AIMS Mathematics, 11(4): 10744–10795.
DOI: 10.3934/math.2026442
Received: 02 March 2026
Revised: 13 April 2026
Accepted: 15 April 2026
Published: 20 April 2026

http://www.aimspress.com/journal/Math

Research article

Exergy-dissipation portfolio optimization: A backtest analysis

Muhammad Hilal Alkhudaydi*

Department of Mathematics and Statistics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

* **Correspondence:** Email: mh.ayedh@tu.edu.sa

Abstract: Exergy-dissipation portfolio optimization (EDPO) is a physics-inspired alternative to mean-variance allocation that treats portfolio formation as a nonequilibrium optimization problem. Rather than maximizing an estimated mean return subject to a variance penalty, EDPO maximizes an exergy rate $X_a(\mathbf{w})$, interpreted as a risk-sensitive certainty equivalent of log-growth, while penalizing an entropy production rate (EPR) $\mathcal{D}(\mathbf{w})$ that quantifies time-irreversibility in the portfolio return process. We propose a tractable long-only implementation with stabilizers including weight caps, Kullback–Leibler (KL), and turnover regularization along with a dissipation-budget thermostat: a primal–dual control law that updates λ to keep the in-sample dissipation $\mathcal{D}_{\mathrm{IS}}$ close to a target budget σ_{bud} without explicit regime labels. The accompanying theory establishes (i) a variational identity $X_a = \mathbb{E}_Q[g] + a^{-1}D_{\mathrm{KL}}(Q\|P)$, linking exergy to a tilted expected log-growth and an information term, and (ii) a constrained efficient frontier in the $(X_a, \mathcal{D})$ plane, with a convex candidate-set dual used as an internal consistency diagnostic. An ablation study confirms that combining exergy and dissipation improves risk-adjusted performance relative to either term in isolation. Empirically, a rolling backtest over 2015–2025 on four exchange-traded fund (ETF) universes shows that EDPO is competitive with standard benchmarks. In the diversified ETF7 universe, EDPO attains the highest Sharpe ratio (0.955) with moderate quarterly turnover (0.251), while remains comparable in more correlated universes. These model-based thermodynamic diagnostics are economically interpretable: stress episodes coincide with elevated $\mathcal{D}$, and the thermostat raises λ when dissipation pressure increases. Finally, $\mathcal{D}$ contains incremental predictive content for realized volatility in three of the four universes, as measured by mean squared error (MSE) ratios below one, although the associated forecast-gain tests remain exploratory once serial dependence and multiple comparisons are taken into account. A compact empirical-evidence package further shows that the conclusions are most sensitive to the coarse-graining level K, moderately sensitive to the lookback length L, and least sensitive to the smoothing parameter α. The strategy remains viable under plausible transaction costs.

Keywords: exergy-dissipation portfolio optimisation; nonequilibrium thermodynamics; entropy production rate; certainty-equivalent log-growth; primal-dual thermostat control

Mathematics Subject Classification: 62P05; 91G10; 91G60, 90C30, 60J10, 94A17, 65K10

1. Introduction

1.1. Portfolio choice beyond mean-variance

Portfolio optimization remains anchored by mean-variance efficiency [1]. That benchmark is indispensable, but two empirical difficulties persist in long-only allocation. First, expected-return estimates are unstable and often dominate the optimization error. Second, variance is time-symmetric: it measures dispersion, but it does not distinguish between return paths that share similar second moments yet differ sharply in drawdown dynamics, volatility bursts, or rapid risk-on/risk-off reversals. Growth-based portfolio criteria address part of this gap by evaluating allocations through long-run compounding under model ambiguity rather than through a single point estimate of expected return [2–5]. Recent work also helps position exergy-dissipation portfolio optimization (EDPO), a physics-inspired alternative to mean-variance allocation that treats portfolio formation as a nonequilibrium optimization problem, more precisely within adjacent strands of the finance literature. Entropy-based portfolio studies show that information-theoretic risk measures can provide a useful alternative or complement to the classical mean-variance framework, although entropy is typically used there as a diversification or return-distribution criterion rather than as a path-dependent measure of market irreversibility [6]. At the same time, regime-aware portfolio models with hidden Markov structures show that latent market states are economically relevant for dynamic asset allocation under changing business-cycle conditions [7]. Complementing these strands, research on financial time-series irreversibility shows that departures from time-reversal symmetry contain information beyond conventional volatility-based summaries [8]. Relative to this literature, EDPO combines a risk-sensitive growth objective with a thermodynamic-inspired, model-based dissipation diagnostic derived from coarse-grained portfolio wealth dynamics, thereby linking portfolio selection not only to dispersion and diversification, but also to the directional structure of return paths.

1.2. Irreversibility as a distinct risk dimension

A second gap is that portfolio risk is not only about dispersion; it is also about temporal asymmetry. Financial return series frequently display time-irreversible behavior around stress episodes, regime transitions, and clustered volatility, and these features are not identified by covariance alone [8–10]. This observation motivates the use of entropy-production-based diagnostics as an additional state variable for portfolio choice. In the present paper, the exergy term $X_a(\boldsymbol{w})$ and the dissipation term $\mathcal{D}(\boldsymbol{w})$ are therefore interpreted as model-based thermodynamic diagnostics of the induced wealth dynamics, whose magnitudes depend on the chosen coarse-graining, sample window, and estimation procedure rather than representing literal physical measurements or universal market constants.

1.3. What EDPO adds relative to the literature

Exergy-dissipation portfolio optimization (EDPO) lies at the intersection of robust growth, entropy-based portfolio design, and regime-aware financial modelling. Existing entropy-based portfolio studies typically use entropy as a diversification or distributional summary statistic [6]. Recent finance papers, however, study portfolio-return prediction through machine-learning factor models [11], market-state classification through ensemble and hidden Markov frameworks [12], and latent information-flow regimes through hidden-state financial models [13].

Recent work further clarifies the niche that EDPO occupies within adjacent strands of the finance literature. Studies that combine economically interpretable factor structures with nonlinear sequence learning show that portfolio-return prediction can be improved when richer dependence patterns and downside-risk characteristics are allowed to enter the modelling stage [11]. Complementary regime-aware approaches show that machine-learning ensembles and hidden Markov state detection can be combined to translate bull, bear, and neutral market classifications into implementable exchange-traded fund (ETF) trading rules [12]. At a broader modelling level, latent information-flow regimes have been shown to help explain volatility clustering, jumps, and heterogeneous responses to shocks, reinforcing the view that hidden-state dynamics may contain economically relevant information beyond conventional moment-based summaries [13].

EDPO is related to this literature in spirit, but it differs in mechanism. Rather than estimating an external forecasting model, classifying market regimes first, or recovering latent states and then passing those outputs into a separate portfolio-construction stage, EDPO embeds the risk-sensitive growth functional $X_a(\boldsymbol{w})$ together with the thermodynamic-inspired, model-based dissipation diagnostic $\mathcal{D}(\boldsymbol{w})$ directly in the allocation objective. The trade-off between these components is then regulated through a dissipation-budget thermostat, so that the shadow price of dissipation remains both economically interpretable and operational in the backtest. In this sense, EDPO is closer to a portfolio rule built from pathwise asymmetry diagnostics than to a forecast-then-allocate architecture.

More broadly, recent optimization research outside finance also underscores the importance of hybrid search and stabilization mechanisms when nonlinear objectives must be implemented under multiple constraints [14]. Although that literature is developed for engineering path-planning problems rather than portfolio allocation, it highlights a methodological point that is relevant here, as well: practical performance depends not only on the economic criterion itself, but also on how the optimization rule is made numerically stable and operational under a constrained feasible set.

1.4. Contributions and roadmap

This paper makes three contributions.

- **Conceptual contribution.** We formulate a long-only EDPO problem in which exergy $X_a(\boldsymbol{w})$ captures risk-sensitive log-growth, and dissipation $\mathcal{D}(\boldsymbol{w})$ captures time-irreversibility in coarse-grained wealth dynamics.
- **Methodological contribution.** We develop a practical rolling-window implementation with stabilizers and a primal–dual thermostat that tracks a dissipation budget without requiring exogenous regime labels.
- **Empirical contribution.** Across four ETF universes, EDPO is competitive with standard

benchmarks, with its strongest evidence in the diversified ETF7 universe. The forecasting gains associated with dissipation are economically suggestive, but the inferential evidence is treated as exploratory.

The remainder of this paper is organized as follows. Section 2 develops the theoretical framework, including the thermodynamic objective, the dissipation-budget formulation, and the connection between the idealized theory and its empirical implementation. Section 3 details the dataset and outlines the rolling-window backtesting methodology. Section 4 presents the main empirical findings, including robustness checks and forecasting diagnostics, to evaluate the proposed framework. Finally, Section 5 concludes with a summary of the findings and a discussion of their implications.

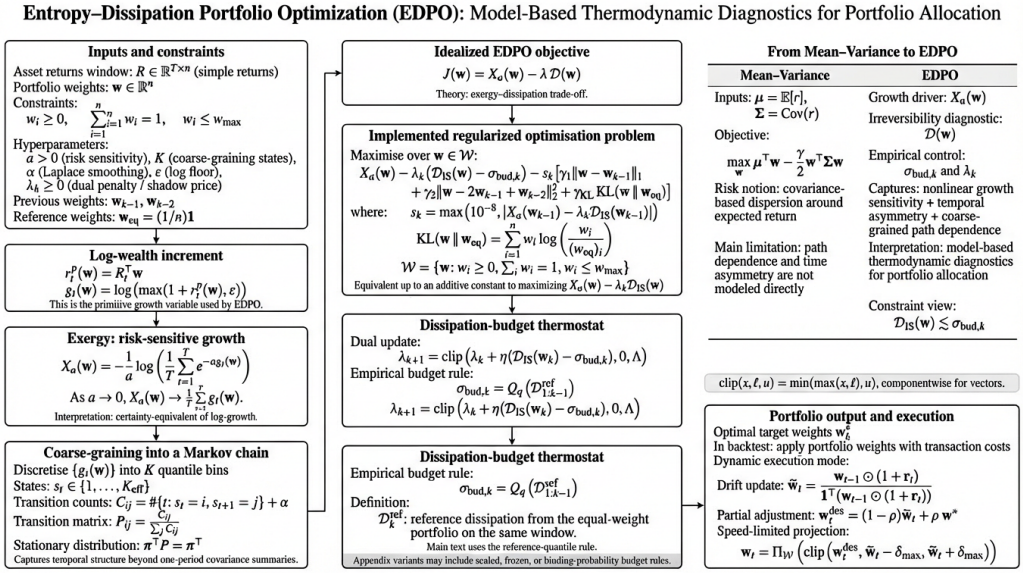


Figure 1. Schematic of the EDPO mechanism. Daily asset returns are mapped into portfolio log-wealth increments, from which exergy $X_a(\boldsymbol{w})$ and dissipation $\mathcal{D}(\boldsymbol{w})$ are computed. The EDPO objective trades off risk-sensitive growth against dissipation subject to long-only simplex constraints and stabilizers, and a primal–dual thermostat updates λ to track a dissipation budget.

2. Thermodynamic portfolio optimization with dissipation budgets

This section develops the theoretical layer underpinning EDPO. We make explicit which objects are treated as modelling primitives (defined under a reference probability measure $\mathbb{P}$) and which are empirical plug-in quantities computed on a rolling window. Throughout, portfolio weights are vectors $w \in \mathbb{R}^n$, and asset returns are $\boldsymbol{r}_t \in \mathbb{R}^n$.

Basic wealth dynamics. Given weights w applied over $(t-1, t]$, the gross portfolio return and log-wealth increment are

$$R_t(w) \;:=\; 1 + w^\top \boldsymbol{r}_t, \qquad g_t(w) \;:=\; \log R_t(w)$$

whenever $R_t(w) > 0$. In the empirical implementation, a small numerical floor is applied to avoid $-\infty$ values; the theory below is stated under a *uniform nonbankruptcy* condition that keeps $\log R_t(w)$

well-defined and preserves the concavity properties used by the optimization arguments.

Mean-variance benchmark. Let $\boldsymbol{r}$ denote a generic one-period return vector with mean $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{r}]$ and covariance $\boldsymbol{\Sigma} = \mathrm{Cov}(\boldsymbol{r})$. Under long-only and full-investment constraints (and optional caps), the classical mean-variance allocation can be written as

$$\max_{w \in \mathcal{W}} \ \boldsymbol{\mu}^\top w - \frac{\gamma}{2} w^\top \boldsymbol{\Sigma} w, \tag{2.1}$$

where $\gamma > 0$ is a risk-aversion parameter, and $\mathcal{W}$ is the feasible set. The model is "memoryless" in the sense that risk is summarized by a second-moment penalty applied to a one-period return distribution.

EDPO objective. EDPO replaces the moment-based proxy with two thermodynamic functionals computed from the induced log-wealth dynamics:

(1) the *exergy rate* $X_a(w)$, a risk-sensitive certainty equivalent of $g_t(w)$ (equivalently, an entropic risk transform);

(2) the *dissipation* $\mathcal{D}(w)$, defined as an entropy production rate (EPR) of a coarse-grained Markov representation of the increment process, quantifying time-irreversibility through deviations from detailed balance.

The exergy admits a classical variational representation (Gibbs/Donsker–Varadhan) with a Kullback–Leibler penalty (Theorem 2.1), whereas the dissipation controls the statistical distinguishability of the time direction via a Pinsker-type bound (Theorem 2.2). These are standard information-theoretic results, but we state them in a form tailored to EDPO and cite the underlying sources.

Regularized long-only EDPO used in the backtest. At rebalance k, the implemented portfolio solves a stabilized, long-only penalized problem of the form

$$\max_{w \in \mathcal{W}} \ \underbrace{\big(X_a(w) - \lambda_k \mathcal{D}(w)\big)}_{\text{physics score}} \ - \ s_k\Big(\gamma_{\text{turnover}}\|w - w_{k-1}\|_1 + \gamma_{\text{accel}}\|w - 2w_{k-1} + w_{k-2}\|_2^2 + \gamma_{\text{KL}} D_{\text{KL}}^{\text{w}}(w\|w^{eq})\Big), \tag{2.2}$$

where $w^{eq} = (1/n)\mathbf{1}$ is the equal-weight reference,

$$s_k \ := \ \max\{10^{-8}, \ |X_a(w_{k-1}) - \lambda_k \mathcal{D}(w_{k-1})|\}$$

is an objective-scale normalization (formalized in Appendix B), and the feasible set is the box-constrained simplex

$$\mathcal{W} = \{w \in \mathbb{R}^n : \ w \geq 0, \ \mathbf{1}^\top w = 1, \ w_i \leq w_{\max}\}.$$

The multiplier λ_k is either fixed or updated by the dissipation-budget "thermostat" in Definition 2.7.

Section roadmap. To keep the exposition focused, Section 2 now moves from the core modelling objects to the budgeted EDPO problem, then to the thermostat control law and the finite-sample estimator used in the empirical study. Standard technical derivations that are useful for completeness but not essential for the first reading, including the projection formula, the objective-scale normalization argument, the detailed thermostat bounds, and the estimator stability and consistency proofs, are collected in Appendix B.

2.1. Market primitives and admissible portfolios

We work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Let $\boldsymbol{r}_t \in \mathbb{R}^n$ denote the (simple) return vector over $(t-1, t]$. For weights $w \in \mathbb{R}^n$, the one-period gross return is $R_t(w) = 1 + w^\top \boldsymbol{r}_t$, and the log-wealth increment is

$$g_t(w) \;\coloneqq\; \log(1 + w^\top \boldsymbol{r}_t). \tag{2.3}$$

Definition 2.1 (Admissible weight set)**.** *Let $\mathcal{W} \subseteq \Delta^{n-1} \coloneqq \{w \in \mathbb{R}^n_{\geq 0} : \mathbf{1}^\top w = 1\}$ be a nonempty, compact set (for example a box-constrained simplex). We call $\mathcal{W}$ the set of* admissible portfolios.

Remark 2.1 (Box-constrained simplex in the empirical implementation)**.** *A practically useful instance of Definition 2.1 is the box-constrained simplex*

$$\mathcal{W} = \left\{w \in \mathbb{R}^n_{\geq 0} :\; \mathbf{1}^\top w = 1,\;\; w_i \leq w_{\max} \text{ for all } i\right\},$$

where $w_{\max} \in (0, 1)$ controls concentration and ensures implementability.

Remark 2.2 (Clipping operator (implementation notation))**.** *For scalars $x \in \mathbb{R}$ and bounds $\ell \leq u$, define*

$$\mathrm{clip}(x, \ell, u) \;\coloneqq\; \min\{\max\{x, \ell\}, u\}.$$

For vectors $x, \ell, u \in \mathbb{R}^n$ with $\ell_i \leq u_i$, define $\mathrm{clip}(x, \ell, u)$ *componentwise by*

$$(\mathrm{clip}(x, \ell, u))_i \;\coloneqq\; \mathrm{clip}(x_i, \ell_i, u_i), \qquad i = 1, \ldots, n.$$

When $\ell = 0$ and $u = \Lambda$, $\mathrm{clip}(\cdot, 0, \Lambda)$ *coincides with Euclidean projection onto the interval* $[0, \Lambda]$.

A computational note on feasibility. Because the box-constrained simplex projection is used only to enforce feasibility numerically, we postpone its explicit formula to Section 2.9, where it enters the rolling algorithm directly, and defer the short Karush–Kuhn–Tucker (KKT) derivation to Appendix B.

Assumption 2.1 (Uniform nonbankruptcy positivity of gross returns)**.** *There exists $\varepsilon \in (0, 1)$ such that*

$$R_t(w) = 1 + w^\top \boldsymbol{r}_t \;\geq\; \varepsilon \qquad \mathbb{P}\text{-a.s. for all } w \in \mathcal{W} \text{ and all } t.$$

Here, "a.s." means almost surely *under* $\mathbb{P}$.

Remark 2.3 (Scope of Assumption 2.1)**.** *Assumption 2.1 is a technical convenience ensuring that $g_t(w)$ is finite and uniformly bounded below by* $\log \varepsilon$*. It is satisfied, for example, if each asset satisfies $1 + r_{t,i} \geq \varepsilon$ almost surely (no single-asset "total loss"), and w is long-only. In the empirical pipeline, we additionally clip $R_t(w)$ below by a small numerical floor ε when evaluating $g_t(w)$; this prevents numerical $-\infty$ values on finite samples but is not used in the theoretical arguments that rely on the concavity of $w \mapsto \log(1 + w^\top \boldsymbol{r}_t)$.*

2.2. Exergy as risk-sensitive log-growth

The first thermodynamic primitive is a risk-sensitive certainty equivalent of the log-increment. The properties used later are its concavity on the admissible set, its variational representation with a Kullback–Leibler penalty, and its local mean-variance expansion for small values of the risk-sensitivity parameter a.

Definition 2.2 (Exergy functional)**.** *Fix $a \geq 0$ (risk sensitivity inverse temperature). For $a > 0$, define*

$$X_a(w) \;:=\; -\frac{1}{a}\log \mathbb{E}[\exp(-a\, g_1(w))], \tag{2.4}$$

and for $a = 0$, define $X_0(w) := \mathbb{E}[g_1(w)]$ (the continuous extension at $a = 0$). Equivalently, $X_a(w) = -\rho_a(-g_1(w))$ where ρ_a is the entropic risk measure; see, for example, [15].

Lemma 2.1 (Concavity of exergy on $\mathcal{W}$)**.** *Under Assumption 2.1, for each $a \geq 0$, the map $w \mapsto X_a(w)$ is concave on $\mathcal{W}$.*

Proof. Concavity follows from two steps. First, under Assumption 2.1, the map $w \mapsto g_1(w) = \log(1 + w^\top \boldsymbol{r}_1)$ is concave on $\mathcal{W}$. Second, after multiplying by $-a$ and exponentiating, Hölder's inequality yields the log-convex bound needed for the expectation in (2.4). Applying $-(1/a)\log(\cdot)$ converts this bound into concavity of X_a; the case $a = 0$ is immediate from linearity of expectation. The full derivation is deferred to Appendix B. □

Theorem 2.1 (Variational representation of exergy)**.** *Let $a > 0$ and $w \in \mathcal{W}$. Then,*

$$X_a(w) \;=\; \inf_{Q \ll \mathbb{P}} \left\{ \mathbb{E}_Q[g_1(w)] \;+\; \frac{1}{a}\, D_{\mathrm{KL}}(Q\|\mathbb{P}) \right\}. \tag{2.5}$$

Moreover, an optimizer is given by the exponentially tilted measure $Q^{w,a}$ with density

$$\frac{dQ^{w,a}}{d\mathbb{P}} \;=\; \frac{\exp(-a\, g_1(w))}{\mathbb{E}[\exp(-a\, g_1(w))]}.$$

Identity (2.5) *is a standard instance of the Gibbs (a.k.a. Donsker–Varadhan) variational principle; see, for example, [16, 17].*

Proof. This is the Gibbs variational principle applied to $Z = -a\, g_1(w)$. Jensen's inequality yields the upper bound $X_a(w) \leq \mathbb{E}_Q[g_1(w)] + a^{-1} D_{\mathrm{KL}}(Q\|\mathbb{P})$ for every $Q \ll \mathbb{P}$, and equality is attained by the exponentially tilted measure with density proportional to $\exp(-a g_1(w))$. The algebra is standard and is recorded in Appendix B. □

Remark 2.4 (Interpretation)**.** *The variational form* (2.5) *shows that maximizing $X_a(w)$ is equivalent to maximizing expected log growth under an adversarial change of measure penalized by relative entropy. In particular, a controls the amount of distributional distortion (in KL units) that is contemplated. This is the discrete-time analog of robust/risk-sensitive growth formulations used in robust control and large deviations.*

Proposition 2.1 (Cumulant expansion and the mean-variance bridge)**.** *Fix $w \in \mathcal{W}$, and assume the moment generating function $\mathbb{E}[\exp(t\, g_1(w))]$ is finite in a neighborhood of $t = 0$. Let $\kappa_m(w)$ denote the mth cumulant of $g_1(w)$. Then, as $a \downarrow 0$,*

$$X_a(w) \;=\; \kappa_1(w) \;-\; \frac{a}{2}\kappa_2(w) \;+\; \frac{a^2}{6}\kappa_3(w) \;+\; O(a^3). \tag{2.6}$$

In particular,

$$X_a(w) = \mathbb{E}[g_1(w)] \;-\; \frac{a}{2}\mathrm{Var}(g_1(w)) \;+\; O(a^2),$$

so X_a recovers a mean-variance-type trade-off to the second order while retaining sensitivity to higher-order cumulants (skewness, kurtosis, tail behavior) for non-infinitesimal a.

Proof. Write $\psi_w(a) = \log \mathbb{E}[\exp(-ag_1(w))]$ and expand ψ_w around $a = 0$ using the cumulant generating function of $g_1(w)$. Dividing by $-a$ gives (2.6). Thus, the leading correction to expected log growth is a variance penalty, and higher-order cumulants enter once a is non-negligible. The complete expansion is collected in Appendix B. □

2.3. Dissipation as the entropy production rate

Exergy is a one-step (marginal) functional. Our second primitive is explicitly *temporal*: it measures the time-irreversibility of coarse-grained portfolio dynamics.

Definition 2.3 (Coarse-grained microstates)**.** *Fix an integer $K \geq 2$ and a measurable map $\phi : \mathbb{R} \to \{1, \ldots, K\}$. For each $w \in \mathcal{W}$, define the microstate process*

$$S_t^{(w)} \coloneqq \phi(g_t(w)) \in \{1, \ldots, K\}.$$

Assumption 2.2 (Stationary Markov microdynamics)**.** *For each $w \in \mathcal{W}$, $(S_t^{(w)})_{t\geq 0}$ is a time-homogeneous, irreducible, aperiodic Markov chain with transition matrix $P^{(w)} = (P_{ij}^{(w)})_{i,j=1}^K$ and unique stationary distribution $\pi^{(w)} = (\pi_i^{(w)})_{i=1}^K$.*

Definition 2.4 (Entropy production rate (EPR))**.** *For $w \in \mathcal{W}$, define the dissipation functional*

$$\mathcal{D}(w) \coloneqq \sum_{i=1}^K \sum_{j=1}^K \pi_i^{(w)} P_{ij}^{(w)} \log\left(\frac{\pi_i^{(w)} P_{ij}^{(w)}}{\pi_j^{(w)} P_{ji}^{(w)}}\right), \tag{2.7}$$

with the conventions $0\log(0/q) = 0$ and $x\log(x/0) = +\infty$ for $x > 0$. When $\pi_i^{(w)} P_{ij}^{(w)} > 0$ implies $\pi_j^{(w)} P_{ji}^{(w)} > 0$ (microreversibility of the stationary flow), the dissipation is finite. In the empirical estimator (Definition 2.8), Laplace smoothing enforces strict positivity. The EPR for Markov chains is standard in stochastic thermodynamics; see, for example, [18, 19].

Lemma 2.2 (Non-negativity and equilibrium characterization)**.** *For each $w \in \mathcal{W}$, $\mathcal{D}(w) \geq 0$. Moreover, $\mathcal{D}(w) = 0$ if and only if the stationary flows satisfy*

$$\pi_i^{(w)} P_{ij}^{(w)} = \pi_j^{(w)} P_{ji}^{(w)} \quad \text{for all } i, j, \tag{2.8}$$

that is, the chain is time-reversible (detailed balance).

Proof. Define the stationary edge-flow $F_{ij}^{(w)} = \pi_i^{(w)} P_{ij}^{(w)}$ and its reversal $(F^{(w)})_{ij}^{\mathrm{rev}} = F_{ji}^{(w)}$. Then, (2.7) is exactly the Kullback–Leibler divergence $D_{\mathrm{KL}}(F^{(w)} \| (F^{(w)})^{\mathrm{rev}})$, which is non-negative and vanishes only under detailed balance. Full details are deferred to Appendix B. □

Proposition 2.2 (Path-space meaning of EPR)**.** *Under Assumption 2.2, $\mathcal{D}(w)$ equals the relative-entropy rate between the stationary forward path measure of $(S_t^{(w)})$ and the stationary time-reversed path measure:*

$$\mathcal{D}(w) = \lim_{T\to\infty} \frac{1}{T} D_{\mathrm{KL}}\Big(\mathbb{P}(S_{0:T}^{(w)} \in \cdot) \,\Big\|\, \mathbb{P}(S_{T:0}^{(w)} \in \cdot)\Big),$$

where $S_{T:0}^{(w)} = (S_T^{(w)}, S_{T-1}^{(w)}, \ldots, S_0^{(w)})$.

Proof. Compare the stationary probability of a forward microstate path with that of its reversal. The boundary term produced by the stationary distribution is $O(1)$, and the transition part contributes an ergodic average of the log forward-reverse transition ratio. Dividing by the path length and passing to the limit yields the relative-entropy rate in (2.7). The detailed computation is given in Appendix B. □

Theorem 2.2 (Arrow-of-time detectability bound)**.** *Let $(S_t)_{t\geq 0}$ be a stationary Markov chain on $\{1,\dots,K\}$ with transition matrix P and stationary distribution π. Define the stationary edge-flow distribution on ordered pairs by $F_{ij} := \pi_i P_{ij}$ and its time reversal by $F^{\mathrm{rev}}_{ij} := F_{ji}$. Assume $F_{ij} > 0 \Rightarrow F^{\mathrm{rev}}_{ij} > 0$. Then, the entropy production rate $\mathcal{D} = D_{\mathrm{KL}}(F\|F^{\mathrm{rev}})$ controls the total variation distance via Pinsker's inequality:*

$$\|F - F^{\mathrm{rev}}\|_{\mathrm{TV}} \;\leq\; \sqrt{\frac{\mathcal{D}}{2}}, \qquad \textit{equivalently} \qquad \|F - F^{\mathrm{rev}}\|_1 \;\leq\; \sqrt{2\mathcal{D}},$$

where

$$\|p - q\|_{\mathrm{TV}} = \frac{1}{2}\sum_{i,j} |p_{ij} - q_{ij}|.$$

Consequently, in the symmetric binary test that observes one transition (I, J) drawn from either F (forward) or F^{rev} (reversed), the success probability of any decision rule φ satisfies

$$\mathbb{P}(\varphi(I, J) = H) \;\leq\; \frac{1}{2}\Big(1 + \sqrt{\frac{\mathcal{D}}{2}}\Big).$$

Pinsker's inequality and the testing characterization are standard; see, for example, [20, 21].

Proof. Apply Pinsker's inequality to the stationary edge-flow distribution F and its reversal F^{rev}. This immediately yields the total-variation and ℓ^1 bounds. The testing statement then follows from the classical characterization of the optimal error rate in a symmetric binary experiment. The short derivation is recorded in Appendix B. □

Remark 2.5 (Multistep detectability)**.** *The one-step bound in Theorem 2.2 is the quantity used in our interpretation of a per-period dissipation budget. For a length-T microstate trajectory, the corresponding path-space KL divergence scales as $D_{\mathrm{KL}}(\mathbb{P}(S_{0:T})\|\mathbb{P}(S_{T:0})) = T\,\mathcal{D} + O(1)$ by Proposition 2.2. Combining with Pinsker yields the coarse bound $\|\mathbb{P}(S_{0:T}) - \mathbb{P}(S_{T:0})\|_{\mathrm{TV}} \leq \sqrt{T\mathcal{D}/2}$ (capped at 1), making explicit that dissipation controls the statistical detectability of time direction even when using longer observed paths.*

Remark 2.6 (Why dissipation is not a moment)**.** *Two processes can share the same one-step marginal distribution (hence identical mean and variance) yet differ in time-reversal asymmetry. $\mathcal{D}(w)$ quantifies this asymmetry at the transition level and therefore carries information unavailable to mean-variance summaries.*

2.4. The dissipation-budgeted formulation

The dissipation-budgeted formulation treats dissipation as a resource that can be budgeted for.

Definition 2.5 (Dissipation-budget EDPO (primal))**.** *Fix $a \geq 0$ and a dissipation budget $\sigma_{\mathrm{bud}} \geq 0$. The exergy under dissipation budget problem is*

$$(\mathrm{P})_{\sigma_{\mathrm{bud}}}: \qquad \max_{w \in \mathcal{W}} X_a(w) \quad \textit{subject to} \quad \mathcal{D}(w) \leq \sigma_{\mathrm{bud}}. \tag{2.9}$$

Let $V(\sigma_{\mathrm{bud}})$ denote the optimal value (possibly $-\infty$ if infeasible).

Assumption 2.3 (Feasibility and convex-analytic regularity)**.** *There exists at least one $w \in \mathcal{W}$ with $\mathcal{D}(w) < \sigma_{\mathrm{bud}}$ (strict feasibility / Slater condition). Moreover, $\mathcal{D}(\cdot)$ is convex and lower semicontinuous on $\mathcal{W}$.*

Remark 2.7 (When is $\mathcal{D}$ convex?)**.** *Convexity of $w \mapsto \mathcal{D}(w)$ is not automatic for an arbitrary coarse-graining map ϕ and arbitrary dependence of the induced transition law on w. However, EPR* is *convex as a functional of the stationary edge-flow distribution F because $\mathcal{D}(F) = D_{\mathrm{KL}}(F\|F^{\mathrm{rev}})$, and KL divergence is jointly convex, whereas $F \mapsto F^{\mathrm{rev}}$ is a linear permutation. Hence, if the modelling layer is such that the stationary edge-flow depends affinely on w (for instance, when the coarse-grained kernel is a convex mixture of base kernels indexed by the portfolio weights), then $w \mapsto \mathcal{D}(w)$ is convex. Assumption 2.3 isolates the convex regime in which Lagrange duality is exact; the empirical pipeline replaces $\mathcal{D}$ by a plug-in estimate $\mathcal{D}_{\mathrm{IS}}$ and uses duality primarily as an internal diagnostic (Remark 2.10 and the subsequent duality-gap discussion).*

Theorem 2.3 (Existence of a primal optimizer)**.** *Under Assumptions 2.1 and 2.3, problem $(\mathrm{P})_{\sigma_{\mathrm{bud}}}$ admits an optimizer $w^\star \in \mathcal{W}$.*

Proof. Under Assumption 2.3, the feasible set $\{w \in \mathcal{W} : \mathcal{D}(w) \leq \sigma_{\mathrm{bud}}\}$ is nonempty and closed and therefore compact because $\mathcal{W}$ is compact. Because X_a is continuous on $\mathcal{W}$, the extreme-value theorem ensures that the constrained maximum is attained. The formal argument is deferred to Appendix B. □

2.5. Duality and the EDPO penalization

The Lagrangian for $(\mathrm{P})_{\sigma_{\mathrm{bud}}}$ is

$$\mathcal{L}(w, \lambda) \;=\; X_a(w) \;-\; \lambda(\mathcal{D}(w) - \sigma_{\mathrm{bud}}), \qquad \lambda \geq 0. \tag{2.10}$$

For fixed λ, maximizing $\mathcal{L}(w, \lambda)$ over w is the penalized EDPO objective with penalty weight λ.

Definition 2.6 (Dual function and dual problem)**.** *Define the dual function*

$$d(\lambda) \;:=\; \sup_{w \in \mathcal{W}} \mathcal{L}(w, \lambda) = \lambda\sigma_{\mathrm{bud}} + \sup_{w \in \mathcal{W}} (X_a(w) - \lambda\mathcal{D}(w)), \qquad \lambda \geq 0,$$

and the dual problem $(\mathrm{D})_{\sigma_{\mathrm{bud}}}$ is

$$(\mathrm{D})_{\sigma_{\mathrm{bud}}}: \qquad \inf_{\lambda \geq 0} d(\lambda).$$

Lemma 2.3 (Weak duality)**.** *For any $\lambda \geq 0$, $V(\sigma_{\mathrm{bud}}) \leq d(\lambda)$. Hence, $V(\sigma_{\mathrm{bud}}) \leq \inf_{\lambda \geq 0} d(\lambda)$.*

Proof. For any feasible portfolio, the constraint term satisfies $-\lambda(\mathcal{D}(w) - \sigma_{\mathrm{bud}}) \geq 0$ whenever $\lambda \geq 0$. Hence, every feasible objective value is bounded above by the Lagrangian value, and therefore by the dual function. Taking the supremum over feasible portfolios yields weak duality. The short algebraic argument is reported in Appendix B. □

Theorem 2.4 (Strong duality and shadow price)**.** *Under Assumptions 2.1 and 2.3,*

$$V(\sigma_{\text{bud}}) \;=\; \inf_{\lambda\geq 0} d(\lambda). \tag{2.11}$$

Moreover, there exists a dual minimizer $\lambda^\star \geq 0$ *and at least one* $w^\star \in \arg\max_{w\in\mathcal{W}} \mathcal{L}(w,\lambda^\star)$ *such that the KKT relations hold as follows:*

$$\mathcal{D}(w^\star) \leq \sigma_{\text{bud}}, \qquad \lambda^\star \geq 0, \qquad \lambda^\star(\mathcal{D}(w^\star) - \sigma_{\text{bud}}) = 0.$$

Strong duality under Slater's condition is standard in convex optimization; see, for example, [22, 23].

Proof. Assumptions 2.1 and 2.3 place the budgeted EDPO problem in the standard convex-programming setting: the feasible set is convex, X_a is concave, and Slater feasibility holds. Strong duality and the KKT conditions therefore follow from the usual convex duality theorem. Appendix B records the argument and the exact optimality relations. □

Remark 2.8 (Penalized EDPO is the Lagrangian relaxation)**.** *Theorem 2.4 provides a clean interpretation of the penalty parameter:* $\lambda^\star$ *is the* shadow price *of dissipation. Tightening the budget* σ_{bud} *typically increases* $\lambda^\star$*.*

Proposition 2.3 (Monotone work–dissipation trade-off along the penalty path)**.** *Fix* $a \geq 0$*. For each* $\lambda \geq 0$*, let* $w_\lambda \in \arg\max_{w\in\mathcal{W}}\{X_a(w) - \lambda\,\mathcal{D}(w)\}$*. Then, for any* $0 \leq \lambda_1 < \lambda_2$ *and any corresponding maximizers* $w_{\lambda_1}, w_{\lambda_2}$*,*

$$\mathcal{D}(w_{\lambda_2}) \;\leq\; \mathcal{D}(w_{\lambda_1}). \tag{2.12}$$

Moreover, the penalized value function

$$\varphi(\lambda) \;:=\; \sup_{w\in\mathcal{W}}\{X_a(w) - \lambda\,\mathcal{D}(w)\}$$

is convex and nonincreasing in λ*. If the maximizer is unique at* λ*, then* φ *is differentiable at* λ*, and*

$$\varphi'(\lambda) \;=\; -\mathcal{D}(w_\lambda).$$

Proof. Write the optimality inequalities for the maximizers at two penalty values $\lambda_1 < \lambda_2$ and add them. The cross terms cancel, leaving $(\lambda_2 - \lambda_1)(\mathcal{D}(w_{\lambda_1}) - \mathcal{D}(w_{\lambda_2})) \geq 0$, which proves monotonicity of dissipation along the penalty path. Convexity of φ follows because it is the supremum of affine functions of λ. The envelope argument is deferred to Appendix B. □

Remark 2.9 (Work–dissipation frontier and supporting hyperplanes)**.** *The set of achievable pairs*

$$\mathcal{A} \;:=\; \{(\mathcal{D}(w), X_a(w)) :\; w \in \mathcal{W}\} \subset \mathbb{R}^2$$

encodes the work–dissipation trade-off*. For a given* $\lambda \geq 0$*, the penalized problem selects a point on the upper boundary of* $\mathcal{A}$ *that maximizes the supporting line* $(x_1, x_2) \mapsto x_2 - \lambda x_1$*. Proposition 2.3 shows that moving along* $\lambda \uparrow$ *traverses this boundary from high-dissipation/high-growth regimes to low-dissipation regimes in a monotone fashion.*

Remark 2.10 (Regularized EDPO solved in the empirical pipeline)**.** *The inner maximization in* (2.14) *is an idealized primal oracle. Empirically, we solve a regularized variant that stabilizes estimation error and accounts for trading frictions:*

$$w_k \in \arg\max_{w\in\mathcal{W}} \Big\{ X_a(w) \;-\; \lambda_k(\mathcal{D}_{\mathrm{IS}}(w) - \sigma_{\mathrm{bud},k}) \\ - s_k\Big(\gamma_{\mathrm{KL}}\, D^{\mathrm{w}}_{\mathrm{KL}}(w \,\|\, w^{\mathrm{ref}}) + \gamma_{\mathrm{turnover}}\, \|w - w_{k-1}\|_1 + \gamma_{\mathrm{accel}}\, \|w - 2w_{k-1} + w_{k-2}\|_2^2\Big) \Big\}. \quad (2.13)$$

Here, $\mathcal{D}_{\mathrm{IS}}(w)$ denotes the in-sample plug-in estimate of dissipation (Definition 2.8); when it is useful to display the window length T and smoothing level α, we write $\mathcal{D}_{\mathrm{IS}}(w) \equiv \widehat{\mathcal{D}}_T(w;\alpha)$ The scalar

$$s_k \;:=\; \max\{\underline{s},\, |\, X_a(w_{k-1}) - \lambda_k\, \mathcal{D}_{\mathrm{IS}}(w_{k-1})\,|\}, \qquad \underline{s} > 0,$$

normalizes regularization strength to the score magnitude, making the stabilizers effectively dimensionless. The additive constant $\lambda_k\sigma_{\mathrm{bud},k}$ does not affect the argmax and is often dropped in code. These stabilizers modify the map $\lambda \mapsto w_\lambda$ but preserve the interpretation of λ as a shadow price for the estimated dissipation budget.

The scale normaliser in (2.13) is introduced so that the stabilizers act in units of the current score magnitude rather than in arbitrary raw objective units. This makes the regularization weights easier to interpret across windows and across universes. A formal invariance statement is reported in Appendix B.

Remark 2.11 (Strong duality versus empirical coarse-graining)**.** *Identity* (2.11) *is exact under the convex-analytic idealization in Assumption 2.3. In the empirical EDPO pipeline, the dissipation functional $\mathcal{D}(\cdot)$ is replaced by an in-sample plug-in estimate $\mathcal{D}_{\mathrm{IS}}(\cdot)$ built from a coarse partition of log-increments and a smoothed transition estimator (Definition 2.8). As a function of the portfolio weights, $w \mapsto \mathcal{D}_{\mathrm{IS}}(w)$ need not be convex, and feasibility may be approximate rather than exact.*

For this reason, the "duality gap" plots are best read as diagnostics computed on a finite candidate set $C \subset \mathcal{W}$ (replace $\sup_{w\in\mathcal{W}}$ by $\max_{w\in C}$ in the dual function). A small candidate-set gap indicates that (i) the chosen portfolio is nearly optimal relative to the tested candidate universe, and (ii) the thermostat multiplier behaves like a shadow price for the estimated dissipation. This is weaker than a theorem about the continuous problem but matches the level of verification available when $\mathcal{D}$ is learned from rolling data.

2.6. Thermostat as a primal–dual algorithm

The dual representation suggests a principled "thermostat" update of λ as a feedback control law that targets a dissipation budget.

Definition 2.7 (Thermostat (projected dual subgradient))**.** *Fix a budget σ_{bud} and step size $\eta > 0$. Given $\lambda_k \geq 0$, compute*

$$w_k \;\in\; \arg\max_{w\in\mathcal{W}} \; \mathcal{L}(w, \lambda_k) = \arg\max_{w\in\mathcal{W}} \; (X_a(w) - \lambda_k(\mathcal{D}(w) - \sigma_{\mathrm{bud}})), \quad (2.14)$$

and update

$$\lambda_{k+1} \;:=\; \Pi_{[0,\Lambda]}(\lambda_k + \eta\,(\mathcal{D}(w_k) - \sigma_{\mathrm{bud}})) = \mathrm{clip}(\lambda_k + \eta\,(\mathcal{D}(w_k) - \sigma_{\mathrm{bud}}),\, 0,\, \Lambda), \quad (2.15)$$

where $\Pi_{[0,\Lambda]}$ denotes Euclidean projection onto $[0,\Lambda]$ (equivalently the clipping map of Remark 2.2); taking $\Lambda = +\infty$ recovers the standard non-negative projection $[x]_+ = \max\{x, 0\}$.

Remark 2.12 (Empirical dissipation budget selection)**.** *In rolling-window applications, the dissipation budget is calibrated from a reference dissipation series built from the equal-weight portfolio on the same asset universe. At rebalance k, let $\mathcal{D}_k^{\mathrm{ref}}$ denote the reference dissipation and maintain the history $\{\mathcal{D}_j^{\mathrm{ref}}\}_{j<k}$. The main empirical analysis uses the reference-quantile rule*

$$\sigma_{\mathrm{bud},k} = Q_q(\mathcal{D}_{1:k-1}^{\mathrm{ref}}),$$

where Q_q denotes the empirical q-quantile of the reference history. We retain this rule in the main text because it is the most transparent calibration: it ties the current budget to a clearly interpretable reference portfolio and a fixed exceedance quantile of its own past dissipation levels. Alternative scaled, binding-probability, and frozen-budget variants are collected in Appendix B.5.1 for completeness, but they are not needed to understand the empirical design reported in Section 4.

The reference-quantile budget is attractive because it targets a chosen exceedance frequency for the reference dissipation process under standard empirical-process conditions, and the multiplier update in Definition 2.7 is the projected dual-subgradient step suggested by the Lagrangian structure. The formal exceedance guarantee, the subgradient identity, the finite-time control bound, and the alternative budget variants are collected in Appendix B; here, we retain only the default calibration rule, the update law itself, and their economic interpretation.

Remark 2.13 (From theory thermostat to bounded "practical thermostat")**.** *In implementation, one may replace the hard projection onto $[0,\Lambda]$ by a smooth saturating map (for example logistic bounding). The theoretical role of projection is to keep λ in a compact set so that subgradient bounds yield explicit rates.*

2.7. Finite-state estimation of the dissipation term

The empirical backtest replaces the idealized dissipation functional by a finite-sample plug-in estimator computed on each rolling window. This subsection retains the estimator definition and the interpretation needed for the empirical sections, and the detailed stability and consistency proofs are deferred to Appendix B.

Definition 2.8 (Quantile microstates and Laplace-smoothed EPR estimator)**.** *Fix $K \geq 2$ and a smoothing level $\alpha \geq 0$. Given a window of log-increments $\{g_t(w)\}_{t=1}^T$, choose thresholds*

$$b_0 = -\infty \;<\; b_1 \;<\; \cdots \;<\; b_{K-1} \;<\; b_K = \infty,$$

for instance, by setting b_k equal to the empirical (k/K)-quantile of $\{g_t(w)\}_{t=1}^T$. Define the (window-dependent) binning map

$$\phi_T(x) \;:=\; \min\{k \in \{1,\dots,K\} :\; x \leq b_k\}, \qquad S_t^{(w)} \;:=\; \phi_T(g_t(w)).$$

Form transition counts

$$C_{ij}^{(w)}(T) \;:=\; \sum_{t=2}^{T} \mathbb{1}\{S_{t-1}^{(w)} = i,\; S_t^{(w)} = j\},$$

and define the Laplace-smoothed transition estimator

$$\hat{P}_{ij}^{(w)}(T;\alpha) \;:=\; \frac{C_{ij}^{(w)}(T)+\alpha}{\sum_{j'=1}^{K} C_{ij'}^{(w)}(T)+K\alpha}. \tag{2.16}$$

Let $\hat{\pi}^{(w)}(T;\alpha)$ denote the stationary distribution of $\hat{P}^{(w)}(T;\alpha)$. The corresponding plug-in EPR estimator is

$$\widehat{\mathcal{D}}_T(w;\alpha) \;:=\; \sum_{i=1}^{K}\sum_{j=1}^{K} \hat{\pi}_i^{(w)}(T;\alpha)\hat{P}_{ij}^{(w)}(T;\alpha)\log\left(\frac{\hat{\pi}_i^{(w)}(T;\alpha)\hat{P}_{ij}^{(w)}(T;\alpha)}{\hat{\pi}_j^{(w)}(T;\alpha)\hat{P}_{ji}^{(w)}(T;\alpha)}\right). \tag{2.17}$$

When $\alpha > 0$, all entries of $\hat{P}^{(w)}(T;\alpha)$ are strictly positive, so the logarithm is always well-defined. In rolling-window applications, we refer to $\widehat{\mathcal{D}}_T(w;\alpha)$ as the in-sample dissipation estimate and write

$$\mathcal{D}_{\mathrm{IS}}(w) := \widehat{\mathcal{D}}_T(w;\alpha).$$

Remark 2.14 (Fixed versus w-dependent partitions)**.** *For theoretical properties of $\mathcal{D}(w)$ as a functional of w, it is convenient to fix the coarse-graining map ϕ independently of w. In the empirical estimator above, the thresholds (b_k) are computed from the sample $\{g_t(w)\}$ and therefore vary with w. This choice improves numerical stability (each microstate is populated) but introduces mild nonsmoothness in w because the bin boundaries move. Laplace smoothing ($\alpha > 0$) mitigates this by ensuring that the estimated transition kernel remains strictly positive even when some empirical counts are small.*

For a fixed portfolio w and a fixed coarse partition, Laplace smoothing makes the estimated transition law strictly positive and therefore keeps the plug-in entropy-production formula well-defined. Under the finite-state ergodic assumptions used in Section 2, the smoothed transition estimator and the associated stationary distribution converge to their population counterparts, so the plug-in dissipation estimate converges to $\mathcal{D}(w)$. Because these are support results for the implementation rather than new modelling primitives, the formal Lipschitz and consistency statements are collected in Appendix B.

Remark 2.15 (Coarse-graining as a conservative dissipation proxy)**.** *If one defines a "full-information" irreversibility level via the KL rate between forward and time-reversed path measures of the raw increment process, then passing to microstates $S_t = \phi(g_t)$ can only decrease that KL divergence by the data-processing inequality. Hence the EPR computed on microstates is naturally interpreted as a coarse-grained (typically conservative) proxy for the irreversibility of wealth dynamics; see, for example, [20] for the data-processing inequality.*

2.8. Beyond mean-variance: time-irreversibility as an independent risk dimension

The next proposition isolates the key point that entropy production is genuinely pathwise and is not determined by one-step marginal information.

Proposition 2.4 (Identical one-step distribution, different dissipation)**.** *There exist two stationary Markov chains (S_t) and $(\tilde{S}_t)$ on $\{1,2,3\}$ and a map $v : \{1,2,3\} \to \mathbb{R}$ such that the scalar time series $Y_t = v(S_t)$ and $\tilde{Y}_t = v(\tilde{S}_t)$ have the same one-step marginal distribution (hence the same mean and variance), but*

$$\mathcal{D}(S) \;=\; 0, \qquad \mathcal{D}(\tilde{S}) \;>\; 0,$$

where $\mathcal{D}$ denotes the entropy production rate associated with the Markov transition laws.

Proof. Take a reversible three-state chain with a uniform stationary law and a second chain with the same stationary law but a cycle-biased transition matrix. Any observable that depends only on the current state therefore has the same one-step marginal distribution under both chains, whereas the reversible chain has zero entropy production, and the cycle-biased chain has strictly positive entropy production. The explicit matrices and calculation are reported in Appendix B. □

Remark 2.16 (Why this matters for portfolios)**.** *The proposition highlights that marginal mean/variance information is insufficient to identify time irreversibility. Within EDPO, dissipation $\mathcal{D}(w)$ governs this missing dynamic dimension: two portfolios may have similar one-period risk summaries but very different exposure to irreversible (nonequilibrium) market dynamics.*

2.9. Computational implementation

This subsection summarizes the numerical procedure used to implement the EDPO allocation rule in rolling windows. Because the box-constrained simplex projection is a computational device rather than a modelling primitive, we record its explicit form here, next to the algorithm that uses it; the full derivation is deferred to Appendix B.

Proposition 2.5 (Euclidean projection onto the box-constrained simplex)**.** *Let $\mathcal{W}$ be the box-constrained simplex from Remark 2.1 with cap $w_{\max} \in (0, 1]$. For any $v \in \mathbb{R}^n$, consider the Euclidean projection*

$$\Pi_{\mathcal{W}}(v) \;:=\; \arg\min_{w \in \mathcal{W}} \; \frac{1}{2}\|w - v\|_2^2.$$

This problem admits a unique solution $w^\star = \Pi_{\mathcal{W}}(v)$, and there exists a scalar $\tau \in \mathbb{R}$ such that

$$w_i^\star \;=\; \mathrm{clip}(v_i - \tau,\, 0,\, w_{\max}) \qquad \textit{for all } i = 1, \ldots, n, \tag{2.18}$$

with τ chosen so that $\sum_{i=1}^n w_i^\star = 1$. Moreover, the map $\tau \mapsto \sum_{i=1}^n \mathrm{clip}(v_i - \tau, 0, w_{\max})$ is continuous and strictly decreasing; hence, τ is computed efficiently by bisection. Efficient projection algorithms (including proofs of the clipping form) are standard; see, for example, [24, 25].

Proof. The feasible set is closed and convex, so the Euclidean projection is uniquely defined. Writing the KKT conditions for the box-constrained simplex shows that each coordinate of the minimiser must equal the clipped value $\mathrm{clip}(v_i - \tau, 0, w_{\max})$ for a common scalar threshold τ, which is chosen to satisfy the budget constraint $\sum_i w_i^\star = 1$. Because the left-hand side is continuous and strictly decreasing in τ, bisection computes τ efficiently. The full derivation is reported in Appendix B. □

At each rebalance date k, weights are obtained by solving the regularised objective in Remark 2.10 using a derivative-free optimiser with multiple random restarts. Feasibility is enforced by mapping unconstrained parameters to a capped simplex: we apply a softmax transformation and then project onto

$$\mathcal{W} = \{w \in \Delta^{n-1} :\; 0 \le w_i \le w_{\max}\} (\textit{Proposition } 2.5).$$

When the thermostat is enabled, the dissipation multiplier is updated by the projected dual step in Definition 2.7, with the budget $\sigma_{\mathrm{bud},k}$ calibrated by the reference-quantile rule in Remark 2.12; alternative calibration variants are deferred to Appendix B.5.1.

Algorithm 1 Rolling-window EDPO optimization with dissipation-budget thermostat (implementation summary)

Require: Training returns $\{\boldsymbol{r}_t\}_{t=k-L+1}^{k}$, previous weights w_{k-1}, w_{k-2}, parameters (a, K, α), cap $w_{\max}$, multiplier λ_k, regularizer weights $(\gamma_{\mathrm{KL}}, \gamma_{\mathrm{turnover}}, \gamma_{\mathrm{accel}})$, reference weights w^{ref}, restarts R, step size η, multiplier cap Λ.

Ensure: Updated weights w_k (and, if used, updated multiplier λ_{k+1}).

Define the feasible set $\mathcal{W} = \{w \in \Delta^{n-1} : \ 0 \le w_i \le w_{\max}\}$.

Parameterize feasible weights by

$$w(\theta) = \Pi_{\mathcal{W}}(\mathrm{softmax}(\theta)), \qquad [\mathrm{softmax}(\theta)]_i = \frac{e^{\theta_i}}{\sum_{j=1}^{n} e^{\theta_j}},$$

where $\Pi_{\mathcal{W}}$ denotes Euclidean projection onto $\mathcal{W}$.

Compute the objective-scale normalizer

$$s_k = \max\{\underline{s}, \ |X_a(w_{k-1}) - \lambda_k \mathcal{D}_{\mathrm{IS}}(w_{k-1})|\}, \qquad \underline{s} > 0.$$

for $r = 1, \ldots, R$ **do**

  Initialize $\theta^{(r)} \leftarrow \log(w_{k-1}) + \xi^{(r)}$ (small random perturbation).

  Solve (derivative-free) for $\theta^{\star(r)} \in \arg\max_\theta \ J_k(w(\theta))$, where

  $J_k(w) = X_a(w) - \lambda_k \mathcal{D}_{\mathrm{IS}}(w)$

  $-s_k\Big(\gamma_{\mathrm{KL}} D_{\mathrm{KL}}^{\mathrm{w}}(w\|w^{\mathrm{ref}}) + \gamma_{\mathrm{turnover}}\|w - w_{k-1}\|_1$

  $+\gamma_{\mathrm{accel}}\|w - 2w_{k-1} + w_{k-2}\|_2^2\Big)$.

end for

Select the best restart $r^\star$ and set $w_k \leftarrow w(\theta^{\star(r^\star)})$.

Compute in-sample dissipation $\mathcal{D}_k^{\mathrm{IS}} \leftarrow \mathcal{D}_{\mathrm{IS}}(w_k)$.

if thermostat enabled **then**

  Compute reference dissipation $\mathcal{D}_k^{\mathrm{ref}}$ from the equal-weight portfolio on the same window.

  Set budget $\sigma_{\mathrm{bud},k} \leftarrow Q_q(\mathcal{D}_{1:k-1}^{\mathrm{ref}})$ by the reference-quantile rule of Remark 2.12.

  Update the multiplier by projected dual ascent:

$$\lambda_{k+1} \leftarrow \mathrm{clip}(\lambda_k + \eta(\mathcal{D}_k^{\mathrm{IS}} - \sigma_{\mathrm{bud},k}), \ 0, \ \Lambda).$$

else

  Set $\lambda_{k+1} \leftarrow \lambda_k$.

end if

return w_k and diagnostics (for example $X_a(w_k)$, $\mathcal{D}_k^{\mathrm{IS}}$, $\sigma_{\mathrm{bud},k}$, λ_{k+1}).

2.10. From theory to empirical implementation

The theoretical development above defines EDPO in terms of idealized objects, most importantly the exergy functional $X_a(w)$, the dissipation functional $\mathcal{D}(w)$, and the dissipation-budgeted problem with multiplier λ. The empirical implementation preserves this structure but replaces the idealized objects by finite-sample, rolling-window plug-in estimates. Table 1 makes this distinction explicit.

Table 1. Key deviations between the idealized EDPO theory and the rolling-window empirical implementation. The table distinguishes exact reference properties from the finite-sample objects used in the backtest and states the acceptable reading of the corresponding empirical results.

Theoretical property / role	Empirical proxy or key deviation	What remains interpretable	Acceptable reading in the results
Exergy $X_a(w)$ as the risk-sensitive growth functional	Rolling-window plug-in evaluation from realized portfolio log-increments $g_t(w)$ on each estimation window	The variational identity and the growth-versus-risk interpretation of X_a	Empirical X_a is a sample-dependent growth diagnostic. Its level is informative within the reported design, but it is not a sample-free constant.
Dissipation $\mathcal{D}(w)$ as the time-irreversibility functional	Laplace-smoothed EPR of a quantile-state Markov chain built from K bins over a finite lookback window	Non-negativity, zero under detailed balance, and the interpretation of larger values as stronger time asymmetry	Empirical $\mathcal{D}$ depends on coarse-graining (K), smoothing (α), and window length (L). Window-to-window variation reflects estimation choices as well as market conditions.
Convex budgeted problem and strong duality	The backtest replaces $\mathcal{D}(\cdot)$ by $\mathcal{D}_{\mathrm{IS}}(\cdot)$, which need not be convex in w, and evaluates the dual on a finite candidate set	The shadow-price logic of the Lagrangian and the comparative-statics benchmark provided by the idealized theory	Candidate-set dual gaps are internal consistency diagnostics. They need not vanish exactly and should not be read as proofs that finite-sample optimization satisfies exact strong duality.
Thermostat as a projected dual update	Projected update based on $\mathcal{D}_k^{\mathrm{IS}}$ and a rolling budget $\sigma_{\mathrm{bud},k}$ calibrated from the reference dissipation history	The directional interpretation of λ_k as a dissipation-pressure response	Imperfect one-window tracking is expected. The empirical question is whether λ_k rises when estimated dissipation pressure rises and relaxes when that pressure recedes.
Smooth stability and frontier geometry	Quantile microstates, Laplace smoothing, sensitivity checks in K, α, and L, and candidate-cloud frontier plots	The qualitative trade-off between exergy and dissipation remains meaningful	Finite windows need not display exact smoothness, monotonicity, or a globally convex frontier. These figures are numerical diagnostics of the trade-off, not exact geometric theorems for the realized sample.

Notes: The table serves as a compact "key deviation–acceptable reading" guide for Sections 4 and Appendix A.4. The theory defines the reference geometry of EDPO, whereas the backtest evaluates finite-sample approximations based on coarse-grained wealth dynamics, rolling-window estimation, and candidate-set diagnostics.

The purpose of Table 1 is interpretive. It can be read as a compact key deviation–acceptable reading guide for the empirical sections. The formal convexity, duality, and stability results above describe the reference geometry of the EDPO problem under fixed coarse-graining and regularity assumptions. The rolling-window backtest then asks whether the same qualitative comparative statics

remain visible when X_a and $\mathcal{D}$ are estimated from finite samples with quantile microstates, Laplace smoothing, and candidate-set diagnostics. Accordingly, empirical departures from exact convexity, exact strong duality, or perfectly smooth thermostat tracking are treated as finite-sample implementation features rather than as contradictions of the model.

3. Data and backtest protocol

This section specifies the empirical protocol used throughout the study and is aligned with the released implementation. All backtests are conducted on daily data over the period January 2015 to December 2025, using a rolling estimation window and piecewise-constant portfolio weights between rebalancing.

3.1. Data source and return construction

Daily closing prices are downloaded from Stooq (tickers in the form *asset.us*) using the daily CSV endpoint and cached locally. For each asset, the "Close" series is filtered to the sample period and merged into a single price panel via an outer join on dates. Assets are first screened by a coverage filter, retaining only series with at least 90% nonmissing observations in the raw merged panel. Remaining missing prices are forward-filled for up to five consecutive trading days, after which any rows containing missing values are dropped. Simple daily returns are then computed as $r_{t,i} = P_{t,i}/P_{t-1,i} - 1$; any nonfinite return values are removed.

3.2. Asset universes

The empirical study evaluates EDPO across four exchange-traded fund (ETF) universes designed to expose the method to different correlation structures: a diversified multiasset mix (ETF7), a set of United States (US) sector ETFs (SECTOR9), a fixed-income sleeve (BOND7), and a broad equity/style mix (EQUITY6). The constituents are listed in Table 2.

Table 2. Asset universes used in the backtests. Constituents are reported in Stooq ticker format.

Universe	Constituents
ETF7	spy.us, qqq.us, iwm.us, efa.us, eem.us, tlt.us, gld.us
SECTOR9	xlf.us, xlk.us, xle.us, xlv.us, xly.us, xlp.us, xli.us, xlu.us, xlb.us
BOND7	shy.us, ief.us, tlt.us, tip.us, agg.us, lqd.us, hyg.us
EQUITY6	spy.us, ijr.us, ive.us, ivw.us, iemg.us, efa.us

3.3. Rolling-window design and rebalancing

Let $L = 252$ denote the lookback length (one trading year), and let $\{t_k\}_{k\geq 1}$ be the rebalance dates, defined as the last available trading days in each calendar period (quarterly frequency in the baseline experiments). At each rebalance t_k, the strategy estimates a weight vector $\boldsymbol{w}_k$ from the training window $\{\boldsymbol{r}_t : t \leq t_k,\ t \text{ among the last } L \text{ trading days}\}$ and then holds $\boldsymbol{w}_k$ constant over the subsequent out-of-sample interval $t \in (t_k, t_{k+1}]$. This design avoids look-ahead bias because the portfolio performance is computed on returns strictly after the rebalance date.

3.4. Transaction costs

We incorporate proportional transaction costs using a linear turnover model. Let $\mathrm{TO}_k = \|\boldsymbol{w}_k - \boldsymbol{w}_{k-1}\|_1 = \sum_i |w_{k,i} - w_{k-1,i}|$ denote gross turnover at rebalance k. For a transaction cost rate c (in decimal form), the implementation cost is $\mathrm{TC}_k = c\,\mathrm{TO}_k$. In the backtest engine, TC_k is deducted from the first out-of-sample portfolio return after the rebalancing. Unless otherwise stated, the baseline cost rate is $c = 2$ basis points, and sensitivity is examined over a broader range in Section 4.8.

3.5. Strategies and hyperparameters

The EDPO objective and thermostat are defined in Section 2. Unless stated otherwise, EDPO uses risk aversion $a = 3.0$, dissipation penalty $\lambda = 0.75$, and EPR coarse-graining level $K = 7$, with Laplace smoothing $\alpha = 0.5$ for transition estimates and a numerical floor $\varepsilon = 10^{-6}$ in log-growth computations. The empirical implementation is long-only and fully invested, and we impose a cap $w_i \le w_{\max}$ (baseline $w_{\max} = 0.35$) to avoid excessive concentrations. To improve implementability, the optimization includes stabilizers: a turnover penalty with coefficient γ_{turnover}, an optional acceleration penalty γ_{accel}, and a KL divergence penalty γ_{KL} for the equal-weight portfolio.

Benchmarks include equal-weighting (EqualWeight), minimum variance (MinVar), risk parity / equal risk contribution (RiskParity), maximum Sharpe (MaxSharpe), inverse volatility (InvVol), and hierarchical risk parity (HRP).

3.6. Performance metrics

Unless otherwise specified, annualized return and volatility are computed from daily log returns $\log(1 + r_t)$ with an annualization factor of 252 trading days, and Sharpe ratios are reported without an explicit risk-free adjustment. The maximum drawdown is computed from the cumulative wealth path of simple returns.

4. Results

This section reports the out-of-sample performance of the exergy-dissipation portfolio optimization (EDPO) framework alongside standard benchmarks, using the rolling-window protocol in Section 3. Unless stated otherwise, EDPO uses the baseline hyperparameters in Section 3.5 and is evaluated across the four ETF universes in Table 2.

This section is structured as follows. We begin with a cross-universe scorecard to situate the EDPO relative to conventional allocations. We then inspect (i) cumulative growth dynamics in the representative ETF7 universe, (ii) the work–dissipation efficiency picture in the $(X_a, \mathcal{D})$ plane, (iii) evidence that the primal–dual thermostat tracks a dissipation budget, (iv) predictive tests assessing whether the reference dissipation contains incremental information for future realized volatility, and (v) sensitivity to proportional transaction costs.

The empirical sections should be read through the theory-to-implementation bridge in Section 2.10. The propositions in Section 2 are established for idealized functionals with fixed coarse-graining and exact dissipation budgets, whereas the backtest uses rolling, quantile-based plug-in estimators, and finite candidate-set diagnostics. For that reason, the empirical goal is not to verify exact dual equalities window by window, but to assess whether the qualitative patterns

predicted by the theory remain economically interpretable: higher λ when dissipation pressure rises, sensible shifts along the $(X_a, \mathcal{D})$ trade-off, and incremental risk information in $\mathcal{D}$.

For readability, Table 3 lists the abbreviations for the strategies and technical terms used throughout the Results section.

Table 3. List of abbreviations for strategies and technical terms used throughout the results section.

Abbreviation	Description
Framework and Metrics	
EDPO	Exergy-dissipation portfolio optimization
EPR	Entropy production rate
OOS	Out-of-sample
RV	Realized volatility
MSE	Mean squared error
KKT	Karush–Kuhn–Tucker
KL	Kullback–Leibler divergence
$w_{\max}$	Maximum single-asset weight (cap constraint)
γ_{KL}	KL-divergence regularization coefficient
$\gamma_{\mathrm{turnover}}$	Turnover regularization coefficient
Benchmark Strategies	
EqualWeight	Equally weighted $(1/N)$ portfolio
MinVar	Minimum variance portfolio
RiskParity	Risk parity portfolio
MaxSharpe	Maximum Sharpe ratio portfolio
InvVol	Inverse volatility portfolio
HRP	Hierarchical risk parity portfolio
MaxDiv	Maximum diversification portfolio
MinCVaR	Minimum conditional value-at-risk portfolio
EDPO Strategy Variants	
EDPO_Thermostat	EDPO with the thermostat mechanism enabled

4.1. Cross-universe scorecard

The primary objective of this subsection is to establish the competitiveness of the EDPO strategy and introduce the thermodynamic lens, characterized by the exergy $X_a(\mathbf{w})$ and the dissipation $\mathcal{D}(\mathbf{w})$, as a meaningful framework for portfolio analysis. Table 4 presents the performance metrics for the EDPO strategy and several benchmarks across four asset universes: ETF7 (a diversified mix of equity, bond, and commodity ETFs), SECTOR9 (US sector ETFs), BOND7 (fixed-income ETFs), and EQUITY6 (US equity ETFs). Figure 2 provides a visual comparison of the Sharpe ratios. In the sense of Table 1, the Sharpe-ratio comparisons reported here should be read as finite-sample diagnostics of the rolling-window, coarse-grained implementation rather than as exact realizations of the idealized convex program of Section 2.

The results in Table 4 indicate that the EDPO strategy is competitive across all four universes,

although it does not dominate the benchmarks. In the ETF7 universe, the EDPO strategy achieves the highest Sharpe ratio of 0.955, outperforming the EqualWeight (0.802), MinVar (0.802), and RiskParity (0.805) strategies. The annualised return of 10.58% is comparable to that of EqualWeight (10.87%), whereas the annualized volatility of 11.08% is lower than that of EqualWeight (13.55%). The maximum drawdown of −26.59% is similar to the benchmarks, suggesting that the EDPO strategy does not sacrifice downside protection for higher returns. The average rebalancing turnover of 0.251 (approximately 25% gross weight change per quarter) is moderate, indicating that the strategy does not require excessive trading.

Table 4. Cross-universe performance metrics for the EDPO strategy and benchmarks. The table reports the annualized return, annualized volatility, Sharpe ratio, maximum drawdown, and average turnover. Turnover is computed at rebalancing dates from changes in target weights (ignoring intraperiod drift), that is, $\mathrm{TO}_k = \sum_{i=1}^{N} |w_{i,k} - w_{i,k-1}|$, and the reported value is the sample mean of TO_k. The model-based thermodynamic diagnostics X_a and $\mathcal{D}$ are reported as per-period (daily) rates computed from the same return sampling frequency as the backtest, and their magnitudes therefore depend on the chosen coarse-graining and estimation window; they are not directly comparable in scale to the annualized return/volatility columns. The sample period is January 2015 to December 2025 with quarterly rebalancing.

Universe	Strategy	Ann. Return	Ann. Vol.	Sharpe	Max DD	Turnover	X_a	$\mathcal{D}$
ETF7	EDPO	0.1058	0.1108	0.955	-0.266	0.251	0.000326	0.0112
ETF7	EDPO_Thermostat	0.1003	0.1117	0.898	-0.263	0.285	0.000305	0.0154
ETF7	EqualWeight	0.1087	0.1355	0.802	-0.269	0.000	0.000300	0.0117
ETF7	MaxSharpe	0.1057	0.1439	0.734	-0.249	0.812	0.000275	0.0101
ETF7	MinVar	0.0798	0.0994	0.802	-0.269	0.274	0.000246	0.0131
ETF7	RiskParity	0.0868	0.1078	0.805	-0.265	0.091	0.000261	0.0099
SECTOR9	EDPO	0.1186	0.1907	0.622	-0.404	0.345	0.000225	0.0211
SECTOR9	EDPO_Thermostat	0.1154	0.1825	0.633	-0.373	0.347	0.000233	0.0203
SECTOR9	EqualWeight	0.1246	0.1723	0.723	-0.369	0.000	0.000287	0.0149
SECTOR9	MaxSharpe	0.1414	0.1993	0.710	-0.334	0.841	0.000287	0.0218
SECTOR9	MinVar	0.0731	0.1491	0.490	-0.359	0.314	0.000146	0.0111
SECTOR9	RiskParity	0.1189	0.1640	0.725	-0.359	0.049	0.000284	0.0216
BOND7	EDPO	0.0109	0.0532	0.204	-0.204	0.291	0.000026	0.0150
BOND7	EDPO_Thermostat	0.0048	0.0547	0.088	-0.243	0.335	0.000001	0.0091
BOND7	EqualWeight	0.0053	0.0603	0.088	-0.236	0.000	-0.000001	0.0091
BOND7	MaxSharpe	-0.0197	0.0650	-0.302	-0.291	0.904	-0.000104	0.0174
BOND7	MinVar	0.0034	0.0169	0.204	-0.066	0.063	0.000012	0.0164
BOND7	RiskParity	0.0036	0.0391	0.093	-0.161	0.089	0.000005	0.0132
EQUITY6	EDPO	0.1003	0.1751	0.573	-0.349	0.357	0.000194	0.0115
EQUITY6	EDPO_Thermostat	0.0981	0.1772	0.554	-0.344	0.332	0.000182	0.0054
EQUITY6	EqualWeight	0.1133	0.1769	0.640	-0.349	0.000	0.000237	0.0058
EQUITY6	MaxSharpe	0.1180	0.1883	0.627	-0.351	0.990	0.000230	0.0136
EQUITY6	MinVar	0.0979	0.1624	0.603	-0.334	0.478	0.000211	0.0101
EQUITY6	RiskParity	0.1128	0.1748	0.645	-0.347	0.041	0.000240	0.0099

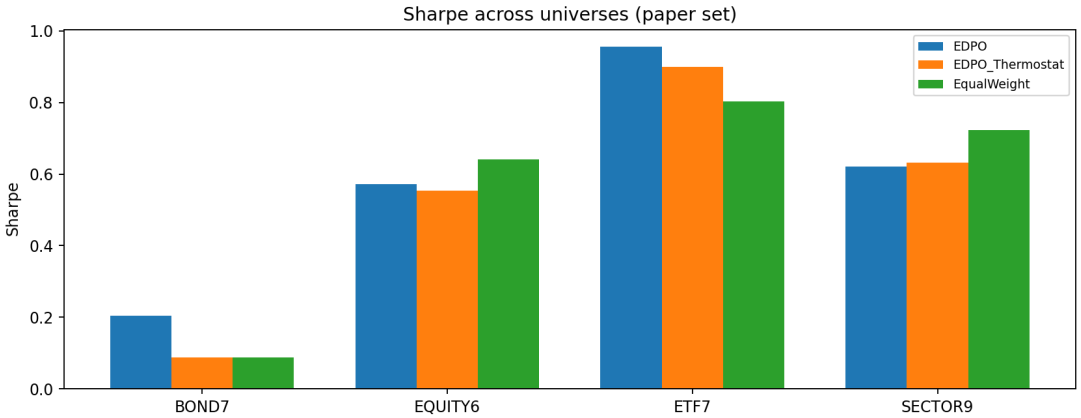


Figure 2. Sharpe ratios for the EDPO strategy and benchmarks across four asset universes. The EDPO strategy achieves the highest Sharpe ratio in the ETF7 universe (0.955), demonstrating its effectiveness in diversified settings. Performance is more nuanced in the SECTOR9, BOND7, and EQUITY6 universes, where the strategy remains competitive but does not universally dominate.

In the SECTOR9 universe, which consists of highly correlated US sector ETFs, the EDPO strategy achieves a Sharpe ratio of 0.622, which is lower than those of EqualWeight (0.723) and RiskParity (0.725). This result is expected because the EDPO framework is designed to exploit diversification benefits, which are limited in a universe of highly correlated assets. The BOND7 universe presents a challenging environment for all strategies because of the low-return, low-volatility nature of fixed-income assets. The EDPO strategy achieves a Sharpe ratio of 0.204, which is comparable to that of MinVar (0.204) but higher than that of EqualWeight (0.088). In the EQUITY6 universe, the EDPO strategy achieves a Sharpe ratio of 0.573, which is lower than the Sharpe ratios of EqualWeight (0.640) and RiskParity (0.645).

The key insight from Table 4 is that the value of the EDPO framework lies not in delivering the highest Sharpe ratio under all market conditions, but in providing an interpretable and controllable trade-off between risk and return. The exergy X_a and the EPR $\mathcal{D}$ provide a thermodynamic lens for understanding portfolio dynamics, which is complementary to the traditional mean-variance framework. The EDPO strategy is particularly effective in diversified universes (for example, ETF7), where thermodynamic constraints can be meaningfully exploited.

4.2. Empirical evidence package: hyperparameter stability and exploratory inference

To address the empirical concerns about sensitivity and inferential robustness, we add a compact empirical-evidence package built around two questions. First, are the main conclusions materially altered when the coarse-graining level K, the Laplace-smoothing parameter α, and the lookback length L are varied around the baseline configuration? Second, do the headline performance and forecasting differences retain signal when serial dependence and multiplicity are treated explicitly? Tables 5 and 6 summarize these additions, and Appendix Figures 13 and 14 provide the corresponding cross-universe visual diagnostics.

Table 5. Baseline hyperparameters and stability summary across universes.

Hyperparameter	Baseline	Tested values	Sharpe range	Main implication
Coarse-graining K	7	$\{5, 7, 9, 11\}$	0.136 (med.), 0.445 (max)	Largest source of variation; the sign of the main empirical message is preserved, but the magnitude is universe dependent.
Laplace smoothing α	0.5	$\{0.5, 1.0\}$	0.036 (med.), 0.152 (max)	Smallest effect; stronger smoothing changes the Sharpe ratios only modestly in most universes.
Lookback L (days)	252	$\{126, 252, 504\}$	0.116 (med.), 0.178 (max)	Moderate effect; a one-year window remains a balanced compromise between responsiveness and stability.

Notes: The summary pools the one-factor-at-a-time sensitivity runs for EDPO and EDPO_Thermostat across ETF7, SECTOR9, BOND7, and EQUITY6. The reported Sharpe range is max(Sharpe) − min(Sharpe) within a given hyperparameter sweep. The coarse-graining level K generates the largest cross-universe variation, followed by the lookback length L, whereas the smoothing parameter α is the least influential. Appendix Figure 13 visualizes the K-sensitivity curves.

Table 6. Exploratory inference summary across universes.

Panel A: paired moving-block bootstrap for $\Delta SR = SR_{\mathrm{EDPO}} - SR_{\mathrm{EqualWeight}}$

Universe	ΔSR	95% CI low	95% CI high	q_{BH}
ETF7	0.114	-0.082	0.317	0.320
SECTOR9	-0.155	-0.323	0.008	0.155
BOND7	0.145	-0.020	0.326	0.155
EQUITY6	-0.017	-0.116	0.069	0.680

Panel B: OOS forecast gain for RV + $\mathcal{D}^{\mathrm{ref}}$ relative to RV

Universe	ΔMSE (%)	HAC p	q_{BH}	MBB p
ETF7	1.99	0.259	0.518	0.190
SECTOR9	1.62	0.553	0.553	0.557
BOND7	2.21	0.235	0.518	0.203
EQUITY6	-0.91	0.462	0.553	0.505

Notes: Panel A reports the prespecified EDPO-versus-EqualWeight contrast in each universe using the manuscript's paired moving-block bootstrap design. The reported q_{BH} values are Benjamini–Hochberg adjustments across the four universe-level Sharpe comparisons. Panel B reports the percentage reduction in out-of-sample mean squared error (MSE) from augmenting the realized-volatility forecast with the reference EPR signal together with HAC and moving-block-bootstrap (MBB) p-values for the loss differential. Positive ΔMSE values favor the augmented model. All p-values are interpreted as exploratory rather than confirmatory.

The sensitivity summary indicates a clear ordering. The coarse-graining level K is the dominant modelling choice, the lookback length L has a moderate effect, and the Laplace-smoothing parameter α is the least influential within the successful runs. In practical terms, the baseline configuration $(K, \alpha, L) = (7, 0.5, 252)$ should be interpreted as a stable reference specification rather than as a finely tuned optimum. The cross-universe K-sensitivity curves reported in Appendix Figure 13 show that the strongest movement occurs in ETF7 and SECTOR9, whereas the bond and equity sleeves are more stable. Taken together, the inference package supports an economically informative but statistically cautious reading of the results. The EDPO strategy remains competitive relative to EqualWeight in ETF7 and BOND7, but the paired bootstrap intervals for the prespecified universe-level contrasts generally include zero, and the Benjamini–Hochberg adjusted q-values remain above conventional significance thresholds. Similarly, the augmented RV + $\mathcal{D}^{\mathrm{ref}}$ forecast lowers the out-of-sample mean squared error in ETF7, SECTOR9, and BOND7, yet the HAC and moving-block-bootstrap tests

remain exploratory and do not support a strong claim of universal predictive superiority.

4.3. Bootstrap inference for Sharpe-ratio differences

The scorecard in Table 4 compares annualized Sharpe ratios across strategies, but standard i.i.d. approximations are ill-suited to backtest returns that exhibit serial dependence and heteroskedasticity. We therefore supplement the main comparison with a paired moving-block bootstrap that resamples blocks of out-of-sample daily returns jointly across the two strategies being compared, thereby preserving both temporal dependence and the contemporaneous dependence created by common market shocks.

For the main text, Table 6 reports the prespecified EDPO-versus-EqualWeight contrast in each universe together with Benjamini–Hochberg adjusted q-values. The point estimates broadly agree with the descriptive ranking in Table 4: ETF7 and BOND7 retain positive ΔSR values, whereas SECTOR9 and EQUITY6 are weaker relative to EqualWeight. However, the percentile confidence intervals generally straddle zero, and the adjusted q-values remain at 0.155 or above. Accordingly, the Sharpe differences are best interpreted as exploratory evidence that the EDPO family is competitive rather than as conclusive proof of universal outperformance.

4.4. Growth dynamics

This subsection examines the cumulative log returns of the EDPO strategy and benchmarks in the ETF7 universe, which serves as the representative universe for the main analysis. Figure 3 displays the cumulative log-returns over the sample period.

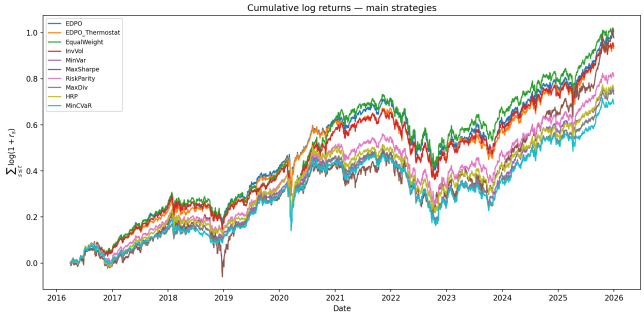


Figure 3. Cumulative log-returns for the EDPO strategy and benchmarks in the ETF7 universe. For a rebalanced strategy with piecewise-constant weights $\{\mathbf{w}_t\}$, the cumulative log-return is $\sum_{t=1}^{T} g_t(\mathbf{w}_t) = \sum_{t=1}^{T} \log(1 + \mathbf{w}_t^\top \mathbf{R}_t)$, where $\mathbf{w}_t$ denotes the weight vector applied to the asset return vector $\mathbf{R}_t$ over period t. The EDPO strategy demonstrates a stable growth trajectory, keeping pace with the benchmarks during normal market conditions and exhibiting resilience during stress periods.

The cumulative log-return (as defined in Figure 3) is aligned with the EDPO focus on log-wealth accumulation under periodic rebalancing. Figure 3 shows that EDPO maintains a stable growth trajectory over the sample period, broadly keeping pace with EqualWeight and RiskParity. Notably, the EDPO path does not exhibit disproportionate drawdowns during well-known stress episodes (for example the COVID-19 sell-off in March 2020), consistent with the interpretation that dissipation control reduces exposure to time-irreversible dynamics during rapid regime transition.

Economically, Figure 3 suggests that the EDPO can navigate changing regimes without requiring accurate short-horizon return forecasts. The emphasis on dissipation control, rather than on point estimates of expected return, naturally favors allocations whose growth profiles are less sensitive to nonstationarity and regime-dependent rotations.

4.5. Thermodynamic efficiency picture

The central contribution of the EDPO framework is the introduction of a thermodynamic lens for portfolio analysis, characterized by the trade-off between exergy $X_a(\mathbf{w})$ and dissipation $\mathcal{D}(\mathbf{w})$. Figure 4 presents the work-dissipation map for the ETF7 universe, which plots each strategy as a point in the $(X_a, \mathcal{D})$ space.

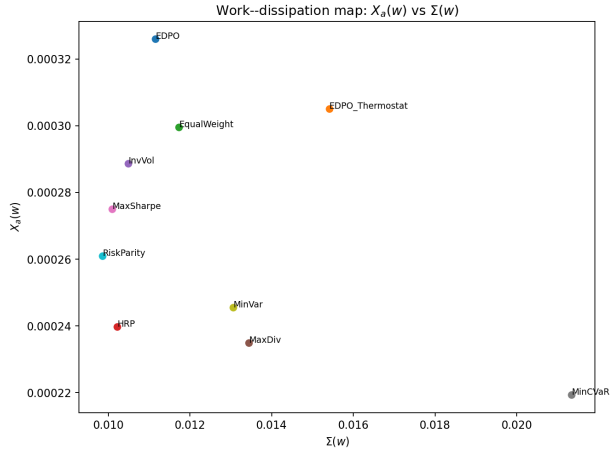


Figure 4. Work–dissipation map for the ETF7 universe. Each strategy is represented as a point in the $(X_a, \mathcal{D})$ space, where X_a denotes exergy (risk-sensitive growth potential), and $\mathcal{D}$ denotes dissipation (irreversibility of the portfolio dynamics). Strategies closer to the upper-left region are preferred because they offer higher exergy for a given level of dissipation. In this universe, EDPO improves X_a relative to several benchmarks at comparable $\mathcal{D}$, whereas other heuristics (for example, RiskParity) can attain lower $\mathcal{D}$ at lower X_a, illustrating a genuine efficiency trade-off rather than universal dominance.

The work dissipation map in Figure 4 provides a visual representation of the thermodynamic efficiency of different portfolio strategies. The horizontal axis represents the dissipation $\mathcal{D}(\mathbf{w})$, which measures the degree of irreversibility in portfolio dynamics. Lower values of $\mathcal{D}$ indicate that the portfolio is less sensitive to the time-irreversible components of the return process, which is desirable from a risk-management perspective. The vertical axis represents the exergy $X_a(\mathbf{w})$, which measures the risk-sensitive growth potential of the portfolio. Higher values of X_a indicate that the portfolio has a higher expected log growth rate, adjusted for risk aversion.

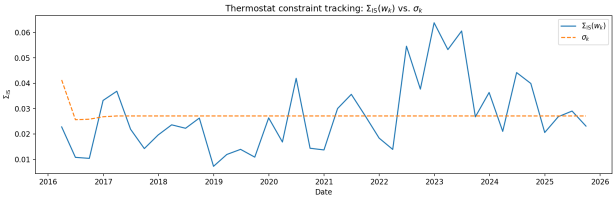
Within the finite set of strategies considered, the EDPO sits near the upper-left (Pareto-efficient) boundary of the point cloud in the $(X_a, \mathcal{D})$ plane, indicating a favorable trade-off between exergy and dissipation in this universe. This is consistent with the objective of the strategy to maximize exergy while penalizing dissipation. Several heuristic benchmarks (for example, EqualWeight and RiskParity) occupy nearby regions but with visibly different efficiency profiles (either lower X_a for a similar $\mathcal{D}$, or lower $\mathcal{D}$ at the cost of lower X_a). By contrast, MaxSharpe, which targets a mean-variance proxy

without regard to dissipation, typically appears at a higher dissipation level.

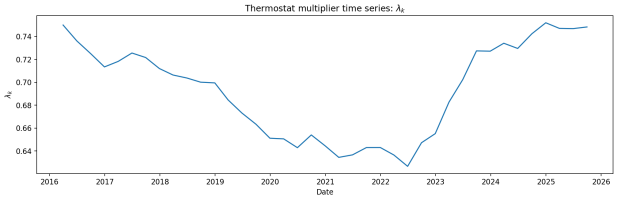
In portfolio terms, the work–dissipation map complements the mean-variance view by separating dispersion from time-irreversibility. Dissipation $\mathcal{D}$ quantifies exposure to time-asymmetric return dynamics that tend to intensify in stress and rotation regimes; penalizing $\mathcal{D}$ therefore targets a distinct risk dimension that is not captured by the variance alone.

4.6. Dissipation-budget control

The EDPO framework incorporates a dissipation-budget control mechanism that adjusts the Lagrange multiplier λ in response to changes in market conditions. This subsection provides evidence for the effectiveness of this mechanism. Figure 5 displays the sigma tracking and the λ time series for the ETF7 universe, and Table 7 summarizes the thermostat Karush–Kuhn–Tucker (KKT) conditions across all universes.



(a) Dissipation budget tracking: the in-sample EPR $\mathcal{D}_{\text{IS}}$ versus the target budget σ_{bud}.



(b) Time series of the Lagrange multiplier λ_t, which adjusts in response to dissipation pressure.

Figure 5. Thermostat evidence for the ETF7 universe. The left panel shows the dissipation budget tracking, comparing the in-sample EPR $\mathcal{D}_{\text{IS}}$ to the target budget σ_{bud}. The right panel shows the time series of the Lagrange multiplier λ_t, updated by the primal–dual rule $\lambda_{k+1} = \Pi[\lambda_k + \eta(\mathcal{D}_{\text{IS}}(\mathbf{w}_k) - \sigma_{\text{bud}})]$, where Π denotes projection onto a feasible interval, and η is the step size.

Table 7. Thermostat KKT summary across universes. The table reports the binding rate (fraction of periods where the dissipation constraint is active), mean Lagrange multiplier λ, mean dissipation budget σ_{bud}, and mean in-sample EPR $\mathcal{D}_{\text{IS}}$.

Universe	Binding rate	Mean λ	Mean σ_{bud}	Mean $\mathcal{D}_{\text{IS}}$
ETF7	0.385	0.693	0.0274	0.0272
SECTOR9	0.179	0.531	0.0464	0.0321
BOND7	0.179	0.578	0.0477	0.0363
EQUITY6	0.333	0.644	0.0492	0.0459

The dissipation-budget control mechanism operates according to the primal–dual update rule,

$$\lambda_{k+1} = \Pi\left[\lambda_k + \eta\left(\mathcal{D}_{\text{IS}}(\mathbf{w}_k) - \sigma_{\text{bud}}\right)\right]. \tag{4.1}$$

Figure 5a shows that the dissipation-budget control mechanism is effective in tracking the target budget σ_{bud}. The in-sample EPR $\mathcal{D}_{\text{IS}}$ fluctuates around the budget with occasional exceedances during periods of market stress. Figure 5b shows the corresponding time series of the Lagrange multiplier λ_t, which increases during periods of high dissipation pressure (for example, the COVID-19 crash in 2020 and the rate-hiking cycle in 2022–2023) and decreases during calmer periods.

Operationally, the multiplier λ acts as a shadow price for the dissipation budget. When in-sample dissipation rises relative to σ_{bud}, the thermostat increases λ, and the optimizer shifts toward lower-$\mathcal{D}$ allocations. This feedback loop offers a transparent "risk budget" interpretation while remaining agnostic about explicit regime labelling. Consistently with Table 1, the path of λ_t is therefore interpreted as a shadow-price-like response of the finite-sample algorithm to estimated dissipation pressure, rather than as an exact dual trajectory of the continuous strong-duality problem.

Table 7 reports that the binding rate (the fraction of rebalances at which the dissipation constraint is active) varies across universes, ranging from 17.9% in SECTOR9 and BOND7 to 38.5% in ETF7. The mean multiplier λ is highest in ETF7 (0.693), indicating that the optimizer requires a stronger dissipation penalty in the diversified setting; correspondingly, the mean in-sample EPR $\mathcal{D}_{\text{IS}}$ is held at low levels (0.027).

4.7. Predictive content of dissipation

This subsection examines whether the entropy production rate (EPR) contains predictive information regarding future risk. The analysis compares an out-of-sample (OOS) realized-volatility (RV) forecast with an augmented specification that adds the reference EPR signal as a state variable. The cross-universe forecast-gain summary is reported in Panel B of Table 6, and Figure 6 displays the forecast time series for the ETF7 universe.

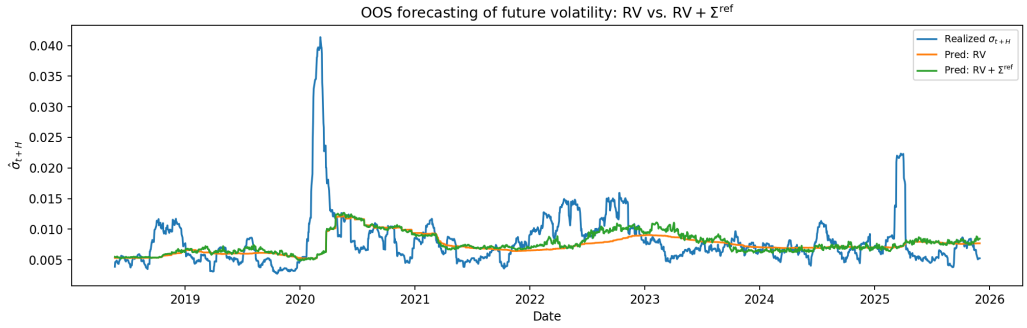


Figure 6. Out-of-sample forecast time series for the ETF7 universe. The figure compares realized volatility with the forecasts from the baseline RV and augmented RV with EPR specifications. The augmented model tracks several volatility spikes more closely, but the figure is intended as an economic illustration only. Formal forecast-gain inference is reported in Table 6 and should be interpreted as descriptive/exploratory rather than confirmatory because the forecast horizon overlaps, and multiple universe-level comparisons are examined.

The augmented specification lowers the OOS mean squared error in ETF7, SECTOR9, and BOND7 by 1.99%, 1.62%, and 2.21%, respectively, whereas it slightly underperforms in EQUITY6 (−0.91%). These forecasting comparisons should nevertheless be read cautiously. Because the target uses overlapping 21-day horizons, and the manuscript reports multiple universe-level comparisons, the HAC and moving-block-bootstrap p-values are reported only as descriptive/exploratory diagnostics and should not be interpreted as confirmatory tests. The empirical message is therefore that the EPR can be a useful complementary state variable for volatility forecasting, but not a uniformly dominant predictor across universes.

One interpretation, consistent with the nonequilibrium lens, is that elevated EPR signals time-irreversible behavior associated with stress and rotation regimes. In such periods, the return distribution is less stationary, and variance-based summaries can become less informative. Treating the EPR as an additional state variable therefore provides a simple way to condition volatility forecasts on "arrow-of-time" dynamics; the modest gains reported in Table 6 should nevertheless be read as suggestive rather than definitive. Figure 6 provides a representative visual comparison for ETF7. The augmented RV with EPR model tracks several realized-volatility spikes more closely than the baseline RV forecast, especially around the major stress episodes in 2020 and 2022–2023, but the figure is intended as an economic illustration rather than a standalone statistical test.

HAC-robust predictive regression. To complement the OOS MSE ratios, we estimate a predictive regression of future realized volatility on contemporaneous realized volatility and the reference EPR signal:

$$\mathrm{RV}_{t,H}^{\mathrm{future}} = \beta_0 + \beta_{\mathrm{RV}}\,\mathrm{RV}_t + \beta_{\mathcal{D}}\,\mathcal{D}_t^{\mathrm{ref}} + u_t, \tag{4.2}$$

where $\mathrm{RV}_{t,H}^{\mathrm{future}}$ is the realized volatility over the next H trading days (here, $H = 21$), and $\mathcal{D}_t^{\mathrm{ref}}$ is the EPR of the equal-weight reference portfolio computed at time t from the same universe. Because both volatility and the constructed EPR series are persistent, and the future-volatility target is built from overlapping windows, inference is based on a heteroskedasticity and autocorrelation consistent (HAC) covariance estimator with five lags. Consistently with the broader inference discussion, the resulting HAC p-values are used only as descriptive diagnostics of conditional association.

Table 8. Predictive regression for future realized volatility (exploratory HAC diagnostics).

Universe	Regressor	Coef.	HAC SE	z-stat	p-value
ETF7					
	Intercept	0.003524	0.0007488	4.706	0.000
	$\mathcal{D}_t^{\mathrm{ref}}$	0.01518	0.002969	5.114	0.000
	RV_t	0.2776	0.06964	3.987	0.000
SECTOR9					
	Intercept	0.001902	0.001301	1.462	0.144
	$\mathcal{D}_t^{\mathrm{ref}}$	0.02807	0.01032	2.720	0.007
	RV_t	0.3414	0.05813	5.872	0.000
BOND7					
	Intercept	0.001987	0.0003208	6.192	0.000
	$\mathcal{D}_t^{\mathrm{ref}}$	-0.005787	0.00125	-4.629	0.000
	RV_t	0.5861	0.05239	11.188	0.000
EQUITY6					
	Intercept	0.006622	0.001218	5.436	0.000
	$\mathcal{D}_t^{\mathrm{ref}}$	-0.001197	0.006709	-0.178	0.858
	RV_t	0.2771	0.06329	4.379	0.000

Notes: The dependent variable is $\mathrm{RV}_{t,H}^{\mathrm{future}}$ and the regressors are contemporaneous realized volatility RV_t and the reference EPR signal $\mathcal{D}_t^{\mathrm{ref}}$. Standard errors are HAC-robust with five lags (reported to three decimals; 0.000 indicates $p < 0.0005$). Because the dependent variable uses overlapping H-day horizons, and the manuscript reports four parallel universe-level comparisons, the reported HAC p-values should be read only as descriptive/exploratory diagnostics and not as confirmatory tests. Each universe regression uses $N = 2492$ observations. The R^2 values are as follows: ETF7: 0.060; SECTOR9: 0.088; BOND7: 0.235; EQUITY6: 0.033.

The coefficient on $\mathcal{D}_t^{\mathrm{ref}}$ is positive, with comparatively small HAC p-values in ETF7 ($\hat{\beta}_{\mathcal{D}} = 0.01518$, $p < 0.001$) and SECTOR9 ($\hat{\beta}_{\mathcal{D}} = 0.02807$, $p = 0.007$), suggesting incremental conditional association with future volatility after controlling for RV_t. In BOND7, the estimated coefficient is negative ($\hat{\beta}_{\mathcal{D}} = -0.005787$, $p < 0.001$), which is consistent with a universe-specific link between the reference irreversibility measure and subsequent volatility. In EQUITY6, $\hat{\beta}_{\mathcal{D}} = -0.001197$ is not statistically distinguishable from zero ($p = 0.858$). These estimates should be read as descriptive conditional correlations rather than causal or confirmatory statistical evidence. In future work, cross-validation, sample-splitting, and permutation or randomization tests would provide a stronger basis for confirmatory assessment under overlapping horizons and multiple comparisons.

Nonparametric cross-check. As a complementary diagnostic, Table 9 reports the mean future volatility and mean future drawdown by quartiles of $\mathcal{D}_t^{\mathrm{ref}}$. The ETF7 and SECTOR9 universes exhibit higher average future volatility in the top EPR quartile than in the bottom quartile, whereas the pattern is reversed in BOND7 and is weak in EQUITY6. This stratification supports the regression evidence while highlighting that the sign and strength of the relationship between irreversibility and future risk can be universe dependent.

Table 9. Future risk by reference EPR quartile.

Universe	EPR quartile	Mean future vol.	Mean future MDD	N
ETF7				
	Q1	0.006492	-0.022793	623
	Q2	0.007636	-0.030743	623
	Q3	0.007532	-0.030712	623
	Q4	0.008017	-0.035151	623
SECTOR9				
	Q1	0.007913	-0.031140	623
	Q2	0.007909	-0.030988	623
	Q3	0.008002	-0.027670	623
	Q4	0.011089	-0.046804	623
BOND7				
	Q1	0.003979	-0.017161	623
	Q2	0.003625	-0.016369	623
	Q3	0.002968	-0.013384	623
	Q4	0.002879	-0.012513	623
EQUITY6				
	Q1	0.009194	-0.038451	623
	Q2	0.009248	-0.036395	623
	Q3	0.009759	-0.037582	623
	Q4	0.009429	-0.036845	623

Notes: Quartiles are computed within each universe using the empirical distribution of the reference EPR signal $\mathcal{D}_t^{\mathrm{ref}}$. "Mean future vol." and "Mean future maximum drawdown (MDD)" are averages of the realized future-risk measures over the next $H = 21$ trading days conditional on the quartile at time t.

4.8. Implementation realism

This subsection examines the sensitivity of the EDPO strategy to transaction costs, which are critical for practical implementation. Figure 7 displays the Sharpe ratio as a function of transaction costs, measured in basis points (bps), for the EDPO and EDPO_Thermostat strategies in the ETF7 universe.

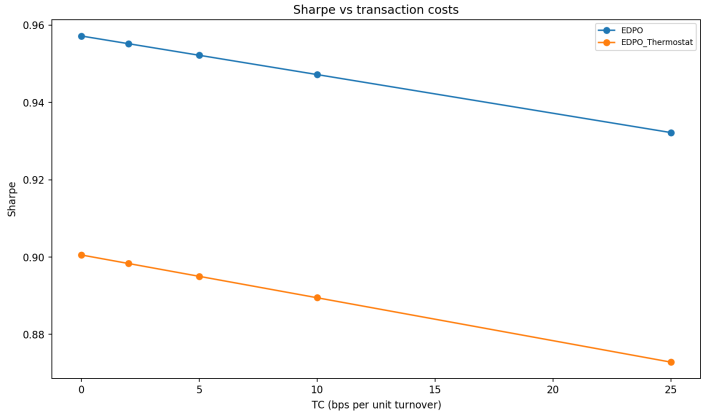


Figure 7. Transaction cost sensitivity for the ETF7 universe. The figure shows the Sharpe ratio as a function of transaction costs (in basis points) for the EDPO and EDPO_Thermostat strategies. The EDPO strategy remains viable within a plausible range of transaction costs (0–25 bps), with a Sharpe ratio decline of approximately 2.5% at 25 bps.

The transaction cost model assumes a linear cost structure, where the total transaction cost is proportional to the portfolio turnover (computed at the rebalancing instants):

$$\mathrm{TC}_t = c \cdot \sum_{i=1}^{N} |w_{i,t} - w_{i,t-1}|, \tag{4.3}$$

where c is the transaction cost rate (in basis points), and $w_{i,t}$ is the *target* weight of asset i immediately after rebalancing t. Under this convention, the gross turnover term $\sum_i |w_{i,t} - w_{i,t-1}|$ is the same object reported (on average) in Table 4. The net return is then computed as $r_t^{\mathrm{net}} = r_t^{\mathrm{gross}} - \mathrm{TC}_t$.

Figure 7 shows that the EDPO strategy remains viable within a plausible range of transaction costs. At zero transaction costs, the EDPO strategy achieves a Sharpe ratio of 0.957. At 25 basis points, the Sharpe ratio declines to 0.932, representing a reduction of approximately 2.6%. The EDPO_Thermostat strategy exhibits a similar pattern, with a Sharpe ratio declining from 0.901 at zero costs to 0.873 at 25 basis points.

The relatively low sensitivity to transaction costs can be attributed to the moderate turnover of the EDPO strategy (0.251 on average), which is controlled by the turnover regularization term in the objective function and the Kullback–Leibler (KL) divergence penalty that encourages smooth weight transitions. The cap constraint ($w_{\max} = 0.35$) also helps reduce extreme weight changes, which would otherwise increase turnover and transaction costs.

Practically, these results suggest that the EDPO is implementable for liquid ETFs under common proportional-cost assumptions. The focus on dissipation control, together with the stabilizers, naturally limits turnover and, therefore, reduces sensitivity to transaction costs.

Limitations and interpretation. These caveats are also the reason the theory in Section 2.10 should be read as a conceptual reference framework for the empirical algorithm, rather than as a claim that every rolling window inherits the exact idealized properties. First, X_a and $\mathcal{D}$ depend on finite-sample estimates (including the coarse-graining level K used for EPR estimation); therefore, their absolute values should be interpreted as model-based thermodynamic diagnostics rather than literal physical measurements or universal market constants. Second, the thermostat and duality diagnostics involve in-sample objects ($\mathcal{D}_{\mathrm{IS}}$ and the dual computed over a candidate set of feasible portfolios) that need not coincide in scale with the realized, backtest-aggregated quantities. Therefore, for publication-quality claims, the emphasis should be placed on robust comparative statements (for example monotone shifts under stress, consistent trade-offs in the $(X_a, \mathcal{D})$ plane, and out-of-sample performance stability) rather than on any single point estimate.

5. Conclusions and future work

EDPO reframes portfolio optimization as a nonequilibrium problem in which growth potential (exergy) and time-irreversible risk (dissipation) are jointly managed. Empirically, the 2015–2025 rolling backtests show that EDPO is competitive with standard benchmarks across four ETF universes and is particularly effective in the diversified ETF7 setting, where it achieves the highest Sharpe ratio (0.955) with moderate quarterly turnover (0.251) (Section 4.1). The thermodynamic lens is also informative beyond performance: stress episodes coincide with elevated $\mathcal{D}$, the thermostat increases λ when the in-sample dissipation pressure rises (Section 4.6), and the EPR carries incremental predictive content for realized volatility in three of four universes (Section 4.7). The new empirical-evidence package adds a more cautious interpretation to these findings: the results are most sensitive to the coarse-graining choice K, less sensitive to the lookback length L, and least sensitive to the smoothing parameter α, whereas the bootstrap and HAC statistics are exploratory rather than confirmatory.

From a contribution standpoint, this study provides (i) a tractable long-only EDPO implementation with stabilizers and a dissipation-budget thermostat, (ii) interpretable model-based thermodynamic diagnostics, such as the work–dissipation map and candidate-set duality checks, and (iii) evidence that incorporating irreversibility as a risk state can improve the risk/return trade-off in settings where diversification benefits are meaningful.

The theoretical results should be read in the same spirit. They provide a reference geometry for interpreting the empirical algorithm—how exergy, dissipation, budgets, and shadow prices should interact under idealized assumptions—but the rolling-window implementation replaces these objects by coarse-grained plug-in estimators and candidate-set diagnostics. Hence, finite-sample discrepancies, such as nonzero candidate-set dual gaps or imperfect thermostat tracking, are informative implementation diagnostics about estimation and execution rather than evidence against the conceptual model. Accordingly, the empirical quantities X_a and $\mathcal{D}$ should be read as model-based thermodynamic diagnostics whose magnitudes depend on the coarse-graining choice, sample window, and estimation procedure, rather than as literal physical measurements or universal constants of markets.

Several future directions follow naturally. **(1) Stronger confirmatory inference.** The paired moving-block bootstrap and HAC tests reported here are informative but exploratory; larger samples,

prespecified comparison sets, and alternative model-comparison procedures would sharpen statistical conclusions under serial dependence and overlapping forecast targets. **(2) EPR estimation and coarse-graining.** Alternative discretizations (adaptive bins, continuous-state estimators, or model-based state-space approaches) could reduce estimation noise and clarify the economic meaning of $\mathcal{D}$ across universes. **(3) Multiperiod and execution-aware control.** Embedding EDPO in a multiperiod setting with explicit transaction-cost and market-impact models would connect the thermostat to realistic execution constraints. **(4) Broader constraint sets and universes.** Extensions to leverage, limited shorting, and larger cross-asset universes would test scalability and robustness while also enabling comparisons to modern robust and distributionally robust optimization methods.

Use of Generative-AI tools declaration

The author declares he has used Large Language Models (LLMs) in the creation of this article. Specifically, OpenAI's ChatGPT (GPT-5.3, 2026 version) and Writefull for Overleaf (version 2025) were used.

The AI tools were employed solely for language polishing, including grammar correction, improving sentence structure, enhancing clarity, and organizing the presentation of content. AI-assisted language refinements were applied across all sections of the manuscript. The author reviewed and approved all modifications.

Acknowledgments

The author would like to acknowledge the Deanship of Graduate Studies and Scientific Research, Taif University for funding this work.

Conflict of interest

The author declares that there is no conflict of interest.

References

1. H. Markowitz, Portfolio selection, *J. Finance*, **7** (1952), 77–91.
2. J. L. Kelly, A new interpretation of information rate, *Bell Syst. Technical J.*, **35** (1956), 917–926.
3. C. Kardaras, S. Robertson, Robust maximization of asymptotic growth, *Ann. Appl. Probab.*, **22** (2012), 1576–1610. http://doi.org/10.1214/11-AAP802
4. N. Rujeerapaiboon, D. Kuhn, W. Wiesemann, Robust growth-optimal portfolios, *Manag. Sci.*, **62** (2016), 2090–2109. https://doi.org/10.1287/mnsc.2015.2228
5. T. Leung, H. Park, H. Yeo, Robust long-term growth rate of expected utility for leveraged ETFs, *Math. Finan. Econ.*, **18** (2024), 671–705. https://doi.org/10.1007/s11579-024-00371-1
6. P. J. Mercurio, Y. Wu, H. Xie, An entropy-based approach to portfolio optimization, *Entropy*, **22** (2020), 332. https://doi.org/10.3390/e22030332

7. L. MacLean, L. Yu, Y. Zhao, A generalized entropy approach to portfolio selection under a hidden markov model, *J. Risk Finan. Manag.*, **15** (2022), 337. https://doi.org/10.3390/jrfm15080337
8. R. Flanagan, L. Lacasa, Irreversibility of financial time series: A graph-theoretical approach, *Phys. Lett. A*, **380** (2016), 1689–1697. https://doi.org/10.1016/j.physleta.2016.03.011
9. M. Zanin, D. Papo, Algorithmic approaches for assessing irreversibility in time series: Review and comparison, *Entropy*, **23** (2021), 1474. https://doi.org/10.3390/e23111474
10. M. Zanin, A. Rodríguez-González, E. Menasalvas Ruiz, D. Papo, Assessing time series reversibility through permutation patterns, *Entropy*, **20** (2018), 665. https://doi.org/10.3390/e20090665
11. T. Afzal, M. A. Afridi, N. J. Muhammad, Integrating lstm with fama-french six factor model for predicting portfolio returns: Evidence from shenzhen stock market china, *Data Sci. Finance Econ.*, **5** (2025), 177–204. https://doi.org/10.3934/DSFE.2025009
12. R. Gupta, S. Kapoor, H. Gupta, S. Natesan, A forest of opinions: A multi-model ensemble-HMM voting framework for market regime shift detection and trading, *Data Sci. Finance Econ.*, **5** (2025), 466–501. https://doi.org/10.3934/DSFE.2025019
13. Z. Li, Z. Chen, Z. Huang, Modelling the data-generating mechanism of china's commodity market by identifying hidden information flow regimes, *Financ. Innov.*, **12** (2026), 28. https://doi.org/10.1186/s40854-025-00804-w
14. H. Zhao, L. Li, W. Deng, Multi-uav path planning using improved artificial hummingbird algorithm based on differential evolution and gradient descent, *IEEE Trans. Consumer Elect.*, **72** (2025), 558–569. https://doi.org/10.1109/TCE.2025.3645430
15. H. Föllmer, A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, 4 Eds., Berlin and Boston: De Gruyter, 2016.
16. A. Dembo, O. Zeitouni, *Large eeviations techniques and applications*, In: *Stochastic Modelling and Applied Probability*, 2Eds, New York: Springer, 1998.
17. P. Dupuis, R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, New York: John Wiley & Sons, 1997.
18. J. L. Lebowitz, H. Spohn, A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics, *J. Stat. Phys.*, **95** (1999), 333–365. https://doi.org/10.1023/A:1004589714161
19. U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.*, **75** (2012), 126001. https://doi.org/10.1088/0034-4885/75/12/126001
20. T. M. Cover, J. A. Thomas, *Elements of Information Theory*, 2 Eds, New York: John Wiley & Sons, 2006.
21. A. B. Tsybakov, *Introduction to Nonparametric Estimation*, New York: Springer, 2009.
22. S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge: Cambridge University Press, 2004.
23. R. T. Rockafellar, *Convex Analysis*, Princeton: Princeton University Press, 1970.
24. W. Wang, M. A. Carreira-Perpiñán, Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application, preprint paper, 2013. https://doi.org/10.48550/arXiv.1309.1541

25. L. Condat, Fast projection onto the simplex and the ℓ_1 ball, *Math. Program.*, **158** (2016), 575–585. https://doi.org/10.1007/s10107-015-0946-6

26. A. Dvoretzky, J. Kiefer, J. Wolfowitz, Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator, *Ann. Math. Stat.*, **27** (1956), 642–669.

27. Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Boston: Kluwer Academic Publishers, 2004.

28. S. Bubeck, Convex optimization: Algorithms and complexity, *Found. Trends Mach. Learning*, **8** (2015), 231–357. https://doi.org/10.1561/2200000050

29. J. R. Norris, *Markov Chains*, Cambridge: Cambridge University Press, 1997.

30. D. A. Levin, Y. Peres, E. L. Wilmer, *Markov Chains and Mixing Times*, Providence: American Mathematical Society, 2009.

Appendix

A. Robustness and supplementary analysis

This appendix provides additional robustness checks and supplementary analyses to support the main results presented in Section 4. The analysis includes: (A.1) an ablation study to identify the contribution of each component of the EDPO objective; (A.2) a time-split stability analysis to assess performance across different market regimes; (A.3) regime conditioning by EPR to evaluate strategy behavior in high- and low-stress environments; (A.4) a theory-support credibility pack to verify the numerical traces of the theoretical identities; (A.5) a stabilizer sensitivity analysis to assess the impact of regularization parameters; and (A.6) the empirical-evidence package for hyperparameter sensitivity and forecast-loss diagnostics.

A.1. Ablation study

The ablation study examined the contribution of each component of the EDPO objective function by systematically removing terms and evaluating the resulting performance. Table 10 presents the performance metrics for the ablation variants, and Figure 8 displays the cumulative log-returns.

Table 10. Ablation metrics for the ETF7 universe. The table compares the full EDPO objective $J(\mathbf{w}) = X_a(\mathbf{w}) - \lambda\mathcal{D}(\mathbf{w})$ with reduced variants: an exergy-only version ($\lambda = 0$), dissipation-only version (objective $-\mathcal{D}$), and log-growth-only version (risk-neutral limit $a \to 0$ with $\lambda = 0$). The last column reports the fixed turnover-regularization coefficient γ_{turnover} used in the optimization (it is not realized turnover).

Strategy	Ann. Return	Ann. Vol.	Sharpe	Max DD	γ_{turnover}
Dissipation_only	0.0987	0.1184	0.833	-0.248	0.050
EDPO_full	0.1058	0.1108	0.955	-0.266	0.050
Exergy-only ($\lambda = 0$)	0.0965	0.1502	0.643	-0.293	0.050
LogGrowth_only(a=0)	0.0907	0.1682	0.539	-0.339	0.050

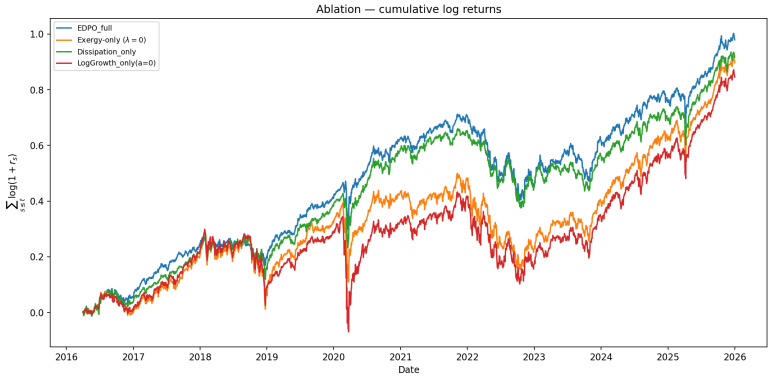


Figure 8. Ablation cumulative log-returns for the ETF7 universe. The figure compares the growth trajectories of the full EDPO model with those of the ablation variants. The full model achieves the highest terminal wealth, demonstrating the value of combining both the exergy and dissipation terms in the objective function.

The ablation study considers four variants: (i) the full EDPO model with both exergy and dissipation terms; (ii) a dissipation-only model, which directly minimizes dissipation (equivalently, maximizes $-\mathcal{D}(\mathbf{w})$) under the same constraints; (iii) an exergy-only model, which sets $\lambda = 0$ in $J(\mathbf{w})$; and (iv) a log-growth-only model, which uses the risk-neutral limit $a \to 0$ of the exergy functional with $\lambda = 0$. The results in Table 10 indicate that the full EDPO model achieved the highest Sharpe ratio (0.955), followed by the dissipation-only model (0.833), exergy-only model (0.643), and log-growth-only model (0.539).

The key insight from the ablation study is that both exergy and dissipation terms contribute to the performance of the EDPO strategy. The dissipation term plays a particularly important role in controlling volatility, as evidenced by the lower annualized volatility of the dissipation-only model (11.84%) compared with the exergy-only model (15.02%). The exergy term contributes to higher returns, as evidenced by the higher annualized return of the full model (10.58%) compared to that of the dissipation-only model (9.87%).

A.2. Time-split stability

The time-split stability analysis evaluates the performance of the EDPO strategy across various market regimes. The sample period is divided into three subperiods: (i) pre-COVID (January 2015 to December 2019), (ii) COVID rebound (January 2020 to December 2021), and (iii) post-2022 (January 2022 to December 2025). Table 11 presents the performance metrics for each subperiod.

The results in Table 11 show a clear regime dependence, which is expected for any systematic allocation rule. In the pre-COVID period, the EDPO attains a Sharpe ratio of 1.276, which is competitive but not the best (RiskParity reaches 1.309). During the COVID rebound, EDPO's Sharpe ratio was 0.864, remaining in the main cluster of benchmark performance. In the post-2022 period, EDPO's Sharpe declines to 0.443, reflecting a difficult environment; several strategies (for example, MaxSharpe) display higher Sharpe in this subperiod, but often with markedly different risk profiles.

Table 11. Time-split metrics for the ETF7 universe. The table reports performance metrics for the EDPO strategy and benchmarks across three subperiods: pre-COVID (2015–2019), COVID rebound (2020–2021), and post-2022 (2022–2025). The purpose is to gauge stability and regime dependence: EDPO remains competitive across subperiods, but absolute performance varies materially with market conditions.

Period	Strategy	Ann. Return	Ann. Vol.	Sharpe	Max DD
Pre-COVID	EDPO	0.1186	0.0930	1.276	-0.147
Pre-COVID	EDPO_Thermostat	0.1209	0.1003	1.205	-0.153
Pre-COVID	EqualWeight	0.1132	0.0936	1.209	-0.149
Pre-COVID	HRP	0.0911	0.0729	1.250	-0.103
Pre-COVID	MaxSharpe	0.0769	0.1242	0.619	-0.242
Pre-COVID	MinCVaR	0.0797	0.0661	1.207	-0.103
Pre-COVID	MinVar	0.0839	0.0670	1.253	-0.090
Pre-COVID	RiskParity	0.0933	0.0713	1.309	-0.110
COVID Rebound	EDPO	0.1648	0.1908	0.864	-0.251
COVID Rebound	EDPO_Thermostat	0.1604	0.1904	0.842	-0.251
COVID Rebound	EqualWeight	0.1679	0.1823	0.921	-0.240
COVID Rebound	HRP	0.0911	0.1316	0.692	-0.174
COVID Rebound	MaxSharpe	0.1035	0.1501	0.689	-0.163
COVID Rebound	MinCVaR	0.0876	0.1283	0.683	-0.180
COVID Rebound	MinVar	0.0897	0.1191	0.753	-0.168
COVID Rebound	RiskParity	0.1111	0.1291	0.861	-0.183
Post-2022	EDPO	0.0696	0.1570	0.443	-0.258
Post-2022	EDPO_Thermostat	0.0637	0.1532	0.415	-0.258
Post-2022	EqualWeight	0.0760	0.1412	0.538	-0.256
Post-2022	HRP	0.0654	0.1183	0.553	-0.241
Post-2022	MaxSharpe	0.1345	0.1575	0.854	-0.245
Post-2022	MinCVaR	0.0607	0.1146	0.530	-0.247
Post-2022	MinVar	0.0708	0.1132	0.625	-0.252
Post-2022	RiskParity	0.0687	0.1235	0.556	-0.249

The key insight from the time-split analysis is that the EDPO strategy is not curve-fitted to a specific market regime. The strategy maintains a competitive performance across different periods, although the absolute level of performance varies with market conditions. This suggests that the thermodynamic framework provides a robust foundation for portfolio optimization that is not overly sensitive to the training period's specific characteristics.

A.3. Regime conditioning by EPR

The regime conditioning analysis evaluated the performance of the EDPO strategy in high- and low-EPR environments. The EPR was used as a state variable to partition the sample into two regimes: (i) low EPR (below the 25th percentile) and (ii) high EPR (above the 75th percentile). Figure 9 displays the reference EPR time series, and Table 12 presents the Sharpe ratios for each regime.

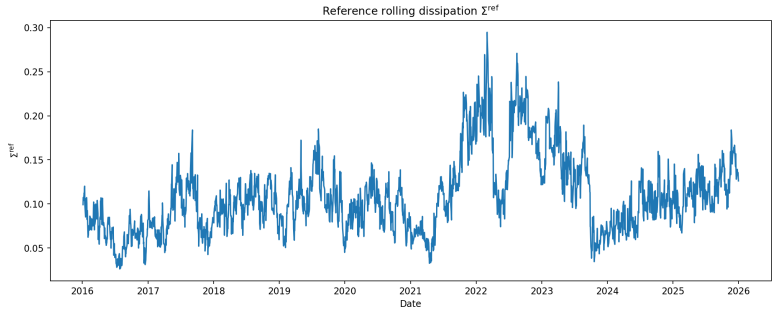


Figure 9. Reference EPR time series for the ETF7 universe. The figure shows the entropy production rate (EPR) over the sample period, with the 25th and 75th percentile thresholds indicated in the figure. High EPR values are associated with periods of market stress and regime changes.

Table 12. Regime Sharpe ratios for the ETF7 universe. The table reports the Sharpe ratio for each strategy in the low EPR (below the 25th percentile) and high EPR (above the 75th percentile) regimes. The results indicate that all strategies perform better in low EPR environments, with the EDPO strategy maintaining competitive performance in both regimes.

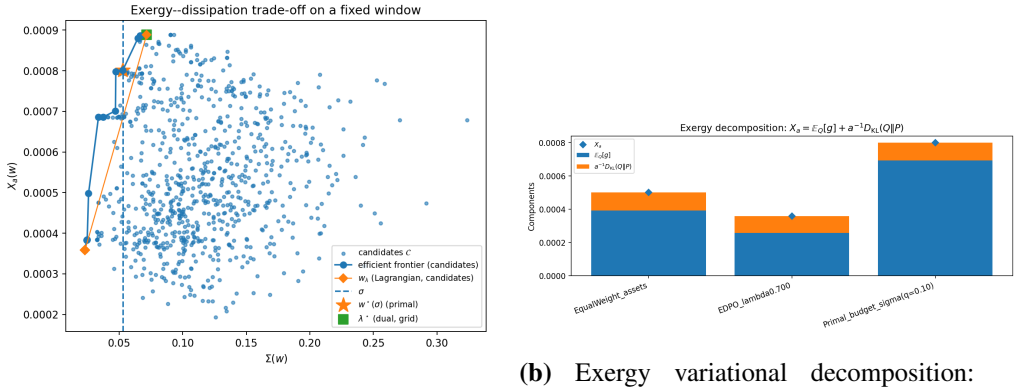
Strategy	Sharpe LOW	Sharpe HIGH
EDPO	1.818	0.702
EDPO_Thermostat	1.654	0.654
EqualWeight	1.861	0.549
HRP	1.198	0.680
InvVol	1.769	0.605
MaxDiv	1.004	0.501
MaxSharpe	1.217	0.528
MinCVaR	1.199	0.527
MinVar	1.148	0.525
RiskParity	1.332	0.638

The results in Table 12 indicate that all strategies perform better in low EPR environments, which are characterized by lower market stress and more stable return distribution. The EDPO strategy achieves a Sharpe ratio of 1.818 in the low EPR regime compared to 0.702 in the high EPR regime. This pattern is consistent across all strategies, suggesting that EPR is a meaningful state variable for characterizing market conditions.

The key insight from the regime conditioning analysis is that the EDPO strategy maintains competitive performance in both low- and high-EPR environments. In the low EPR regime, the EDPO strategy achieves the second-highest Sharpe ratio (after EqualWeight), whereas in the high EPR regime, the EDPO strategy achieves the highest Sharpe ratio among all strategies. This suggests that the dissipation-budget control mechanism is effective in adapting portfolio allocation to different market conditions.

A.4. Theory-support credibility pack

This subsection provides additional theoretical support for the EDPO framework, including the verification of the duality gap, exergy variational identity, and efficient frontier. Figure 10 displays the frontier exergy versus EPR and the exergy variational decomposition, Figure 11 displays the dual function, and Table 13 and Table 14 present the corresponding numerical results.



(a) Frontier exergy versus EPR: the efficient frontier in the $(X_a, \mathcal{D})$ space.

(b) Exergy variational decomposition: verification of the identity $X_a = \mathbb{E}_Q[g] + a^{-1}D_{\mathrm{KL}}(Q\|P)$.

Figure 10. Theory-support figures for the ETF7 universe. The left panel shows the efficient frontier in the $(X_a, \mathcal{D})$ space, which characterizes the optimal trade-off between exergy and dissipation. The right panel verifies the exergy variational identity, which decomposes exergy into a Gibbs-tilted expected log-growth $\mathbb{E}_Q[g]$ and an information term $a^{-1}D_{\mathrm{KL}}(Q\|P)$.

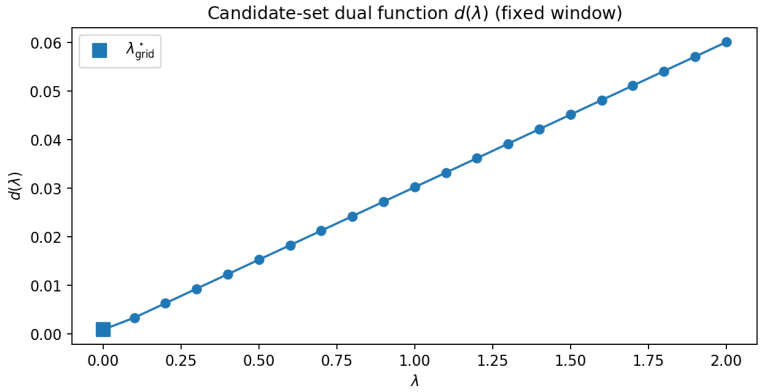


Figure 11. Dual function for the ETF7 universe. For the constrained problem $\max_{\mathbf{w}\in C} X_a(\mathbf{w})$ subject to $\mathcal{D}(\mathbf{w}) \le \sigma_{\mathrm{bud}}$, the Lagrangian dual can be written as $d(\lambda) = \max_{\mathbf{w}\in C}\{X_a(\mathbf{w}) - \lambda(\mathcal{D}(\mathbf{w}) - \sigma_{\mathrm{bud}})\} = \lambda\sigma_{\mathrm{bud}} + \max_{\mathbf{w}\in C}\{X_a(\mathbf{w}) - \lambda\mathcal{D}(\mathbf{w})\}$. As a pointwise maximum of affine functions, $d(\lambda)$ is *convex* in λ. The dual problem is $\inf_{\lambda\ge 0} d(\lambda)$, and for each $\lambda \ge 0$, the value $d(\lambda)$ provides an upper bound on the primal optimum; in practice, the curve is evaluated over the finite candidate set used for the frontier computation.

Table 13. Duality gap summary for the ETF7 universe. The table reports the primal exergy, primal EPR, dual λ, dual value, and duality gap for the final optimization window. The small duality gap indicates that the optimization is near-optimal.

σ_{bud}	Primal X_a	Primal $\mathcal{D}$	Dual λ	Dual Value	Duality Gap	Window End
0.0530	0.000801	0.0528	0.0000	0.000889	8.85e-05	2025-09-30

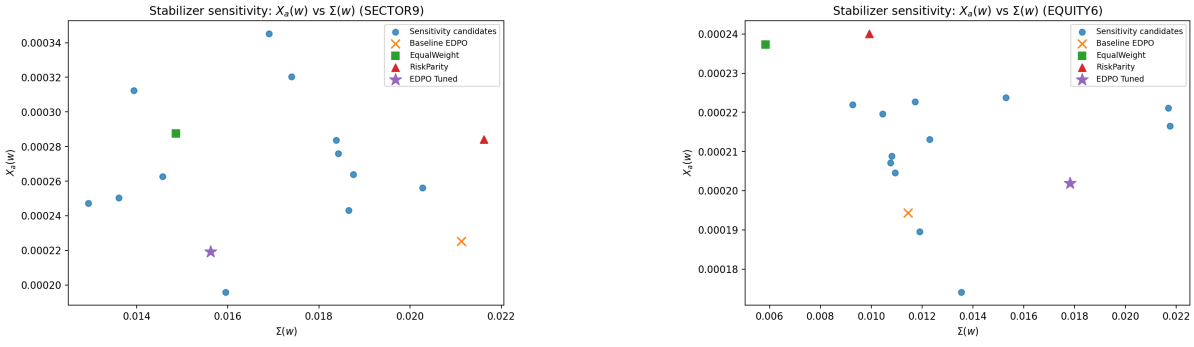
Table 14. Exergy variational identity check for the ETF7 universe. The table verifies the identity $X_a = \mathbb{E}_Q[g] + a^{-1} D_{\text{KL}}(Q\|P)$ for different portfolio configurations, where Q is the Gibbs-tilted measure attaining the variational representation, and P is the empirical reference measure. The small residuals indicate that the identity holds to numerical precision.

Label	X_a	$\mathbb{E}_Q[g]$	$D_{\text{KL}}(Q\|P)$	RHS	Residual
EqualWeight_assets	0.000502	0.000392	0.000329	0.000502	2.91e-17
EDPO_lambda0.700	0.000358	0.000258	0.000300	0.000358	7.59e-18
Primal_budget_sigma(q=0.10)	0.000801	0.000693	0.000324	0.000801	4.34e-18

The results in Table 13 show a small numerical duality gap (8.85×10^{-5}) at the final window. In the implementation, the dual is evaluated over the same finite candidate set used to trace the frontier; therefore, the gap should be read as an internal consistency diagnostic (rather than a proof of strong duality in the continuous problem). The results in Table 14 verify the exergy variational identity $X_a = \mathbb{E}_Q[g] + a^{-1} D_{\text{KL}}(Q\|P)$, with residuals on the order of 10^{-17} to 10^{-18}, indicating that the identity holds with numerical precision.

A.5. Stabilizer sensitivity

The stabilizer sensitivity analysis evaluates the impact of the regularization parameters (weight cap $w_{\max}$ and KL-divergence penalty γ_{KL}) on the work-dissipation map. Figure 12 displays the sensitivity analysis for the SECTOR9 and EQUITY6 universes.



(a) SECTOR9 universe.

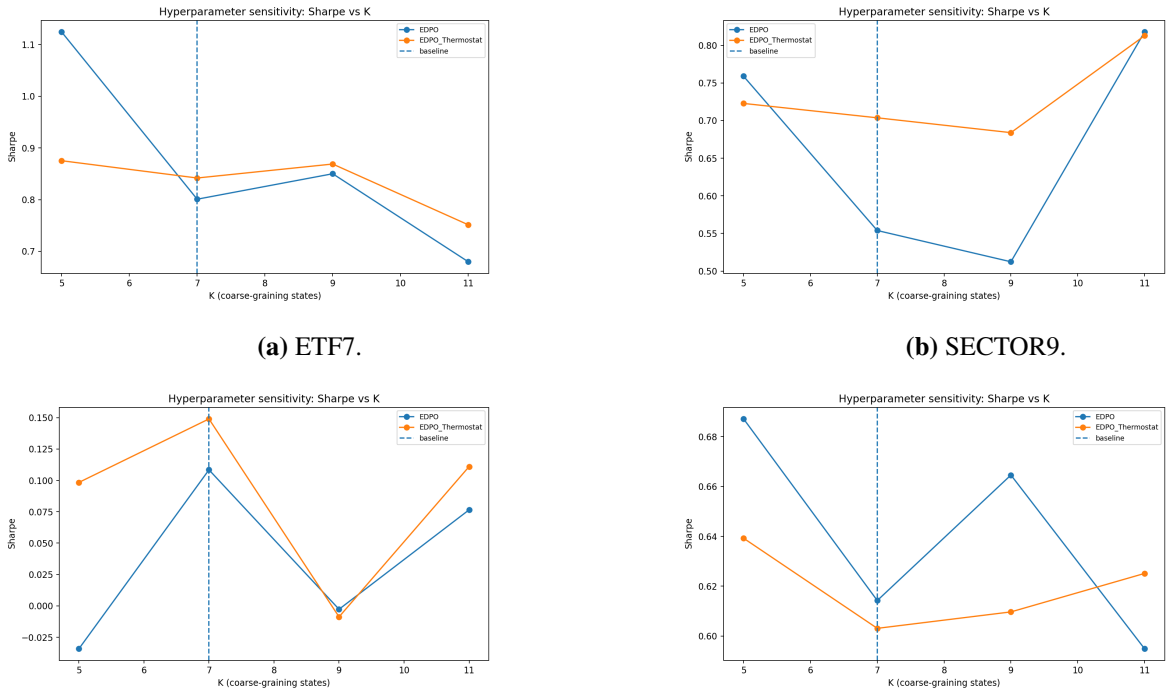
(b) EQUITY6 universe.

Figure 12. Sensitivity of the work-dissipation map to the choice of stabilizers for the SECTOR9 and EQUITY6 universes. The figures show how the location of the EDPO strategy in the $(X_a, \mathcal{D})$ space varies with different values of the weight cap $w_{\max}$ and KL-divergence penalty γ_{KL}.

The results in 12 indicate that the choice of stabilizer moderately affects the work-dissipation map. Tighter weight caps (lower $w_{\max}$) tend to reduce both exergy and dissipation as the portfolio is forced to be more diversified. Higher KL divergence penalties (higher γ_{KL}) also reduce dissipation, as the portfolio is encouraged to stay closer to the reference allocation. The key insight is that the stabilizers provide a mechanism for controlling the trade-off between exergy and dissipation, allowing practitioners to tune the strategy to their specific risk preferences and implementation constraints.

A.6. Empirical-evidence package

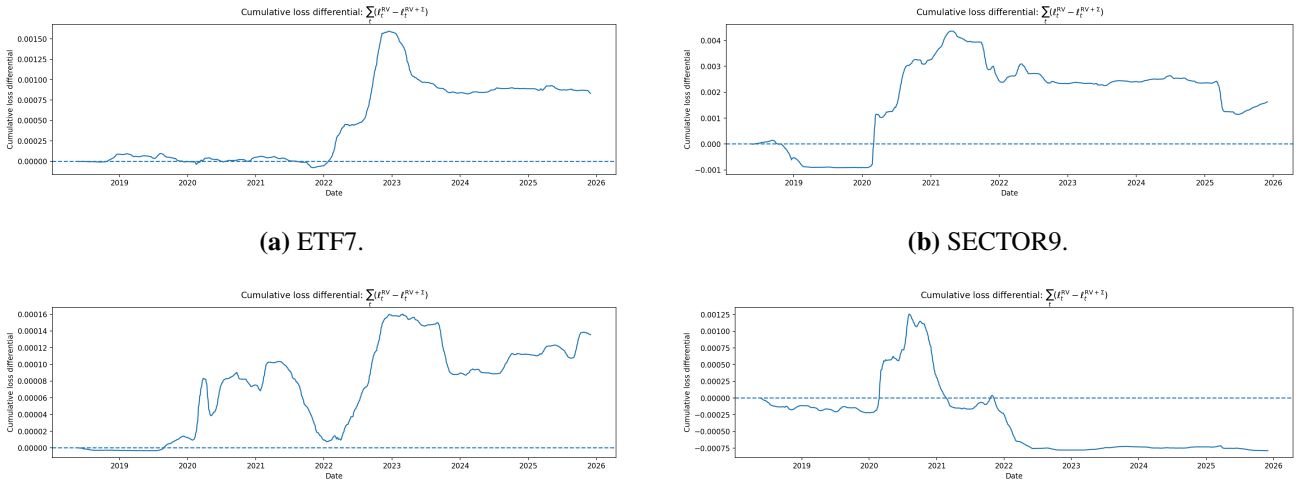
This subsection reports the two cross-universe visual diagnostics that accompany the main-text empirical-evidence package. Figure 13 focuses on the coarse-graining level K because it generated the largest sensitivity ranges in Table 5; the figure therefore provides the clearest graphical view of how the EDPO and EDPO_Thermostat Sharpe ratios move under one-factor-at-a-time hyperparameter changes. Figure 14 plots the cumulative OOS loss differential $\sum_t(\ell_t^{\mathrm{RV}} - \ell_t^{\mathrm{RV}+\mathcal{D}})$ for the four universes. Upward-sloping paths favor the augmented volatility forecast, whereas flat or downward paths indicate limited or negative incremental value from the EPR signal. Both figures are descriptive and should be interpreted jointly with the exploratory inference reported in Table 6.



(a) ETF7. **(b)** SECTOR9.

(c) BOND7. **(d)** EQUITY6.

Figure 13. Cross-universe sensitivity of the EDPO-family Sharpe ratio to the coarse-graining level K. Each panel reports the one-factor-at-a-time sensitivity curve for the relevant universe while holding the remaining hyperparameters at their baseline values. The figure confirms the pattern summarized in Table 5: K is the most influential of the three hyperparameters considered in the empirical-evidence package, with the largest movement in ETF7 and SECTOR9 and more moderate variation in BOND7 and EQUITY6.



(a) ETF7.

(b) SECTOR9.

(c) BOND7.

(d) EQUITY6.

Figure 14. Cumulative out-of-sample loss differential for the volatility-forecast comparison across universes. The plotted series are cumulative sums of $\ell_t^{\mathrm{RV}} - \ell_t^{\mathrm{RV}+\mathcal{D}}$, so upward drift favors the augmented model. ETF7, SECTOR9, and BOND7 show mostly positive cumulative gains, whereas EQUITY6 is flatter to negative. This visual pattern is consistent with the sign of the forecast-gain estimates in Table 6, while also illustrating why the formal HAC and moving-block-bootstrap tests should be interpreted cautiously.

B. Technical proofs and computational details

This appendix collects the technical derivations and implementation details that were streamlined from Section 2 in order to improve the flow of the main text. Results that remain stated in the main body are proved here in full, while secondary technical statements are moved here entirely.

B.1. Exergy functional

Proof of Lemma 2.1. Fix $a > 0$ and $w_0, w_1 \in \mathcal{W}$, $\theta \in [0,1]$. Because $x \mapsto \log x$ is concave, and $w \mapsto 1 + w^\top \boldsymbol{r}_1$ is affine, the composition $w \mapsto g_1(w) = \log(1 + w^\top \boldsymbol{r}_1)$ is concave on $\mathcal{W}$ (Assumption 2.1 keeps the logarithm well-defined). Thus,

$$g_1(\theta w_0 + (1-\theta)w_1) \;\geq\; \theta g_1(w_0) + (1-\theta) g_1(w_1).$$

Multiplying by $-a$ reverses the inequality, and exponentiating preserves order:

$$\exp\big(-a\, g_1(\theta w_0 + (1-\theta)w_1)\big) \;\leq\; \exp\big(-a\theta g_1(w_0)\big) \exp\big(-a(1-\theta) g_1(w_1)\big).$$

Taking expectation and applying Hölder's inequality yields

$$\mathbb{E}\big[e^{-a g_1(\theta w_0 + (1-\theta)w_1)}\big] \leq \Big(\mathbb{E}\big[e^{-a g_1(w_0)}\big]\Big)^{\theta} \Big(\mathbb{E}\big[e^{-a g_1(w_1)}\big]\Big)^{1-\theta}.$$

Apply log and multiply by $-1/a$ to obtain concavity as follows:

$$X_a(\theta w_0 + (1-\theta)w_1) \;\geq\; \theta X_a(w_0) + (1-\theta) X_a(w_1).$$

The case $a = 0$ follows because $X_0(w) = \mathbb{E}[g_1(w)]$ is an expectation of a concave function of w. □

Proof of Theorem 2.1. Set $Z := -a\, g_1(w)$, and fix any $Q \ll \mathbb{P}$ with the Radon–Nikodym derivative $L := (dQ/d\mathbb{P})$. Rewrite the $\mathbb{P}$-expectation through Q as

$$\mathbb{E}_{\mathbb{P}}[e^Z] \;=\; \mathbb{E}_Q\left[\frac{e^Z}{L}\right].$$

Because log is concave, Jensen's inequality gives

$$\log \mathbb{E}_Q\left[\frac{e^Z}{L}\right] \;\geq\; \mathbb{E}_Q\left[\log\left(\frac{e^Z}{L}\right)\right] \;=\; \mathbb{E}_Q[Z] \;-\; \mathbb{E}_Q[\log L].$$

Recognizing $\mathbb{E}_Q[\log L] = D_{\mathrm{KL}}(Q\|\mathbb{P})$ and substituting $Z = -ag_1(w)$ yields

$$-\frac{1}{a}\log \mathbb{E}_{\mathbb{P}}[e^Z] \;\leq\; \mathbb{E}_Q[g_1(w)] \;+\; \frac{1}{a}D_{\mathrm{KL}}(Q\|\mathbb{P}).$$

Because Q is arbitrary, we obtain (2.5) as an upper bound on $X_a(w)$.

To see tightness, take $Q^{w,a}$ with density proportional to e^Z. For this choice, $\log L = Z - \log \mathbb{E}_{\mathbb{P}}[e^Z]$, and therefore,

$$D_{\mathrm{KL}}(Q^{w,a}\|\mathbb{P}) = \mathbb{E}_{Q^{w,a}}[\log L] = \mathbb{E}_{Q^{w,a}}[Z] - \log \mathbb{E}_{\mathbb{P}}[e^Z].$$

Rearranging yields

$$-\frac{1}{a}\log \mathbb{E}_{\mathbb{P}}[e^Z] = \mathbb{E}_{Q^{w,a}}[g_1(w)] + \frac{1}{a}D_{\mathrm{KL}}(Q^{w,a}\|\mathbb{P}),$$

so the bound is achieved, and the infimum is attained at $Q^{w,a}$. □

Proof of Proposition 2.1. Let

$$\psi_w(a) := \log \mathbb{E}[\exp(-a\, g_1(w))]$$

for a near 0. Because

$$\psi_w(a) = \log \mathbb{E}[\exp(t\, g_1(w))]$$

with $t = -a$, its Taylor expansion is the cumulant series

$$\psi_w(a) = \sum_{m=1}^{\infty} \frac{\kappa_m(w)}{m!}(-a)^m = -a\,\kappa_1(w) + \frac{a^2}{2}\kappa_2(w) - \frac{a^3}{6}\kappa_3(w) + O(a^4).$$

By Definition 2.2, $X_a(w) = -(1/a)\psi_w(a)$ for $a > 0$. Dividing the expansion by $-a$ yields (2.6). □

B.2. Entropy production and time irreversibility

Proof of Lemma 2.2. ,Define the stationary *edge-flow* matrix $F^{(w)}$ by $F^{(w)}_{ij} := \pi^{(w)}_i P^{(w)}_{ij}$. Then, $\sum_{i,j} F^{(w)}_{ij} = \sum_i \pi^{(w)}_i = 1$, so $F^{(w)}$ is a probability distribution on ordered pairs. Define its time reversal by $(F^{(w)})^{\mathrm{rev}}_{ij} := F^{(w)}_{ji}$. With the conventions in Definition 2.4, (2.7) is

$$\mathcal{D}(w) = \sum_{i,j} F^{(w)}_{ij} \log\left(\frac{F^{(w)}_{ij}}{(F^{(w)})^{\mathrm{rev}}_{ij}}\right) = D_{\mathrm{KL}}\Big(F^{(w)} \,\Big\|\, (F^{(w)})^{\mathrm{rev}}\Big) \;\geq\; 0.$$

KL divergence is non-negative and equals 0 if and only if its arguments coincide, that is, $F^{(w)}_{ij} = F^{(w)}_{ji}$ for all i, j, which is precisely (2.8). □

Proof of Proposition 2.2. Consider a microstate path $s_{0:T} = (s_0, \dots, s_T)$. Under stationarity, its forward probability is $\pi^{(w)}_{s_0} \prod_{t=1}^{T} P^{(w)}_{s_{t-1}s_t}$. The reversed path $s_{T:0}$ has probability $\pi^{(w)}_{s_T} \prod_{t=1}^{T} P^{(w)}_{s_t s_{t-1}}$. Their log-ratio is

$$\log \frac{\pi_{s_0} \prod_{t=1}^{T} P_{s_{t-1}s_t}}{\pi_{s_T} \prod_{t=1}^{T} P_{s_t s_{t-1}}} = \log \frac{\pi_{s_0}}{\pi_{s_T}} + \sum_{t=1}^{T} \log \frac{P_{s_{t-1}s_t}}{P_{s_t s_{t-1}}},$$

where we suppress (w) to lighten notation. Taking expectation under the stationary chain, the boundary term $\mathbb{E}[\log(\pi_{S_0}/\pi_{S_T})]$ is $O(1)$ and vanishes after division by T. The remaining sum is a stationary average of $\log(P_{ij}/P_{ji})$ over transitions, yielding

$$\lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^{T} \log \frac{P_{S_{t-1}S_t}}{P_{S_t S_{t-1}}}\right] = \sum_{i,j} \pi_i P_{ij} \log \frac{P_{ij}}{P_{ji}}.$$

Adding and subtracting $\log(\pi_i/\pi_j)$ inside the logarithm recovers (2.7). □

Proof of Theorem 2.2. Pinsker's inequality gives

$$\|F - F^{\mathrm{rev}}\|_{\mathrm{TV}} \le \sqrt{\tfrac{1}{2} D_{\mathrm{KL}}(F \| F^{\mathrm{rev}})} = \sqrt{\tfrac{\mathcal{D}}{2}},$$

and $\|p - q\|_1 = 2\|p - q\|_{\mathrm{TV}}$ yields the ℓ^1 bound. In the symmetric binary testing problem with equal priors, the optimal success probability equals $(1/2)(1 + \|F - F^{\mathrm{rev}}\|_{\mathrm{TV}})$, so combining with Pinsker yields the stated upper bound. □

B.3. Budgeted EDPO and duality

Proof of Theorem 2.3. By the compactness of $\mathcal{W}$ and lower semicontinuity of $\mathcal{D}$, the feasible set $\mathcal{F}_{\sigma_{\mathrm{bud}}} := \{w \in \mathcal{W} : \mathcal{D}(w) \le \sigma_{\mathrm{bud}}\}$ is compact and nonempty. Under Assumption 2.1, the map $w \mapsto e^{-a g_1(w)} = (1 + w^\top \boldsymbol{r}_1)^{-a}$ is bounded by ε^{-a} and continuous in w, so dominated convergence implies that $w \mapsto \mathbb{E}[e^{-a g_1(w)}]$ is continuous. Therefore, $X_a(w) = -(1/a) \log \mathbb{E}[e^{-a g_1(w)}]$ is continuous on $\mathcal{W}$ (for $a = 0$, $X_0(w) = \mathbb{E}[g_1(w)]$ is continuous by dominated convergence using the uniform lower bound). Hence, the continuous function X_a attains its maximum over the compact set $\mathcal{F}_{\sigma_{\mathrm{bud}}}$. □

Proof of Lemma 2.3. If w is feasible, then $\mathcal{D}(w) - \sigma_{\mathrm{bud}} \le 0$, so for any $\lambda \ge 0$,

$$X_a(w) \;\le\; X_a(w) - \lambda(\mathcal{D}(w) - \sigma_{\mathrm{bud}}) \;=\; \mathcal{L}(w, \lambda) \;\le\; \sup_{u \in \mathcal{W}} \mathcal{L}(u, \lambda) = d(\lambda).$$

Taking the supremum over feasible, w yields $V(\sigma_{\mathrm{bud}}) \le d(\lambda)$, and then minimizing over λ gives the claim. □

Proof of Theorem 2.4. Problem $(\mathrm{P})_{\sigma_{\mathrm{bud}}}$ is a convex optimization problem in the standard sense: it maximizes a concave objective X_a over a convex set subject to a convex inequality constraint $\mathcal{D}(w) \le \sigma_{\mathrm{bud}}$. By Assumption 2.3 (Slater), strong duality holds, and the dual optimum is attained; see [22, Ch. 5]. The stated KKT relations are the standard optimality conditions for convex programs with inequality constraints. □

Proof of Proposition 2.3. Let $w_1 \in \arg\max_w\{X_a(w) - \lambda_1\mathcal{D}(w)\}$ and $w_2 \in \arg\max_w\{X_a(w) - \lambda_2\mathcal{D}(w)\}$. Optimality gives

$$X_a(w_1) - \lambda_1\mathcal{D}(w_1) \;\geq\; X_a(w_2) - \lambda_1\mathcal{D}(w_2), \qquad X_a(w_2) - \lambda_2\mathcal{D}(w_2) \;\geq\; X_a(w_1) - \lambda_2\mathcal{D}(w_1).$$

Adding yields $(\lambda_2 - \lambda_1)(\mathcal{D}(w_1) - \mathcal{D}(w_2)) \geq 0$, which implies (2.12). The function φ is the supremum of affine functions of λ (one for each w), hence convex, and is nonincreasing because each affine function decreases in λ. If the maximizer is unique at λ, the envelope theorem for pointwise suprema yields $\varphi'(\lambda) = -\mathcal{D}(w_\lambda)$. □

B.4. Computational projection and score normalization

Proof of Proposition 2.5. The feasible set $\mathcal{W}$ is nonempty, closed, and convex, and the objective $w \mapsto (1/2)\|w - v\|_2^2$ is strictly convex; therefore, a unique minimizer exists.

Introduce a multiplier $\tau \in \mathbb{R}$ for the equality constraint $\mathbf{1}^\top w = 1$ and vectors $\mu, \nu \in \mathbb{R}^n_{\geq 0}$ for the inequality constraints $w \geq 0$ and $w \leq w_{\max}\mathbf{1}$. The KKT conditions for optimality read

$$w^\star - v + \tau\mathbf{1} - \mu + \nu = 0, \qquad \mathbf{1}^\top w^\star = 1, \qquad 0 \leq w^\star \leq w_{\max}\mathbf{1},$$

together with complementary slackness, $\mu_i w_i^\star = 0$ and $\nu_i(w_i^\star - w_{\max}) = 0$ for each i. From stationarity, we have $w_i^\star = v_i - \tau + \mu_i - \nu_i$. If $0 < w_i^\star < w_{\max}$, then $\mu_i = \nu_i = 0$, and $w_i^\star = v_i - \tau$. If $w_i^\star = 0$, then $\nu_i = 0$, and $\mu_i = v_i - \tau \geq 0$; hence, $v_i - \tau \leq 0$. If $w_i^\star = w_{\max}$, then $\mu_i = 0$, and $\nu_i = v_i - \tau - w_{\max} \geq 0$; hence, $v_i - \tau \geq w_{\max}$. These cases yield the clipping relation (2.18). Finally, τ must satisfy $\sum_i w_i^\star = 1$. The sum is continuous and strictly decreasing in τ; hence, the solution is unique and can be found by bisection. □

Proposition B.1 (Objective-scale normalization yields scale invariance)**.** *Fix a rebalance index k, and write the unregularized score as*

$$S_k(w) \;\coloneqq\; X_a(w) - \lambda_k\, \mathcal{D}_{\mathrm{IS}}(w),$$

and the (non-negative) stabilizer as

$$P_k(w) \;\coloneqq\; \gamma_{\mathrm{KL}}\, D^{\mathrm{w}}_{\mathrm{KL}}(w \,\|\, w^{\mathrm{ref}}) + \gamma_{\mathrm{turnover}}\, \|w - w_{k-1}\|_1 + \gamma_{\mathrm{accel}}\, \|w - 2w_{k-1} + w_{k-2}\|_2^2.$$

Let $s_k = \max\{\underline{s}, |S_k(w_{k-1})|\}$ with $\underline{s} > 0$, and consider the regularized objective $J_k(w) = S_k(w) - s_k P_k(w)$. If the score is rescaled by a constant $c > 0$, $\widetilde{S}_k(w) = c\, S_k(w)$, and the normalizer is updated to $\widetilde{s}_k = c\, s_k$ (when $|S_k(w_{k-1})| \geq \underline{s}$), then

$$\arg\max_{w\in\mathcal{W}} \{\widetilde{S}_k(w) - \widetilde{s}_k P_k(w)\} \;=\; \arg\max_{w\in\mathcal{W}} \{S_k(w) - s_k P_k(w)\}.$$

Proof. Assume $|S_k(w_{k-1})| \geq \underline{s}$, so that $\widetilde{s}_k = c\, s_k$. For any $w \in \mathcal{W}$,

$$\widetilde{S}_k(w) - \widetilde{s}_k P_k(w) = c\Big(S_k(w) - s_k P_k(w)\Big).$$

Multiplying an objective by a positive constant does not change the argmax set. □

B.5. Thermostat details

B.5.1. Alternative budget calibration rules

The main text uses the reference-quantile rule

$$\sigma_{\mathrm{bud},k} = Q_q(\mathcal{D}^{\mathrm{ref}}_{1:k-1})$$

as the canonical empirical calibration. For completeness, the implementation also admits a small family of closely related variants. First, one may rescale the reference quantile,

$$\sigma_{\mathrm{bud},k} = s\, Q_q(\mathcal{D}^{\mathrm{ref}}_{1:k-1}),$$

where $s > 0$ acts as an explicit risk dial. Second, one may equivalently specify a target binding probability $\beta \in (0,1)$ and define

$$\sigma_{\mathrm{bud},k} = Q_{1-\beta}(\mathcal{D}^{\mathrm{ref}}_{1:k-1}), \qquad \sigma_{\mathrm{bud},k} = s\, Q_{1-\beta}(\mathcal{D}^{\mathrm{ref}}_{1:k-1}).$$

Finally, after a warm-up period of m rebalances, the rolling budget may be frozen by setting $\sigma_{\mathrm{bud},k} \equiv \sigma_{\mathrm{bud},m}$ for all $k \geq m$. These variants are useful as implementation checks, but they are not required for the empirical results in the main text, which rely on the reference-quantile rule throughout.

Proposition B.2 (Quantile budgets target an exceedance probability)**.** *Let $(Z_j)_{j\geq 1}$ be an i.i.d. sequence with continuous cumulative distribution function (CDF) F, and fix a target binding probability $\beta \in (0,1)$. For $m \in \mathbb{N}$, let $\widehat{\sigma_{\mathrm{bud},m}} := Q_{1-\beta}(Z_{1:m})$ denote the empirical $(1-\beta)$-quantile (smallest x with empirical CDF $\widehat{F}_m(x) \geq 1-\beta$). Then, for any $\delta \in (0,1)$, with probability at least $1-\delta$,*

$$\left|\, \mathbb{P}(Z_{m+1} > \widehat{\sigma_{\mathrm{bud},m}}) - \beta \,\right| \;\leq\; \sqrt{\frac{\log(2/\delta)}{2m}} \;+\; \frac{1}{m}. \tag{B.1}$$

In particular, $\mathbb{P}(Z_{m+1} > \widehat{\sigma_{\mathrm{bud},m}}) \to \beta$ as $m \to \infty$. This is a direct consequence of the Dvoretzky–Kiefer–Wolfowitz inequality [26].

Proof. By the Dvoretzky–Kiefer–Wolfowitz inequality, with probability at least $1-\delta$,

$$\sup_{x\in\mathbb{R}} \left|\widehat{F}_m(x) - F(x)\right| \leq \varepsilon_m(\delta) \;:=\; \sqrt{\frac{\log(2/\delta)}{2m}}.$$

On this event, at $x = \widehat{\sigma_{\mathrm{bud},m}}$, we have

$$F(\widehat{\sigma_{\mathrm{bud},m}}) \in \left[\widehat{F}_m(\widehat{\sigma_{\mathrm{bud},m}}) - \varepsilon_m(\delta),\ \widehat{F}_m(\widehat{\sigma_{\mathrm{bud},m}}) + \varepsilon_m(\delta)\right].$$

By definition of the empirical quantile, $\widehat{F}_m(\widehat{\sigma_{\mathrm{bud},m}}) \geq 1-\beta$, andb because the empirical CDF jumps in steps of size $1/m$, $\widehat{F}_m(\widehat{\sigma_{\mathrm{bud},m}}) \leq 1-\beta+1/m$. Therefore,

$$F(\widehat{\sigma_{\mathrm{bud},m}}) \in \left[1-\beta-\varepsilon_m(\delta),\ 1-\beta+\varepsilon_m(\delta)+1/m\right].$$

Because Z_{m+1} is independent of $Z_{1:m}$ and distributed according to F, $\mathbb{P}(Z_{m+1} > \widehat{\sigma_{\mathrm{bud},m}}) = 1 - F(\widehat{\sigma_{\mathrm{bud},m}})$, which yields (B.1). □

Lemma B.1 (A subgradient identity)**.** *Assume w_k solves* (2.14)*. Then,*

$$g_k \;\coloneqq\; \sigma_{\text{bud}} - \mathcal{D}(w_k)$$

is a subgradient of the convex function $d(\lambda)$ at $\lambda = \lambda_k$.

Proof. By definition, $d(\lambda) = \sup_{w\in\mathcal{W}}\{X_a(w) - \lambda(\mathcal{D}(w) - \sigma_{\text{bud}})\}$ is the supremum of affine functions of λ, and hence, is convex. Fix λ_k and a maximizer w_k. For any λ,

$$d(\lambda) \;\geq\; X_a(w_k) - \lambda(\mathcal{D}(w_k) - \sigma_{\text{bud}}) = d(\lambda_k) + (\lambda - \lambda_k)\,(\sigma_{\text{bud}} - \mathcal{D}(w_k)),$$

which is precisely the subgradient inequality. □

Theorem B.1 (Finite-time control of dual suboptimality and dissipation violation)**.** *Assume that there exists $\Lambda > 0$ such that an optimal dual multiplier satisfies $\lambda^\star \in [0, \Lambda]$. Assume that $|\mathcal{D}(w) - \sigma_{\text{bud}}| \leq G$ for all $w \in \mathcal{W}$ (bounded constraint residual). Run the projected thermostat with a projection onto*

$$[0, \Lambda] : \lambda_{k+1} = \Pi_{[0,\Lambda]}(\lambda_k + \eta(\mathcal{D}(w_k) - \sigma_{\text{bud}})).$$

Then, for constant step size $\eta = \Lambda/(G\sqrt{T})$, we have

$$\min_{0\leq k<T} d(\lambda_k) - d(\lambda^\star) \;\leq\; \frac{\Lambda G}{\sqrt{T}}, \qquad \frac{1}{T}\sum_{k=0}^{T-1}(\mathcal{D}(w_k) - \sigma_{\text{bud}}) \;\leq\; \frac{\Lambda}{\eta T} = \frac{G}{\sqrt{T}}.$$

Such bounds are standard for projected subgradient methods; see, for example, [27, 28].

Proof. Let Π denote Euclidean projection onto $[0, \Lambda]$, and define $\tilde{\lambda}_{k+1} = \lambda_k - \eta g_k$, where $g_k = \sigma_{\text{bud}} - \mathcal{D}(w_k)$ as in Lemma B.1. Our update is $\lambda_{k+1} = \Pi(\tilde{\lambda}_{k+1})$. Projection is nonexpansive, so

$$(\lambda_{k+1} - \lambda^\star)^2 \;\leq\; (\tilde{\lambda}_{k+1} - \lambda^\star)^2 = (\lambda_k - \lambda^\star)^2 - 2\eta g_k(\lambda_k - \lambda^\star) + \eta^2 g_k^2.$$

Rearranging and using $g_k^2 \leq G^2$:

$$2\eta g_k(\lambda_k - \lambda^\star) \;\leq\; (\lambda_k - \lambda^\star)^2 - (\lambda_{k+1} - \lambda^\star)^2 + \eta^2 G^2.$$

Sum from $k = 0$ to $T - 1$ to telescope:

$$2\eta\sum_{k=0}^{T-1} g_k(\lambda_k - \lambda^\star) \;\leq\; (\lambda_0 - \lambda^\star)^2 - (\lambda_T - \lambda^\star)^2 + \eta^2 G^2 T \;\leq\; \Lambda^2 + \eta^2 G^2 T.$$

Now, use the subgradient inequality from Lemma B.1: $d(\lambda_k) - d(\lambda^\star) \leq g_k(\lambda_k - \lambda^\star)$. Thus,

$$2\eta\sum_{k=0}^{T-1}(d(\lambda_k) - d(\lambda^\star)) \;\leq\; \Lambda^2 + \eta^2 G^2 T.$$

Divide by $2\eta T$ and take a minimum over k to obtain

$$\min_{0\leq k<T} d(\lambda_k) - d(\lambda^\star) \;\leq\; \frac{\Lambda^2}{2\eta T} + \frac{\eta G^2}{2}.$$

Choosing $\eta = \Lambda/(G\sqrt{T})$ balances the two terms and yields the first bound.

For the average dissipation violation, note that the update is $\lambda_{k+1} \geq \lambda_k + \eta(\mathcal{D}(w_k) - \sigma_{\text{bud}})$ up to projection. Summing and using $\lambda_T \leq \Lambda$ yields

$$\eta \sum_{k=0}^{T-1} (\mathcal{D}(w_k) - \sigma_{\text{bud}}) \;\leq\; \lambda_T - \lambda_0 \;\leq\; \Lambda;$$

hence, $(1/T)\sum_{k=0}^{T-1}(\mathcal{D}(w_k) - \sigma_{\text{bud}}) \leq \Lambda/(\eta T) = G/\sqrt{T}$. □

B.6. Finite-state estimator stability and consistency

Lemma B.2 (Strict positivity and Lipschitz stability under Laplace smoothing)**.** *Fix $K \geq 2$, a window length $T \geq 2$, and a smoothing level $\alpha > 0$. Let $\hat{P}(T;\alpha)$ and $\hat{\pi}(T;\alpha)$ be constructed from transition counts as in* (2.16)*, and let $\hat{F}(T;\alpha)$ denote the estimated stationary edge-flow, $\hat{F}_{ij}(T;\alpha) := \hat{\pi}_i(T;\alpha)\hat{P}_{ij}(T;\alpha)$.*

(i) Uniform positivity. *All entries satisfy the explicit lower bounds*

$$\hat{P}_{ij}(T;\alpha) \;\geq\; p_{\min}(T,\alpha) \;:=\; \frac{\alpha}{(T-1)+K\alpha}, \qquad \hat{\pi}_i(T;\alpha) \;\geq\; p_{\min}(T,\alpha);$$

hence, $\hat{F}_{ij}(T;\alpha) \geq m(T,\alpha) := p_{\min}(T,\alpha)^2$ for all i, j.

(ii) Flow-level Lipschitz continuity of EPR. *For any two edge-flows F, G on $\{1,\dots,K\}^2$ with $F_{ij}, G_{ij} \geq m$ for some $m \in (0,1)$ and $\sum_{i,j} F_{ij} = \sum_{i,j} G_{ij} = 1$, define*

$$\mathcal{D}(F) \;:=\; \sum_{i=1}^{K}\sum_{j=1}^{K} F_{ij}\log\!\left(\frac{F_{ij}}{F_{ji}}\right).$$

Then,

$$|\mathcal{D}(F) - \mathcal{D}(G)| \;\leq\; L(m)\,\|F-G\|_1, \qquad L(m) \;:=\; 2\max\Big\{1+\log\Big(\frac{1}{m}\Big),\; \frac{1}{m}\Big\}. \tag{B.2}$$

Proof. **(i)** From (2.16), we have

$$\hat{P}_{ij}(T;\alpha) = \frac{C_{ij}(T)+\alpha}{\sum_{j'} C_{ij'}(T) + K\alpha} \;\geq\; \frac{\alpha}{(T-1)+K\alpha},$$

because $\sum_{j'} C_{ij'}(T) \leq T-1$. For the stationary law, use

$$\hat{\pi}_i = \sum_{j=1}^{K} \hat{\pi}_j \hat{P}_{ji} \;\geq\; p_{\min}(T,\alpha)\sum_{j=1}^{K}\hat{\pi}_j = p_{\min}(T,\alpha),$$

which yields $\hat{F}_{ij} = \hat{\pi}_i\hat{P}_{ij} \geq p_{\min}^2$.

(ii) Write $\mathcal{D}(F) = \sum_{i,j}\phi(F_{ij}, F_{ji})$ with $\phi(x,y) = x\log(x/y)$. On the domain $x, y \in [m,1]$, the partial derivatives satisfy

$$\left|\frac{\partial\phi}{\partial x}(x,y)\right| = \big|\log(x/y)+1\big| \leq 1+\log(1/m), \qquad \left|\frac{\partial\phi}{\partial y}(x,y)\right| = \frac{x}{y} \leq \frac{1}{m}.$$

Hence, ϕ is Lipschitz on $[m,1]^2$ with constant $\max\{1+\log(1/m), 1/m\}$ in the ℓ^1-metric: $|\phi(x,y) - \phi(x',y')| \leq \max\{1+\log(1/m), 1/m\}(|x-x'| + |y-y'|)$. Summing over (i,j) and noting that each difference $|F_{ij} - G_{ij}|$ appears at most twice yields (B.2). □

Proposition B.3 (Consistency of plug-in EPR estimator)**.** *Under Assumption 2.2, for each fixed $w \in \mathcal{W}$ and any fixed smoothing level $\alpha \geq 0$,*

$$\hat{P}^{(w)}(T;\alpha) \to P^{(w)} \quad a.s.\ as\ T \to \infty,$$

and if $\hat{\pi}^{(w)}(T;\alpha)$ denotes the stationary distribution of $\hat{P}^{(w)}(T;\alpha)$, then the plug-in estimator $\widehat{\mathcal{D}}_T(w;\alpha)$ from (2.17) *satisfies*

$$\widehat{\mathcal{D}}_T(w;\alpha) \longrightarrow \mathcal{D}(w) \quad a.s.$$

(Ergodic consistency for finite-state Markov chains is standard; see, for example, [29, 30].)

Proof. By ergodicity of a finite irreducible Markov chain, empirical transition frequencies converge almost surely to their expectations: $C_{ij}^{(w)}(T)/T \to \pi_i^{(w)} P_{ij}^{(w)}$ a.s., and $\sum_{j'} C_{ij'}^{(w)}(T)/T \to \pi_i^{(w)}$ a.s. For fixed $\alpha \geq 0$,

$$\hat{P}_{ij}^{(w)}(T;\alpha) = \frac{C_{ij}^{(w)}(T)/T + \alpha/T}{\sum_{j'} C_{ij'}^{(w)}(T)/T + K\alpha/T} \longrightarrow \frac{\pi_i^{(w)} P_{ij}^{(w)}}{\pi_i^{(w)}} = P_{ij}^{(w)} \quad \text{a.s.}$$

For irreducible stochastic matrices, the stationary distribution depends continuously on the matrix entries (the eigenvalue 1 is simple); hence $\hat{\pi}^{(w)}(T;\alpha) \to \pi^{(w)}$ a.s. When $\alpha > 0$, strict positivity holds for every finite T, and the mapping $(P,\pi) \mapsto \sum_{ij} \pi_i P_{ij} \log(\pi_i P_{ij}/\pi_j P_{ji})$ is continuous. Therefore, $\widehat{\mathcal{D}}_T(w;\alpha) \to \mathcal{D}(w)$ a.s. by the continuous mapping principle. □

B.7. Time irreversibility beyond one-step moments

Proof of Proposition 2.4. Let the stationary distribution be uniform $\pi = (1/3, 1/3, 1/3)$, and choose any distinct values of $v(1), v(2), v(3)$. Then, both chains share the same marginal distribution of Y_t as long as they share π.

Step 1 (reversible chain). Take a symmetric transition matrix

$$P = \begin{pmatrix} 1-2p & p & p \\ p & 1-2p & p \\ p & p & 1-2p \end{pmatrix}, \qquad p \in (0, 1/2).$$

Then, $\pi_i P_{ij} = \pi_j P_{ji}$, so detailed balance holds and, $\mathcal{D}(S) = 0$ by Lemma 2.2.

Step 2 (irreversible chain with the same stationary law). Fix $\delta \in (0, p)$, and define the cycle-biased transition matrix

$$\tilde{P} = \begin{pmatrix} 1-2p & p+\delta & p-\delta \\ p-\delta & 1-2p & p+\delta \\ p+\delta & p-\delta & 1-2p \end{pmatrix}.$$

Each row sums to 1, and each column sums to 1, so $\tilde{P}$ is doubly stochastic, and $\pi\tilde{P} = \pi$. Thus, $\tilde{S}_t$ is stationary with marginal π, and $\tilde{Y}_t$ has the same one-step distribution as Y_t.

However, detailed balance fails when $\delta \neq 0$ (for example $\pi_1 \tilde{P}_{12} \neq \pi_2 \tilde{P}_{21}$). Because π is uniform, the EPR simplifies to

$$\mathcal{D}(\tilde{S}) = \frac{1}{3} \sum_{i<j} (\tilde{P}_{ij} - \tilde{P}_{ji}) \log \frac{\tilde{P}_{ij}}{\tilde{P}_{ji}} = 2\delta \log\left(\frac{p+\delta}{p-\delta}\right) > 0.$$

□

Mathematical notation

Table 15. Summary of mathematical notation used in the EDPO theoretical framework, including a reader guide for easily confused symbols.

Symbol	Definition / interpretation
Probability space and basic operators	
$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$	Filtered probability space supporting the return process.
$\mathbb{P}$; $\mathbb{E}$	Reference probability measure and expectation under $\mathbb{P}$.
$\mathbb{R}$; $\mathbb{N}$	Real numbers; positive integers.
$\langle x, y\rangle$	Euclidean inner product $\langle x, y\rangle = \sum_i x_i y_i$.
$\|x\|_1$; $\|x\|_2$	ℓ^1-norm and ℓ^2-norm.
$\|p - q\|_{\mathrm{TV}}$	Total variation distance: $\|p - q\|_{\mathrm{TV}} = \frac{1}{2}\|p - q\|_1$ for distributions.
$\mathbb{1}$	All-ones vector when used as $\langle \mathbb{1}, w\rangle$; indicator when used as $\mathbb{1}\{A\}$ (context dependent).
$\mathrm{clip}(x, \ell, u)$	Clipping: $\mathrm{clip}(x, \ell, u) = \min\{\max\{x, \ell\}, u\}$ (componentwise for vectors).
arg max; arg min	Argmax/argmin set operators.
Frequently confused symbols (reader guide)	
$\boldsymbol{\Sigma}$	Covariance matrix of asset returns in the mean–variance benchmark.
$\mathcal{D}(w)$	Dissipation / entropy production rate used in EDPO.
σ_{bud}; $\sigma_{\mathrm{bud},k}$	Scalar dissipation budget and its rolling-window version in the thermostat control law.
$\widehat{\sigma}_{t+H}$ or σ_{t+H}	Forecasted or realized volatility at horizon H in the predictive exercises.
Market primitives and admissible portfolios	
n	Number of assets.
$R_t \in \mathbb{R}^n$	Simple return vector of the n assets over $(t-1, t]$.
$w \in \mathbb{R}^n$	Portfolio weights (long-only in the baseline theory).
$\Delta^{n-1} = \Delta^{n-1}$	Probability simplex: $\{w \in \mathbb{R}^n_{\geq 0} : \langle \mathbb{1}, w\rangle = 1\}$.
$\mathcal{W} \subseteq \Delta^{n-1}$	Compact admissible set of portfolios.
$w_{\max} \in (0, 1]$	Per-asset cap for the box-constrained simplex.
$\Pi_{\mathcal{W}}(v)$	Euclidean projection of $v \in \mathbb{R}^n$ onto $\mathcal{W}$.
$v \in \mathbb{R}^n$	Generic vector to be projected onto the capped simplex.
$\tau \in \mathbb{R}$	Scalar shift ensuring $\sum_i \mathrm{clip}(v_i - \tau, 0, w_{\max}) = 1$ in the projection formula.
$\mu, \nu \in \mathbb{R}^n_{\geq 0}$	KKT multipliers for constraints $w \geq 0$ and $w \leq w_{\max}\mathbb{1}$ in the projection problem.
$\varepsilon \in (0, 1)$	Uniform nonbankruptcy constant: $1 + \langle w, R_t\rangle \geq \varepsilon$ almost surely.
Log-wealth increments and exergy	
$1 + \langle w, R_t\rangle$	One-period portfolio gross return under weights w.
$g_t(w)$	Log-wealth increment: $g_t(w) = \log(1 + \langle w, R_t\rangle)$.
$a \geq 0$	Risk sensitivity (inverse-temperature) parameter in the exergy.
$X_a(w)$	Exergy rate: $X_a(w) = -(1/a)\log \mathbb{E}[\exp(-a g_1(w))]$ for $a > 0$; $X_0(w) = \mathbb{E}[g_1(w)]$.
$Q \ll \mathbb{P}$	Absolute continuity of a measure Q with respect to $\mathbb{P}$.
$D_{\mathrm{KL}}(Q\|\mathbb{P})$	Kullback–Leibler divergence: $\mathbb{E}_Q[\log(dQ/d\mathbb{P})]$.

Continued on next page

Table 15 (continued)

Symbol	Definition / interpretation
$Q^{w,a}$	Exponentially tilted measure attaining the variational representation of $X_a(w)$.
Z	Shorthand variable $Z = -a\, g_1(w)$ used in the proof of the variational identity.
$L = dQ/d\mathbb{P}$	Radon–Nikodym derivative of Q with respect to $\mathbb{P}$.
$\kappa_m(w)$	mth cumulant of $g_1(w)$ (for example $\kappa_1 = \mathbb{E}[g_1(w)]$, $\kappa_2 = \mathrm{Var}(g_1(w))$).
$\psi_w(a)$	Cumulant generating function $\psi_w(a) = \log \mathbb{E}[\exp(-ag_1(w))]$.
t	Scalar argument of the moment generating function $\mathbb{E}[\exp(t\, g_1(w))]$.
$\theta \in [0, 1]$	Convex-combination weight used in concavity arguments.
Microstates, Markov dynamics, and dissipation	
$K \geq 2$	Number of microstates (bins) in coarse-grained dynamics.
$\phi : \mathbb{R} \to \{1, \ldots, K\}$	Coarse-graining map from log-increments to microstates.
$S_t^{(w)}$	Microstate process: $S_t^{(w)} = \phi(g_t(w))$.
$P^{(w)} = (P_{ij}^{(w)})$	Transition matrix of $(S_t^{(w)})$ under the Markov idealization.
$\pi^{(w)} = (\pi_i^{(w)})$	Stationary distribution of $P^{(w)}$.
$\mathcal{D}(w)$	Entropy production rate (EPR) of the microstate chain under w.
$F^{(w)}$	Stationary edge-flow: $F_{ij}^{(w)} = \pi_i^{(w)} P_{ij}^{(w)}$.
$(F^{(w)})^{\mathrm{rev}}$	Time-reversed edge-flow: $(F^{(w)})_{ij}^{\mathrm{rev}} = F_{ji}^{(w)}$.
F; F^{rev}	Edge-flow distribution and its time reversal in the arrow of time theorem.
$s_{0:T}$	Microstate trajectory $(s_0, \ldots, s_T)$ used in path-space representation.
$H \in \{+, -\}$	Latent time-direction label in the binary arrow of time test.
(I, J)	A single observed transition drawn from F (forward) or F^{rev} (reverse).
φ	Decision rule $\varphi : \{1, \ldots, K\}^2 \to \{+, -\}$ in the hypothesis test.
Dissipation-budget optimization and duality	
$\sigma_{\mathrm{bud}} \geq 0$	Dissipation budget (constraint level for $\mathcal{D}(w)$).
$(\mathrm{P})_{\sigma_{\mathrm{bud}}}$	Primal problem: maximize $X_a(w)$ subject to $\mathcal{D}(w) \leq \sigma_{\mathrm{bud}}$ over $w \in \mathcal{W}$.
$V(\sigma_{\mathrm{bud}})$	Optimal value of the primal problem $(\mathrm{P})_{\sigma_{\mathrm{bud}}}$.
$\mathcal{F}_{\sigma_{\mathrm{bud}}}$	Feasible set $\{w \in \mathcal{W} : \mathcal{D}(w) \leq \sigma_{\mathrm{bud}}\}$.
$\lambda \geq 0$	Lagrange multiplier/shadow price for the dissipation constraint.
$\mathcal{L}(w, \lambda)$	Lagrangian: $\mathcal{L}(w, \lambda) = X_a(w) - \lambda(\mathcal{D}(w) - \sigma_{\mathrm{bud}})$.
$d(\lambda)$	Dual function: $d(\lambda) = \sup_{w \in \mathcal{W}} \mathcal{L}(w, \lambda)$.
$(\mathrm{D})_{\sigma_{\mathrm{bud}}}$	Dual problem: $\inf_{\lambda \geq 0} d(\lambda)$.
$\lambda^\star$	Optimal dual multiplier (if it exists).
$w^\star$	Primal optimizer; under strong duality also maximizes $\mathcal{L}(w, \lambda^\star)$.
w_λ	Maximizer of the penalized problem $\arg\max_{w \in \mathcal{W}}\{X_a(w) - \lambda\mathcal{D}(w)\}$.
$\varphi(\lambda)$	Penalized value function: $\sup_{w \in \mathcal{W}}\{X_a(w) - \lambda\mathcal{D}(w)\}$.
$\mathcal{A}$	Achievable set $\mathcal{A} = \{(\mathcal{D}(w), X_a(w)) : w \in \mathcal{W}\} \subset \mathbb{R}^2$.
C	Finite candidate set used in the empirical duality gap diagnostics.

Continued on next page

Table 15 (continued)

Symbol	Definition / interpretation
Thermostat (primal–dual update) and finite-time bounds	
k	Rebalance/iteration index.
w_k	Portfolio chosen at iteration k (primal step).
λ_k	Dual iterate at iteration k (shadow price estimate).
$\eta > 0$	Thermostat step size in dual update.
$\Pi_{[0,\Lambda]}$	Projection onto $[0, \Lambda]$ (equivalently clipping to $[0, \Lambda]$).
Λ	Upper bound used to keep $\lambda_k \in [0, \Lambda]$.
g_k	Dual subgradient: $g_k = \sigma_{\text{bud}} - \mathcal{D}(w_k)$ of $d(\lambda)$ at λ_k.
$\tilde{\lambda}_{k+1}$	Unprojected dual update $\tilde{\lambda}_{k+1} = \lambda_k - \eta g_k$ used in the proof.
T	Number of thermostat iterations in the finite-time guarantee.
G	Bound on constraint residual: $\lvert\mathcal{D}(w) - \sigma_{\text{bud}}\rvert \le G$ for all $w \in \mathcal{W}$.
Data-driven dissipation budgets (rolling-window calibration)	
$\sigma_{\text{bud},k}$	Time-varying dissipation budget applied at rebalance k.
$\mathcal{D}_k^{\text{ref}}$	Reference dissipation series (for example equal-weight portfolio) at rebalance k.
$Q_q(\cdot)$	Empirical quantile functional at level $q \in (0, 1)$.
q	Quantile level used in the reference-quantile budget rule.
$s > 0$	Scale factor in scaled budget rules (distinct from the objective normalizer s_k).
$\beta \in (0, 1)$	Target binding probability for the binding probability budget rule.
m	Warm-up length after which the budget may be frozen, $\sigma_{\text{bud},k} \equiv \sigma_{\text{bud},m}$ for $k \ge m$.
Quantile-binding analysis (reference process)	
$(Z_j)_{j\ge1}$	i.i.d. reference dissipation sequence for theoretical calibration analysis.
F	Distribution function of Z_1.
$\widehat{F}_m$	Empirical CDF based on $Z_{1:m}$.
$\widehat{\sigma_{\text{bud},m}}$	Empirical $(1-\beta)$-quantile: $\widehat{\sigma_{\text{bud},m}} = Q_{1-\beta}(Z_{1:m})$.
$\delta \in (0, 1)$	Confidence level in the DKW inequality.
$\varepsilon_m(\delta)$	DKW radius $\sqrt{\log(2/\delta)/(2m)}$.
Regularized EDPO objective (empirical pipeline)	
$\mathcal{D}_{\text{IS}}(w)$	In-sample dissipation estimate; $\mathcal{D}_{\text{IS}}(w) \equiv \widehat{\mathcal{D}}_T(w; \alpha)$.
$\widehat{\mathcal{D}}_T(w; \alpha)$	Plug-in EPR estimator computed on a window of length T with smoothing α.
γ_{KL}	Weight on the KL-to-reference regularizer $D_{\text{KL}}^{\text{w}}(w\|\|w^{\text{ref}})$.
γ_{turnover}	Weight on turnover penalty $\|w - w_{k-1}\|_1$.
γ_{accel}	Weight on acceleration penalty $\|w - 2w_{k-1} + w_{k-2}\|_2^2$.
$w^{\text{ref}} \in \Delta^{n-1}$	Reference allocation (typically equal-weight).
w_{k-1}; w_{k-2}	Previous and two-steps-previous weights.
$D_{\text{KL}}^{\text{w}}(w\|\|u)$	KL divergence between weights: $D_{\text{KL}}^{\text{w}}(w\|\|u) = \sum_i w_i \log(w_i/u_i)$.
s_k	Objective-scale normalizer: $s_k = \max\{\underline{s},\ \lvert X_a(w_{k-1}) - \lambda_k \mathcal{D}_{\text{IS}}(w_{k-1})\rvert\}$.

Continued on next page

Table 15 (continued)

Symbol	Definition / interpretation
$\underline{s} > 0$	Positive floor in s_k (for example 10^{-8}) to avoid degeneracy.
$S_k(w)$	Unregularized score: $S_k(w) = X_a(w) - \lambda_k \mathcal{D}_{\mathrm{IS}}(w)$.
$P_k(w)$	Stabilizer combining KL, turnover, and acceleration penalties.
$J_k(w)$	Regularized objective: $J_k(w) = S_k(w) - s_k P_k(w)$.
Finite-state EPR estimation and smoothing	
$T \geq 2$	Window length used for estimating the transition law and the EPR.
$b_0, \ldots, b_K$	Thresholds defining microstate bins ($b_0 = -\infty$, $b_K = +\infty$).
ϕ_T	Window-dependent binning map defined by thresholds (b_k).
$C_{ij}^{(w)}(T)$	Transition counts $\sum_{t=2}^{T} \mathbb{1}\{S_{t-1}^{(w)} = i, S_t^{(w)} = j\}$.
$\alpha \geq 0$	Laplace smoothing level applied to the transition counts.
$\hat{P}_{ij}^{(w)}(T;\alpha)$	Laplace-smoothed transition estimator.
$\hat{\pi}^{(w)}(T;\alpha)$	Stationary distribution of $\hat{P}^{(w)}(T;\alpha)$.
$\hat{F}_{ij}(T;\alpha)$	Estimated edge-flow: $\hat{F}_{ij} = \hat{\pi}_i \hat{P}_{ij}$.
$p_{\min}(T,\alpha)$	Lower bound on $\hat{P}_{ij}(T;\alpha)$ and $\hat{\pi}_i(T;\alpha)$.
$m(T,\alpha)$	Lower bound on edge-flow entries: $m(T,\alpha) = p_{\min}(T,\alpha)^2$.
$\mathcal{D}(F)$	EPR functional of an edge-flow F: $\mathcal{D}(F) = \sum_{i,j} F_{ij} \log(F_{ij}/F_{ji})$.
F, G	Two edge-flow distributions used in Lipschitz stability analysis.
m	Generic positive lower bound on flow entries in the Lipschitz bound.
$L(m)$	Lipschitz constant for $\mathcal{D}(\cdot)$ on flows bounded below by m.
Illustrative example separating moments from dissipation	
(S_t); $(\tilde{S}_t)$	Two stationary Markov chains with identical stationary laws but different EPR.
$v : \{1,2,3\} \to \mathbb{R}$	Mapping from microstates to a scalar observable (unrelated to the projection vector $v \in \mathbb{R}^n$).
$Y_t = v(S_t)$; $\tilde{Y}_t = v(\tilde{S}_t)$	Derived scalar processes with identical one-step distributions.
P; $\tilde{P}$	Transition matrices of (S_t) and $(\tilde{S}_t)$.
π	Stationary distribution in the construction (uniform on $\{1,2,3\}$).
$p \in (0, 1/2)$	Base transition intensity parameter in P and $\tilde{P}$.
$\delta \in (0, p)$	Cycle-bias parameter in $\tilde{P}$ controlling irreversibility.