s \gamma]' p + \frac{1}{2} \text{tr}(\Lambda_s \Lambda_s' M) + \theta g(s, x, h, \gamma; \theta) \right\}. \quad (2.31)$$

For notational convenience, let

$$\mathcal{P}_s^+(\theta) := I_d + \theta \Sigma_s' (\Sigma_s \Sigma_s')^{-1} \Sigma_s, \quad (2.32)$$

$$\mathcal{P}_s^-(\theta) := I_d - \frac{\theta}{\theta + 1} \Sigma_s' (\Sigma_s \Sigma_s')^{-1} \Sigma_s. \quad (2.33)$$

If we define $\Pi_s := \Sigma_s' (\Sigma_s \Sigma_s')^{-1} \Sigma_s$, then $\Pi_s^2 = \Pi_s$, so $\mathcal{P}_s^-(\theta) \mathcal{P}_s^+(\theta) = I_d$. Hence,

$$(\mathcal{P}_s^+(\theta))^{-1} = \mathcal{P}_s^-(\theta). \quad (2.34)$$

Proposition 2.7. *The Bellman–Isaacs equation (2.29) yields a unique pair of candidate fed-*

back policies \hat{h} and $\hat{\gamma}$, given equivalently by either of the following two representations:

$$\hat{h}(s, x, Du(s, x)) = \frac{1}{\theta + 1} (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s + \Sigma_s \Lambda_s' Du(s, x)) \quad (2.35)$$

$$\hat{\gamma}(s, x, Du(s, x)) = \Lambda_s' Du(s, x) - \theta \left(\Sigma_s' \hat{h}(s, x, Du(s, x)) - \Xi_s \right) \quad (2.36)$$

$$= \mathcal{P}_s^-(\theta) \Lambda_s' Du(s, x) - \frac{\theta}{\theta + 1} \Sigma_s' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x) + \theta \mathcal{P}_s^-(\theta) \Xi_s, \quad (2.37)$$

or

$$\hat{h}(s, x, Du(s, x)) = (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x) + (\Sigma_s \Sigma_s')^{-1} \Sigma_s \hat{\gamma}(s, x, Du(s, x)) \quad (2.38)$$

$$\hat{\gamma}(s, x, Du(s, x)) = \mathcal{P}_s^-(\theta) \left[-\theta \Sigma_s' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda_s' Du(s, x) \right] \quad (2.39)$$

Moreover, the two Hamiltonians coincide: $\mathcal{H}^+(s, x, p, M) = \mathcal{H}^-(s, x, p, M)$. Therefore, the Isaacs condition holds, and the stochastic differential game is well defined.

Technically, we only need to solve the problem (2.29) with the Hamiltonian \mathcal{H}^+ given at (2.30). Therefore, Proposition 2.7 gives us a stronger result than required and justifies the interpretation of the regularized problem as a stochastic differential game.

Proof. See Appendix A.2. □

Remark 5. The candidate policies satisfy the following local saddle-point condition:

$$\begin{aligned} & [b_s + B_s x + \Lambda_s \gamma]' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' D^2 u(s, x)) + \theta g \left(s, x, \hat{h}(s, x, Du(s, x)), \gamma; \theta \right) \\ & \leq [b_s + B_s x + \Lambda_s \hat{\gamma}(s, x, Du(s, x))]' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' D^2 u(s, x)) \\ & \quad + \theta g \left(s, x, \hat{h}(s, x, Du(s, x)), \hat{\gamma}(s, x, Du(s, x)); \theta \right) \\ & \leq [b_s + B_s x + \Lambda_s \hat{\gamma}(s, x, Du(s, x))]' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' D^2 u(s, x)) + \theta g \left(s, x, h, \hat{\gamma}(s, x, Du(s, x)); \theta \right), \end{aligned} \quad (2.40)$$

for all $h \in \mathbb{R}^m$ and $\gamma \in \mathbb{R}^d$. We use this pointwise saddle-point property later in the proof of Verification Theorem 2.9 to establish the corresponding global saddle-point condition.

Remark 6 (Measure $\mathbb{P}^{\hat{\Gamma}}$). Since the candidate control strategies in Proposition 2.7 admit equivalent representations, we use here the first pair (2.35)–(2.36). The corresponding density increment over $[t, T]$ is

$$\begin{aligned} \chi_{[t, T]}^{\hat{\Gamma}} := & \exp \left\{ -\frac{1}{2} \int_t^T \left\| \Lambda_s' Du(s, X_s) - \theta \left(\Sigma_s' \hat{h}(s, X_s, Du(s, X_s)) - \Xi_s \right) \right\|^2 ds \right. \\ & \left. + \int_t^T \left[\Lambda_s' Du(s, X_s) - \theta \left(\Sigma_s' \hat{h}(s, X_s, Du(s, X_s)) - \Xi_s \right) \right]' dW_s \right\}. \end{aligned} \quad (2.41)$$

Substituting (2.37) yields the equivalent expression

$$\begin{aligned} \chi_{[t,T]}^{\hat{\Gamma}} = & \exp \left\{ -\frac{1}{2} \int_t^T \left\| -\frac{\theta}{\theta+1} \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s X_s) + \theta \mathcal{P}_s^-(\theta) \Xi_s + \mathcal{P}_s^-(\theta) \Lambda'_s Du(s, X_s) \right\|^2 ds \right. \\ & \left. + \int_t^T \left[-\frac{\theta}{\theta+1} \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s X_s) + \theta \mathcal{P}_s^-(\theta) \Xi_s + \mathcal{P}_s^-(\theta) \Lambda'_s Du(s, X_s) \right]' dW_s \right\}. \end{aligned}$$

Remark 7 (Special case of the Kelly portfolio). Formally, as $\theta \rightarrow 0$, one recovers the Kelly or growth-optimal portfolio. In that limit, the density increment over $[t, T]$ simplifies to

$$\chi_{[t,T]}^{\hat{\Gamma}} = \exp \left\{ -\frac{1}{2} \int_t^T \|\Lambda'_s Du(s, X_s)\|^2 ds + \int_t^T [\Lambda'_s Du(s, X_s)]' dW_s \right\}. \quad (2.42)$$

Theorem 2.8. *The Bellman–Isaacs equation (2.29), associated with the candidate strategy pair $(\hat{H}, \hat{\Gamma})$ defined in Proposition 2.7, admits a solution of the form*

$$u(t, x) = -\theta \left(\frac{1}{2} x' Q_t x + q'_t x + k_t \right), \quad (2.43)$$

where $Q : [0, T] \rightarrow \mathbb{R}^{n \times n}$, $q : [0, T] \rightarrow \mathbb{R}^n$, and $k : [0, T] \rightarrow \mathbb{R}$ are deterministic functions such that:

1. Q is the unique symmetric positive semidefinite solution to the matrix Riccati equation

$$\begin{aligned} \dot{Q}_s - \theta Q_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s Q_s + \left(B'_s - \frac{\theta}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s \right) Q_s \\ + Q_s \left(B_s - \frac{\theta}{\theta+1} \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} A_s \right) + \frac{1}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} A_s = 0, \quad Q(T) = 0; \end{aligned} \quad (2.44)$$

2. q solves the linear ODE

$$\begin{aligned} \dot{q}_s + \left(B'_s - \frac{\theta}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s \right) q_s - \theta Q_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s \\ + Q_s \left[b_s - \frac{\theta}{\theta+1} \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) + \theta \Lambda_s \Xi_s \right] \\ - C'_s + \frac{1}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) = 0, \quad q(T) = 0. \end{aligned} \quad (2.45)$$

3. k is given by the integral

$$\begin{aligned} k_t = \int_t^T \left\{ \frac{\theta}{2} q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s - \left[b'_s + \theta \Xi'_s \Lambda'_s - \frac{\theta}{\theta+1} (a_s + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s \right] q_s - \frac{1}{2} \text{tr} (\Lambda_s \Lambda'_s Q_s) \right. \\ \left. - \frac{1}{2(\theta+1)} (a_s + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) + c_s + \frac{\theta-1}{2} \Xi'_s \Xi_s \right\} ds. \end{aligned} \quad (2.46)$$

Proof. See Appendix A.3. □

Remark 8. The term $\Lambda'_s Du(s, x)$ in the definition of $\hat{\gamma}$ at (2.36) generates, after optimization over γ , a quadratic gradient term in the Bellman–Isaacs PDE. In the present benchmarked risk-sensitive setting, this term appears through the matrix $\Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s$. Consequently, the equations (2.44)–(2.46) contain terms that are characteristic of risk-sensitive control problems and absent from standard LQG problems, namely, $-\theta Q_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s Q_s$ in (2.44), $-\theta Q_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s$ in (2.45), and $\frac{\theta}{2} q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s$ in (2.46).

Theorem 2.9 (Verification). *Let u be the value function defined at (2.27) and having the quadratic expression in (2.43). Let $H^* = (h^*(s, X_s))_{s \in [t, T]} := \left(\hat{h}(s, X_s, Du(s, X_s)) \right)_{s \in [t, T]}$ and $\Gamma^* = (\gamma^*(s, X_s))_{s \in [t, T]} := \left(\hat{\gamma}(s, X_s, Du(s, X_s)) \right)_{s \in [t, T]}$, where $\hat{h}(\cdot)$ and $\hat{\gamma}(\cdot)$ are the candidate policies given in Proposition 2.7, and let*

$$h_s^* := h^*(s, X_s), \quad \gamma_s^* := \gamma^*(s, X_s).$$

Then,

1. *The pair of strategies (H^*, Γ^*) is admissible in the sense that they belong to the classes \mathcal{A}^H and \mathcal{A}^Γ respectively. Moreover, their defining policies $h^*(\cdot)$ and $\gamma^*(\cdot)$ are Borel-measurable maps.*
2. *The value function u solves the stochastic differential game with the objective (2.26), in the sense that the saddle point condition*

$$\mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\theta \int_t^T g(s, X_s, h_s^*, \gamma_s; \theta) ds \right] \leq u(t, x) \leq \mathbf{E}_{t,x}^{\mathbb{P}^{\Gamma^*}} \left[\theta \int_t^T g(s, X_s, h_s, \gamma_s^*; \theta) ds \right] \quad (2.47)$$

holds for any admissible strategies $H = (h_s)_{s \in [t, T]} \in \mathcal{A}^H$ and $\Gamma = (\gamma_s)_{s \in [t, T]} \in \mathcal{A}^\Gamma$, and that

$$u(t, x) = \mathbf{E}_{t,x}^{\mathbb{P}^{\Gamma^*}} \left[\theta \int_t^T g(s, X_s, h_s^*, \gamma_s^*; \theta) ds \right]. \quad (2.48)$$

3. *The strategies H^* and Γ^* are optimal.*

Proof. See Appendix A.4. □

Remark 9. The pair of strategies (H^*, Γ^*) forms a *saddle point* of the stochastic differential game, as evidenced by Remark 5 and equation (2.47). Equivalently, it is a *Nash equilibrium in pure strategies* for the associated stochastic game.

We conclude this section with a well-known decomposition³ of the optimal investment strategy as a fractional Kelly strategy. The strategy combines three constituent portfolios: the Kelly or log-optimal portfolio, a benchmark-tracking portfolio, and an intertemporal hedging portfolio.

Proposition 2.10 (Fractional Kelly strategy). *Let $f := \frac{1}{\theta+1}$. Then, for $s \in [t, T]$, the optimal investment strategy admits the decomposition*

$$h^*(s, X_s) = f h^K(s, X_s) + (1-f) h_s^{\text{Bench}} - (1-f) h^{\text{IHP}}(s, X_s), \quad (2.49)$$

³We refer the reader to (Davis & Lleo, 2008) and Chapter 3 in (Davis & Lleo, 2014), among others.

where

$$h^K(s, X_s) = (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s X_s), \quad (2.50)$$

$$h_s^{\text{Bench}} = (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Xi_s, \quad (2.51)$$

$$h^{\text{HP}}(s, X_s) = (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s (q_s + Q_s X_s). \quad (2.52)$$

Proof. Using (2.35) together with $Du(s, X_s) = -\theta (Q_s X_s + q_s)$, we obtain

$$h^*(s, X_s) = \frac{1}{\theta + 1} (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s X_s) + \frac{\theta}{\theta + 1} (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Xi_s - \frac{\theta}{\theta + 1} (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s (q_s + Q_s X_s),$$

which is exactly (2.49) for the definitions (2.50)–(2.52). □

The free energy–entropy duality also yields a complementary interpretation of the optimal allocation. In addition to the fractional Kelly decomposition above, the optimal strategy can be written as the Kelly portfolio corrected by a term induced by the optimal change-of-measure control γ^* , as shown in the next corollary.

Corollary 2.11 (Optimal Allocation as a Regularized Kelly Strategy). *The optimal investment strategy satisfies*

$$h^*(s, X_s) = h^K(s, X_s) + (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \gamma^*(s, X_s). \quad (2.53)$$

Proof. The identity follows immediately from (2.38) and the definition of h^K in (2.50). □

Equation (2.53) mirrors the structure of the free energy–entropy duality. At the criterion level, the duality rewrites the original risk-sensitive problem as an LQG problem under a transformed probability measure, with an additional quadratic penalty on the change-of-measure control γ . At the control level, the optimal allocation is the Kelly (or log-optimal) strategy corrected by the term $(\Sigma_s \Sigma'_s)^{-1} \Sigma_s \gamma^*(s, X_s)$ induced by the optimal antagonistic control γ^* . This correction also admits a natural geometric interpretation: it is the image of the optimal antagonistic control $\gamma^*(s, X_s) \in \mathbb{R}^d$ under the linear map $(\Sigma_s \Sigma'_s)^{-1} \Sigma_s$.

3 Understanding the Free Energy–Entropy Duality Through the Lens of the Kuroda–Nagai Change of Probability Measure

This section analyzes the role of the change of measure in the application of the free energy–entropy duality to the benchmarked risk-sensitive investment management problem. To this end, we compare the direct one-step approach developed in the previous section with a two-step route in which one first performs a change of measure *à la* Kuroda & Nagai (2002) and then applies the duality under the transformed probability measure. Because the Kuroda–Nagai

measure change underlies many of the standard solution techniques in the risk-sensitive investment management literature (see, e.g., Davis & Lleo, 2014, 2021, 2024, and the references therein), it provides a natural benchmark against which to interpret the direct duality-based construction. The comparison reveals how the change of measure induced by the duality in the approach proposed in Section 2 relates to the classical Kuroda–Nagai transformation, and thereby clarifies the role of the entropy penalty in (2.16).

3.1 A Two-Step Kuroda–Nagai/Duality Route

This subsection shows that the direct solution derived in Section 2.3 is equivalent to a sequential two-step procedure in which one first performs the Kuroda–Nagai change of measure and then applies the free energy–entropy duality under the transformed measure. Consistently with the derivations in subsections 2.3.2–2.3.3, and in light of the results proved in Section 2.3.3, we will focus here on classes of Markov controls.

3.1.1 Kuroda–Nagai Change of Measure from \mathbb{P} to \mathbb{P}^H

We start by expressing the risk-sensitive benchmarked criterion in (2.23) as follows:

$$J(t, x; H, \theta) := -\frac{1}{\theta} \ln \mathbf{E}_{t,x} \left[e^{-\theta R_T} \right] = -\frac{1}{\theta} \ln \mathbf{E}_{t,x} \left[\exp \left\{ \theta \int_t^T g_1(s, X_s, h_s; \theta) ds \right\} \chi_{[t,T]}^H \right], \quad (3.1)$$

where

$$g_1(s, x, h; \theta) := \frac{\theta + 1}{2} h' \Sigma_s \Sigma_s' h - h' (a_s + A_s x) - \theta h' \Sigma_s \Xi_s + (c_s + C_s x) + \frac{\theta - 1}{2} \Xi_s' \Xi_s, \quad (3.2)$$

and the Doléans–Dade exponential $\chi_{[t,T]}^H$ is defined via

$$\chi_{[t,T]}^H := \exp \left\{ -\theta \int_t^T (h_s' \Sigma_s - \Xi_s') dW_s - \frac{1}{2} \theta^2 \int_t^T (h_s' \Sigma_s - \Xi_s') (\Sigma_s' h_s - \Xi_s) ds \right\}. \quad (3.3)$$

Here, we need to assume that the investor’s strategy $H = (h_s)_{s \in [t,T]}$ is in class \mathcal{A}^{HN} defined below.

Definition 3.1. $H = (h_s)_{s \in [t,T]}$ with values in \mathbb{R}^m is in class \mathcal{A}^{HN} if the control h_s satisfies the following conditions:

- (i) h_s is a Markov control;
- (ii) $P \left(\int_t^T |h_s|^2 ds < +\infty \right) = 1$;
- (iii) the Doléans exponential $\chi_{[t,T]}^H$ defined as (3.3) is an exponential martingale.

For $H \in \mathcal{A}^{HN}$, let \mathbb{P}^H be the probability measure on (Ω, \mathcal{F}_T) defined via the Radon–Nikodym derivative

$$\left. \frac{d\mathbb{P}^H}{d\mathbb{P}} \right|_{\mathcal{F}_T} := \chi_{[t,T]}^H. \quad (3.4)$$

Moreover,

$$W_s^H = W_s + \theta \int_t^s (\Sigma'_u h_u - \Xi_u) du \quad (3.5)$$

is a standard Wiener process under \mathbb{P}^H . The \mathbb{P}^H -dynamics of the state process X_s is

$$dX_s = [b_s + B_s X_s - \theta \Lambda_s (\Sigma'_s h_s - \Xi_s)] ds + \Lambda_s dW_s^H, \quad s \in [t, T]. \quad (3.6)$$

Finally, the \mathbb{P}^H -risk-sensitive criterion J^H and exponentially-transformed criterion I^H are

$$J^H(t, x; H, \theta) := -\frac{1}{\theta} \ln \mathbf{E}_{t,x}^H \left[\exp \left\{ \theta \int_t^T g_1(s, X_s, h_s; \theta) ds \right\} \right] \quad (3.7)$$

$$I^H(t, x; H, \theta) := e^{-\theta J^H(t, x; H, \theta)} = \mathbf{E}_{t,x}^H \left[\exp \left\{ \theta \int_t^T g_1(s, X_s, h_s; \theta) ds \right\} \right], \quad (3.8)$$

where $\mathbf{E}_{t,x}^H[\cdot]$ denotes the expectation taken with respect to the probability measure \mathbb{P}^H and with initial condition $(t, X_t = x)$.

Remark 10. Under the probability measure \mathbb{P} , the benchmarked investment management problem is not a standard (risk-sensitive) stochastic control problem due to two important features. First, the state dynamics at (2.2) is uncontrolled. Second, the terminal reward $-\theta R_T$ in the exponential contains an Itô integral with respect to the \mathbb{P} -Brownian motion W_s . The main interest of the Kuroda–Nagai change of measure is to transform this non-standard control problem into a standard linear exponential-of-quadratic Gaussian (LEQG) control problem with a controlled state process under the probability measure \mathbb{P}^H (see Kuroda & Nagai, 2002; Davis & Lleo, 2008, 2014).

3.1.2 Applying the Free Energy–Entropy Duality Under \mathbb{P}^H

To apply the free energy–entropy argument, we introduce the probability measure \mathbb{P}^N defined from the probability measure \mathbb{P}^H via the Radon-Nikodym derivative

$$\frac{d\mathbb{P}^N}{d\mathbb{P}^H} \Big|_{\mathcal{F}_T} = \exp \left\{ -\frac{1}{2} \int_t^T \|\nu_s\|^2 ds + \int_t^T \nu'_s dW_s^H \right\} =: \chi_{[t,T]}^N, \quad (3.9)$$

for a strategy N in class \mathcal{A}^N defined as:

Definition 3.2. A strategy $N = (\nu_s)_{s \in [t, T]}$ with values in \mathbb{R}^d is in class \mathcal{A}^N if the control ν_s satisfies the following conditions:

- (i) ν_s is a Markov control;
- (ii) $P^H \left(\int_t^T |\nu_s|^2 ds < +\infty \right) = 1$;
- (iii) $\chi_{[t,T]}^N$, defined at (3.9), is an exponential martingale.

The \mathbb{P}^N -standard Wiener process is

$$W_s^N = W_s^H - \int_t^s \nu_u du \quad (3.10)$$

and the state process X_s satisfies

$$dX_s = \{b_s + B_s X_s - \Lambda_s [\theta (\Sigma'_s h_s - \Xi_s) - \nu_s]\} ds + \Lambda_s dW_s^N, \quad s \in [t, T]. \quad (3.11)$$

The *free energy–entropy duality relation* establishes the following relation between the \mathbb{P}^H -criterion I^H introduced at (3.8) and the \mathbb{P}^N -expectation of the function g_1 plus a regularization penalty:

$$\ln I^H(t, x; H, \theta) = \sup_{N \in \mathcal{A}^N} \mathbf{E}_{t,x}^N \left[\int_t^T \left\{ \theta g_1(s, X_s, h_s; \theta) - \frac{1}{2} \|\nu_s\|^2 \right\} ds \right]. \quad (3.12)$$

As in Section 2.3, we focus on the case $\theta > 0$. By Lemma A.1 in Appendix A.1,

$$\begin{aligned} \inf_{H \in \mathcal{A}^{HN}} I^H(t, x; H, \theta) &= \inf_{H \in \mathcal{A}^{HN}} \exp \left\{ \sup_{N \in \mathcal{A}^N} \mathbf{E}_{t,x}^N \left[\int_t^T \left\{ \theta g_1(s, X_s, h_s; \theta) - \frac{1}{2} \|\nu_s\|^2 \right\} ds \right] \right\} \\ &= \exp \left\{ \inf_{H \in \mathcal{A}^{HN}} \sup_{N \in \mathcal{A}^N} \mathbf{E}_{t,x}^N \left[\int_t^T \left\{ \theta g_1(s, X_s, h_s; \theta) - \frac{1}{2} \|\nu_s\|^2 \right\} ds \right] \right\}, \end{aligned} \quad (3.13)$$

where the state process X is given at (3.11). Using relation (3.12), we express $J^H(t, x; H, \theta)$ as

$$J^H(t, x; H, \theta) = -\frac{1}{\theta} \ln I^H(t, x; H, \theta) = \inf_{N \in \mathcal{A}^N} \mathbf{E}_{t,x}^N \left[\int_t^T \left\{ -g_1(s, X_s, h_s; \theta) + \frac{1}{2\theta} \|\nu_s\|^2 \right\} ds \right]. \quad (3.14)$$

Hence,

$$\sup_{H \in \mathcal{A}^{HN}} J^H(t, x; H, \theta) = \sup_{H \in \mathcal{A}^{HN}} \inf_{N \in \mathcal{A}^N} \mathbf{E}_{t,x}^N \left[\int_t^T \left\{ -g_1(s, X_s, h_s; \theta) + \frac{1}{2\theta} \|\nu_s\|^2 \right\} ds \right]. \quad (3.15)$$

We have therefore arrived again at a stochastic differential game formulation, now under the probability measure \mathbb{P}^N , in which the state process is controlled jointly by h_s and ν_s . As in subsection 2.3.3, we shall obtain optimal strategies defined by policies that are affine functions of the state variable X .

We may then define a value function $U(t, x)$ as

$$U(t, x) := \sup_{H \in \mathcal{A}^{HN}} \inf_{N \in \mathcal{A}^N} \mathbf{E}_{t,x}^N \left[\int_t^T \left\{ -g_1(s, X_s, h_s; \theta) + \frac{1}{2\theta} \|\nu_s\|^2 \right\} ds \right], \quad (3.16)$$

to which we associate the Bellman–Isaacs PDE

$$\frac{\partial U(s, x)}{\partial s} + \mathcal{H}^+(s, x, DU(s, x), D^2U(s, x)) = 0, \quad U(T, x) = 0, \forall x \in \mathbb{R}^n, \quad (3.17)$$

where DU and D^2U are defined as in Section 2.3.3, and

$$\begin{aligned} & \mathcal{H}^+(s, x, p, M) \\ := & \sup_{h \in \mathbb{R}^m} \inf_{\nu \in \mathbb{R}^d} \left\{ \{b_s + B_s x - \Lambda_s [\theta (\Sigma'_s h - \Xi_s) - \nu]\}' p + \frac{1}{2} \text{tr} (\Lambda_s \Lambda'_s M) - g_1(s, x, h; \theta) + \frac{1}{2\theta} \nu' \nu \right\} \end{aligned} \quad (3.18)$$

for $p \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$.

Proceeding as in the proof of Proposition 2.7, we obtain the candidate policies $\hat{h}(s, x, DU(s, x))$ and $\hat{\nu}(s, x, DU(s, x))$ as

$$\begin{cases} \hat{h}(s, x, DU(s, x)) = \frac{1}{\theta+1} (\Sigma_s \Sigma'_s)^{-1} [a_s + A_s x + \theta \Sigma_s \Xi_s - \theta \Sigma_s \Lambda'_s DU(s, x)], \\ \hat{\nu}(s, x, DU(s, x)) = -\theta \Lambda'_s DU(s, x). \end{cases} \quad (3.19)$$

Then, we define the candidate strategies $\hat{H} := \left(\hat{h}_s = \hat{h}(s, X_s, DU(s, X_s)) \right)_{s \in [t, T]}$ and $\hat{N} := \left(\hat{\nu}_s = \hat{\nu}(s, X_s, DU(s, X_s)) \right)_{s \in [t, T]}$.

Since the Hamiltonian in (3.18) is strictly concave in h and strictly convex in ν , these candidate policies define the pointwise maximizer and minimizer, respectively. Moreover, we can check that

$$\begin{aligned} & \mathcal{H}^+(s, x, p, M) \\ = & \inf_{\nu \in \mathbb{R}^d} \sup_{h \in \mathbb{R}^m} \left\{ \{b_s + B_s x - \Lambda_s [\theta (\Sigma'_s h - \Xi_s) - \nu]\}' p + \frac{1}{2} \text{tr} (\Lambda_s \Lambda'_s M) - g_1(s, x, h; \theta) + \frac{1}{2\theta} \nu' \nu \right\} \\ =: & \mathcal{H}^-(s, x, p, M), \end{aligned} \quad (3.20)$$

along the same lines as in the Proof of Proposition 2.7. Therefore, the Bellman–Isaacs minimax condition is satisfied, and U is well defined as the value function for the stochastic differential game. In light of the structure of risk-sensitive control problems, the following Lemma and Proposition show, respectively, that the value function U is quadratic in the state, and that the candidate policies are optimal.

Lemma 3.3. *The value function $U(t, x)$ is of the form*

$$U(t, x) = \frac{1}{2} x' Q_t x + q'_t x + k_t, \quad (3.21)$$

where $Q : [t, T] \rightarrow \mathbb{R}^{n \times n}$, $q : [t, T] \rightarrow \mathbb{R}^n$, and $k : [t, T] \rightarrow \mathbb{R}$ solve (2.44), (2.45), and (2.46) respectively.

Sketch of proof. The proof follows along the lines of that of Theorem 2.8. Substitute the candidate policies at (3.19) and the quadratic form (3.21) into (3.17). Reorganize as a quadratic form in x to recover equations (2.44), (2.45), and (2.46) for Q_t , q_t , and k_t .

□

Proposition 3.4 (Optimality of the Candidate Policies). *Let U be the value function defined at (3.16) and having the quadratic expression in (3.21). Let $H^* = (h^*(s, X_s))_{s \in [t, T]} := (\hat{h}(s, X_s, DU(s, X_s)))_{s \in [t, T]}$ and $N^* = (\nu^*(s, X_s))_{s \in [t, T]} := (\hat{\nu}(s, X_s, DU(s, X_s)))_{s \in [t, T]}$, where $\hat{h}(\cdot)$ and $\hat{\nu}(\cdot)$ are the candidate policies given at (3.19). Then, the strategies H^* and N^* are admissible according to Definitions 3.1 and 3.2 respectively and are optimal.*

Sketch of proof. The proof follows along the lines of that of Theorem 2.9. Showing that H^* is admissible also requires showing that $\chi_{[t, T]}^{H^*}$, defined at (3.3) for $h = h^*$, is an exponential martingale. To see this, note that U is quadratic in x by Lemma 3.3, so $\hat{h}(s, x, DU(s, x))$ is affine in x . Moreover, under the candidate strategy H^* , the state process X with \mathbb{P}^H dynamics (3.6) is Gaussian. Since $h^*(s, X_s)$ is affine in the Gaussian state X_s , it is square-integrable on $[t, T]$. Under the standing C^1 assumptions on the deterministic coefficients and the finite horizon, the coefficients are bounded and the Novikov condition holds; therefore, $\chi_{[t, T]}^{H^*}$ is a true martingale. Hence, condition (iii) in Definition 3.1 is satisfied. It remains to show that N^* is admissible in the sense of Definition 3.2, and specifically, that $\chi_{[t, T]}^{N^*}$, defined at (3.9) for $\nu = \nu^*$, is an exponential martingale. This is shown in full analogy to the proof of Theorem 2.9, in particular, where it is shown that $\Gamma^* \in \mathcal{A}^\Gamma$. □

The following theorem shows that the sequential route consisting of the Kuroda–Nagai change of measure followed by the free energy–entropy duality yields exactly the same value function and optimal investment strategy as the direct one-step approach of Sections 2.3.2–2.3.3.

Theorem 3.5.

(i) *The value functions u defined at (2.27) and U defined at (3.16) are related by*

$$u(t, x) = -\theta U(t, x); \quad (3.22)$$

(ii) *The definitions of the asset allocation $\hat{h}(s, x, Du(s, x))$ at (2.35) and $\hat{h}(s, x, DU(s, x))$ at (3.19) are equivalent, that is,*

$$\hat{h}(s, x, Du(s, x)) = \hat{h}(s, x, DU(s, x)) =: \hat{h}(s, x), \quad \forall (s, x) \in [t, T] \times \mathbb{R}^n. \quad (3.23)$$

Hence, the strategy $H^ := (\hat{h}(s, X_s))_{s \in [t, T]}$ is optimal for the original problem (2.27) and the two-step problem (3.16).*

(iii) *The policies $\hat{\gamma}$ defined at (2.36) and $\hat{\nu}$ defined at (3.19) are related by*

$$\hat{\gamma}(s, x, Du(s, x)) = \hat{\nu}(s, x, DU(s, x)) - \theta \left(\Sigma'_s \hat{h}(s, x, DU(s, x)) - \Xi_s \right), \quad \forall (s, x) \in [t, T] \times \mathbb{R}^n. \quad (3.24)$$

Proof.

(i) From Theorem 2.8 and Lemma 3.3,

$$u(t, x) = -\theta \left(\frac{1}{2} x' Q_t x + q_t' x + k_t \right) = -\theta U(t, x) \quad (3.25)$$

where $Q : [t, T] \rightarrow \mathbb{R}^{n \times n}$, $q : [t, T] \rightarrow \mathbb{R}^n$, and $k : [t, T] \rightarrow \mathbb{R}$ solve (2.44), (2.45), and (2.46) respectively.

(ii) From (3.19),

$$\begin{aligned} \hat{h}(s, x, DU(s, x)) &= \frac{1}{\theta + 1} (\Sigma_s \Sigma_s')^{-1} [a_s + A_s x + \theta \Sigma_s \Xi_s - \theta \Sigma_s \Lambda_s' DU(s, x)] \\ &\stackrel{\text{by (3.22)}}{=} \frac{1}{\theta + 1} (\Sigma_s \Sigma_s')^{-1} [a_s + A_s x + \theta \Sigma_s \Xi_s + \Sigma_s \Lambda_s' Du(s, x)], \end{aligned} \quad (3.26)$$

which is (2.35). Therefore, the strategy $H^* := \left(\hat{h}(s, X_s, Du(s, X_s)) \right)_{s \in [t, T]} = \left(\hat{h}(s, X_s, DU(s, X_s)) \right)_{s \in [t, T]}$ is optimal for the original problem (2.27) and the two-step problem (3.16).

(iii) Start from (2.36):

$$\begin{aligned} \hat{\gamma}(s, x, Du(s, x)) &= -\theta \Lambda_s' (Q_s x + q_s) \\ &\quad - \theta \left[\frac{1}{\theta + 1} \Sigma_s' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s - \theta \Sigma_s \Lambda_s' (Q_s x + q_s)) - \Xi_s \right] \\ &\stackrel{\text{by (3.19) and Lemma 3.3}}{=} \hat{\nu}(s, x, DU(s, x)) - \theta \left(\Sigma_s' \hat{h}(s, x, DU(s, x)) - \Xi_s \right). \end{aligned} \quad (3.27)$$

□

3.2 Analysis of the Change of Measure in the Two-Step Approach

The candidate policy $\hat{\gamma}(s, x, Du(s, x))$ defined at (2.36) induces a change of measure from \mathbb{P} to a probability measure $\mathbb{P}^{\hat{\Gamma}}$ via the Radon-Nikodym derivative at (2.13). Proposition 3.6 below decomposes this change of measure into two sequential changes of measure. The first is the Kuroda–Nagai change of measure introduced at (3.4). The second is equivalent to the change of measure (3.9) induced by the duality under \mathbb{P}^H , as shown in Corollary 3.7 below.

Proposition 3.6. *For the pair of candidate strategies $(\hat{H}, \hat{\Gamma})$, the Radon-Nikodym derivative $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T}$ admits the following decomposition:*

$$\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = \frac{d\mathbb{P}^{\hat{H}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T} \times \frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}} \Big|_{\mathcal{F}_T}, \quad (3.28)$$

where (i) $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T}$ is, for $\Gamma = \hat{\Gamma}$, defined at (2.13) and (2.41), (ii) $\frac{d\mathbb{P}^{\hat{H}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T}$ is defined at (3.4) via the Doléans-Dade exponential $\chi_{[t, T]}^H$ at (3.3), with H evaluated at \hat{H} , and (iii) $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}} \Big|_{\mathcal{F}_T}$ is defined as

$$\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}} \Big|_{\mathcal{F}_T} = \exp \left\{ \int_t^T Du(s, X_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\Lambda_s' Du(s, X_s)\|^2 ds \right\}. \quad (3.29)$$

Proof. Start from the definition of $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}}|_{\mathcal{F}_T}$ at (2.41) and use the compact notation $\hat{h}_s := \hat{h}(s, X_s, DU(s, X_s))$, $s \in [t, T]$ introduced in Theorem 3.5 (ii):

$$\begin{aligned}
& \frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T} \\
&= \exp \left\{ -\frac{1}{2}\theta^2 \int_t^T \|\Sigma'_s \hat{h}_s - \Xi_s\|^2 ds - \frac{1}{2} \int_t^T \|\Lambda'_s Du(s, X_s)\|^2 ds + \theta \int_t^T Du(s, X_s)' \Lambda_s (\Sigma'_s \hat{h}_s - \Xi_s) ds \right. \\
&\quad \left. + \int_t^T Du(s, X_s)' \Lambda_s dW_s - \theta \int_t^T (\hat{h}'(s) \Sigma_s - \Xi'_s) dW_s \right\} \\
&= \exp \left\{ -\theta \int_t^T (\hat{h}'(s) \Sigma_s - \Xi'_s) dW_s - \frac{1}{2}\theta^2 \int_t^T \|\Sigma'_s \hat{h}_s - \Xi_s\|^2 ds \right. \\
&\quad \left. + \int_t^T Du(s, X_s)' \Lambda_s \underbrace{\left[\theta (\Sigma'_s \hat{h}_s - \Xi_s) ds + dW_s \right]}_{=dW_s^{\hat{H}} \text{ by (3.5)}} \right\} - \frac{1}{2} \int_t^T \|\Lambda'_s Du(s, X_s)\|^2 ds \\
&= \frac{d\mathbb{P}^{\hat{H}}}{d\mathbb{P}} \Big|_{\mathcal{F}_T} \times \exp \left\{ \int_t^T Du(s, X_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\Lambda'_s Du(s, X_s)\|^2 ds \right\} \tag{3.30}
\end{aligned}$$

Since $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}}|_{\mathcal{F}_T} = \frac{d\mathbb{P}^{\hat{H}}}{d\mathbb{P}}|_{\mathcal{F}_T} \times \frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}}|_{\mathcal{F}_T}$, we conclude that $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}}|_{\mathcal{F}_T} = \exp \left\{ \int_t^T Du(s, X_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\Lambda'_s Du(s, X_s)\|^2 ds \right\}$. □

Corollary 3.7. *At the pair of candidate strategies $(\hat{H}, \hat{\Gamma})$ defined via (2.35) and (2.37), the change of measure induced by the Radon-Nikodym derivatives $\frac{d\mathbb{P}^{\hat{N}}}{d\mathbb{P}^{\hat{H}}}|_{\mathcal{F}_T}$, defined according to (3.9), and $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}}|_{\mathcal{F}_T}$, given at (3.29), induce the same probability measure on (Ω, \mathcal{F}_T) , that is $\mathbb{P}^{\hat{N}} = \mathbb{P}^{\hat{\Gamma}}$.*

Proof. Start from (3.29):

$$\begin{aligned}
\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}}|_{\mathcal{F}_T} &= \exp \left\{ \int_t^T Du(s, X_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\Lambda'_s Du(s, X_s)\|^2 ds \right\} \\
&\stackrel{\text{by Theo. 2.8}}{=} \exp \left\{ \int_t^T -\theta(Q_s X_s + q_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\theta \Lambda'_s(Q_s X_s + q_s)\|^2 ds \right\} \\
&\stackrel{\text{by (3.21)}}{=} \exp \left\{ \int_t^T -\theta DU(s, X_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\theta \Lambda'_s DU(s, X_s)\|^2 ds \right\} \\
&\stackrel{\text{by (3.19)}}{=} \exp \left\{ \int_t^T \hat{\nu}(s, X_s, DU(s, X_s))' dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\hat{\nu}(s, X_s, DU(s, X_s))\|^2 ds \right\}. \tag{3.31}
\end{aligned}$$

The conclusion follows from the definition of $\frac{d\mathbb{P}^{\hat{N}}}{d\mathbb{P}^{\hat{H}}}|_{\mathcal{F}_T}$ via (3.9) evaluated at $N = \hat{N}$. □

The first change of measure, via the Radon-Nikodym derivative $\frac{d\mathbb{P}^{\hat{H}}}{d\mathbb{P}}|_{\mathcal{F}_T}$ is used to rewrite the non-standard risk-sensitive benchmarked investment management problem into an LEQG problem, as shown in Section 3.1.1.

The second change of measure, via the Radon-Nikodym derivative $\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}}\Big|_{\mathcal{F}_T} = \frac{d\mathbb{P}^{\hat{N}}}{d\mathbb{P}^{\hat{H}}}\Big|_{\mathcal{F}_T}$, is specific to the free energy–entropy duality, which recasts the LEQG control problem under the probability measure $\mathbb{P}^{\hat{H}}$ into an LQG game under the probability measure $\mathbb{P}^{\hat{\Gamma}}$. From a PDE perspective, plugging $\hat{\nu}_s$ into the Bellman–Isaacs equation (3.17) yields the quadratic gradient term $-\frac{\theta}{2}DU(s, x)' \Lambda_s \Lambda_s' DU(s, x)$, which characterizes the HJB PDE for the LEQG control problem (see the last paragraph of Section 2.4.2 in Davis & Lleo, 2014). Consequently, at $\nu_s = \hat{\nu}_s$, the Bellman–Isaacs equation for the stochastic differential game reduces to the HJB PDE for the LEQG stochastic control problem. This interpretation also explains the observation in Remark 8.

Performing the two changes of measure sequentially leads to a surprisingly simple situation. The terms inside the sup inf of the Hamiltonian \mathcal{H}^+ at (3.18) are still quadratic in the controls, but without interaction, that is, without a term of the form $h'M\nu$ for some matrix M . Why did we get this result? The control h_s enters the dynamics of the state process X through the first change of measure from \mathbb{P} to \mathbb{P}^H . So, h_s only affects the drift of X , not its diffusion. Consequently, in the second change of measure, the additional drift component depends only on ν_s , so no bilinear interaction term between h_s and ν_s appears in the Hamiltonian.

Corollary 3.8 (Special Case of the Kelly Portfolio). *Taking the limit as $\theta \rightarrow 0$ yields the Kelly or growth-optimal portfolio. In this case, notice that the probability measure $\mathbb{P}^{\hat{H}}$ is identical to the probability measure \mathbb{P} , but the Radon–Nikodym derivative defining $\mathbb{P}^{\hat{\Gamma}}$ remains of the same form. Hence,*

$$\frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}^{\hat{H}}}\Big|_{\mathcal{F}_T} = \exp \left\{ \int_t^T Du(s, X_s)' \Lambda_s dW_s^{\hat{H}} - \frac{1}{2} \int_t^T \|\Lambda_s' Du(s, X_s)\|^2 ds \right\} = \frac{d\mathbb{P}^{\hat{\Gamma}}}{d\mathbb{P}}\Big|_{\mathcal{F}_T}, \quad (3.32)$$

consistently with Remark 7.

4 Implementation

This section provides a simple computational illustration of the model using market data. We consider a U.S. equity portfolio manager with a five-year investment horizon and risk sensitivity parameter $\theta = 1$. The benchmark allocates 90% to the S&P 500, which comprises the 500 largest publicly traded firms by market capitalization, and 10% to the S&P 400, which comprises 400 mid-capitalization firms. The manager uses a six-factor model consisting of the five Fama–French factors (Fama & French, 2015) together with Momentum (Carhart, 1997).

The investment universe contains $m = 13$ U.S. exchange-traded funds (ETFs). The first 11 ETFs track the S&P 500 sectors defined by the Global Industry Classification Standard (GICS), so the S&P 500 can be replicated by holding these ETFs at their sector weights. The twelfth ETF is the iShares Core S&P 400 Mid-Cap ETF (ticker: IJH), which tracks 400 mid-cap stocks. The thirteenth ETF is the iShares Core S&P 600 Small-Cap ETF (ticker: IJR), which tracks the S&P 600 Small-Cap Index. Since the S&P 600 is not included in the benchmark, any allocation to small-capitalization stocks is purely tactical.

Model parameters are estimated from daily returns over June 20, 2018, to December 31, 2024, giving $\mathcal{T} = 1644$ observations and $\Delta t := \frac{1}{252}$. June 20, 2018, is the earliest date for which

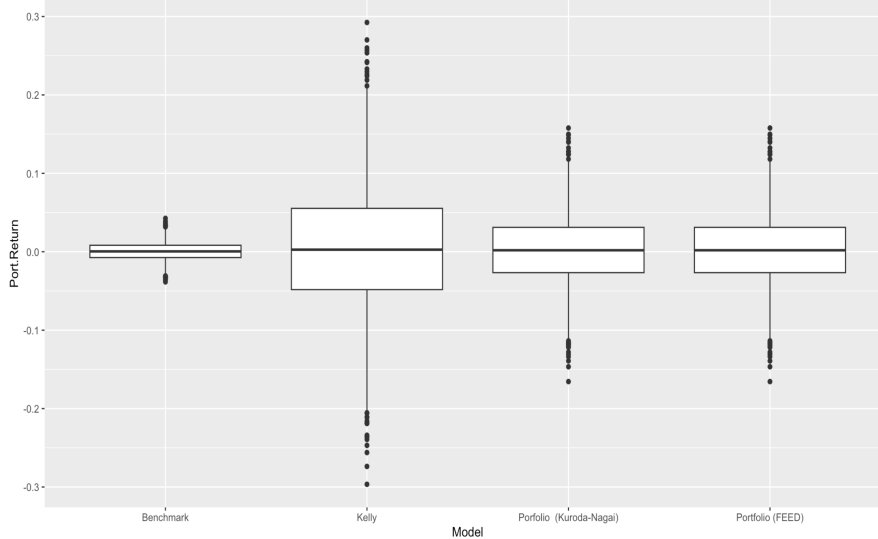


Figure 1: Box-and-whisker plot of the distribution of daily log excess returns over the money market rate for the benchmark, the Kelly portfolio, the portfolio obtained via the Kuroda–Nagai change-of-measure approach, and the portfolio obtained via the Free Energy–Entropy Duality.

all series are available, because a GICS revision changed sector definitions and composition. The sample therefore spans a volatile period in the U.S. equity market, including the rise in 2019, the COVID-19 crash in March 2020, and the subsequent rally. Factor data and money market returns are obtained from Kenneth French’s data library⁴, while ETF returns are taken from the CRSP database⁵.

Because the objective of this section is illustrative rather than predictive, we keep the implementation deliberately simple and assume constant model parameters throughout. Drift coefficients are estimated by applying standard regression methods to a discretized version of the SDE for X_t . Diffusion matrices are estimated from quadratic and cross variations using resampling methods. Here, the benchmark is itself a combination of the m traded assets, so L_t is modeled as a fixed-weight portfolio corresponding to the benchmark weights.

Using these estimates, we run 5,000 simulations over 1,260 daily time steps. This yields terminal outcomes for the portfolio and the benchmark, together with 6,300,000 simulated daily log excess returns over the money market rate, which are used to estimate return distributions and performance measures.

Figure 1 reports the distribution of daily log excess returns over the money market rate for the benchmark, the Kelly portfolio, the portfolio obtained via the Kuroda–Nagai change-of-measure approach, and the portfolio obtained via the Free Energy–Entropy Duality. Table 1 reports summary statistics, tail-risk measures, and risk-adjusted performance measures for the same portfolios. Consistent with the equivalence results established in Section 3, the Free Energy–Entropy Duality and the Kuroda–Nagai change-of-measure approach yield identical simulated performance.

⁴The data are available online at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html.

⁵©2026 Center for Research in Security Prices (CRSP), LLC.

	Benchmark	Portfolio (KN)	Portfolio (FEED)	Kelly
Mean	0.0507 %	0.2437 %	0.2437 %	0.3078 %
Standard deviation	1.1682 %	4.3154 %	4.3154 %	7.8217 %
Semideviation	0.8569 %	3.0500 %	3.0500 %	5.5284 %
Skewness	-0.5895	0.0035	0.0035	0.0030
Kurtosis	17.4858	0.0401	0.0401	0.0450
VaR 95%	1.6773 %	7.0890 %	7.0890 %	12.8507 %
CVaR 95%	2.8334 %	8.9186 %	8.9186 %	16.1727 %
Sharpe	0.0434	0.0565	0.0565	0.0394
Sortino	0.0607	0.0836	0.0836	0.0575
Mean-to-VaR 95%	0.0302	0.0344	0.0344	0.0240
Mean-to-CVaR 95%	0.0179	0.0273	0.0273	0.0190

Table 1: Performance measures for daily log excess returns over the money market rate for the benchmark, the portfolio obtained via the Free Energy–Entropy Duality, the portfolio obtained via the Kuroda–Nagai change-of-measure approach, and the Kelly portfolio. The reported measures include summary statistics (mean, standard deviation, semideviation, skewness, and kurtosis), tail-risk measures (VaR and CVaR at the 95% confidence level, expressed as percentage losses relative to the mean return), and risk-adjusted performance measures (Sharpe ratio, Sortino ratio, mean-to-VaR, and mean-to-CVaR).

5 Conclusion

This paper develops a new formulation for the solution of risk-sensitive benchmarked investment management through the free energy–entropy duality. The new formulation yields a direct, explicit solution to the benchmarked risk-sensitive portfolio problem, embedding the classical Kuroda–Nagai change-of-measure approach within a more general framework. Specifically, it transforms a non-standard risk-sensitive control problem into a LQG game problem under a suitable equivalent probability measure, with an entropic regularization. When the model parameters are known or can be estimated with high accuracy, this leads to an implementation that is both tractable and implementable. In a companion paper, we relax the assumption of known model parameters and show that the present framework is naturally suited to a reinforcement learning implementation.

References

- Bensoussan, A. (1992). *Stochastic Control of Partially Observable Systems*. Cambridge University Press.
- Bielecki, T. & Pliska, S. (1999). Risk-sensitive dynamic asset management. *Applied Mathematics and Optimization*, 39, 337–360.
- Bielecki, T. & Pliska, S. (2000). Risk sensitive asset management with transaction costs. *Finance and Stochastics*, 4, 1–33.

- Bielecki, T. & Pliska, S. (2003). Economic properties of the risk sensitive criterion for portfolio management. *The Review of Accounting and Finance*, 2(2), 3–17.
- Bielecki, T. & Pliska, S. (2004). Risk sensitive intertemporal CAPM. *IEEE Transactions on Automatic Control*, 49(3), 420–432.
- Bielecki, T. R., Chen, T., & Cialenco, I. (2022). Risk-Sensitive Markov Decision Problems under Model Uncertainty: Finite Time Horizon Case. In G. Yin & T. Zariphopoulou (Eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization: A Commemorative Volume to Honor Mark H. A. Davis's Contributions* (pp. 33–52). Cham: Springer International Publishing.
- Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57–82.
- Dai Pra, P., Meneghini, L., & Runggaldier, W. J. (1996). Connections between stochastic control and dynamic games. *Mathematics of Control, Signals and Systems*, 9(4), 303–326.
- Davis, M. & Lleo, S. (2008). Risk-sensitive benchmarked asset management. *Quantitative Finance*, 8(4), 415–426.
- Davis, M. & Lleo, S. (2011). Jump-diffusion risk-sensitive asset management I: Diffusion factor model. *SIAM Journal on Financial Mathematics*, 2, 22–54.
- Davis, M. & Lleo, S. (2013a). Black-Litterman in continuous time: The case for filtering. *Quantitative Finance Letters*, 1(1), 30–35.
- Davis, M. & Lleo, S. (2013b). Jump-diffusion risk-sensitive asset management ii: Jump-diffusion factor model. *SIAM Journal on Control and Optimization*, 51(2), 1441.
- Davis, M. & Lleo, S. (2013c). Jump-diffusion risk-sensitive benchmarked asset management. In H. Gassmann & W. Ziemba (Eds.), *Stochastic Programming: Applications in Finance, Energy, Planning and Logistics* (pp. 97–128).: World Scientific Publishing.
- Davis, M. & Lleo, S. (2014). *Risk-Sensitive Investment Management*, volume 19 of *Advanced Series on Statistical Science and Applied Probability*. World Scientific Publishing.
- Davis, M. & Lleo, S. (2015). Jump-diffusion asset-liability management via risk-sensitive control. *OR Spectrum*, 37(3), 655–675.
- Davis, M. & Lleo, S. (2021). Risk-sensitive benchmarked asset management with expert forecasts. *Mathematical Finance*, 31(4), 1162–1189.
- Davis, M. & Lleo, S. (2024). Jump-diffusion risk-sensitive benchmarked asset management with traditional and alternative data. *Annals of Operations Research*, 336, 661–689.
- Fama, E. & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22.

- Fleming, W. & Soner, H. (2006). *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modeling and Applied Probability*. Springer-Verlag, 2nd edition.
- Hata, H. (2017). Risk-sensitive asset management in a general diffusion factor model: Risk-seeking case. *Japan Journal of Industrial and Applied Mathematics*, 34(1), 59–98.
- Hata, H. (2018). Risk-sensitive portfolio optimization problem for a large trader with inside information. *Japan Journal of Industrial and Applied Mathematics*, 35(3), 1037–1063.
- Hata, H. (2021). Risk-Sensitive Asset Management with Lognormal Interest Rates. *Asia-Pacific Financial Markets*, 28(2), 169–206.
- Hata, H. & Iida, Y. (2006). A risk-sensitive stochastic control approach to an optimal investment problem with partial information. *Finance and Stochastics*, 10(3), 395–426.
- Kuroda, K. & Nagai, H. (2002). Risk-sensitive portfolio optimization on infinite time horizon. *Stochastics and Stochastics Reports*, 73, 309–331.
- Liptser, R. & Shiryaev, A. (2004). *Statistics of Random Processes: I. General Theory*. Probability and Its Applications. Springer-Verlag, 2 edition.
- Lleo, S. & MacLean, L. C. (2025). Dual dominance: how Harry Markowitz and William Ziemba impacted portfolio management. *Annals of Operations Research*, 346(1), 181–216.
- Lleo, S. & Runggaldier, W. (2024). On the separation of estimation and control in risk-sensitive investment problems under incomplete observation. *European Journal of Operational Research*, 316(1), 200–214.
- Lleo, S. & Runggaldier, W. (2026). Reinforcement learning for risk-sensitive investment management: A free energy–entropy duality approach. Preprint.
- Meneghini, L. (1994). *Modelli Risolvibili per Problemi di Controllo di Sistemi Dinamici Imprecisi Multivariati*. PhD thesis, Università Degli Studi Di Padova.
- Nagai, H. & Peng, S. (2002). Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon. *The Annals of Applied Probability*, 12(1), 173–195.

A Proofs

A.1 Recalling Lemma 5.3.1 from Meneghini (1994) along with its proof

Lemma A.1. *Let $f : X \rightarrow \mathbb{R}^+$ be a function such that $\inf_{x \in X} f(x) > 0$. Then*

$$\log \left(\inf_{x \in X} f(x) \right) = \inf_{x \in X} \log(f(x))$$

Proof. i) First we prove the one-sided inequality (\leq)

For any $\bar{x} \in X$ we have $\inf_{x \in X} f(x) \leq f(\bar{x})$ which, by the monotonicity of $\log(\cdot)$, implies

$$\log \left(\inf_{x \in X} f(x) \right) \leq \log (f(\bar{x})) \quad \text{and, with it,} \quad \log \left(\inf_{x \in X} f(x) \right) \leq \inf_{x \in X} \log (f(x))$$

ii) Inverse inequality (\geq)

By the definition of \inf , for each $n \in \mathbb{N}$ there exists $x_n \in X$ such that

$$\inf_{x \in X} f(x) \leq f(x_n) \leq \inf_{x \in X} f(x) + \frac{1}{n}$$

By the monotonicity of $\log(\cdot)$ it then follows

$$\log f(x_n) \leq \log \left(\inf_{x \in X} f(x) + \frac{1}{n} \right) \quad \text{and, with it,} \quad \inf_{x \in X} \log f(x) \leq \log \left(\inf_{x \in X} f(x) + \frac{1}{n} \right)$$

Passing to the limit when $n \rightarrow \infty$ and thereby using the continuity of $\log(\cdot)$,

$$\inf_{x \in X} \log f(x) \leq \log \left(\inf_{x \in X} f(x) \right)$$

□

A.2 Proof of Proposition 2.7

We express the Bellman–Isaacs equation (2.29) as

$$\begin{aligned} & \frac{\partial u(s, x)}{\partial s} + \inf_{h \in \mathbb{R}^m} \sup_{\gamma \in \mathbb{R}^d} \left\{ [b_s + B_s x + \Lambda_s \gamma]' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' D^2 u(s, x)) + \theta g(s, x, h, \gamma; \theta) \right\} \\ &= \frac{\partial u(s, x)}{\partial s} + \left[(b_s + B_s x)' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' D^2 u(s, x)) + \theta (c_s + C_s x) - \frac{\theta}{2} \Xi_s' \Xi_s \right] \\ &+ \inf_{h \in \mathbb{R}^m} \sup_{\gamma \in \mathbb{R}^d} \theta F(s, x, h, \gamma; \theta, Du(s, x)), \end{aligned} \quad (\text{A.1})$$

where

$$F(s, x, h, \gamma; \theta, p) := \frac{1}{2} h' \Sigma_s \Sigma_s' h - h' (a_s + A_s x) - \frac{1}{2\theta} \gamma' \gamma - \gamma' (\Sigma_s' h - \Xi_s) + \frac{1}{\theta} \gamma' \Lambda_s' p. \quad (\text{A.2})$$

To derive the two equivalent representations of the candidate policies, we optimize in the two possible orders over h and γ .

For the first representation, (2.35)–(2.36), we first solve the supremum problem in γ . The function F is quadratic in γ . Its unique maximizer corresponds to the candidate policy

$$\hat{\gamma}(s, x, h, Du(s, x)) = \Lambda_s' Du(s, x) - \theta (\Sigma_s' h - \Xi_s). \quad (\text{A.3})$$

Substituting into F , we get

$$\begin{aligned}
& F(s, x, h, \hat{\gamma}(s, x, h, Du(s, x)); \theta, Du(s, x)) \\
&= \frac{\theta+1}{2} h' \Sigma_s \Sigma_s' h - h' (a_s + A_s x + \Sigma_s \Lambda_s' Du(s, x) + \theta \Sigma_s \Xi_s) + \frac{\theta}{2} \Xi_s' \Xi_s + \Xi_s' \Lambda_s' Du(s, x) \\
&\quad + \frac{1}{2\theta} Du(s, x)' \Lambda_s \Lambda_s' Du(s, x). \\
&= \frac{\theta+1}{2} \left[h - \frac{1}{\theta+1} (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s + \Sigma_s \Lambda_s' Du(s, x)) \right]' (\Sigma_s \Sigma_s') \\
&\quad \cdot \left[h - \frac{1}{\theta+1} (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s + \Sigma_s \Lambda_s' Du(s, x)) \right] \\
&\quad - \frac{1}{2(\theta+1)} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s) \\
&\quad - \frac{1}{\theta+1} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' Du(s, x) \\
&\quad + \frac{\theta}{2} \Xi_s' \Xi_s + \Xi_s' \Lambda_s' Du(s, x) + \frac{1}{2\theta} Du'(s, x) \Lambda_s \mathcal{P}_s^-(\theta) \Lambda_s' Du(s, x)
\end{aligned} \tag{A.4}$$

Next, we solve the infimum problem in h , given the maximizer in γ obtained above. The function F is quadratic in h . Its unique minimizer corresponds to the candidate policy

$$\hat{h}(s, x, Du(s, x)) = \frac{1}{\theta+1} (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s + \Sigma_s \Lambda_s' Du(s, x)). \tag{A.5}$$

Hence, at the minimizer \hat{h} , F takes the value

$$\begin{aligned}
& F\left(s, x, \hat{h}(s, x, Du(s, x)), \hat{\gamma}(s, x, Du(s, x)); \theta, Du(s, x)\right) \\
&= -\frac{1}{2(\theta+1)} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s) \\
&\quad - \frac{1}{\theta+1} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' Du(s, x) \\
&\quad + \frac{\theta}{2} \Xi_s' \Xi_s + \Xi_s' \Lambda_s' Du(s, x) + \frac{1}{2\theta} Du'(s, x) \Lambda_s \mathcal{P}_s^-(\theta) \Lambda_s' Du(s, x)
\end{aligned} \tag{A.6}$$

Finally, we substitute (A.5) into (A.3) to get

$$\begin{aligned}
\hat{\gamma}(s, x, Du(s, x)) &= \Lambda_s' Du(s, x) - \theta \left[\frac{1}{\theta+1} \Sigma_s' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s + \Sigma_s \Lambda_s' Du(s, x)) - \Xi_s \right] \\
&= \mathcal{P}_s^-(\theta) \Lambda_s' Du(s, x) - \frac{\theta}{\theta+1} \Sigma_s' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x) + \theta \mathcal{P}_s^-(\theta) \Xi_s.
\end{aligned} \tag{A.7}$$

Next, we reverse the order of optimization and begin with the infimum problem in h . The function F is quadratic in h . Its unique minimizer corresponds to the candidate policy

$$\hat{h}(s, x, \gamma) = (\Sigma_s \Sigma_s')^{-1} [(a_s + A_s x) + \Sigma_s \gamma]. \tag{A.8}$$

Substituting into F , we get

$$\begin{aligned}
& F\left(s, x, \hat{h}(s, x, \gamma), \gamma; \theta, Du(s, x)\right) \\
&= \frac{1}{2} [(a_s + A_s x) + \Sigma_s \gamma]' \underbrace{(\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1}}_{I_m} \\
&\quad \cdot [(a_s + A_s x) + \Sigma_s \gamma] - (a_s + A_s x)' (\Sigma_s \Sigma'_s)^{-1} [(a_s + A_s x) + \Sigma_s \gamma] \\
&\quad - \frac{1}{2\theta} \gamma' \gamma - \gamma' \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} [(a_s + A_s x) + \Sigma_s \gamma] + \gamma' \Xi_s + \frac{1}{\theta} \gamma' \Lambda'_s Du(s, x) \\
&= -\frac{1}{2\theta} \left\{ \gamma - [\mathcal{P}_s^+(\theta)]^{-1} \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right] \right\}' \\
&\quad \mathcal{P}_s^+(\theta) \left\{ \gamma - [\mathcal{P}_s^+(\theta)]^{-1} \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right] \right\} \\
&\quad + \frac{1}{2\theta} \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right]' \\
&\quad [\mathcal{P}_s^+(\theta)]^{-1} \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right] - \frac{1}{2} (a_s + A_s x)' (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x)
\end{aligned} \tag{A.9}$$

We then solve the supremum for γ , given the value we found for \hat{h} . The function F is quadratic and strictly concave in γ . Its unique maximizer corresponds to the candidate policy

$$\hat{\gamma}(s, x, Du(s, x)) = [\mathcal{P}_s^+(\theta)]^{-1} \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right] \tag{A.10}$$

Substituting the maximizer (A.10) into (A.8) recovers (2.38), and hence, at the maximizer $\hat{\gamma}$, F takes the value

$$\begin{aligned}
& F\left(s, x, \hat{h}(s, x, Du(s, x)), \hat{\gamma}(s, x, Du(s, x)); \theta, Du(s, x)\right) \\
&= \frac{1}{2\theta} \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right]' [\mathcal{P}_s^+(\theta)]^{-1} \\
&\quad \cdot \left[-\theta \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x) + \theta \Xi_s + \Lambda'_s Du(s, x) \right] - \frac{1}{2} (a_s + A_s x)' (\Sigma_s \Sigma'_s)^{-1} (a_s + A_s x)
\end{aligned} \tag{A.11}$$

The final step is to apply (2.34) to (A.11) and to develop the expression to obtain (A.6).

Thus, the same saddle value is obtained under either order of optimization. Since F at (A.6) and (A.11) are equal, we conclude that \mathcal{H}^+ given at (2.30) and \mathcal{H}^- given at (2.31) are equal, $\mathcal{H}^+(s, x, p, M) = \mathcal{H}^-(s, x, p, M)$, so the stochastic differential game is well defined.

A.3 Proof of Theorem 2.8

With (A.6), the Bellman–Isaacs equation at (A.1) becomes

$$\begin{aligned} & \frac{\partial u(s, x)}{\partial s} + \left[(b_s + B_s x)' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' D^2 u(s, x)) + \theta (c_s + C_s x) - \frac{\theta}{2} \Xi_s' \Xi_s \right] \\ & + \theta \left[-\frac{1}{2(\theta + 1)} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s) \right. \\ & - \frac{1}{\theta + 1} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' Du(s, x) \\ & \left. + \frac{\theta}{2} \Xi_s' \Xi_s + \Xi_s' \Lambda_s' Du(s, x) + \frac{1}{2\theta} Du'(s, x) \Lambda_s \mathcal{P}_s^-(\theta) \Lambda_s' Du(s, x) \right] = 0. \end{aligned} \quad (\text{A.12})$$

Next, substitute the derivatives of the ansatz (2.43), namely

$$Du(s, x) = -\theta(Q_s x + q_s), \quad D^2 u(s, x) = -\theta Q_s,$$

to obtain

$$\begin{aligned} & -\theta \left(\frac{1}{2} x' \dot{Q}_s x + x' \dot{q}_s + \dot{k}_s \right) + \left[-\theta (b_s + B_s x)' (Q_s x + q_s) - \frac{\theta}{2} \text{tr} (\Lambda_s \Lambda_s' Q_s) + \theta (c_s + C_s x) - \frac{\theta}{2} \Xi_s' \Xi_s \right] \\ & + \theta \left[-\frac{1}{2(\theta + 1)} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} (a_s + A_s x + \theta \Sigma_s \Xi_s) \right. \\ & + \frac{\theta}{\theta + 1} (a_s + A_s x + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' (Q_s x + q_s) \\ & \left. + \frac{\theta}{2} \Xi_s' \Xi_s - \theta \Xi_s' \Lambda_s' (Q_s x + q_s) + \frac{\theta}{2} (Q_s x + q_s)' \Lambda_s \mathcal{P}_s^-(\theta) \Lambda_s' (Q_s x + q_s) \right] = 0. \end{aligned} \quad (\text{A.13})$$

where we used the notation $\dot{Q}_s = \frac{dQ}{ds}$, $\dot{q}_s = \frac{dq}{ds}$, $\dot{k}_s = \frac{dk}{ds}$.

Divide by $-\theta$, rearrange, and factor into a quadratic form in x :

$$\begin{aligned} & \frac{1}{2} x' \left\{ \dot{Q}_s + 2B_s' Q_s + \frac{1}{(\theta + 1)} A_s' (\Sigma_s \Sigma_s')^{-1} A_s - \frac{2\theta}{(\theta + 1)} A_s' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' Q_s - \theta Q_s' \Lambda_s \mathcal{P}_s^-(\theta) \Lambda_s' Q_s \right\} x \\ & + x' \left\{ \dot{q}_s + Q_s' b_s + B_s' q_s - C_s' + \frac{1}{\theta + 1} A_s' (\Sigma_s \Sigma_s')^{-1} (a_s + \theta \Sigma_s \Xi_s) - \frac{\theta}{\theta + 1} A_s' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' q_s \right. \\ & \left. - \theta Q_s' \Lambda_s \mathcal{P}_s^-(\theta) \Lambda_s' q_s - \frac{\theta}{\theta + 1} Q_s' \Lambda_s \Sigma_s' (\Sigma_s \Sigma_s')^{-1} (a_s + \theta \Sigma_s \Xi_s) + \theta Q_s' \Lambda_s \Xi_s \right\} \\ & + \left\{ \dot{k}_s + b_s' q_s - \frac{\theta}{2} q_s' \Lambda_s \Lambda_s' q_s + \theta \Xi_s' \Lambda_s' q_s + \frac{1}{2} \text{tr} (\Lambda_s \Lambda_s' Q_s) - c_s - \frac{\theta - 1}{2} \Xi_s' \Xi_s \right. \\ & \left. + \frac{1}{2(\theta + 1)} (a_s + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} (a_s + \theta \Sigma_s \Xi_s) + \frac{\theta^2}{2(\theta + 1)} q_s' \Lambda_s \Sigma_s' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' q_s \right. \\ & \left. - \frac{\theta}{\theta + 1} (a_s + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma_s')^{-1} \Sigma_s \Lambda_s' q_s \right\} = 0. \end{aligned} \quad (\text{A.14})$$

Equation (A.14) holds if and only if the coefficients of the quadratic term in x , of the linear

term in x , and the deterministic term all equal 0. We now address each of these terms to express them into a more convenient form, starting with the quadratic term in x :

$$\begin{aligned}
& \dot{Q}_s + 2B'_s Q_s + \frac{1}{(\theta+1)} A'_s (\Sigma_s \Sigma'_s)^{-1} A_s - \frac{2\theta}{(\theta+1)} A'_s (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s Q_s - \theta Q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s Q_s = 0 \\
\Leftrightarrow & \dot{Q}_s - \theta Q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s Q_s + \left(B'_s - \frac{\theta}{(\theta+1)} A'_s (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s \right) Q_s \\
& + Q'_s \left(B_s - \frac{\theta}{(\theta+1)} \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} A_s \right) + \frac{1}{(\theta+1)} A'_s (\Sigma_s \Sigma'_s)^{-1} A_s = 0. \tag{A.15}
\end{aligned}$$

This final expression is a Riccati equation for the function Q_s .

Moving on to the linear term in x , we get

$$\begin{aligned}
& \dot{q}_s + Q'_s b_s + B'_s q_s - C'_s + \frac{1}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) - \frac{\theta}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s q_s \\
& - \theta Q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s - \frac{\theta}{\theta+1} Q'_s \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) + \theta Q'_s \Lambda_s \Xi_s = 0 \\
\Leftrightarrow & \dot{q}_s + \left(B'_s - \frac{\theta}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s \right) q_s - \theta Q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s \\
& + Q'_s \left[b_s - \frac{\theta}{\theta+1} \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) + \theta \Lambda_s \Xi_s \right] \\
& - C'_s + \frac{1}{\theta+1} A'_s (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) = 0, \tag{A.16}
\end{aligned}$$

which is a linear ODE for q_s .

Finally,

$$\begin{aligned}
k_t = \int_t^T & \left\{ \frac{\theta}{2} q'_s \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s q_s - \left[b'_s + \theta \Xi'_s \Lambda'_s - \frac{\theta}{\theta+1} (a_s + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma'_s)^{-1} \Sigma_s \Lambda'_s \right] q_s \right. \\
& \left. - \frac{1}{2} \text{tr} (\Lambda_s \Lambda'_s Q_s) - \frac{1}{2(\theta+1)} (a_s + \theta \Sigma_s \Xi_s)' (\Sigma_s \Sigma'_s)^{-1} (a_s + \theta \Sigma_s \Xi_s) + c_s + \frac{\theta-1}{2} \Xi'_s \Xi_s \right\} ds. \tag{A.17}
\end{aligned}$$

This yields (2.44), (2.45), and (2.46), and hence proves the theorem.

A.4 Proof of Theorem 2.9

1. We start by establishing that the candidate strategies H^* and Γ^* from Proposition 2.7 are Borel-measurable and Markov. For the value function $u(t, x)$ in Theorem 2.8, the candidate strategies H^* and Γ^* induced by (2.35) and (2.37) simplify to $H^* = (h^*(s, X_s))_{s \in [t, T]}$ and $\Gamma^* = (\gamma^*(s, X_s))_{s \in [t, T]}$ with

$$h^*(s, x) = \frac{1}{\theta+1} (\Sigma_s \Sigma'_s)^{-1} [(A_s - \theta \Sigma_s \Lambda'_s Q_s) x + a_s + \theta \Sigma_s (\Xi_s - \Lambda'_s q_s)] \tag{A.18}$$

$$\gamma^*(s, x) = \left\{ -\frac{\theta}{\theta+1} \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} A_s - \theta \mathcal{P}_s^-(\theta) \Lambda'_s Q_s \right\} x - \frac{\theta}{\theta+1} \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} a_s + \theta \mathcal{P}_s^-(\theta) (\Xi_s - \Lambda'_s q_s). \tag{A.19}$$

Therefore, the maps defining $h^*(s, X_s)$ at (A.18) and $\gamma^*(s, X_s)$ at (A.19) are Borel-measurable. Therefore, the defining policies $h^*(\cdot)$ and $\gamma^*(\cdot)$ are Borel-measurable, and the induced strategies H^* and Γ^* are Markov.

Moreover, under \mathbb{P}^{Γ^*} , the state process $X = (X_s)_{s \in [t, T]}$ remains Gaussian. To see this, substitute (A.19) into (2.14) to obtain the following dynamics:

$$dX_s = \left\{ \left[b_s - \frac{\theta}{\theta+1} \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} a_s + \theta \Lambda_s \mathcal{P}_s^-(\theta) (\Xi_s - \Lambda'_s q_s) \right] + \left[B_s - \frac{\theta}{\theta+1} \Lambda_s \Sigma'_s (\Sigma_s \Sigma'_s)^{-1} A_s - \theta \Lambda_s \mathcal{P}_s^-(\theta) \Lambda'_s Q_s \right] X_s \right\} ds + \Lambda_s dW_s^{\Gamma^*}, \quad (\text{A.20})$$

From (A.18) and (A.20), by which $h^*(s, x)$ is affine in x and X is a Gaussian process, we can conclude that the strategy H^* is a Markov strategy in class \mathcal{A}^H . Since $h^*(s, X_s)$ is affine in the Gaussian state X_s , it is square-integrable on $[t, T]$. Hence $H^* \in \mathcal{A}^H$. To show that Γ^* is Markov and in class \mathcal{A}^Γ recall that, by Theorem 2.8, $u(t, x)$ is quadratic in x and so $\hat{\gamma}(s, x, Du(s, x))$ in (2.36) of Proposition 2.7 is affine in x . Moreover, under \mathbb{P}^{Γ^*} , the state process X given by (A.20) is Gaussian.

For convenience, define the density increment process associated with Γ^* on $[t, T]$ by

$$\chi_{t,s}^{\Gamma^*} := \exp \left\{ -\frac{1}{2} \int_t^s \|\gamma^*(u, X_u)\|^2 du + \int_t^s \gamma^*(u, X_u)' dW_u \right\}, \quad s \in [t, T]. \quad (\text{A.21})$$

Equivalently, $\chi_{t,s}^{\Gamma^*} = \chi_s^{\Gamma^*} / \chi_t^{\Gamma^*}$, where χ^{Γ^*} is defined in (2.13).

Since $\gamma^*(s, X_s)$ is affine in the Gaussian state X_s , it is square-integrable on $[t, T]$. Under the standing C^1 assumptions on the deterministic coefficients and the finite time horizon, the coefficients are bounded and the Novikov condition is satisfied. Therefore, the process $(\chi_{t,s}^{\Gamma^*})_{s \in [t, T]}$ is a true martingale. Hence, condition (iii) in Definition 2.3 holds. Since $\gamma^*(s, x)$ is affine in x , conditions (i) and (ii) in Definition 2.3 also hold. Therefore, $\Gamma^* \in \mathcal{A}^\Gamma$.

2. We focus on the first inequality. For an admissible strategy $\Gamma = (\gamma_s)_{s \in [t, T]} \in \mathcal{A}^\Gamma$, Remark 5 implies that

$$\frac{\partial u(s, x)}{\partial s} + [b_s + B_s x + \Lambda_s \gamma_s]' Du(s, x) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda'_s D^2 u(s, x)) + \theta g(s, x, h^*(s, x), \gamma_s; \theta) \leq 0. \quad (\text{A.22})$$

By Dynkin's formula,

$$\begin{aligned} & \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} [u(T, X_T)] \\ &= u(t, x) + \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\int_t^T \left\{ \frac{\partial u(s, X_s)}{\partial s} + [b_s + B_s X_s + \Lambda_s \gamma_s]' Du(s, X_s) + \frac{1}{2} \text{tr} (\Lambda_s \Lambda'_s D^2 u(s, X_s)) \right\} ds \right] \\ &\leq u(t, x) - \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\int_t^T \theta g(s, X_s, h^*(s, X_s), \gamma_s; \theta) ds \right] \end{aligned} \quad (\text{A.23})$$

Recall that $u(T, y) = 0 \forall y \in \mathbb{R}^n$, hence

$$u(t, x) \geq \mathbf{E}_{t,x}^{\mathbb{P}^\Gamma} \left[\int_t^T \theta g(s, X_s, h^*(s, X_s), \gamma_s; \theta) ds \right]. \quad (\text{A.24})$$

If $\Gamma = \Gamma^*$ then equality holds, that is,

$$u(t, x) = \mathbf{E}_{t,x}^{\mathbb{P}^{\Gamma^*}} \left[\int_t^T \theta g(s, X_s, h^*(s, X_s), \gamma^*(s, X_s); \theta) ds \right]. \quad (\text{A.25})$$

Similarly, for any admissible strategy $H = (h_s)_{s \in [t, T]} \in \mathcal{A}^H$, the local saddle-point condition implies

$$\frac{\partial u}{\partial s}(s, x) + [b_s + B_s x + \Lambda_s \gamma^*(s, x)]' Du(s, x) + \frac{1}{2} \text{tr}(\Lambda_s \Lambda_s' D^2 u(s, x)) + \theta g(s, x, h_s, \gamma^*(s, x); \theta) \geq 0. \quad (\text{A.26})$$

Applying Dynkin's formula under \mathbb{P}^{Γ^*} yields the second inequality in (2.47).

3. This result follows directly from 1. and 2. above.