

Basis Immunity: Isotropy as a Regularizer for Uncertainty

Florent Ségonne*

[80]

November 2025

Diversification is a cornerstone of robust portfolio construction, yet its application remains fraught with challenges due to model uncertainty and estimation errors. Practitioners often rely on sophisticated, proprietary heuristics to navigate these issues. Among recent advancements, Agnostic Risk Parity [2] introduces eigenrisk parity (ERP), an innovative approach that leverages isotropy to evenly allocate risk across eigenmodes, enhancing portfolio stability.

In this paper, we review and extend the isotropy-enforced philosophy of ERP proposing a versatile framework that integrates mean-variance optimization with an isotropy constraint acting as a geometric regularizer against signal uncertainty. The resulting allocations decompose naturally into *canonical portfolios* [8], smoothly interpolating between full isotropy (closed-form *isotropic-mean* allocation) and pure *mean-variance* through a tunable isotropy penalty.

Beyond methodology, we revisit fundamental concepts and clarify foundational links between isotropy, canonical portfolios [8], principal portfolios [13], primal versus dual representations, and intrinsic basis-invariant metrics for returns, risk, and isotropy. Applied to sector trend-following [11], the isotropy constraint systematically induces *negative average-signal exposure*—a structural, parameter-robust crash hedge.

This work offers both a practical, theoretically grounded tool for resilient allocation under signal uncertainty and a pedagogical synthesis of modern portfolio concepts.

Notations

$\Omega \in \mathcal{R}^{n \times n}$	return covariance $\Omega = E[rr^T]$
$\Xi \in \mathcal{R}^{m \times m}$	signal covariance $\Xi = E[ss^T]$
$\Pi \in \mathcal{R}^{n \times m}$	return/signal cross-covariance $\Pi = E[rs^T]$
$\tilde{\Pi} = \Omega^{-\frac{1}{2}} \Pi \Xi^{-\frac{1}{2}}$	normalized predictability
$r = \beta s + \epsilon$	regressing -assumption joint normal
	$\beta = E[rs^T] E[ss^T]^{-1} = \Pi \Xi^{-1}$
	$E[r s] = \beta s = \Pi \Xi^{-1} s$
$w = L^T s$	positions: $w \in \mathcal{R}^n, L \in \mathcal{R}^{m \times n}$
$w^T r$	next-step PnL: $w^T r = s^T L r$
	$E[w^T r] = \text{Tr}(L \Pi)$
	$\text{Var}[w^T r] = \text{Tr}(\Xi L \Omega L^T) + \text{Tr}(\Pi L \Pi L)$
$M_{\rightarrow n}$	The first-left n -vector columns of matrix M
$M_{\leftarrow m}$	The last-right m -vector columns of matrix M

$\{e_i^r\}$	natural basis for returns
$\{e_i^s\}$	natural basis for signals
$\{\bar{b}\} \& \{\bar{u}\}$	pca basis $\Omega = \bar{B} \Sigma \bar{B}^T \& \Xi = \bar{U} \Lambda \bar{U}^T$
$\{b_i\}$	Riccati basis $b_i = \Omega^{-\frac{1}{2}} e_i^r$
$\{u_i\}$	Riccati basis $u_i = \Xi^{-\frac{1}{2}} e_i^s$
$\{\hat{b}_i\}$	Isotropic basis $\hat{b}_i = \Omega^{-\frac{1}{2}} \mathbb{R}_{\hat{b}} e_i^r$
$\{\hat{u}_i\}$	Isotropic basis $\hat{u}_i = \Xi^{-\frac{1}{2}} \mathbb{R}_{\hat{u}} e_i^s$

Used Singular Value Decompositions $m \geq n$, $M \in \mathcal{R}^{m \times n}$

$\Pi_{bu} = \tilde{\Pi}$	$\tilde{B} \tilde{\Psi} \tilde{U}^T = \tilde{B} \tilde{\Psi}_{\rightarrow n} \tilde{U}_{\rightarrow n}^T$
$\Pi_{\hat{b}\hat{u}} = \mathbb{R}_{\hat{b}}^T \Pi_{bu} \mathbb{R}_{\hat{u}}$	$(\mathbb{R}_{\hat{b}}^T \tilde{B}) \tilde{\Psi} (\mathbb{R}_{\hat{u}}^T \tilde{U})$
$\Omega^{-\frac{1}{2}} \Xi^{+\frac{1}{2}}$	$\hat{B} \hat{\Psi} \hat{U}^T$
$\Omega^{-\frac{1}{2}} (M^T \Xi M)^{+\frac{1}{2}}$	$\hat{B} \hat{\Psi} \hat{U}^T$ (same as $\hat{B} \hat{\Psi} \hat{U}^T$ when $M^T = \mathbb{I}_d$)
$\Omega^{-\frac{1}{2}} M^T \Xi^{+\frac{1}{2}}$	$\hat{B} \hat{\Psi} \hat{U}^T$ (same as $\tilde{B} \tilde{\Psi} \tilde{U}^T$ when $M^T = \beta$)

* The author would like to thank the 80 Technologies research team, especially M. Lapides and L. Jeannerod.
Contact email: quantbasics101@gmail.com — Affiliation: WorkMotion

1 Introduction

1.1 Motivation

Risk management is a fundamental pillar of quantitative finance, with diversification serving as a primary strategy to reduce portfolio volatility and safeguard capital against market uncertainties. Traditional diversification methods, such as Markowitz’s mean-variance optimization [16, 17], rely on precise estimates of expected returns and covariances—assumptions that often fail in practice due to market non-stationarity and estimation errors.

These well-known limitations (see Section 2.2.3), consistently highlighted in the literature (e.g. [7, 5, 19]), can lead to suboptimal risk allocations and significant losses when market conditions shift, especially during periods of market stress. This underscores the need for more robust, uncertainty-aware approaches.

Specifically, the mean-variance optimization is very sensitive to errors in the input parameters, as expected returns are typically scaled up by the inverse covariance of the returns (see [23]). Small changes can lead to significant portfolio variations, resulting in unstable or extreme weights (e.g. corner solutions where the portfolio heavily concentrates on a few assets, defeating the diversification goal).

Agnostic Risk Parity, introduced by Benichou et al [2], aims to address some of these challenges through the concept of eigenrisk parity (ERP), allocating risk equally across uncorrelated factors¹. At its core, the approach enforces *isotropy* in both return and signal spaces to prevent error compounding across correlated dimensions. This isotropic framework enables balanced risk contributions with minimal distortion, offering resilience to both known risks and “unknown unknowns,” and proving particularly effective in strategies like trend-following.

In this paper, we review and extend the isotropy philosophy beyond ERP, examining a broader class of portfolio allocation schemes that operate under uncertainty. Our focus is narrow and precise: within a stochastic mean-variance setting (asset returns \mathbf{r} and predictors \mathbf{s} both random), we treat *signal uncertainty* as the dominant threat and use *isotropy* as a geometric regularizer — a principle we frame as *Basis Immunity* (BI).

Signal errors compound when correlated: “bad things go together” in the signal basis, and mean-variance optimization amplifies them by exploiting return correlations to reduce variance. To break this dual compounding, we penalize anisotropy in both spaces, decoupling all directions. The resulting *Isotropy-Regularized Mean-Variance* (IRMV) allocations decompose naturally into *canonical components* [8] with:

- Closed-form *Isotropic-Mean* (IM) solutions for full isotropy,
- A *tunable isotropy penalty* yielding cubic equations that smoothly interpolate between mean-variance (MV) and isotropic-mean (IM).

The paper is organized as follows:

- First, we define notations (Section 2.1) and review the general mean-variance framework when asset returns and signals are stochastic (Section 2.2). The theoretical MV solution serves as the starting point for constructing isotropy-regularized allocations.

Before proceeding, we introduce the concept of *isotropic bases* (Section 2.3.3). This allows us to reinterpret the ERP approach of [2]—where equal risk per eigenvector is a *consequence* of enforced isotropy, not the objective—and extend it systematically. Canonical portfolios [8], key building blocks of MV, are defined in Section 2.3.6.

- In Section 3, we construct exact isotropy-enforced allocations in two steps: first the balanced case (as in [2]), then the general case with more signals than assets.
- In Section 4, we depart from “pure” isotropy and augment mean-variance with a penalty on anisotropy. This is the core of the paper, unifying isotropy, canonical portfolios, and basis-invariant risk design.
- A compact illustration using sector trend-following [11] appears in Section 5; isotropy systematically induces *negative average-signal exposure*—a structural crash hedge.

We deliberately omit empirical studies. As any practitioner knows, the success of an investment strategy depends not only on the framework, but on countless implementation details—context-dependent, proprietary, and beyond the scope of this work.

However, by providing a comprehensive theoretical foundation, we aim to equip portfolio managers with tools to navigate the complexities of modern financial markets with greater confidence and resilience.

¹ All the while using cleaned covariance matrices to mitigate the impact of noisy data [6]

2 Setting up the scene

2.1 Notations

We consider the natural basis $\{e_i^r\}$ of n assets. A vector $\mathbf{w} \in \mathcal{R}^n$ represents a portfolio allocation $\sum w_i e_i^r$ across all assets, where w_i is the percentage weight into asset S^i . The positions are derived from some signals $\mathbf{s} \in \mathcal{R}^m$. Over the next interval, the allocation \mathbf{w} generates a PnL $\sum w_i r_i$ where r_i is the next-step return of S^i .

We work in an idealized framework where the stochastic variables of interest, i.e. the asset returns $\mathbf{r} \in \mathcal{R}^n$ and the signals $\mathbf{s} \in \mathcal{R}^m$, are centered (i.e. of null unconditional expectation $E[\mathbf{r}] = E[\mathbf{s}] = \mathbf{0}$) and jointly normal. Furthermore, we assume that the quantities, such as conditional expectations or second-order moments, are well-estimated (potentially through regularization methods, such as linear shrinkage [15], or other techniques, e.g. correlation cleaning [6], factor models [18, 20]).

The natural basis $\{e_i^r\}$ is embedded with an inner product \bullet defined by the assets' covariance structure:

$$e_i^r \bullet e_j^r = E[r_i r_j] = \Omega_{i,j}$$

where r_i is the return of the i^{th} asset. A given position \mathbf{w} generates a PnL $\mathbf{w}^\top \mathbf{r}$ with unconditional variance:

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] = \mathbf{w}^\top \boldsymbol{\Omega} \mathbf{w} = \mathbf{w} \bullet \mathbf{w},$$

where $\boldsymbol{\Omega} = E[\mathbf{r}\mathbf{r}^\top]$ is the covariance matrix of assets' returns ($\boldsymbol{\Omega}$ is symmetric definite positive). This defines a Hilbert space that we denote \mathcal{H}_r .

The signals $\mathbf{s} \in \mathcal{R}^m$ used to predict future returns are known on time for trading, that is before the realization of \mathbf{r} . The information up to that time is captured by the filtration and denoted \mathcal{F} (e.g. the conditional expectation $E[\mathbf{r}|\mathcal{F}] = E[\mathbf{r}|\mathcal{F}]$). We denote by $\boldsymbol{\Xi} = E[\mathbf{s}\mathbf{s}^\top]$ the signals' covariance, which we also assume to be definite positive² (the signal Hilbert space is denoted \mathcal{H}_s).

In full generality, we do not assume $m = n$. Several situations can be considered:

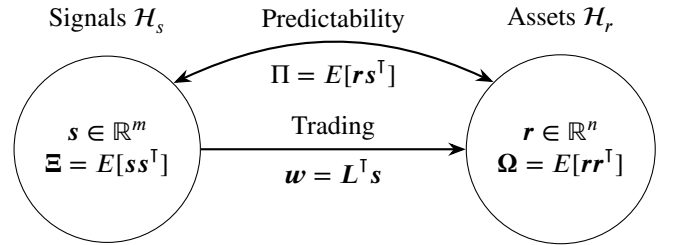
- $m < n$: less signals than assets. The features are typically aggregated factors, common to all assets, like macroeconomic variables (e.g. market volatility, unemployment rates, GDP growth rate, interest rate changes or yield curve slopes) or sector-level metrics (e.g. average sector valuation). We do not consider this case.
- $m = n$: when the number of signals equals the number of assets, each signal s_i is often specifically "designed" to predict the future return r_i of a corresponding asset (so that

$E[s_i r_i] \geq 0$). We note that the case where signals are linearly combined as $\mathbf{z} = \mathbf{M}^\top \mathbf{s}$ with $\mathbf{M} \in \mathcal{R}^{n \times m}$ a given matrix, could be similarly tackled by working with the signals \mathbf{z}_i directly.

- $m > n$: this typical scenario where signals outnumber assets leverages high-dimensional datasets, including technical indicators (e.g. trends [10], volume changes, Bollinger bands, carry [1, 14]), alternative data (e.g. social media sentiment), and machine learning-derived features. In this general setting, common aggregated factors could also be included.

In this work, we only focus on the more common scenario $m \geq n$.

The cross-covariance between returns and signals is denoted by $\boldsymbol{\Pi} = E[\mathbf{r}\mathbf{s}^\top]$. It is also termed the predictability matrix since it is a measure of the signal-return predictability. We note that it is typically not symmetric (even when $m = n$), as the predictive strength of a signal i on asset j may be different from that of signal j on asset i . The accurate estimation of $\boldsymbol{\Pi}$ is difficult, where the source of uncertainty mainly lies.



2.2 Trading: Mean-Variance Framework

We suggest to trade the assets with some positions $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$ where the matrix \mathbf{L} is of size $m \times n$. At \mathbf{L} fixed and given, the positions \mathbf{w} become stochastic variables, functions of the signal realizations \mathbf{s} . The operator \mathbf{L} is typically chosen so as to maximize some objective function over the joint dynamics of signals and returns.

In this work, we consider a standard mean-variance framework where the functional to optimize is expressed as:

$$E[\mathbf{w}^\top \mathbf{r}] - \frac{\gamma}{2} \text{Var}[\mathbf{w}^\top \mathbf{r}], \quad (1)$$

with γ a Lagrange coefficient used to set an expected level of risk. Thanks to our Gaussian assumptions (i.e. \mathbf{r} and \mathbf{s} being jointly normal), the different expectations (conditional and unconditional) can be computed efficiently in closed-form.

²When the signals are not linearly independent, we pre-process them and remove the linear dependencies.

Some straight-forward calculations show that:

$$E[\mathbf{w}^\top \mathbf{r}] = E[\mathbf{s}^\top \mathbf{L} \mathbf{r}] = \text{Tr}(\mathbf{L} \mathbf{\Pi}) \quad (2)$$

The expectation is taken over asset returns \mathbf{r} and signals \mathbf{s} , which are both stochastic variables (again, assumed to be centered and jointly normal). This can be compared to the conditional expectation at signal fixed:

$$E[\mathbf{w}^\top \mathbf{r} | \mathbf{s}] = \mathbf{w}^\top E[\mathbf{r} | \mathbf{s}] = \mathbf{w}^\top \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{s}$$

We quickly verify that:

$$E[\mathbf{w}^\top \mathbf{r}] = E[\mathbf{w}^\top E[\mathbf{r} | \mathbf{s}]] = E[\mathbf{s}^\top \mathbf{L} \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{s}] = \text{Tr}(\mathbf{L} \mathbf{\Pi})$$

The variance is slightly more challenging to compute. As we assume that all the variables of interest are centered Gaussian vectors, we can use the following identity for centered Gaussian variables (known as Wick's theorem or Isserlis' theorem):

$$E[z_1 z_2 z_3 z_4] = E[z_1 z_2] E[z_3 z_4] + E[z_1 z_3] E[z_2 z_4] + E[z_1 z_4] E[z_2 z_3]$$

We find that:

$$\begin{aligned} \text{Var}[\mathbf{w}^\top \mathbf{r}] &= E[(\mathbf{w}^\top \mathbf{r})^2] - E[\mathbf{w}^\top \mathbf{r}]^2 \\ &= \sum_{i,j,k,l} L_{i,j} L_{k,l} E[s_i s_k r_j r_l] - \text{Tr}(\mathbf{L} \mathbf{\Pi})^2 \\ &= \sum_{i,j,k,l} L_{i,j} L_{k,l} (E[s_i s_k] E[r_j r_l] + E[s_i r_l] E[r_j s_k]) \\ &= \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top) + \text{Tr}(\mathbf{\Pi} \mathbf{L} \mathbf{\Pi} \mathbf{L}^\top) \end{aligned} \quad (3)$$

The second term is typically much smaller than the first one (as it contains squared cross-correlations). This is almost always the case but would need to be checked in practice (see Section 5 in the case of a simple trend-following model, particularly Figure 5). Ignoring it is usually a sensible choice, while having the great advantage of leading to interpretable close-form solutions. In this work, we neglect it and focus only on the first part:

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] \approx \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top) \quad (4)$$

The conditional variance could also be computed as:

$$\text{Var}[\mathbf{w}^\top \mathbf{r} | \mathbf{s}] = \mathbf{w}^\top (\mathbf{\Omega} - \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{\Pi}^\top) \mathbf{w}$$

and we can easily verify the law of total variance:

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] = E[\text{Var}[\mathbf{w}^\top \mathbf{r} | \mathbf{s}]] + \text{Var}[E[\mathbf{w}^\top \mathbf{r} | \mathbf{s}]]$$

using $E[E[\mathbf{w}^\top \mathbf{r} | \mathbf{s}]^2] = \text{Tr}(\mathbf{L} \mathbf{\Pi})^2 + \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{\Pi}^\top \mathbf{L}^\top + \mathbf{\Xi} \mathbf{L} \mathbf{\Xi} \mathbf{L}^\top)$.

2.2.1 Mean-Variance Functional and Solution

The standard mean-variance functional can be written as:

$$\arg_L \max \text{Tr}(\mathbf{L} \mathbf{\Pi}) - \frac{\gamma}{2} \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top), \quad (5)$$

with first-order condition:

$$\mathbf{\Pi} = \gamma \mathbf{\Omega} \mathbf{L}^\top \mathbf{\Xi}$$

This leads to the general solution $\mathbf{L}^\top = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1}$ and we finally obtain:

General Mean-Variance

$$\mathbf{w} = \mathbf{L}^\top \mathbf{s} = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{s} \quad (6)$$

The risk is generally calibrated through the Lagrange coefficient γ to a target variance σ^2 , so that:

$$\gamma^2 = \frac{1}{\sigma^2} \text{Tr}(\mathbf{\Xi}^{-1} \mathbf{\Pi}^\top \mathbf{\Omega}^{-1} \mathbf{\Pi}) = \frac{1}{\sigma^2} \text{Tr}(\tilde{\mathbf{\Pi}}^\top \tilde{\mathbf{\Pi}}) \quad (7)$$

where $\tilde{\mathbf{\Pi}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}}$, the normalized predictability matrix (a key element of the framework).

Even though we worked in the natural asset basis, the mean-variance framework could be expressed anywhere. The resulting solution Eq 6 is totally invariant to the choice of basis. This is obviously the case because the definition of expected returns in Eq 2 and the variance in Eq. 3 is intrinsic, that is independent from the choice of coordinates.

2.2.2 The Regression Angle

Eq. 6 does not appear out of nowhere. There is a clear link between this approach and a standard regression problem where one tries to regress the returns \mathbf{r} onto a set of predictors \mathbf{s} :

$$\mathbf{r} = \boldsymbol{\beta} \mathbf{s} + \boldsymbol{\epsilon}$$

Under standard Gaussian assumptions, we find that:

$$\boldsymbol{\beta} = E[\mathbf{r} \mathbf{s}^\top] E[\mathbf{s} \mathbf{s}^\top]^{-1} = \mathbf{\Pi} \mathbf{\Xi}^{-1} \quad \text{and} \quad E[\mathbf{r} | \mathcal{F}] = E[\mathbf{r} | \mathbf{s}] = \boldsymbol{\beta} \mathbf{s} \quad (8)$$

and the solution of Eq. 6:

$$\mathbf{w} = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \boldsymbol{\beta} \mathbf{s} = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{s}$$

appears naturally. The closed-form expression of Eq. 6, which we typically express as:

$$\mathbf{w} = \frac{1}{\gamma} \mathbf{\Omega}^{-1} E[\mathbf{r} | \mathcal{F}] \quad (9)$$

is the starting point for constructing several isotropy-enforced portfolio allocations. Before we do so, we briefly review some of the limitations of the mean-variance framework.

2.2.3 Mean-Variance Limitations

The formulations Eq. 6-9 can be used to review some of the well-documented limitations and challenges of the mean-variance approach. This also allows us to set the stage and explore how the concept of isotropy can be used to address some of these issues, particularly their robustness to signal uncertainty.

1. Challenges with Covariance Estimation and Inversion

Covariance matrices are notoriously difficult to estimate. Not enough samples and we are dealing with too much noise; too many samples and we are probably mixing different market dynamics. Inverting these matrices (see Eq. 6) amplifies errors, especially for ill-conditioned matrices (e.g. highly correlated assets or small sample sizes). This can lead to numerical instability and unrealistic portfolio weights (for an enlightening discussion see [23]).

Recent advances in the field of random matrix theory [4, 3] have been proposed to mitigate those limitations [6]. In this work, we assume that the matrices $\mathbf{\Omega}$ and $\mathbf{\Xi}$ are accurate, well-estimated.

2. Sensitivity to Input Estimates

Mean-variance optimization is highly sensitive to errors in the covariance matrix $\mathbf{\Omega}$ and in the estimated expected returns $E[\mathbf{r}|\mathcal{F}]$ (that is implicitly in $\mathbf{\Pi}$ and $\mathbf{\Xi}$, see Eq. 6). Small changes in these inputs can lead to significantly different portfolio allocations, resulting in unstable or extreme weights (e.g. corner solutions).

Some (recent) techniques can greatly help with the estimates of covariances (e.g. linear shrinkage [15], correlation cleaning [6], factor models [18, 20]), yet estimating expected returns and the predictability matrix $\mathbf{\Pi}$ remains problematic.

3. Stability/Market Regime

Market conditions evolve rapidly, undermining the stability of in-sample estimates. This is a core challenge in quantitative finance, and the mean-variance framework is particularly vulnerable. Diversified portfolios may be less affected than concentrated ones, but resilience to uncertainty remains critical.

4. Model Risk and Distributional Assumptions

The mean-variance model relies on simplistic assumptions, including normally distributed returns, ignoring fat tails, skewness, and kurtosis prevalent in real-world markets. It also overlooks transaction costs, constraints, and parameter uncertainty. This leads to overly optimistic risk-return trade-offs and underestimation of extreme risks (e.g. black swan events).

Practitioners address these well-known limitations by integrating mean-variance principles with proprietary practical adjustments informed by years of experience. Rigorous implementation is vital for real-world success.

Our approach, named Isotropy-Regularized Mean-Variance (IRMV), does not aim at resolving all mean-variance limitations but specifically targets sensitivity to input estimates and out-of-sample instability (mostly point 2 and arguably point 3). By emphasizing resilience to uncertainty—unmeasurable randomness distinct from quantifiable risk—they reduce dependence on mis-specified signals.

Built on the concept of isotropic bases, in the spirit of [2], they offer a pathway to stable portfolio construction in unpredictable markets. To explore this alternative, we first need a bit of algebra to understand how to change perspective.

2.3 Changing Perspective

The natural basis $\{\mathbf{e}_i^r\}$ of \mathcal{H}_r is not orthonormal for the inner product \bullet (except if the covariance matrix $\mathbf{\Omega}$ is the identity matrix). Nothing prevents us from working in a different basis. In the following, we denote the belonging to a basis by the corresponding subscript (except at times for the natural basis when there is no ambiguity).

2.3.1 Change of Basis

We consider a general basis $\{\mathbf{y}_i\}$ of \mathcal{H}_r defined by an invertible transformation \mathbf{Y} : the automorphism $\mathbf{w}_y \mapsto \mathbf{Y}\mathbf{w}_y$ is the change of coordinate operator that takes us from the basis $\{\mathbf{y}_i\}$ into the natural basis $\{\mathbf{e}_i\}$, i.e. a vector with coordinates \mathbf{w}_y in $\{\mathbf{y}_i\}$ has coordinates $\mathbf{w}_e = \mathbf{Y}\mathbf{w}_y$ in $\{\mathbf{e}_i\}$. With an abuse of notation³, we say that the vector \mathbf{y}_i whose coordinates in $\{\mathbf{e}_i\}$ are the i^{th} -column of \mathbf{Y} is defined by $\mathbf{y}_i = \mathbf{Y}\mathbf{e}_i$.

It is important to understand how our variables transform under changes of coordinates. First, we note that we are dealing with two distinct Hilbert spaces, the space \mathcal{H}_r of asset returns \mathbf{r} with inner

³One needs to be careful with this (abuse of) notation, particularly when working with more than 2 bases. For example, if $\mathbf{f}_i = \mathbf{F}\mathbf{e}_i$ and $\mathbf{g}_i = \mathbf{G}\mathbf{f}_i$ (i.e. the basis vector \mathbf{g}_i has coordinates the i^{th} -column of \mathbf{G} in the basis $\{\mathbf{f}_i\}$), then we have $\mathbf{g}_i = (\mathbf{F}\mathbf{G})\mathbf{e}_i$ (and certainly not $\mathbf{g}_i = \mathbf{G}\mathbf{F}\mathbf{e}_i$ as a mis-interpretation of the abuse of notations could imply!).

product defined by Ω and the space \mathcal{H}_s of s with inner product defined by Ξ . The natural bases of \mathcal{H}_r and of \mathcal{H}_s are denoted by $\{e_i^r\}$ and $\{e_i^s\}$ respectively, although we often drop the subscript for notional convenience.

The positions w are contravariant vectors of \mathcal{H}_r , regular vectors of $\{e_i\}$, whereas returns r and signals s are covectors of \mathcal{H}_r and \mathcal{H}_s , i.e. they belong to the corresponding duals denoted \mathcal{H}_r^* and \mathcal{H}_s^* , with basis $\{e_i^{r*}\}$ and $\{e_i^{s*}\}$ (in the case where $m = n$, both dual spaces can be identified together $\mathcal{H}_r^* \sim \mathcal{H}_s^*$). To summarize the change of basis operations, we consider $y_i = Y e_i^r$ of \mathcal{H}_r and $x_i = X e_i^s$ of \mathcal{H}_s where Y and X are change of coordinate operators (i.e. invertible matrices):

e_i^r, e_i^s	$y_i = Y e_i^r, x_i = X e_i^s$
w	$w_y = Y^{-1} w$
r, s	$r_y = Y^T r, s_x = X^T s$
$\Omega = E[rr^T]$	$\Omega_y = Y^T \Omega Y$
$\Xi = E[ss^T]$	$\Xi_x = X^T \Xi X$
$w = L^T s$	$w_y = L_{xy}^T s_x$ with $L_{xy} = X^{-1} L Y^{-1}$
$\Pi = E[rs^T]$	$\Pi_{yx} = E[r_y s_x^T] = Y^T \Pi X$

As a sanity check, one can easily verify the following equalities: $E[w_y^T r_y] = E[w^T r]$, $\text{Tr}(L_{xy} \Pi_{yx}) = \text{Tr}(L \Pi)$, or $\text{Tr}(\Xi_x L_{xy} \Omega_y L_{xy}^T) = \text{Tr}(\Xi L \Omega L^T)$.

2.3.2 Operator L

The operator L^T takes us from the signal dual space $\mathcal{H}_s^* \sim \mathcal{R}^m$ to the natural vector space $\mathcal{H}_r \sim \mathcal{R}^n$. It is enlightening to think of it as the combination of two steps:

$$L^T = \frac{1}{\gamma} P M^T \quad (10)$$

- A mapping M^T takes us from the dual $\mathcal{H}_s^* \sim \mathcal{R}^m$ (where the signals live) to the dual $\mathcal{H}_r^* \sim \mathcal{R}^n$ (where the returns live and where the positions are derived) with $z = M^T s$. The linear operator M is determined so that the mapped signals z are as predictive as possible of future returns r . The vector z , which is linearly constructed from the set of all signals s , is our best⁴ estimate/guess for $E[r|\mathcal{F}]$.

Many options are possible to obtain the mapping M^T . As our best guess, it is not necessarily the best mapping in

⁴Because we work in an idealized Gaussian setting, the best linear estimator is also the best estimator over all linear and non-linear operators (in the sense of the least-square distance).

⁵For any vector z of the dual, representing our best estimate of future returns, we have the equality $w^T E[r|\mathcal{F}] = w^T z = w \bullet (\Omega^{-1} z)$.

⁶The addition of constraints, typically used in trading (e.g. limits on margin, on max trading, on max absolute positions, ...), would obviously break this invariance property.

absolute and/or even within our framework. Many source of errors could creep in and the signals s could be mis-specified (known unknowns or unknown unknowns as described in [2]).

To derive it, one could imagine using e.g. some deterministic relationships where some features s are explicitly designed/tailored for some assets (e.g. the carry of an asset), or some statistical estimation (typically through standard linear regressions/conditional expectations), or by directly integrating the unknown mapping M^T into a general (e.g. mean-variance) functional as Eq. 1 (as described in Section 2.2).

- This first step is then followed by a transformation of the covector $z \in \mathcal{H}_r^*$ (the space of returns) into a vector of tradable positions $w = \frac{1}{\gamma} P z \in \mathcal{H}_r$. This step depends on our choice of functional, which links dual and primal space together.

Working within the mean-variance framework corresponding to Eq. 1, the operator P is the decorrelation operator⁵ $P = \Omega^{-1}$, while M^T is a standard beta $M^T = \beta$. The typical mean-variance allocation, which we use as a starting point, can be then expressed as:

<p>Mean-Variance</p> $w_e = \frac{1}{\gamma} \Omega^{-1} E[r \mathcal{F}] = \frac{1}{\gamma} \Omega^{-1} M^T s_e \quad (11)$

Note that it could also be phrased in different bases of \mathcal{H}_s and \mathcal{H}_r without difficulty. For instance, in the two bases $y_i = Y e_i^r$ of \mathcal{H}_r and $x_i = X e_i^s$ of \mathcal{H}_s , we can easily check that:

$$M_{xy} = X^{-1} M Y \text{ and } P_y = Y^{-1} P Y^{-1} \quad (12)$$

so that we have:

$$w_y = L_{xy}^T s_x = \frac{1}{\gamma} P_y M_{xy}^T s_x \quad (13)$$

As we already discussed, the mean in Eq 2 and variance in Eq. 3 are intrinsic quantities and the mean-variance framework (in its simplest form, as in Eq 5) does not depend on the choice of basis⁶.

2.3.3 Isotropic Bases

Some bases possess noticeable attractive properties. For example, let's consider the one defined by $b_i = \Omega^{-\frac{1}{2}} e_i^r$ (also known as the Riccati root of Ω). It is easy to see that $\{b_i\}$ is orthonormal (for the asset returns r_b). From a variance perspective, it means that

all directions are equivalent and carry the same risk: the space has become isotropic. This choice of basis is useful when aggregating signals together, since risk (as measured by the variance of the assets) is now the same in any direction. This is where the term ‘‘Eigenrisk Parity’’ comes from in [2].

The Riccati basis is not the only isotropic basis since any rotation of the basis would have the same property. In fact, one can show that all the isotropic basis are of the form $\hat{\mathbf{b}}_i = \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \mathbf{e}_i^r$ with $\mathbb{R}_{\hat{\mathbf{b}}}$ a rotation operator, i.e. $\mathbb{R}_{\hat{\mathbf{b}}} \mathbb{R}_{\hat{\mathbf{b}}}^T = \mathbb{R}_{\hat{\mathbf{b}}}^T \mathbb{R}_{\hat{\mathbf{b}}} = \mathbb{Id}$. The operator $\mathbb{R}_{\hat{\mathbf{b}}}$ belongs to the Special Orthogonal group, the set of rotations of \mathbb{R}^n and denoted $SO(n)$, itself part of the orthogonal group, which includes rotations and symmetries and denoted $O(n)$.

For example, let’s consider the Cholesky decomposition of the covariance matrix $\mathbf{\Omega} = \mathbf{L}_{\mathbf{\Omega}} \mathbf{L}_{\mathbf{\Omega}}^T$ where $\mathbf{L}_{\mathbf{\Omega}}$ is a lower triangular matrix with positive coefficients on the diagonal. The Cholesky decomposition is unique and defines an isotropic basis $\hat{\mathbf{b}}_i = \mathbf{L}_{\mathbf{\Omega}}^{-T} \mathbf{e}_i$. One can easily show that $\mathbf{L}_{\mathbf{\Omega}} = \mathbf{\Omega}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}}$ where $\mathbb{R}_{\hat{\mathbf{b}}}$ is indeed a rotation:

$$\begin{aligned} \text{Cholesky} \quad & \mathbf{\Omega} = \mathbf{L}_{\mathbf{\Omega}} \mathbf{L}_{\mathbf{\Omega}}^T \\ & \mathbf{L}_{\mathbf{\Omega}} = \mathbf{\Omega}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \text{ and } \hat{\mathbf{b}}_i = \mathbf{L}_{\mathbf{\Omega}}^{-T} \mathbf{e}_i \end{aligned} \quad (14)$$

Among all isotropic bases, the Riccati basis has the good behavior⁷ of being the one that is the closest in the sense of the Mahalanobis distance $D_{\mathbf{\Omega}}$, as discussed in [2]. To show that, they consider the stochastic variable $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega})$, a centered Gaussian vector with covariance $\mathbf{\Omega}$, and compare its expression across both bases as $\mathbf{r}_e = \mathbf{r}$ in $\{\mathbf{e}_i\}$ and $\mathbf{r}_{\hat{\mathbf{b}}} = \mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r}$ in $\{\hat{\mathbf{b}}_i\}$.

We define the following generic distance $D_{\mathbf{\Omega}}^{\eta}$ between \mathbf{r} and $\mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r}$ (from the reference point of $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega})$):

$$\begin{aligned} D_{\mathbf{\Omega}}^{\eta} &= \text{Dist}_{\mathbf{\Omega}}^{\eta} \left(\mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r}, \mathbf{r} \right) \\ &= E \left[\left(\mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r} - \mathbf{r} \right)^T \mathbf{\Omega}^{\eta} \left(\mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r} - \mathbf{r} \right) \right] \\ &= E \left[\mathbf{r}^T \left(\mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} - \mathbb{Id} \right)^T \mathbf{\Omega}^{\eta} \left(\mathbb{R}^T \mathbf{\Omega}^{-\frac{1}{2}} - \mathbb{Id} \right) \mathbf{r} \right] \\ &= E \left[\mathbf{u}^T \left(\mathbb{R} \mathbf{\Omega}^{-\frac{1}{2}} - \mathbb{Id} \right) \mathbf{\Omega}^{1+\eta} \left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}^T - \mathbb{Id} \right) \mathbf{u} \right] \\ &= \text{Tr} \left[\left(\mathbb{R} \mathbf{\Omega}^{-\frac{1}{2}} - \mathbb{Id} \right) \mathbf{\Omega}^{1+\eta} \left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}^T - \mathbb{Id} \right) \right] \\ &= \text{Tr} \left[\mathbf{\Omega}^{\eta} + \mathbf{\Omega}^{1+\eta} - 2 \mathbb{R} \mathbf{\Omega}^{\frac{1}{2}+\eta} \right] \end{aligned}$$

⁷The Cholesky basis might be preferred for a variety of reasons: slight computational efficiency, numerical stability, memory efficiency.

⁸Interestingly, this result is valid for any choice of η (since the correlation is definite positive and \mathbb{R} is a rotation operator, hence with diagonal elements smaller than one).

⁹The term ‘‘Mahalanobis distance’’ is misleading as it is non-symmetric.

¹⁰Think of the main modes of the covariance matrix.

where we have expressed $\mathbf{r} = \sqrt{\mathbf{\Omega}} \mathbf{u}$ with $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbb{Id})$. The Mahalanobis distance $D_{\mathbf{\Omega}}$ corresponds to the value $\eta = -1$. Minimizing the distance amounts to maximizing $\text{Tr} \left[\mathbb{R} \mathbf{\Omega}^{\frac{1}{2}+\eta} \right]$. By working in the basis of $\mathbf{\Omega}$, we can then easily conclude that the minimum is reached when $\mathbb{R} = \mathbb{Id}$ (see [12])⁸.

The Mahalanobis metric⁹ quantifies the proximity of a basis $\{\mathbf{z}_i\}$ from a reference one $\{\mathbf{y}_i\}$, where $\mathbf{z}_i = \mathbf{T}_y \mathbf{y}_i$ by measuring the following:

$$D_{\mathbf{\Omega}_y}(\mathbf{r}_z, \mathbf{r}_y) = D_{\mathbf{\Omega}_y} \left(\mathbf{T}_y^T \mathbf{r}_y, \mathbf{r}_y \right) \text{ with } \mathbf{r}_y \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_y)$$

This proximity property is often used to build isotropic allocations, that is allocations which are less dependent on the risk that is naturally embedded in a specific basis through its inner product (more details in Section 3). This is the premise of the eigenrisk parity (ERP) allocations defined in [2].

As an example, let’s consider a fixed allocation $\mathbf{w}_e = \mathbf{w}$ defined in the natural basis $\{\mathbf{e}_i\}$ where \mathbf{w} has been randomly chosen on the unit sphere of \mathcal{R}^n , that is such that $\|\mathbf{w}\|^2 = \sum w_i^2 = 1$. It generates a PnL $\mathbf{w}^T \mathbf{r}_e$ where the expected total variance $\mathbf{w}^T \mathbf{\Omega} \mathbf{w}$ depends explicitly on the realized coefficients w_i on each basis vector \mathbf{e}_i through the covariance $\mathbf{\Omega}$. Large, significant (absolute) covariances generate pockets of risk¹⁰ that we would want to avoid when invested in erroneous positions (e.g. constructed from inaccurate signal estimates). The cost of being wrong is embedded in the natural asset basis $\{\mathbf{e}_i\}$ through the inner product \bullet defined by $\mathbf{\Omega}$.

Now, if the Riccati basis $\{\mathbf{b}_i\}$ is close enough from $\{\mathbf{e}_i\}$, one can hope that the realized PnL $\mathbf{w}^T \mathbf{r}_{\hat{\mathbf{b}}}$ will be similar to $\mathbf{w}^T \mathbf{r}_e$. Yet, the basis risk would disappear, as no single coefficient would be exposed to excessive level of risk (the basis being isotropic), and the variance would then become $\|\mathbf{w}\|^2 = 1$.

Clearly, everything that has been discussed so far can also be applied to the signal space and the associated bilinear form $\mathbf{\Xi}$. We can similarly define the Riccati basis $\{\mathbf{u}_i\}$ of the signals, defined by $\mathbf{u}_i = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{e}_i^s$. It is also the closest isotropic signal basis among all isotropic basis $\hat{\mathbf{u}}_i = \mathbf{\Xi}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} \mathbf{e}_i^s$ (where $\mathbb{R}_{\hat{\mathbf{u}}} \in SO(m)$) from the perspective of $D_{\mathbf{\Xi}}$.

\mathbf{b}_i	$\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{e}_i^r$	Riccati Root of \mathcal{H}_r , \mathbf{r} – Isotropic
$\hat{\mathbf{b}}_i$	$\mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \mathbf{e}_i^r$	\mathbf{r} – Isotropic
\mathbf{u}_i	$\mathbf{\Xi}^{-\frac{1}{2}} \mathbf{e}_i^s$	Riccati Root of \mathcal{H}_r , \mathbf{s} – Isotropic
$\hat{\mathbf{u}}_i$	$\mathbf{\Xi}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} \mathbf{e}_i^s$	\mathbf{s} – Isotropic

2.3.4 Risk Decompositions in Dual Eigenbases

We consider a general allocation $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$ with $\mathbf{L} \in \mathbb{R}^{m \times n}$. The portfolio variance is:

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] = \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top) + \text{Tr}(\mathbf{\Pi} \mathbf{L} \mathbf{H} \mathbf{L}^\top)$$

We consider the eigenvalue decompositions $\mathbf{\Omega} = \mathbf{\tilde{B}} \mathbf{\Sigma} \mathbf{\tilde{B}}^\top$ and $\mathbf{\Xi} = \mathbf{\tilde{U}} \mathbf{\Lambda} \mathbf{\tilde{U}}^\top$. We express \mathbf{L} into the dual eigenbasis $\{\tilde{\mathbf{b}}\}, \{\tilde{\mathbf{u}}\}$:

$$\tilde{\mathbf{L}} = \mathbf{L}_{\tilde{\mathbf{v}}\tilde{\mathbf{u}}} = \mathbf{\tilde{U}}^\top \mathbf{L} \mathbf{\tilde{B}}$$

so that $\tilde{L}_{ji} = \mathbf{\tilde{U}}_j^\top \mathbf{L} \mathbf{\tilde{B}}_i$ represents exposure to the “dual cross-mode” $\tilde{s}_j^\top \tilde{\mathbf{r}}_i$ with $\tilde{\mathbf{r}}_i = \mathbf{\tilde{B}}_i^\top \mathbf{r}$ in $\{\mathbf{\tilde{B}}\}$ and $\tilde{s}_j = \mathbf{\tilde{U}}_j^\top \mathbf{s}$ in $\{\mathbf{\tilde{U}}\}$ with respective variance Σ_{ii} and Λ_{jj} . The $n \times m$ crossmodes are orthogonal (approximately, up to cross-covariances that we neglect) with:

$$E[\tilde{s}_j^\top \tilde{\mathbf{r}}_i] = \tilde{\Pi}_{ij} \quad \text{and} \quad \text{CoVar}(\tilde{s}_j^\top \tilde{\mathbf{r}}_i, \tilde{s}_l^\top \tilde{\mathbf{r}}_k) = \delta_{i=k} \delta_{j=l} \Sigma_{ii} \Lambda_{jj} + \tilde{\Pi}_{ij} \tilde{\Pi}_{kl}$$

The risk $\tilde{\mathcal{R}}_{ij}$ associated with each mode $\tilde{s}_j^\top \tilde{\mathbf{r}}_i$ is:

$$\tilde{\mathcal{R}}_{ij} = \tilde{L}_{ji}^2 (\Sigma_{ii} \Lambda_{jj} + \tilde{\Pi}_{ij}^2) \approx \tilde{L}_{ji}^2 \Sigma_{ii} \Lambda_{jj}$$

and the total variance decomposes as:

$$\sum_{ij} \tilde{\mathcal{R}}_{ji} \approx \sum_{ij} \Sigma_{ii} \tilde{L}_{ji}^2 \Lambda_{jj} = \text{Tr}(\mathbf{\tilde{L}} \mathbf{\Sigma} \mathbf{\tilde{L}}^\top) \approx \text{Var}[\mathbf{w}^\top \mathbf{r}]$$

where we have neglected all n^4 -covariance terms $\tilde{\Pi} \tilde{\mathbf{L}} \tilde{\Pi} \tilde{\mathbf{L}}^\top$.

Marginal risks per return or per signal eigenmode are:

$$\begin{cases} \tilde{\mathcal{R}}(\tilde{\mathbf{b}}_i) & \approx \Sigma_{ii} \sum_j \tilde{L}_{ji}^2 \Lambda_{jj} \\ \tilde{\mathcal{R}}(\tilde{\mathbf{u}}_j) & \approx \Lambda_{jj} \sum_i \tilde{L}_{ji}^2 \Sigma_{ii} \end{cases}$$

Now, consider the Riccati basis $\{\hat{\mathbf{b}}_i\}$ and $\{\hat{\mathbf{u}}_i\}$ defined by rotations $\mathbb{R}_{\hat{\mathbf{b}}}, \mathbb{R}_{\hat{\mathbf{u}}}$ from the whitened spaces. The transformed operator is:

$$\mathbf{L}_{\hat{\mathbf{u}}\hat{\mathbf{b}}} = \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{L}_{ub} \mathbb{R}_{\hat{\mathbf{b}}} = \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbf{L} \mathbf{\Omega}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} = \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\tilde{U}} \mathbf{\Lambda}^{\frac{1}{2}} \mathbf{\tilde{L}} \mathbf{\Sigma}^{\frac{1}{2}} \mathbf{\tilde{B}}^\top \mathbb{R}_{\hat{\mathbf{b}}}$$

Variances simplifies to:

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] \approx \text{Tr}(\mathbf{L}_{\hat{\mathbf{u}}\hat{\mathbf{b}}} \mathbf{L}_{\hat{\mathbf{u}}\hat{\mathbf{b}}}^\top) = \text{Tr}(\mathbf{L}_{ub} \mathbf{L}_{ub}^\top) = \sum_{ij} L_{ub,ij}^2$$

while marginal risks per Riccati direction $\hat{\mathbf{b}}_i$ and $\hat{\mathbf{u}}_j$ become:

$$\begin{cases} \tilde{\mathcal{R}}(\hat{\mathbf{b}}_i) & \approx \sum_j L_{\hat{\mathbf{u}}\hat{\mathbf{b}},ij}^2 = \sum_j (\mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbf{L} \mathbf{\Omega}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}})_{ji}^2 \\ \tilde{\mathcal{R}}(\hat{\mathbf{u}}_j) & \approx \sum_i L_{\hat{\mathbf{u}}\hat{\mathbf{b}},ij}^2 = \sum_i (\mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbf{L} \mathbf{\Omega}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}})_{ji}^2 \end{cases}$$

Marginal risks expressed isotropic bases $\{\hat{\mathbf{b}}_i\}$ and $\{\hat{\mathbf{u}}_i\}$ serve as essential, basis-invariant metrics for enforcing isotropy across signal and return spaces. Those are exactly the Euclidean squared-norm of the column and row vectors of $\mathbf{L}_{\hat{\mathbf{u}}\hat{\mathbf{b}}}$ respectively.

2.3.5 Isotropic Mappings Between Isotropic Bases

Isotropic bases admit no privileged directions. A signal \mathbf{s} expressed as $s_{\hat{\mathbf{u}}}$ in an isotropic signal basis $\{\hat{\mathbf{u}}_i\}$ carries no *additional* risk from embedded correlations. Likewise, an isotropic return basis $\{\hat{\mathbf{b}}_i\}$ imposes no structural bias: all directions are equivalent. Working within such bases ensures *transparency*.

Yet, basis transformation is merely a computational tool, not a panacea. While certain bases may better withstand signal uncertainty, none are inherently superior.

Operating exclusively between isotropic bases $\{\hat{\mathbf{b}}_i\}, \{\hat{\mathbf{u}}_j\}$ eliminates default hidden basis bias in both signal and return spaces. However, this is insufficient: an arbitrary position $\mathbf{w}_{\hat{\mathbf{b}}} = \mathbf{L}_{\hat{\mathbf{u}}\hat{\mathbf{b}}} s_{\hat{\mathbf{u}}}$ defined via a linear mapping $\mathbf{L}_{\hat{\mathbf{u}}\hat{\mathbf{b}}}$ between such bases can reintroduce anisotropy in the output. After all, this is just a change of perspective.

The critical question is: *which linear operators preserve dual isotropy?* These form the cornerstone of our approach. Allocations that enforce *basis immunity* by construction must rely on an *isotropic linear application* \mathbf{L}_{ub} such that marginal risk is uniform across all Riccati directions.

- **Balanced** ($m = n$): The only matrices satisfying both conditions $\mathcal{R}(\mathbf{b}_i) = \mathcal{R}(\mathbf{u}_j) = \sigma^2/n$ are *scaled orthogonal matrices*:

$$\mathbf{L}_{ub} = \kappa \cdot \mathbb{R}, \quad \mathbb{R}^\top \mathbb{R} = \mathbb{I}, \quad \kappa = \sigma/\sqrt{n}.$$

In natural asset bases: $\mathbf{L} \propto \mathbf{\Xi}^{-\frac{1}{2}} \mathbb{R} \mathbf{\Omega}^{-\frac{1}{2}}$

- **Unbalanced** ($m > n$): Only return-side isotropy $\mathcal{R}(\mathbf{b}_i) = \sigma^2/n$ can be enforced everywhere. In signal space, there exist $m - n$ dimensions (i.e. $n \times (m - n)$ crossmodes) that have no contribution. The solution is a *scaled partial isometry*:

$$\mathbf{L}_{ub} = \kappa \hat{\mathbf{U}} \begin{bmatrix} \mathbb{I}_{\mathbb{d}_n} \\ 0_{(m-n),n} \end{bmatrix} \hat{\mathbf{B}}^\top,$$

with $\hat{\mathbf{B}} \in \mathbb{R}^{n,n}$, $\hat{\mathbf{U}} \in \mathbb{R}^{m,m}$ orthogonal, $\kappa = \sigma/\sqrt{n}$.

In natural bases: $\mathbf{L} \propto \mathbf{\Xi}^{-\frac{1}{2}} \hat{\mathbf{U}}_n^\top \hat{\mathbf{B}}^\top \mathbf{\Omega}^{-\frac{1}{2}}$ where $\hat{\mathbf{U}}_n^\top$ are the first- n left vectors of the matrix $\hat{\mathbf{U}}$. The remaining $m - n$ directions $\hat{\mathbf{U}}_{m-n}^\top$ span the kernel and do not contribute.

In conclusion: isotropic linear applications are scaled orthogonal when $m = n$ and scaled partial isometries when $m > n$. This structure is the geometric foundation of *Basis Immunity*.

This orthogonal (or partial isometry) form induces *uniform risk across return eigenmodes* in any isotropic basis $\{\hat{\mathbf{b}}_i\}$, but also in the eigenbasis $\{\hat{\mathbf{b}}_i\}$ —hence the term “eigenrisk parity” in [2]. This equality is a *consequence* of enforced dual isotropy, not its objective.

2.3.6 Canonical Portfolios

The mean-variance framework is agnostic to the choice of bases one decide to work with. It can be derived anywhere and will lead to the same solution (see Eq. 13) when expressed in the natural asset and signal bases.

It is enlightening to rephrase the general closed-form solution within the perspective of the isotropic basis $\{b_i\}$ and $\{u_i\}$ (or any other isotropic basis $\{\hat{b}_i\}$ and $\{\hat{u}_i\}$):

General Mean-Variance

$$w_e = \frac{1}{\gamma} \Omega^{-1} \Pi \Xi^{-1} s_e = \frac{1}{\gamma} \Omega^{-\frac{1}{2}} \downarrow \left(\overbrace{\left(\Omega^{-\frac{1}{2}} \Pi \Xi^{-\frac{1}{2}} \right)}^{\{b_i^*\} \leftarrow \{u_i^*\} \text{ in } \{u_i^*\}} \Xi^{-\frac{1}{2}} s_e \right) \quad (15)$$

The above expression involves the correlation matrix:
 $\tilde{\Pi} = \Pi_{bu} = \Omega^{-\frac{1}{2}} \Pi \Xi^{-\frac{1}{2}},$

The matrix $\tilde{\Pi}$ is cross-correlation between normalized assets and normalized signals expressed into their corresponding Riccati basis $\{b_i\}$ and $\{u_i\}$. It is also referred to as the canonical correlation matrix or just as the normalized predictability matrix.

The cross-correlation matrix $\tilde{\Pi}$, of size $n \times m$, plays an important role through its singular value decomposition (SVD):

$$\tilde{\Pi} = \tilde{B} \tilde{\Psi} \tilde{U}^T = \tilde{B} \tilde{\Psi}_n \tilde{U}_n^T \quad (16)$$

where \tilde{B} and \tilde{U} are the left and right singular vectors, of size $n \times n$ and $m \times m$ respectively, and $\tilde{\Psi}$ is the matrix of singular values, of size $n \times m$.

Because $m \geq n$, $\tilde{\Psi}$ and \tilde{U} can be respectively block-decomposed as $\tilde{\Psi} = [\tilde{\Psi}_n^T, 0_{n, m-n}]$ and $\tilde{U} = [\tilde{U}_n^T, \tilde{U}_{m-n}^T]$, where the notation M_n^{\rightarrow} to refer to the n first-left column vectors of a matrix M , while M_n^{\leftarrow} are the n last-right vectors.

Here, $\tilde{\Psi}_n^{\rightarrow}$ of size $n \times n$ is diagonal positive (corresponding to the positive singular values), while $0_{n, m-n}$ is the null matrix of size $n \times (m - n)$. $\tilde{U}_n^{\rightarrow}$ are the eigenvectors corresponding to $\tilde{\Psi}_n^{\rightarrow}$ and $\tilde{U}_{m-n}^{\leftarrow}$ is in the kernel of $\tilde{\Pi}$.

Through its singular value decomposition, we have the following:

- The singular values $\tilde{\Psi}$ can be used to compute the Sharpe ratio of the mean-variance allocation $L = \frac{1}{\gamma} \Xi^{-1} \Pi^T \Omega^{-1}$ as:

$$\begin{aligned} E[w^T r] &= \frac{1}{\gamma} \text{Tr}(\Xi^{-1} \Pi^T \Omega^{-1} \Pi) = \frac{1}{\gamma} \text{Tr}(\tilde{\Pi}^T \tilde{\Pi}) \\ \text{Var}[w^T r] &\approx \text{Tr}(\Xi L \Omega L^T) = \frac{1}{\gamma^2} \text{Tr}(\tilde{\Pi}^T \tilde{\Pi}) \\ \text{Sharpe} &= \sqrt{\text{Tr}(\tilde{\Pi}^T \tilde{\Pi})} = \sqrt{\text{Tr}(\tilde{\Psi}^2)} \quad (17) \end{aligned}$$

Because the canonical eigenvalues are bounded by one, that is we have $0 \leq \tilde{\Psi}_i \leq 1$, the Sharpe ratio is strictly bounded¹¹ by n . The cap is very large and not relevant in practice.

- In addition, Eq. 16 clearly shows that the general mean-variance allocation of Eq. 15 can be decomposed into a set of n orthogonal portfolios (out of $n \times m$ crossmodes), defined as canonical portfolios in [8]:

$$L^T = \frac{1}{\gamma} \Omega^{-\frac{1}{2}} \tilde{\Pi} \Xi^{-\frac{1}{2}} = \frac{1}{\gamma} \Omega^{-\frac{1}{2}} \tilde{B} \tilde{\Psi}_n \tilde{U}_n^T \Xi^{-\frac{1}{2}}$$

Those form a set of uncorrelated allocations that combine assets and signals so as to optimize their join predictive power (in the sense of maximizing the Sharpe ratio).

Canonical Portfolios [8]

$$\begin{aligned} w_e &= L_\star^T s_e = \frac{1}{\gamma} \sum_{k=1}^n \tilde{\Psi}_k \tilde{w}_k \\ L_\star &= \arg_L \max E[s^T L r] - \frac{\gamma}{2} \text{Var}[s^T L r] \\ L_\star &= \frac{1}{\gamma} \Xi^{-1} \Pi^T \Omega^{-1} \\ \tilde{\Pi} &= \Omega^{-\frac{1}{2}} \Pi \Xi^{-\frac{1}{2}} = \tilde{B} \tilde{\Psi} \tilde{U}^T \\ \tilde{w}_k &= \Omega^{-\frac{1}{2}} \tilde{B}_k \tilde{U}_k^T \Xi^{-\frac{1}{2}} s_e \end{aligned} \quad (18)$$

We note that one could have used any other isotropic basis without having any impact on the canonical portfolios \tilde{w}_k . For instance, with bases $\hat{b}_i = \Omega^{-\frac{1}{2}} \mathbb{R}_{\hat{b}} e_i$ and $\hat{u}_i = \Xi^{-\frac{1}{2}} \mathbb{R}_{\hat{u}} e_i$, we would have:

$$\Pi_{\hat{b}\hat{u}} = \mathbb{R}_{\hat{b}}^T \Pi_{bu} \mathbb{R}_{\hat{u}} = \mathbb{R}_{\hat{b}}^T \tilde{B} \tilde{\Psi} \tilde{U}^T \mathbb{R}_{\hat{u}}$$

The eigenvalues $\tilde{\Psi}$ are unchanged, while the eigenvectors have only been rotated, i.e. $\tilde{B} \mapsto \mathbb{R}_{\hat{b}}^T \tilde{B}$ and $\tilde{U} \mapsto \mathbb{R}_{\hat{u}}^T \tilde{U}$, leaving canonical portfolios unimpaired. This concept of rotational invariance is significant and will be revisited multiple times in this work.

The canonical portfolios, which are defined by the allocations \tilde{w}_k in Eq. 18, are leveraged and ordered by their canonical correlations $\tilde{\Psi}_1 \geq \tilde{\Psi}_2 \geq \dots \tilde{\Psi}_n \geq 0$, and as such, by their amount of linear predictability. The risk of overfitting is clearly visible in the decomposition into canonical portfolios. Assuming stable covariances Ω and Ξ , the danger lurks within the predictability matrix $\tilde{\Pi}$, precisely in the eigenspectrum $\tilde{\Psi}$. Isotropy-enforced allocations that we explore next might offer some interesting alternative.

¹¹This comes from our definition of Sharpe ratio as $E[w^T r] / \sqrt{\text{Var}[w^T r]}$. For example for any random variable x_t , we have, we have the property $\frac{1}{n} \sum |x_t| \leq \frac{1}{\sqrt{n}} \sqrt{\sum x_t^2}$.

3 Basis Immunity: Pure Isotropic Allocations

Mean-variance allocations are notoriously sensitive to input estimates, particularly the conditional expected returns $E[\mathbf{r}|\mathcal{F}]$, where small perturbations can induce dramatic shifts in portfolio weights (see Section 2.2.3).

Basis Immunity (BI) addresses the fragility of forecast signals that may be misaligned, spurious, or correlated in ways that amplify error. The central concern is not estimation noise per se, but the *compounding of uncertainty* across dimensions.

To achieve resilience, we construct allocations that minimize dependence on the implicit structure of the asset and signal covariances $\mathbf{\Omega}$ and $\mathbf{\Xi}$ (both assumed well-estimated). The goal is to prevent inevitable forecast errors from propagating—either through signal clustering (“when it rains, it pours”) or through return-side hedging that exploits fragile correlations.

The starting point is the standard mean-variance solution (Eq. 11):

$$\mathbf{w} = \frac{1}{\gamma} \mathbf{\Omega}^{-1} E[\mathbf{r}|\mathcal{F}] = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{M}^\top \mathbf{s} \quad (19)$$

where $E[\mathbf{r} | \mathcal{F}] = \mathbf{M}^\top \mathbf{s}$ captures our best estimate of future returns as a linear function of the signals \mathbf{s} , and γ is fixed via the variance constraint (Eq. 7).

BI minimally perturbs Eq. 19 while *strictly enforcing isotropy* in both signal and return spaces. That is BI follows exactly the isotropy philosophy of ERP introduced in [2].

The objective is not to eliminate risk, but to neutralize the *basis risk* arising from privileged coordinate systems—such as the natural asset and signal bases. Robustness to uncertainty is achieved *through enforced isotropy*. The difficulty comes from the impossibility of finding an optimal transformation that fits both asset and signal perspectives simultaneously.

To ensure transparency, we work in *isotropic bases*. As defined in Section 2.3.3, the set of isotropic asset bases $\mathcal{S}_{\mathbf{\Omega}} \subset \mathcal{H}_r$ consists of all coordinate systems of the form $\hat{\mathbf{b}}_i = \mathbf{\Omega}^{-1/2} \mathbb{R}_{\hat{\mathbf{b}}} \mathbf{e}_i$, where $\mathbb{R}_{\hat{\mathbf{b}}}$ is a rotation. In such a basis, the asset covariance becomes $\mathbf{\Omega}_{\hat{\mathbf{b}}} = \mathbb{I}$, eliminating privileged risk directions. From the return viewpoint, any predictive signal generates no additional structural risk. Similarly, isotropic signal bases $\mathcal{S}_{\mathbf{\Xi}} \subset \mathcal{H}_s$ are given by $\hat{\mathbf{u}}_i = \mathbf{\Xi}^{-1/2} \mathbb{R}_{\hat{\mathbf{u}}} \mathbf{e}_i$, yielding $\mathbf{\Xi}_{\hat{\mathbf{u}}} = \mathbb{I}$.

Under full isotropy, the allocation must satisfy dual symmetry: risk is spherical in *both* whitened return and signal spaces — the underlying principle of Basis Immunity.

We describe the high-level principles on which such allocations are constructed:

- The original signals \mathbf{s} carry some risk through their covariance $\mathbf{\Xi}$. Unavoidable (and frequent) errors in \mathbf{s} could be magnified due to their covariance $\mathbf{\Xi}$ (bad things come together). The idea is then to adjust them through a small transformation \mathbf{T} designed such that $\mathbf{T}\mathbf{s}$ becomes isotropic, that is $\mathbf{T}\mathbf{\Xi}\mathbf{T}^\top = \mathbb{I}$. This is equivalent to expressing $\mathbf{T} = \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{-\frac{1}{2}}$ and the deformation effectively amounts to replacing the original signals \mathbf{s} by some isotropic signals $\mathbf{s}_{\hat{\mathbf{u}}} = \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}$.

If the deformation \mathbf{T} is small enough, one can hope to retain the predictive power of the original signals, while being less exposed to all the unavoidable errors that will arise time after time (known unknowns and unknown unknowns).

- In return space, a similar situation occurs. The positions \mathbf{w} in $\{\mathbf{e}_i\}$, which are derived from the expected returns $E[\mathbf{r}|\mathcal{F}]$, might carry some covariance risk¹² in the form of $\mathbf{w}^\top \mathbf{\Omega} \mathbf{w}$. Errors in the position vector \mathbf{w} could get magnified because of the non-diagonal covariance terms of $\mathbf{\Omega}$. Clearly, if the natural basis were to be isotropic, this risk would naturally disappear.

To emulate this desired behavior, the expected return $E[\mathbf{r}|\mathcal{F}]$, originally derived in $\{\mathbf{e}_i^\star\}$, are used as such in a different isotropic basis $\{\hat{\mathbf{b}}_i^\star\}$. The newly derived positions $\mathbf{w}_{\hat{\mathbf{b}}}$, which would not carry any covariance risk since $\mathbf{\Omega}_{\hat{\mathbf{b}}} = \mathbb{I}$, would then be expressed back into the original basis (where trading takes place), leading to $\mathbf{w}_e = \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \mathbf{w}_{\hat{\mathbf{b}}}$.

Bu there is no free-lunch. As we are now using $E[\mathbf{r}_e|\mathcal{F}]$ instead of $E[\mathbf{r}_{\hat{\mathbf{b}}}|\mathcal{F}]$ in the isotropic basis $\{\hat{\mathbf{b}}_i^\star\}$, we are modifying our expected PnL. Now, if the isotropic basis $\{\hat{\mathbf{b}}_i\}$ is close enough from $\{\mathbf{e}_i\}$, one can hope that the difference is small enough, while the basis risk would disappear.

In a nutshell, the rough idea behind BI allocations is:

1. replace the original signals \mathbf{s} by some better-behaved ones $\mathbf{s}_{\hat{\mathbf{u}}}$ (better-behaved from a risk-perspective),
2. replace the natural basis $\{\mathbf{e}_i^\star\}$ where expected returns $E[\mathbf{r}|\mathcal{F}]$ are computed by an isotropic basis $\{\hat{\mathbf{b}}_i^\star\}$ with the approximation $E[\mathbf{r}_{\hat{\mathbf{b}}}|\mathcal{F}] \approx E[\mathbf{r}_e|\mathcal{F}]$.

¹²In the mean-variance framework, where the variance is constrained, this leads to the decorrelation operator $\mathbf{\Omega}^{-1}$ being applied to $E[\mathbf{r}|\mathcal{F}]$ in Eq. 19.

Both approximations must obviously be done in a controlled way so that Eq. 19, our departing point, is not completely “destroyed”. This is not a trivial task as both approximations are rarely compatible with a given optimization problem.

To investigate, we start below with the simpler balanced case $m = n$. The general case where $m \geq n$ will be explored in Section 3.2.

3.1 The Balanced Case $m = n$ and $E[r|\mathcal{F}] \propto s$

It is usual to work with as many signals as there are assets, where each signal s_i has been designed specifically for a corresponding asset S_i , with $E[r_i|\mathcal{F}] \propto s_i$. In that case, we can link the two dual bases $\{e_i^{\star}\}$ and $\{e_i^{\star\star}\}$, setting the mapping operator to the identity $\mathbf{M} = \mathbb{I}$.

Within this setup (identical to the one described in [2]), the goal is to minimally disrupt the mean-variance (MV) allocation:

$$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{s}_e \quad (20)$$

while ensuring isotropy on both sides (asset returns and signals).

3.1.1 Previous Approaches

As we discussed above, the asset Riccati basis $\mathbf{b}_i = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{e}_i$ and the signal Riccati basis $\mathbf{u}_i = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{e}_i$ that we previously defined are the closest to the natural basis $\{\mathbf{e}_i\}$ for the Mahalanobis distances $\mathbf{D}_{\mathbf{\Omega}}$ and $\mathbf{D}_{\mathbf{\Xi}}$ defined above.

◦ From this proximity property and using a symmetry argument, Benichou and al. advocate in [2] for an allocation of the form:

$$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \quad (21)$$

The signals $\mathbf{s}_e = \mathbf{s}$ are replaced by the closest isotropic transformation $\mathbf{s}_u = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$ (in the sense of $\mathbf{D}_{\mathbf{\Xi}}$):

$$\mathbf{s}_e \longleftarrow \mathbf{s}_u = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e,$$

while the natural assets is substituted for its closest isotropic asset basis (in the sense of $\mathbf{D}_{\mathbf{\Omega}}$), thereby replacing $\mathbf{w}_e = \mathbf{\Omega}^{-1} E[r|\mathcal{F}]$ by $\mathbf{w}_e = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{w}_b$ where the expected returns in $\{\mathbf{b}_i^{\star}\}$ are interchanged by the ones in $\{\mathbf{e}_i^{\star}\}$, that is:

$$\mathbf{w}_e = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{w}_b \text{ where } E[\mathbf{r}_b|\mathcal{F}] \longleftarrow E[\mathbf{r}_e|\mathcal{F}]$$

Issued from a simple argument of symmetry (in return and signal spaces) and proximity (in the sense of the Mahalanobis distance), the solution Eq. 21 is elegant, practical, and effective.

◦ However, this is not the only approach. Segonne et al. advocate in [22] to directly work in the isotropic basis $\{\mathbf{b}_i\}$ where the targeted MV allocation of Eq. 20 takes the simple form:

$$\mathbf{w}_b = \frac{1}{\gamma} \mathbf{s}_b$$

In $\{\mathbf{b}_i\}$, only the signal approximation is needed (since $\{\mathbf{b}_i\}$ is already return-isotropic, no return approximation is required). Using a similar argument of proximity in $\{\mathbf{b}_i\}$ (and not in $\{\mathbf{e}_i\}$), the closest isotropic signal basis of \mathcal{H}_s is not $\mathbf{u}_i = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{e}_i$ but $\hat{\mathbf{u}}_i = \mathbf{\Xi}_b^{-\frac{1}{2}} \mathbf{b}_i$ where $\mathbf{\Xi}_b = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi} \mathbf{\Omega}^{-\frac{1}{2}}$ (that is using $\mathbf{D}_{\mathbf{\Xi}_b} \neq \mathbf{D}_{\mathbf{\Xi}}$). Therefore, from the perspective of the signal distance $\mathbf{D}_{\mathbf{\Xi}_b}$, one should replace \mathbf{s}_b by $\mathbf{\Xi}_b^{-\frac{1}{2}} \mathbf{s}_b$, leading to a different allocation:

$$\begin{aligned} \mathbf{w}_e &= \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{w}_b && \text{change of basis } \mathbf{b}_i \mapsto \mathbf{e}_i \\ &= \mathbf{\Omega}^{-\frac{1}{2}} \left(\frac{1}{\gamma} \mathbf{\Xi}_b^{-\frac{1}{2}} \mathbf{s}_b \right) && \text{approximation } \mathbf{s}_b \longleftarrow \mathbf{s}_{\hat{u}} \\ &= \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi} \mathbf{\Omega}^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{s}_e \end{aligned} \quad (22)$$

The solution Eq 22 coincides with Eq. 21 when the two covariances $\mathbf{\Omega}$ and $\mathbf{\Xi}$ commute¹³. However, when this is not the case, Eq 22 is “closer” (in the Mahalanobis sense) to the initial mean-variance solution Eq. 20 than the allocation Eq. 21 proposed in [2]. As we discuss below in Section 3.1.3, the difference is small.

Interestingly, the solution does not depend on the specific isotropic basis $\{\mathbf{b}_i\}$. It would be identical in any other isotropic asset basis $\{\hat{\mathbf{b}}_i\}$ (that is of the form $\hat{\mathbf{b}}_i = \mathbf{\Omega}_u^{-\frac{1}{2}} \mathbb{R}_b \mathbf{e}_i$).

Finally, we note that we can rewrite Eq. 22 as:

$$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_b^{\star} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \quad (23)$$

where we can verify that $\mathbb{R}_b^{\star} = \left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi} \mathbf{\Omega}^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}}$ is a rotation. By explicitly working in the isotropic basis $\{\mathbf{b}_i\}$, only a signal proximity argument is needed and the symmetry argument not required anymore. From the perspective of the Mahalanobis distance, Eq. 23 is less disruptive than Eq. 21.

◦ Now, a symmetrical argument could also be constructed by considering the isotropic signal basis $\{\mathbf{u}_i\}$ and slightly adjusting the asset basis. In this scenario, no signal approximation would be needed, only the approximation in return space. We substitute the targeted solution of Eq. 20:

$$\mathbf{w}_u = \mathbf{\Omega}_u^{-1} E[\mathbf{r}_u|\mathcal{F}] = \mathbf{\Omega}_u^{-1} \mathbf{s}_u,$$

¹³This would be the case if the covariance $\mathbf{\Xi}$ is chosen as $\mathbf{\Xi} \propto \varphi \mathbf{\Omega} + (1 - \varphi) \mathbb{I}$ as advocated in [2].

by its closest isotropic asset allocation:

$$\mathbf{w}_u = \mathbf{\Omega}_u^{-\frac{1}{2}} \mathbf{w}_{\hat{b}} \text{ where } E[\mathbf{r}_{\hat{b}}|\mathcal{F}] \approx E[\mathbf{r}_u|\mathcal{F}] = \mathbf{s}_u,$$

where we have $\mathbf{\Omega}_u = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega} \mathbf{\Xi}^{-\frac{1}{2}}$. This would lead to the solution:

$$\begin{aligned} \mathbf{w}_e &= \mathbf{\Xi}_u^{-\frac{1}{2}} \mathbf{w}_u && \text{change of basis } \mathbf{u}_i \mapsto \mathbf{e}_i \\ &= \mathbf{\Xi}_u^{-\frac{1}{2}} \mathbf{\Omega}_u^{-\frac{1}{2}} \mathbf{w}_{\hat{b}} && \text{change of basis } \hat{\mathbf{b}}_i \mapsto \mathbf{u}_i \\ &= \frac{1}{\gamma} \mathbf{\Xi}_u^{-\frac{1}{2}} \mathbf{\Omega}_u^{-\frac{1}{2}} \mathbf{s}_u && \text{approx } E[\mathbf{r}_{\hat{b}}|\mathcal{F}] \leftarrow \mathbf{s}_u \\ &= \frac{1}{\gamma} \mathbf{\Xi}^{-\frac{1}{2}} \left(\mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega} \mathbf{\Xi}^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \end{aligned}$$

Similarly, the solution is valid for all isotropic signal basis $\{\hat{\mathbf{u}}_i\}$ (that is not only for $\{\mathbf{u}_i\}$) and could be rewritten as:

$$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{R}_u^* \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \quad (24)$$

where $\mathbf{R}_u^* = \mathbf{\Omega}^{\frac{1}{2}} \mathbf{\Xi}^{-\frac{1}{2}} \left(\mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega} \mathbf{\Xi}^{-\frac{1}{2}} \right)^{-\frac{1}{2}}$ is a rotation.

3.1.2 Balanced Isotropic Allocations

One should quickly realize that the three above allocations Eq. 21 (from [2]), Eq 23 (from [22]), and Eq 24 are not unique and that the space of potential solutions is infinite. Specifically, any isotropic solution can be written as (see Section 2.3.5):

Balanced Isotropy-Enforced Allocation Form

$$m = n, \mathbf{M} = \mathbb{I}$$

$$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{R} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \underbrace{\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{R}_{\hat{b}}}_{\text{return}} \downarrow \underbrace{\mathbf{R}_{\hat{u}}^{\top} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e}_{\text{signal}} \quad (25)$$

where $\mathbf{R}_{\hat{b}}, \mathbf{R}_{\hat{u}}$ are rotation operators associated with isotropic bases $\hat{\mathbf{b}}_i = \mathbf{\Omega}_u^{-\frac{1}{2}} \mathbf{R}_{\hat{b}} \mathbf{e}_i$ and $\hat{\mathbf{u}}_i = \mathbf{\Xi}_u^{-\frac{1}{2}} \mathbf{R}_{\hat{u}} \mathbf{e}_i$. The down-arrow has been added to explicitly indicate a transformation from \mathcal{H}_r^* to \mathcal{H}_r in the basis $\{\hat{\mathbf{b}}_i\}$. Plugging Eq. 25 into the variance constraint, the leverage coefficient γ can be computed as $\gamma\sigma = \sqrt{n}$.

As discussed in Section 2.3.5, isotropic allocations, i.e. of the form Eq. 25, carry the same risk in all directions (both in signal and return spaces). In particular, each eigenmode, in return and signal spaces, carries the same risk, equal to $1/n$ of the total variance σ^2 . This justifies the term ‘‘Eigenrisk Parity’’ coined in [2].

Within the manifold of isotropic allocations, a selection problem remains: we must choose the orthogonal operator $\mathbf{R} = \mathbf{R}_{\hat{b}} \mathbf{R}_{\hat{u}}^{\top}$ that connects the return and signal bases. Not all choices are equal—some induce excessive deformation of the original mean-variance signal, failing to preserve predictive structure.

Although the Mahalanobis distances $D_{\mathbf{\Omega}}$ and $D_{\mathbf{\Xi}}$ could be used to rank candidate solutions, we adopt an alternative, provably equivalent approach that generalizes naturally to the over-determined case $m \geq n$.

To do so, we rephrase the mean-variance solution of Eq. 20 (our starting point) as:

$$\mathbf{w} = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{s} = \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{R}_{\hat{b}} \left(\mathbf{R}_{\hat{b}}^{\top} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} \mathbf{R}_{\hat{u}} \right) \mathbf{R}_{\hat{u}}^{\top} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s} \quad (26)$$

By comparing Eq. 25 with Eq. 26, we then define the optimal allocation by selecting the rotations $\mathbf{R}_{\hat{b}}$ and $\mathbf{R}_{\hat{u}}$ that best aligns the linear mapping $\mathbf{R}_{\hat{b}}^{\top} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} \mathbf{R}_{\hat{u}}$ with \mathbb{I}_n . That is we search for the closest isotropy. To do so, we minimize:

$$\|\mathbf{R}_{\hat{b}}^{\top} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} \mathbf{R}_{\hat{u}} - \mathbb{I}_n\|_{\mathbb{F}}^2 = \|\mathbf{R} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} - \mathbb{I}_n\|_{\mathbb{F}}^2 \quad (27)$$

where $\|\mathbf{A}\|_{\mathbb{F}} = \sqrt{\text{Tr}(\mathbf{A}^{\top} \mathbf{A})}$ is the usual Frobenius norm. The optimal rotation \mathbf{R}^* is the one maximizing the following:

$$\mathbf{R}^* = \arg_{\mathbf{R}} \max \text{Tr} \left(\mathbf{R} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} \right) \quad (28)$$

Using the singular value decomposition¹⁴ of $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}}$:

$$\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} = \hat{\mathbf{B}} \hat{\mathbf{\Psi}} \hat{\mathbf{U}}^{\top}, \quad (29)$$

the solution can be expressed as:

Balanced Dual-Isotropy Allocation
 $m = n, \mathbf{M} = \mathbb{I}$

$$\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{R}^* \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \text{ with } \mathbf{R}^* = \hat{\mathbf{B}} \hat{\mathbf{U}}^{\top}$$

$$\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} = \hat{\mathbf{B}} \hat{\mathbf{\Psi}} \hat{\mathbf{U}}^{\top}$$

¹⁴Please note that $\hat{\mathbf{\Psi}}^2$, the squared singular values of $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}}$, are the eigenvalues of $\mathbf{\Xi}_b = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi} \mathbf{\Omega}^{-\frac{1}{2}}$ while $\hat{\mathbf{\Psi}}^{-2}$ are the eigenvalues of $\mathbf{\Omega}_b = \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega} \mathbf{\Xi}^{-\frac{1}{2}}$.

3.1.3 Discussion

The solution only depends on the product $\mathbb{R}^* = \mathbb{R}_{\hat{b}} \mathbb{R}_{\hat{a}}^\top$, not on specific realizations of $\mathbb{R}_{\hat{b}}$ and $\mathbb{R}_{\hat{a}}$ (a similar observation was made with the solutions Eq. 23 and Eq. 24). Because $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{+\frac{1}{2}}$ is not symmetric (except in some special cases, such as having $\mathbf{\Xi}$ and $\mathbf{\Omega}$ commute), $\hat{\mathbf{B}} \neq \hat{\mathbf{U}}$ and \mathbb{R}^* is not the identity matrix.

The closed-form solution of Eq. 30 is the isotropic allocation that minimizes the deformation between the set of all isotropic bases $S_{\mathbf{\Omega}}$ and $S_{\mathbf{\Xi}}$. It finds the rotation \mathbb{R}^* that align best the two isotropic subspaces (see [21, 9]).

Equivalence with the Distance Approach [2, 22]

Interestingly, we can show that the above approach is equivalent to the distance minimization. Focusing on Eq. 25, we follow the same reasoning that led to Eq. 23 and Eq. 24. Two views are then possible, depending if we chose to work from the return or the signal angles:

- From the asset perspective, working within the isotropic asset basis $\hat{\mathbf{b}}_i = \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{b}} \mathbf{e}_i$, we are replacing the signal vector $\mathbf{s}_{\hat{b}} = \mathbb{R}_{\hat{b}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{s}_e$ by another one $\mathbf{s}_{\hat{a}} = \mathbb{R}_{\hat{a}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$ that has the desired property of being signal isotropic¹⁵:

$$\mathbf{s}_{\hat{a}} = \mathbb{R}_{\hat{a}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e = \mathbb{R}_{\hat{a}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega}^{+\frac{1}{2}} \mathbb{R}_{\hat{b}} \mathbf{s}_{\hat{b}}$$

The distance associated with the approximation is:

$$\begin{aligned} D_{\mathbf{\Xi}_{\hat{b}}}(\mathbf{s}_{\hat{a}}, \mathbf{s}_{\hat{b}}) &= D_{\mathbf{\Xi}_{\hat{b}}}(\mathbb{R}_{\hat{a}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega}^{+\frac{1}{2}} \mathbb{R}_{\hat{b}} \mathbf{s}_{\hat{b}}, \mathbf{s}_{\hat{b}}) \\ &= D_{\mathbf{\Xi}}(\mathbf{\Omega}^{+\frac{1}{2}} \mathbb{R}_{\hat{a}} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e, \mathbf{s}_e) \\ &= \|\mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega}^{\frac{1}{2}} \mathbb{R} - \mathbb{Id}\|_{\mathbb{F}}^2 \end{aligned} \quad (31)$$

- Conversely, we could be working from the signal perspective using the isotropic signal basis $\hat{\mathbf{u}}_i = \mathbf{\Xi}^{-\frac{1}{2}} \mathbb{R}_{\hat{a}} \mathbf{e}_i$ as a starting point. That is, instead of working in the return-isotropic basis $\{\hat{\mathbf{b}}_i\}$ and choosing the closest isotropic signals from $\mathbf{s}_{\hat{b}}$ (based on $D_{\mathbf{\Xi}_{\hat{b}}}$), we would now fix the used signals as $\mathbf{s}_{\hat{a}}$ and pick the closest isotropic basis from $\{\hat{\mathbf{u}}_i\}$ (based on $D_{\mathbf{\Omega}_{\hat{a}}}$).

By doing that, we are effectively replacing the expected position $\mathbf{w}_{\hat{a}} = \frac{1}{\gamma} \mathbf{\Omega}_{\hat{a}}^{-1} \mathbf{s}_{\hat{a}}$, which carries some return covariance risk, by another one $\mathbf{w}_{\hat{b}} = \frac{1}{\gamma} \mathbf{s}_{\hat{a}}$ that is now return-isotropic:

$$\begin{aligned} \mathbf{w}_e &= \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{b}} \mathbf{w}_{\hat{b}} \\ &= \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{b}} \mathbf{s}_{\hat{a}} \\ &= \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{b}} \mathbb{R}_{\hat{a}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \end{aligned}$$

The approximation $E[\mathbf{r}_{\hat{b}} | \mathcal{F}] \leftarrow E[\mathbf{r}_{\hat{a}} | \mathcal{F}] = \mathbf{s}_{\hat{a}}$ is measured as:

$$\begin{aligned} D_{\mathbf{\Omega}_{\hat{a}}}(\mathbf{r}_{\hat{b}}, \mathbf{r}_{\hat{a}}) &= D_{\mathbf{\Omega}_{\hat{a}}}(\mathbb{R}_{\hat{b}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r}_e, \mathbb{R}_{\hat{a}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{r}_e) \\ &= D_{\mathbf{\Omega}}(\mathbf{\Xi}^{+\frac{1}{2}} \mathbb{R}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{r}_e, \mathbf{r}_e) \\ &= \|\mathbb{R} \mathbf{\Xi}^{\frac{1}{2}} \mathbf{\Omega}^{-\frac{1}{2}} - \mathbb{Id}\|_{\mathbb{F}}^2 \end{aligned} \quad (32)$$

Using the singular value decomposition of $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{+\frac{1}{2}}$ in Eq. 29, one can quickly see that both Eq. 31 and Eq. 32 lead to the same solution $\mathbb{R}^* = \hat{\mathbf{B}} \hat{\mathbf{U}}^\top$.

Besides, one can easily show that both allocations Eq. 23 and Eq. 24 are the same and equal to the optimal allocation of Eq. 30. This is not surprising, since both were designed to minimize $D_{\mathbf{\Xi}}$ and $D_{\mathbf{\Omega}}$ respectively:

$$\begin{aligned} \mathbb{R}^* &= \mathbb{R}_b^* = \mathbb{R}_u^* = \hat{\mathbf{B}} \hat{\mathbf{U}}^\top \\ \mathbb{R}_b^* &= \left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi} \mathbf{\Omega}^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{\frac{1}{2}} \\ \mathbb{R}_u^* &= \mathbf{\Omega}^{\frac{1}{2}} \mathbf{\Xi}^{-\frac{1}{2}} \left(\mathbf{\Xi}^{-\frac{1}{2}} \mathbf{\Omega} \mathbf{\Xi}^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \end{aligned}$$

On the other hand, the ERP allocation of Eq. 21 proposed in [2], which amounts to $\mathbb{R} = \mathbb{Id}$, does not generally correspond to an optimum of our metric in Eq. 27, but should most of the time be close (except for rare pathological cases that would not make sense in practice, see below). The distortion of the entry signals \mathbf{s} is still controlled, although to a lesser extent, thanks to the proximity of the two Riccati basis $\{\mathbf{b}_i\}$ and $\{\mathbf{u}_i\}$. We can show (as noticed in [22]) that when both covariances commute $\mathbf{\Xi} \mathbf{\Omega} = \mathbf{\Omega} \mathbf{\Xi}$, then $\mathbb{R}^* = \mathbb{Id}$ and all solutions collapse to Eq. 21.

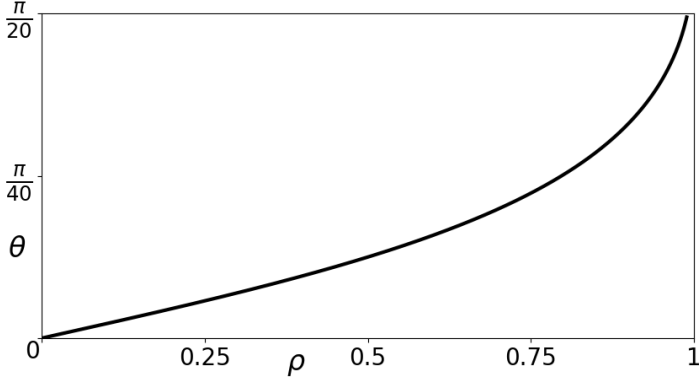
BI versus ERP [2]: Study of a Pathological Case

To illustrate some differences, we consider a pathological case. We consider 3 assets, where only asset 1 and 2 two are return-correlated at ρ , whereas asset 2 and 3 are signal-correlated at $-\rho$.

$$\mathbf{\Omega} = \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathbf{\Xi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\rho \\ 0 & -\rho & 1 \end{pmatrix}$$

As ρ varies from 0 to 1, the optimal rotation \mathbb{R}_ρ^* deviates more and more from the identity matrix. As an element of $SO(3)$, we display below the minimal rotation angle $\theta = \cos^{-1} \left(\frac{\text{Tr}(\mathbb{R}_\rho^*) - 1}{2} \right)$ as a function of ρ . Even for extreme values $\rho \rightarrow 1$, the distortion remains rather small.

¹⁵Because $\{\hat{\mathbf{b}}_i\}$ is return isotropic, the transformation associated with the bilinear form \bullet from covectors space $\hat{\mathbf{b}}_i^*$ into $\hat{\mathbf{b}}_i$ is the identity.



Agnostic Risk Parity

Agnostic Risk Parity is derived in [2] as a special case of the ERP framework. The authors note that Ξ is difficult to reliably estimate and propose a simple parametric form:

$$\Xi = \varphi \Omega + (1 - \varphi) \mathbb{I},$$

which has proven effective in trend-following applications.

Under this assumption, Ξ and Ω commute, and in the balanced case ($m = n$), the ARP allocation exactly coincides with *Basis Immunity*.

Conclusion

Allocations such as Eq. 21 or Eq. 30 are pure isotropic allocations. They assume that the signals s are predictive of future returns, with $E[r|F] \propto s$, and seek to *minimally perturb* the mean-variance benchmark $\Omega^{-1} E[r | F] \propto \Omega^{-1} s$, while enforcing dual isotropy.

However, there is *no guarantee* that the transformed signals retain sufficient predictive power, as the construction is driven *solely by risk considerations*. Even when isotropy is optimally aligned, the final allocation may deviate significantly from the original, particularly when Ω and Ξ are poorly conditioned, potentially leading to *unintended concentration or catastrophic underperformance*.

Controlling the amount of lost predictability while aiming to be as isotropic as possible (in signal space and/or asset space) is therefore essential. We explore this *tunable regularization* in Section 4 via the *Isotropy-Regularized Mean-Variance* framework.

Before doing so, we extend the “pure” isotropic construction to the general case where the mapping \mathbf{M} is not the identity, while the number of signals differ from the number of assets $m \neq n$.

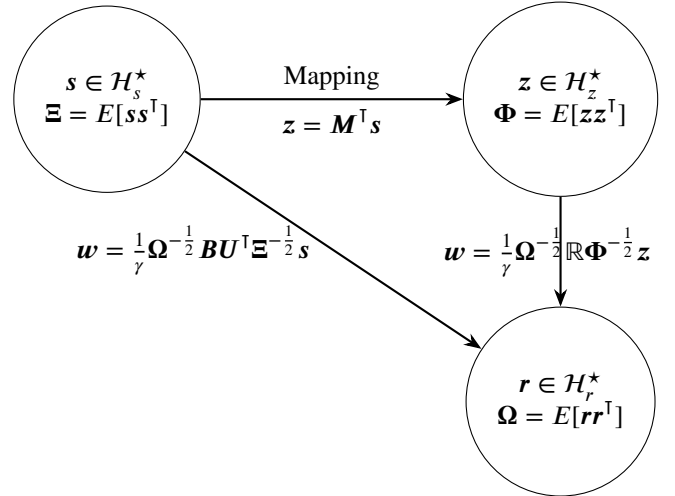
3.2 Unbalanced Case $m \geq n$ and $E[r|F] \propto \mathbf{M}^\top s$

The general case where the number of signals differs from the number of assets requires a mapping, assumed to be a linear application, from the space of signals into the space of assets. This is achieved through the operator \mathbf{M}^\top from signal dual space \mathcal{H}_s^* into return dual space \mathcal{H}_r^* :

$$s \in \mathcal{H}_s^* \mapsto z = \mathbf{M}^\top s \in \mathcal{H}_z^* \sim \mathcal{H}_r^*$$

Each signal z_i , assumed to be predictive of future return r_i , is constructed as a linear aggregation (fixed given weights) of several signals $z_i = \sum_j M_{ji} s_j$ (potentially all of them):

$$E[r|F] \propto z = \mathbf{M}^\top s$$



As we already discussed, the mapping \mathbf{M}^\top can be obtained through several options, for instance e.g. based on explicit deterministic relationships, or by using statistical linear regression with $\mathbf{M}^\top = \beta$, or even directly within a mean-variance optimization leading to $\mathbf{M}^\top = \Pi \Xi^{-1}$ (see section 2.2). The important point is that \mathbf{M} is known and given.

The matrix $\Phi = E[zz^\top] = \mathbf{M}^\top \Xi \mathbf{M}$ is then the covariance of the mapped signals $z = \mathbf{M}^\top s$ in $\mathcal{H}_z^* \sim \mathcal{H}_r^*$; it is typically invertible when $m \geq n$ (except in pathological cases)¹⁶.

We extend our definition of isotropic allocation in Eq. 25 to accommodate/link signal and return spaces with potentially different dimensions (as $m \geq n$) as (see Section 2.3.5):

Unbalanced Isotropy-Enforced Allocation Form

$$\underbrace{w_e}_{\text{asset}} = \underbrace{\frac{1}{\gamma} \Omega^{-\frac{1}{2}} \hat{\mathbf{B}}}_{\text{signal}} \downarrow \hat{\mathbf{U}}_n^\top \Xi^{-\frac{1}{2}} s_e \quad \text{with} \quad \hat{\mathbf{B}}^\top \hat{\mathbf{B}} = \hat{\mathbf{U}}_n^\top \hat{\mathbf{U}}_n = \mathbb{I}_{\mathbb{I}_n} \quad (33)$$

¹⁶When $m < n$, the inverse of Φ does not exist, but we do not consider this case here.

where the two matrices $\hat{\mathbf{B}}$ (of size $n \times n$) and $\hat{\mathbf{U}}_n$ (of size $m \times n$) encode n orthonormal vectors of \mathcal{H}_r^\star and \mathcal{H}_s^\star respectively, i.e. $\hat{\mathbf{B}}^\top \hat{\mathbf{B}} = \hat{\mathbf{U}}_n^\top \hat{\mathbf{U}}_n = \mathbb{I}_n$. The linear application $\hat{\mathbf{B}}\hat{\mathbf{U}}_n^\top$ is the equivalent of the operator $\mathbb{R} = \mathbb{R}_b \mathbb{R}_a^\top$ in Eq. 25. Those are partial isometry (see Section 2.3.5).

When $m > n$, Eq. 33 extracts n orthogonal directions $\hat{\mathbf{U}}_n$ of the isotropic signals $\Xi^{-\frac{1}{2}} s_e$ (hence the subscript n), thereby focusing on a submanifold ($\sim \mathbb{R}^n \subset \mathbb{R}^m$) of an isotropic basis $\{\hat{\mathbf{u}}_i\}$. The n features are then mapped to an isotropic basis $\{\hat{\mathbf{b}}_i\}$ through $\hat{\mathbf{B}}$.

Compared to the balanced case where $m = n$, unbalanced isotropic allocations of the form Eq. 25 cannot carry the same risk in all signal directions. As $m > n$, any linear operator of the form $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$ will have $m - n$ directions that be in the kernel. However, as in the balanced case, each return eigenmode in the Riccati basis carries equal risk—precisely $1/n$ of the total portfolio variance σ^2 . The term “eigenrisk parity” remains descriptively valid.

The mapped signals $\mathbf{z} = \mathbf{M}^\top \mathbf{s}$ provides us with some estimates of future returns as $E[\mathbf{r}|\mathcal{F}] = \mathbf{z}$. Two approaches are possible depending if we prefer to work in the space of mapped signals \mathcal{H}_z^\star or in the original space \mathcal{H}_s^\star .

3.2.1 Working with mapped signals $\mathbf{z} = \mathbf{M}^\top \mathbf{s}$

Working from the perspective of the signals \mathbf{z} in \mathcal{H}_z^\star , we directly apply the previous results of section 3.1. For instance, using Eq. 23, we obtain:

$$\begin{aligned} \mathbf{w}_e &= \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \Phi_b^{-\frac{1}{2}} \mathbf{z}_b \\ &= \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \left(\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi \mathbf{M} \Omega^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{s}_e \\ &= \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \mathbb{R}_b \left(\mathbf{M}^\top \Xi \mathbf{M} \right)^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{s}_e \end{aligned} \quad (34)$$

where $\mathbb{R}_b = \left(\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi \mathbf{M} \Omega^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} \left(\mathbf{M}^\top \Xi \mathbf{M} \right)^{\frac{1}{2}}$ is a rotation.

From our previous discussion, we know that it can expressed from the singular value decomposition of $\Omega^{-\frac{1}{2}} \Phi^{+\frac{1}{2}} = \Omega^{-\frac{1}{2}} \left(\mathbf{M}^\top \Xi \mathbf{M} \right)^{+\frac{1}{2}}$ as:

$$\mathbb{R}_b = \check{\mathbf{B}} \check{\mathbf{U}}^\top \text{ with } \Omega^{-\frac{1}{2}} \left(\mathbf{M}^\top \Xi \mathbf{M} \right)^{+\frac{1}{2}} = \check{\mathbf{B}} \check{\mathbf{U}} \check{\mathbf{U}}^\top$$

3.2.2 Working in \mathcal{H}_s^\star

We can also work directly from the space of signals \mathbf{s} in \mathcal{H}_s^\star . The mean-variance framework can be rephrased as:

$$\begin{aligned} \mathbf{w}_e &= \frac{1}{\gamma} \Omega^{-1} \mathbf{M}^\top \mathbf{s}_e \\ &= \frac{1}{\gamma} \Omega^{-\frac{1}{2}} \mathbf{U} \left(\mathbf{U}^\top \Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{\frac{1}{2}} \mathbf{V} \right) \mathbf{V}^\top \Xi^{-\frac{1}{2}} \mathbf{s}_e \end{aligned} \quad (35)$$

Comparing Eq. 35 with Eq. 33, we define the “best” isotropic allocation as the two orthonormal bases, encoded by the matrices $\check{\mathbf{B}}$ and $\check{\mathbf{U}}$ of size $n \times n$ and size $m \times n$ respectively, that aligns best $\check{\mathbf{B}} \Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{\frac{1}{2}} \check{\mathbf{U}}$ with the identity matrix \mathbb{I}_n . The basis $\check{\mathbf{U}}$ spans a linear space $\sim \mathbb{R}^n$ that is strictly included in $\mathcal{H}_s^\star \sim \mathbb{R}^m$ as soon as $n < m$.

This can be easily determined through a singular value decomposition of $\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{+\frac{1}{2}}$ (and keeping for $\check{\mathbf{U}}$ only the first n right-singular vectors $\check{\mathbf{U}}_n$):

$$\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{+\frac{1}{2}} = \check{\mathbf{B}} \check{\mathbf{U}} \check{\mathbf{U}}_n^\top = \check{\mathbf{B}} \check{\mathbf{U}}_n \check{\mathbf{U}}_n^\top \quad (36)$$

We end up with a generalization of Eq. 30:

Isotropy-Enforced Allocation
 $m \geq n, \quad E[\mathbf{r}|\mathcal{F}] \propto \mathbf{M}^\top \mathbf{s}$

$$\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \check{\mathbf{B}} \check{\mathbf{U}}_n^\top \Xi^{-\frac{1}{2}} \mathbf{s}_e \text{ with } \Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{+\frac{1}{2}} = \check{\mathbf{B}} \check{\mathbf{U}} \check{\mathbf{U}}_n^\top$$

(37)

As $m \geq n$, the rank of $\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{+\frac{1}{2}}$ is n (except in pathological cases). We note that we have the following:

$$\left(\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi \mathbf{M} \Omega^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi^{+\frac{1}{2}} = \check{\mathbf{B}} \check{\mathbf{U}}_n^\top$$

This leads to the same solution as Eq. 34 above, that is:

$$\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \left(\Omega^{-\frac{1}{2}} \mathbf{M}^\top \Xi \mathbf{M} \Omega^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{s}_e$$

It is worth mentioning that one could have chosen to work from the perspective of any other isotropic basis. This would not change the theoretical results, but might be recommended for some efficiency reasons, e.g. numerical stability.

For instance, we might want to use the Cholesky decompositions $\Omega = \mathbf{L}_\Omega \mathbf{L}_\Omega^\top$ and $\Xi = \mathbf{L}_\Xi \mathbf{L}_\Xi^\top$ (see Eq. 14). We rephrase Eq. 37 as:

$$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{L}_\Omega^{-\top} \check{\mathbf{B}} \check{\mathbf{U}}_n^\top \mathbf{L}_\Xi^{-1} \mathbf{s}_e \text{ with } \mathbf{L}_\Omega^{-1} \mathbf{M}^\top \mathbf{L}_\Xi = \check{\mathbf{B}} \check{\mathbf{U}} \check{\mathbf{U}}^\top$$

Because we have $\mathbf{L}_\Omega = \Omega^{\frac{1}{2}} \mathbb{R}_b$ and $\mathbf{L}_\Xi = \Xi^{\frac{1}{2}} \mathbb{R}_a$, we can easily see that the solution is identical with:

$$\check{\mathbf{U}} = \check{\mathbf{U}}, \quad \check{\mathbf{B}} = \mathbb{R}_b^\top \check{\mathbf{B}}, \quad \check{\mathbf{U}} = \mathbb{R}_a^\top \check{\mathbf{U}}$$

3.2.3 Isotropic-Mean Allocation

Let's look at the main case of interest, the general mean-variance solution of Eq. 6 where $\mathbf{M}^\top = \boldsymbol{\beta} = \boldsymbol{\Pi}\boldsymbol{\Xi}^{-1}$. We can compare the mean-variance solution, our departing point:

$$\mathbf{w}_e = \frac{1}{\gamma} \boldsymbol{\Omega}^{-1} \boldsymbol{\Pi} \boldsymbol{\Xi}^{-1} \mathbf{s}_e = \frac{1}{\gamma} \boldsymbol{\Omega}^{-\frac{1}{2}} \tilde{\boldsymbol{\Pi}} \boldsymbol{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$$

with the allocation of Eq. 34 that we express as:

$$\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \boldsymbol{\Omega}^{-\frac{1}{2}} \left(\tilde{\boldsymbol{\Pi}} \tilde{\boldsymbol{\Pi}}^\top \right)^{-\frac{1}{2}} \tilde{\boldsymbol{\Pi}} \boldsymbol{\Xi}^{-\frac{1}{2}} \mathbf{s}_e \quad (38)$$

The effect of the isotropic allocation is to replace the normalized canonical correlation $\tilde{\boldsymbol{\Pi}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^\top$ (see Eq. 16) by another version $\left(\tilde{\boldsymbol{\Pi}} \tilde{\boldsymbol{\Pi}}^\top \right)^{-\frac{1}{2}} \tilde{\boldsymbol{\Pi}} = \tilde{\boldsymbol{\Pi}} \left(\tilde{\boldsymbol{\Pi}}^\top \tilde{\boldsymbol{\Pi}} \right)^{-\frac{1}{2}} = \tilde{\mathbf{B}} \tilde{\mathbf{U}}_n^\top$, leading to a concept of eigenrisk parity across canonical portfolios (see Section 2.3.6). We refer to this allocation as Isotropic-Mean (IM).

Isotropic-Mean Allocation

$$m \geq n, \quad E[\mathbf{r}|F] = \boldsymbol{\Pi} \boldsymbol{\Xi}^{-1} \mathbf{s}$$

$$\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \boldsymbol{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}} \tilde{\mathbf{U}}_n^\top \boldsymbol{\Xi}^{-\frac{1}{2}} \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \sum_{k=1}^N \tilde{\mathbf{w}}_k \quad (39)$$

$$\tilde{\boldsymbol{\Pi}} = \boldsymbol{\Omega}^{-\frac{1}{2}} \boldsymbol{\Pi} \boldsymbol{\Xi}^{-\frac{1}{2}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^\top \text{ and } \tilde{\mathbf{w}}_k = \boldsymbol{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_k \tilde{\mathbf{U}}_k^\top \boldsymbol{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$$

The canonical portfolios $\tilde{\mathbf{w}}_k$ (still ordered by their canonical correlations $\tilde{\Psi}_1 \geq \tilde{\Psi}_2 \geq \dots \geq 0$) are equally invested, leading to a realized Sharpe (measured in-sample)¹⁷:

$$\text{Sharpe} = \frac{1}{\sqrt{n}} \text{Tr}(\tilde{\Psi}) \geq 0 \quad (40)$$

The general allocation of Eq. 37 and its reduced version Isotropic-Mean of Eq. 37 (when $\mathbf{M}^\top = \boldsymbol{\beta} = \boldsymbol{\Pi}\boldsymbol{\Xi}^{-1}$) extend and generalize the concept of ERP allocation introduced in [2] and [22].

It is the best isotropy optimally aligned in the MV direction as encoded by the normalized predictability matrix $\tilde{\boldsymbol{\Pi}}$.

Isotropic-Mean as expressed as Eq. 38 shares some striking similarities with the principal portfolio allocation derived in [13]; both approaches have been designed to manage some form of uncertainty, although through two different perspectives.

3.2.4 Principal Portfolios [13] and Invariance

Principal Portfolios, introduced in [13], have been designed to add robustness to the inference problem by introducing a different risk measure. The original formulation assumes the same number of signals as assets (i.e. $m = n$), where each signal s_i has been designed for a particular asset r_i (as discussed in Section 3.1).

The main idea is to deviate replace the variance estimation by a more robust measure of risk that is independent of the distribution of \mathbf{s} . Instead of estimating risk as the variance $\text{Var}[\mathbf{s}^\top \mathbf{L} \mathbf{r}]$ computed over the joint distribution of \mathbf{s} and \mathbf{r} (see Eq. 4), [13] suggests to use a worst-case scenario, estimating risk as the maximum variance realized across a universe of bounded signals:

$$\max_{\|\mathbf{s}\| \leq 1} \text{Var}[\mathbf{s}^\top \mathbf{L} \mathbf{r}] = \max_{\mathbf{s} \neq 0} \frac{1}{\|\mathbf{s}\|^2} \text{Var}[\mathbf{s}^\top \mathbf{L} \mathbf{r}]$$

The definition avoids the integral over the signal distribution (\mathbf{s} is not considered as a stochastic variable), using only a bounding Euclidean sphere. Consequently, it does not require estimating the signal correlation $\boldsymbol{\Xi}$ (those are usually harder to estimate than $\boldsymbol{\Omega}$ and tends to be less stable). It is also a more robust measure since it requires the variance to be bounded independently of the realization of the signal \mathbf{s} (this is a worst-case scenario).

The above risk measure depends on the basis one is working with (where the bounding sphere is defined and where the variance is computed). This means that there is ambiguity but also flexibility. To avoid being implicitly impacted by the asset correlation/covariance $\boldsymbol{\Omega}$, the norm should be defined in an isotropic basis [13].

For instance, working in the Riccati basis $\{\mathbf{b}_i\}$, we have the following equality $\forall \mathbf{s} \quad \text{Var}[\mathbf{s}^\top \mathbf{L}_{bb} \mathbf{r}_b] = \|\mathbf{L}_{bb}^\top \mathbf{s}\|^2$, and the risk constraint becomes a straight-forward constraint on the (triple) norm of the operator \mathbf{L}_{bb} (expressed in $\{\mathbf{b}_i\}$ with $\mathbf{L}_{bb} = \boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{L} \boldsymbol{\Omega}^{\frac{1}{2}}$ - recall that $m = n$ in the original publication [13]):

$$\|\mathbf{L}_{bb}\|^2 = \max_{\mathbf{s} \neq 0} \frac{\|\mathbf{L}_{bb}^\top \mathbf{s}\|^2}{\|\mathbf{s}\|^2}$$

The allocation problem can then be expressed (in $\{\mathbf{b}_i\}$) as:

$$\mathbf{L}_{bb} = \arg_L \max_{\|\mathbf{L}\| \leq \sigma} \text{Tr}(\mathbf{L} \boldsymbol{\Pi}_{bb}) \quad (41)$$

with solution:

$$\mathbf{L}_{bb} = \frac{\sigma}{\sqrt{n}} (\boldsymbol{\Pi}_{bb}^\top \boldsymbol{\Pi}_{bb})^{-\frac{1}{2}} \boldsymbol{\Pi}_{bb}^\top = \frac{\sigma}{\sqrt{n}} \boldsymbol{\Pi}_{bb}^\top (\boldsymbol{\Pi}_{bb} \boldsymbol{\Pi}_{bb}^\top)^{-\frac{1}{2}}, \quad (42)$$

which exhibits a similar form as Eq. 39. The concept of principal portfolios follows nicely from another singular value decomposition of the predictability matrix $\boldsymbol{\Pi}_{bb}$ expressed in $\{\mathbf{b}_i\}$.

¹⁷Note that we have $0 \leq \frac{1}{\sqrt{n}} \text{Tr}(\tilde{\Psi}) \leq \sqrt{\text{Tr}(\tilde{\Psi}^2)}$, so that in-sample $0 \leq \text{Sharpe}(\text{Eigenrisk}) \leq \text{Sharpe}(\text{Mean-Variance})$. We have equality when all singular values are identical.

To explore further the similarities between both solutions Eq. 39 and Eq. 42, we depart from the original spirit of [13] and suggest to choose different bases, searching directly for an operator L_{ub} defined between isotropic bases $\{u_i\}$ and $\{b_i\}$ and under the constraint $\|L_{ub}\| \leq \sigma$. We also do not assume that $m = n$ anymore and put ourselves in the general setting $m \geq n$.

The allocation problem can be rephrased as:

$$L_{ub} = \arg_L \max_{\|L\| \leq \sigma} \text{Tr}(L\tilde{\Pi})$$

with solution:

$$\begin{aligned} L &= \Xi^{-\frac{1}{2}} L_{ub} \Omega^{-\frac{1}{2}} \\ &= \frac{\sigma}{\sqrt{n}} \Xi^{-\frac{1}{2}} \tilde{\Pi}^T (\tilde{\Pi} \tilde{\Pi}^T)^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} \end{aligned} \quad (43)$$

This is the same solution as Eq. 38 above. This is puzzling at first. Although both approaches are different in spirit (one is force-aligning an isotropic form Eq. 30 in the direction of the normalized predictability, the other is maximizing expected returns under a robust risk metric defined between isotropic bases), we end up with the same solution. Both are trying to build resilience.

The reason is structural: by working in isotropic bases $\{b_i\}$, $\{u_i\}$, we constrain the solution space to isotropic linear applications. This is explicit in BI; implicit in principal portfolios [13].

Crucially, the allocation is *invariant under rotations* of the isotropic bases¹⁸. Only isotropy itself, and not orientation, conditions the solution¹⁹.

This perspective is particularly relevant in the context of principal portfolio methodology, where the choice of basis might obscure this invariance (to rotations) at first glance. The BI approach makes the invariance clearer, as the isotropic constraint directly shapes the solution space.

Rotational invariance is not exclusive to isotropic bases: the triple-norm objective and expected return are both invariant under $SO(n) \times SO(m)$ rotations of the anchor bases in \mathcal{H}_r and \mathcal{H}_s .

The principal portfolio optimization thus defines solutions on a principal bundle with $SO(m) \times SO(n)$ symmetry: entire orbits of equivalent allocations collapse to a single geometric configuration.

Only when anchors are isotropic do principal portfolios reduce to canonical portfolios:

$$\tilde{w}_k = \Omega^{-1/2} \tilde{B}_k \tilde{U}_k^T \Xi^{-1/2} s,$$

with equal weighting across modes, reproducing the isotropic-mean allocation (Eq. 39). Isotropy is the symmetry that unifies.

3.3 Take-Aways

- Given a general allocation expressed as $w \propto \Omega^{-1} M^T s$, a “pure” isotropic allocation with “minimal” distortion can be achieved by enforcing isotropy in the direction most aligned with the matrix $\Omega^{-\frac{1}{2}} M^T \Xi^{+\frac{1}{2}}$, that is by identifying the orthogonal transformations U and V so that $U^T \Omega^{-\frac{1}{2}} M^T \Xi^{+\frac{1}{2}} V$ is as close as possible to the identity matrix \mathbb{I} (in the sense of the Frobenius norm). We obtain:

$$w = \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \dot{B} \dot{U}_n^T \Xi^{-\frac{1}{2}} s \text{ where } \Omega^{-\frac{1}{2}} M^T \Xi^{+\frac{1}{2}} = \dot{B} \dot{\Psi} \dot{U}^T$$

- When the allocation is issued from a general mean-variance optimization $w \propto \Omega^{-1} \Pi \Xi^{-1} s$, the resulting **isotropic-mean** solution is equally allocated along canonical portfolios \tilde{w}_k , built from the singular vectors of the normalized predictability matrix $\tilde{\Pi} = \Omega^{-\frac{1}{2}} \Pi \Xi^{-\frac{1}{2}}$:

$$w = \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} (\tilde{\Pi} \tilde{\Pi}^T)^{-\frac{1}{2}} \tilde{\Pi} \Xi^{-\frac{1}{2}} s = \frac{\sigma}{\sqrt{n}} \sum_{k=1}^N \tilde{w}_k$$

- In the simple setup where $E[r|F] \propto s$ (that is when $m = n$ and $M^T = \mathbb{I}$), the isotropy-enforced allocation takes the form $L^T = \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} (\Omega^{-\frac{1}{2}} \Xi \Omega^{-\frac{1}{2}})^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} s$, slightly different from the ERP approach of [2].
- Principal portfolios [13] have the same solution when the initial choice of basis is isotropic (i.e. when the triple norm is expressed between isotropic bases). Therefore, similar techniques of principal exposure portfolios and principal alpha portfolios could be applied (see [13]), and will be explored in further work.
- Although the solution does not depend on the specific choice of isotropic bases, one could employ alternative ones, such as those designed for enhanced stability (e.g. Cholesky or others).

The above methodology enforces **strictly** isotropy on both return and signal sides. This strong constraint might deform significantly the initial mean-variance allocation and the approach lacks direct control over portfolio deformation. We address this issue next.

¹⁸Let $\hat{b}_i = \Omega^{-1/2} \mathbb{R}_{\hat{b}} e_i$, $\hat{u}_i = \Xi^{-1/2} \mathbb{R}_{\hat{u}} e_i$. Then $(\Pi_{\hat{b}\hat{u}} \Pi_{\hat{u}\hat{b}}^T)^{-1/2} = \mathbb{R}_{\hat{b}}^T (\tilde{\Pi} \tilde{\Pi}^T)^{-1/2} \mathbb{R}_{\hat{b}}$, so $L = \Xi^{-1/2} L_{ub} \Omega^{-1/2}$ is rotation-invariant.

¹⁹This rotational invariance mirrors gauge symmetry in physics: the constraint, not the coordinate, defines the physics.

4 Isotropy-Regularized Mean-Variance: A Geometric Regularizer for Signal Uncertainty

The “pure” (or exact) isotropic allocations of Section 3 have been achieved by enforcing isotropy in the direction most aligned with the normalized predictability matrix $\tilde{\mathbf{\Pi}}$. However, the resulting allocations could deviate significantly from the original mean-variance solution, as no direct control is offered²⁰. By working from a risk perspective only, the resulting solution might deviate significantly from the departing mean-variance allocation (especially when the covariances $\mathbf{\Omega}$ and $\mathbf{\Xi}$ become large).

We suggest to augment the mean-variance framework by adding an isotropy constraint, thereby offering an adjustable trade-off between return maximization, variance minimization, and isotropic control. To naturally integrate some notion of isotropy within the mean-variance framework, we decompose our generic portfolio allocation $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$ in the following form:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{T} \mathbf{\Xi}^{-\frac{1}{2}}, \quad (44)$$

where $\mathbf{T} \in \mathcal{R}^{n \times m}$ is the unknown mapping from $\mathcal{H}_s^\star \sim \mathcal{R}^m$ into $\mathcal{H}_r^\star \sim \mathcal{R}^n$. The linear operator \mathbf{T} is our unknown.

Using the formulation Eq. 44, we have the following:

$$\begin{aligned} E[\mathbf{w}^\top \mathbf{r}] &= \text{Tr}(\mathbf{L} \mathbf{\Pi}) = \frac{\sigma}{\sqrt{n}} \text{Tr}(\mathbf{T}^\top \tilde{\mathbf{\Pi}}) \\ \text{Var}[\mathbf{w}^\top \mathbf{r}] &\approx \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top) = \frac{\sigma^2}{n} \text{Tr}(\mathbf{T} \mathbf{T}^\top) \end{aligned}$$

where $\tilde{\mathbf{\Pi}}$ is the cross-correlation between normalized assets and normalized signals expressed into their corresponding Riccati basis $\{\mathbf{b}_i\}$ and $\{\mathbf{u}_i\}$:

$$\tilde{\mathbf{\Pi}} = \mathbf{\Pi}_{bu} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}}$$

Our isotropy-regularized mean-variance (IRMV) framework aims to maximize the returns under two constraints:

- a standard volatility constraint where we cap the total variance at σ^2 :

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] \leq \sigma^2 \quad (45)$$

- an isotropy (i.e. orthogonality) constraint through:

$$\frac{1}{n} \|\mathbf{T} \mathbf{T}^\top - \eta \mathbf{I}_n\|_{\text{F}}^2 \leq 2\tau \quad (46)$$

where τ controls the amount of isotropy we desire. The positive parameter $\eta \leq 1$, typically chosen close to 1, tilts slightly the variance down to take care of the convexity. This implicitly bounds the variance as we discuss below²¹.

²⁰To be exact, we noticed an equivalence with the principal portfolios methodology, so there exist nonetheless an element of control through the triple norm of the operator \mathbf{L} . However, this requires anchoring the solution exactly between two isotropic bases, a forced implicit constraint on which no control exist.

²¹We could have used an orthogonality penalty of the form $\frac{1}{n} \|\mathbf{T} \mathbf{T}^\top - \frac{\text{Tr}(\mathbf{T} \mathbf{T}^\top)}{n} \mathbf{I}_n\|_{\text{F}}^2$; our choice allows for a volatility control even in the absence of volatility constraint.

4.1 Isotropy Penalty and Participation Ratio

To better understand our additional isotropy penalty, we consider a general portfolio allocation $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$. Without loss of generality, we express the allocation as:

$$\mathbf{w} = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-1} \mathbf{M}^\top \mathbf{s}$$

the volatility σ is set at the current level $\sigma = \sqrt{\text{Tr}(\mathbf{\Omega} \mathbf{L} \mathbf{\Xi} \mathbf{L}^\top)}$ and $\mathbf{M}^\top = \frac{\sqrt{n}}{\sigma} \mathbf{\Omega} \mathbf{L}^\top \in \mathcal{R}^{n \times m}$.

We follow the same decomposition as Eq. 44, we have:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \left(\mathbb{R}_{\hat{\mathbf{b}}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} \right) \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \quad (47)$$

where $\mathbb{R}_{\hat{\mathbf{b}}}$ and $\mathbb{R}_{\hat{\mathbf{u}}}$ are two rotations, corresponding to the isotropic bases $\{\hat{\mathbf{b}}\}$ and $\{\hat{\mathbf{u}}\}$ respectively.

As we saw above, the linear operator:

$$\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}} = \mathbb{R}_{\hat{\mathbf{b}}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} \in \mathcal{R}^{n \times m}$$

facilitates the computation of a few important metrics:

$$\left\{ \begin{array}{ll} \text{return} & E[\mathbf{w}^\top \mathbf{r}] = \frac{\sigma}{\sqrt{n}} \text{Tr}(\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}}^\top \mathbf{\Pi}_{\hat{\mathbf{b}}\hat{\mathbf{u}}}) \\ \text{variance} & \text{Var}[\mathbf{w}^\top \mathbf{r}] = \frac{\sigma^2}{n} \text{Tr}(\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}} \mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}}^\top) \\ \text{anisotropy} & \frac{1}{n} \|\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}} \mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}}^\top - \eta \mathbf{I}_n\|_{\text{F}}^2 \end{array} \right. \quad (48)$$

where the parameter η_T could be chosen as $\frac{\eta}{n} \text{Tr}(\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}} \mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}}^\top)$ or as constant $\eta_T = \eta$. We can easily check that those metrics are intrinsic as they do not depend on the specifics of $\mathbb{R}_{\hat{\mathbf{b}}}$ and $\mathbb{R}_{\hat{\mathbf{u}}}$.

The singular value decomposition of the matrix $\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}}$ is:

$$\mathbf{T}_{\hat{\mathbf{b}}\hat{\mathbf{u}}} = \mathbb{R}_{\hat{\mathbf{b}}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} = \mathbb{R}_{\hat{\mathbf{b}}}^\top \mathbf{B} \mathbf{\Psi} \left(\mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{U} \right)^\top$$

where $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} = \mathbf{B} \mathbf{\Psi} \mathbf{U}^\top$. This is the same singular value decomposition that we already saw in Eq. 36.

By setting $\sigma = \text{Tr}(\mathbf{\Omega} \mathbf{L} \mathbf{\Xi} \mathbf{L}^\top)$, we are in a situation where the variance of the general portfolio $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$ is set at the variance cap in Eq. 45. In this scenario, $\text{Tr}(\tilde{\Psi}^2) = n$ by construction and the isotropy metric takes the following simplified form:

$$\text{anisotropy} = \frac{1}{n} \|\tilde{\Psi}^2 - \eta \mathbf{I}\|_{\text{F}}^2 = \frac{1}{n} \text{Tr}(\tilde{\Psi}^4) - \eta(2 - \eta) \quad (49)$$

The isotropy metric measures the variability of the eigenspectrum $\tilde{\Psi}$ of the matrix $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{+\frac{1}{2}}$. This is achieved through its participation ratio ψ :

$$\text{variance constraint} \quad : \quad \text{Tr}(\tilde{\Psi}^2) = n \quad (50)$$

$$\text{participation ratio } \psi = \frac{\frac{1}{n} \text{Tr}^2(\tilde{\Psi}^2)}{\text{Tr}(\tilde{\Psi}^4)} = \frac{n}{\text{Tr}(\tilde{\Psi}^4)} \quad (51)$$

$$\text{anisotropy} = \frac{1}{\psi} - \eta(2 - \eta) \quad (52)$$

When $\sigma = \text{Tr}(\mathbf{\Omega} \mathbf{L} \mathbf{\Xi} \mathbf{L}^\top)$, our isotropy penalty act as geometric regularizer on the eigenspectrum of $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{+\frac{1}{2}}$ through the inverse of its participation ratio ψ . It is purely intrinsic as it does not depend on the previous choice of isotropic bases.

Let's look at the special case of the mean-variance framework. We have:

$$\mathbf{w} = \frac{\sigma}{\sqrt{\text{Tr}(\tilde{\mathbf{\Pi}} \tilde{\mathbf{\Pi}}^\top)}} \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{\Pi}} \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}$$

We can easily see that $\text{Var}[\mathbf{w}^\top \mathbf{r}] = \sigma^2$ by construction. With variance saturated at σ , the isotropy metric becomes a function of the participation ratio ψ of the normalized predictability matrix $\tilde{\mathbf{\Pi}}$. It measures the variability of the eigenspectrum $\tilde{\Psi}$ through its inverse of the participation ratio.

The above analysis sheds some light on our isotropy-regularized mean-variance framework and the constraints applied to the operator \mathbf{T} through Eq. 45 and Eq. 46. As the solution tries to line up on $\tilde{\mathbf{\Pi}}$ through return maximization, a concentrated eigenspectrum $\tilde{\Psi}$ with a low participation ratio ψ will generate a conflicting situation, as both constraints won't be able to be saturated. As variance approaches saturation, the isotropy metric becomes exactly — ignoring the constant $-\eta(2 - \eta)$ — the inverse participation ratio.

As we work between isotropic bases $\{\mathbf{b}_i\}$ and $\{\mathbf{u}_i\}$, our formulation naturally penalizes situations where the uncertainty loads onto too few modes, reconciling the pure isotropic allocations (e.g. isotropic-mean) defined in Section 3 with the mean-variance framework [16, 17]. The approach is intrinsic, as the solution do not depend on the specific choice of $\{\mathbf{b}_i\}$ and $\{\mathbf{u}_i\}$.

4.2 Functional Formulation

Introducing Lagrange coefficient $\frac{\gamma}{2}$ and $\frac{\lambda}{4}$, the IMV functional to optimize takes the following form:

$$\frac{1}{\sqrt{n}} \text{Tr}(\mathbf{T}^\top \tilde{\mathbf{\Pi}}) - \frac{\gamma}{2n} \text{Tr}(\mathbf{T} \mathbf{T}^\top) - \frac{\lambda}{4n} \|\mathbf{T} \mathbf{T}^\top - \eta \mathbf{I}\|_{\text{F}}^2$$

with first-order condition:

$$\sqrt{n} \tilde{\mathbf{\Pi}}^\top = \gamma \mathbf{T}^\top + \lambda \mathbf{T}^\top (\mathbf{T} \mathbf{T}^\top - \eta \mathbf{I})$$

We can gain some insight by working the bases $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{U}}$ obtained through the singular value decomposition of the predictability matrix $\tilde{\mathbf{\Pi}}$:

$$\tilde{\mathbf{\Pi}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^\top = \tilde{\mathbf{B}} \tilde{\Psi}_n \tilde{\mathbf{U}}_n^\top \quad (53)$$

We express the operator \mathbf{T} in the bases $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{U}}$ as:

$$\mathbf{T} = \tilde{\mathbf{B}} \mathbf{\Theta} \tilde{\mathbf{U}}^\top \text{ with } \mathbf{\Theta} \in \mathcal{R}^{n \times m}$$

We obtain the following equality:

$$\sqrt{n} \tilde{\Psi}^\top = \gamma \mathbf{\Theta}^\top + \lambda \mathbf{\Theta}^\top (\mathbf{\Theta} \mathbf{\Theta}^\top - \eta \mathbf{I}) \quad (54)$$

We can show that the solution $\mathbf{\Theta}$ takes the form $\mathbf{\Theta} = [\mathbf{\Theta}_n, \mathbf{0}_{n, m-n}]$ where $\mathbf{\Theta}_n$ is diagonal with elements θ_i verifying:

$$c_i = \sqrt{n} \tilde{\Psi}_i = \gamma \theta_i + \lambda \theta_i (\theta_i^2 - \eta) = (\gamma - \eta \lambda) \theta_i + \lambda \theta_i^3 \quad (55)$$

where we have defined $c_i = \sqrt{n} \tilde{\Psi}_i$. The optimized allocation can be decomposed along a set of n canonical portfolios (as in [8]):

$$\mathbf{T} = \tilde{\mathbf{B}} \mathbf{\Theta} \tilde{\mathbf{U}}^\top = \tilde{\mathbf{B}} \mathbf{\Theta}_n \tilde{\mathbf{U}}_n^\top \quad (56)$$

$$\begin{aligned} \mathbf{L}^\top &= \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{T} \mathbf{\Xi}^{-\frac{1}{2}} \\ &= \sum_{i=1}^n \theta_i \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_i \tilde{\mathbf{U}}_i^\top \mathbf{\Xi}^{-\frac{1}{2}} \end{aligned} \quad (57)$$

The expected returns and variance are computed as:

$$E[\mathbf{w}^\top \mathbf{r}] = \frac{\sigma}{\sqrt{n}} \text{Tr}(\mathbf{\Theta}^\top \tilde{\Psi}) = \frac{\sigma}{\sqrt{n}} \sum \theta_k \tilde{\Psi}_k = \frac{\sigma}{n} \sum \theta_k c_k$$

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] \approx \frac{\sigma^2}{n} \text{Tr}(\mathbf{\Theta} \mathbf{\Theta}^\top) = \frac{\sigma^2}{n} \sum \theta_k^2$$

with corresponding Sharpe ratio expressed as:

$$\text{Sharpe}(\theta, c) = \frac{1}{n} \sum \theta_k c_k / \sqrt{\frac{1}{n} \sum \theta_k^2}$$

The two constraints can be expressed as:

$$\boxed{\begin{aligned} \frac{1}{n} \sum \theta_k^2 &\leq 1 && \text{variance} && (58) \\ \frac{1}{n} \sum (\theta_k^2 - \eta)^2 &\leq 2\tau && \text{isotropy} && (59) \end{aligned}}$$

The Jensen inequality shows that the isotropy constraint offers an implicit control of volatility:

$$\frac{1}{n} \sum (\theta_k^2 - \eta)^2 \geq \left(\frac{1}{n} \sum \theta_k^2 - \eta \right)^2$$

The variance will be bounded by $\sigma^2(\eta \pm \sqrt{2\tau})$ as:

$$\eta - \sqrt{2\tau} \leq \frac{1}{n} \sum \theta_k^2 \leq \eta + \sqrt{2\tau}$$

However, depending on the choice of η , the solution might be lower and settle around our desired value of 1.

We note that the solution θ does not depend on the magnitude of the eigencurve c , only on its shape. Solving the set of coupled equations of Eq. 55 subject to the constraints Eq. 58 and Eq. 59 can only be done numerically (without difficulty), but we can gain some insight by looking into some simple scenarios.

4.3 Special Cases

We first study some simplifying scenarios and limit properties.

• **No Isotropy Constraint** when $\tau \rightarrow +\infty$ (or $\lambda \rightarrow 0$)

The orthogonality penalization becomes insignificant and we end up with the standard mean-variance solution:

$$\theta_i = \frac{c_i}{\sqrt{\frac{1}{n} \sum c_k^2}} \quad \text{mean-variance}$$

The solution can be decomposed into canonical portfolios $\frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \tilde{\mathbf{B}}_i \tilde{\mathbf{U}}_i^T \Xi^{-\frac{1}{2}}$ (each with weight $\propto \theta_i$), leading to an in-sample Sharpe ratio of $\sqrt{\frac{1}{n} \sum c_k^2}$ (see Eq. 17). We can also check that the operator \mathbf{T} is $\mathbf{T} = \frac{1}{\gamma} \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^T = \frac{1}{\gamma} \tilde{\mathbf{\Pi}}$ as expected, and where the leverage coefficient is $\gamma = \sqrt{\frac{\text{Tr}(\tilde{\mathbf{\Pi}} \tilde{\mathbf{\Pi}}^T)}{n}}$.

There is a value of τ_η^+ for which, the isotropy constraints kicks-in as defined by Eq. 59. The region where only the variance constraint

matters is defined by:

$$\begin{aligned} \tau \geq \tau_\eta^+ &= \frac{1}{2} \left(\frac{1}{n} \sum \theta_k^4 - 2\eta + \eta^2 \right) \\ &= \frac{1}{2} \left(\frac{\frac{1}{n} \sum c_k^4}{(\frac{1}{n} \sum c_k^2)^2} - 2\eta + \eta^2 \right) \\ &\geq \frac{1}{2} (1 - \eta)^2 \end{aligned} \quad (60)$$

We note that the term $\frac{1}{n} \sum c_k^4 / (\frac{1}{n} \sum c_k^2)^2$ is one over the participation ratio $\tilde{\psi}$ of the eigenspectrum of $\tilde{\mathbf{\Pi}}$ as defined in Eq. 51.

When the eigenspectrum is more concentrated (as $\tilde{\psi} \rightarrow 0$), the limit τ_η^+ increases and the isotropy constraint would be harder and harder to avoid (as desired). On the other hand, when the eigenspectrum becomes flatter (i.e. less concentrated, as $\tilde{\psi} \rightarrow 1$), although τ_η^+ would decrease towards its lower limit $\frac{1}{2} (1 - \eta)^2$, the isotropy constraint would be less and less required (since the eigenspectrum is more and more isotropic).

• **Full Isotropy** when $\tau \rightarrow 0$ (or $\lambda \rightarrow +\infty$)

We end up with the same isotropic-mean allocation of Eq. 39:

$$\theta_i \approx \sqrt{\eta} \quad \text{isotropic-mean}$$

The solution can still be decomposed into (the same) canonical portfolios $\sqrt{\eta} \frac{\sigma}{\sqrt{n}} \Omega^{-\frac{1}{2}} \tilde{\mathbf{B}}_i \tilde{\mathbf{U}}_i^T \Xi^{-\frac{1}{2}}$, but with unit weights. We obtain a lower in-sample Sharpe ratio of $\frac{1}{n} \sum c_k$ (see Eq. 40). We can verify that $\mathbf{T} = \sqrt{\eta} \tilde{\mathbf{B}} \tilde{\mathbf{U}}_n^T$.

The variance constraint of Eq. 58 is verified (and saturated if and only if $\eta = 1$):

$$\frac{1}{n} \sum \theta_k^2 = \eta \leq 1$$

• **No Variance Constraint** when $0 < \tau \ll 1$ and $\gamma = 0$

As isotropy implicitly bounds the variance at $\sigma^2(\eta + \sqrt{2\tau})$ thanks to the Jensen inequality, we consider here the case where the variance constraint is fully ignored (setting $\gamma = 0$) while the isotropy constraint is tight (i.e. τ small but strictly positive, η only slightly below one).

This leads to an interesting allocation where the coefficients θ_i are the largest zeros of the third-order polynomials (see Figure):

$$\theta_i(\theta_i^2 - \eta) = \frac{c_i}{\lambda} \quad (61)$$

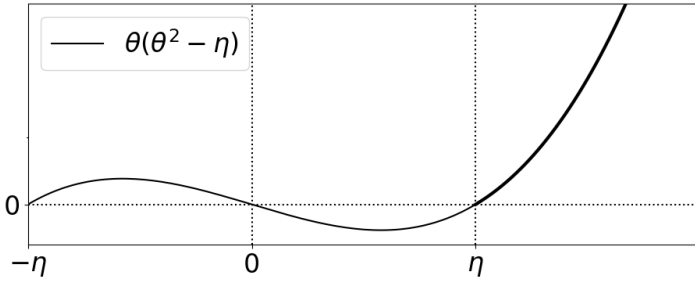
subject to the isotropy constraint only (again neglecting the variance constraint).

We can use Cardano's method to express the general solution as:

$$\begin{aligned} \eta \neq 0 \quad \theta_i &= 2\sqrt{\frac{\eta}{3}} \cos\left(\frac{1}{3} \arccos\left[\frac{3\sqrt{3}}{2\eta^{\frac{3}{2}}} \frac{c_i}{\lambda}\right]\right) \\ \eta = 0 \quad \theta_i &= \left(\frac{c_i}{\lambda}\right)^{\frac{1}{3}} \end{aligned} \quad (62)$$

with λ verifying the isotropy condition, which can be written as:

$$\frac{1}{n} \sum_{c_i > 0} \frac{c_i^2}{\theta_i^2} = 2\tau\lambda^2$$



Although the cubic equation of Eq. 61 hints at a power-law scaling in $\sqrt[3]{\frac{c_i}{\lambda}}$ for large c_i , the orthogonal constraint (on λ) introduces a non-linear balancing act that caps deviations, pushing the dominant modes to bounded values (typically larger than 1), while the smaller modes hover near isotropy around $\sqrt{\eta}$. Recall that the solution θ only depends on the shape of the eigencurve c , and not on its magnitude.

The isotropy constraint, while partially controlling the overall variance, with an implicit hard constraint at the cap $\sigma^2(\eta + \sqrt{2\tau})$, adjust the weighting of the eigenmodes of $\tilde{\mathbf{H}}$, tilting the allocation in favor of the most predictable ones, while retaining some level of isotropy overall.

When $\eta < 1$, there is a value τ_η^- (that depends on the shape of the eigenspectrum) under which the variance cap might not be met (saturation or not might depend on the eigenspectrum $\tilde{\Psi}$). From the bounded variance, we know that:

$$\frac{1}{2}(1 - \eta)^2 \leq \tau_\eta^-$$

For instance, when η is zero, one can easily see that:

$$\tau_{\eta=0}^- = \frac{1}{2n} \sum \tilde{\Psi}_k^{\frac{4}{3}} / \left(\frac{1}{n} \sum \tilde{\Psi}_k^{\frac{2}{3}}\right)^2 \geq \frac{1}{2}$$

4.4 Parameter Selection & Regions

The limit scenarios we discussed above clearly displayed some distinct regions where the constraints (variance and isotropy) might be active or not. Those would depend on the shape of eigenspectrum.

In most application, the number of significant eigenvalues would be small, as computed by the effective rank or the participation ratio, denoted $\tilde{\psi}$, to determine the variance concentration:

$$\begin{aligned} \text{effective rank} \quad \frac{1}{n} \frac{(\sum c_i)^2}{\sum c_i^2} &= \frac{1}{n} \frac{\text{Tr}^2(\tilde{\Psi})}{\text{Tr}(\tilde{\Psi}^2)} = \frac{1}{n} \frac{\text{Tr}^2(\tilde{\mathbf{H}})}{\text{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^T)} \\ \text{participation ratio} \quad \frac{1}{n} \frac{(\sum c_i^2)^2}{\sum c_i^4} &= \frac{1}{n} \frac{\text{Tr}^2(\tilde{\Psi}^2)}{\text{Tr}(\tilde{\Psi}^4)} = \frac{1}{n} \frac{\text{Tr}^2(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^T)}{\text{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\tilde{\mathbf{H}}^T)} \end{aligned} \quad (63)$$

As we discussed above in Section 4.3, the upper limit τ_η^+ in Eq. 60 where the isotropic constraint kicks-in can be expressed as:

$$\tau_\eta^+ = \frac{1}{2} \left(\frac{1}{\tilde{\psi}} - 2\eta + \eta^2 \right)$$

where $\tilde{\psi}$ is the participation ratio. The more concentrated an eigenspectrum is (as the participation ratio $\tilde{\psi}$ decreases towards 0), the more impacted it becomes by the isotropy constraint. Conversely, a flat eigenspectrum (with $\tilde{\psi} \rightarrow 1$) would require less isotropy constraint, as the problem is naturally more isotropic to start with.

The consequence is that choosing τ in the 1 – 2 range should offer a good balance between isotropy and mean-variance, independently of the shape of the eigenspectrum.

The lower bound τ_η^- where variance is not saturated also depends on the shape of the eigenspectrum. Low values of η would bound the variance strictly (through the Jensen equality) with the isotropic constraint dominating, but values close to 1 would avoid such scenarios, setting the framework within the influence of the two constraints.

To illustrate the boundaries, we consider two idealized examples, two different eigencurve shapes (recall that the magnitude has no impact on the solution θ), where the participation ratio $\tilde{\psi}$ is identical in both cases.

- A two-mode case where the m first eigenmodes are constant and equal to $c_{i \leq m} = c_{\max}$, while the remaining ones are constant equal to a value significantly smaller $c_{i > m} = c_{\min} \ll c_{\max}$. For simplicity, we set $c_{\min} = 0$ so that $\tilde{\psi} = \frac{m}{n}$.
- An exponential eigenspectrum $c_i \approx \exp(-i/\tilde{\psi}n)$.

Solving the system for different values of (η, τ) , we find the following regions:

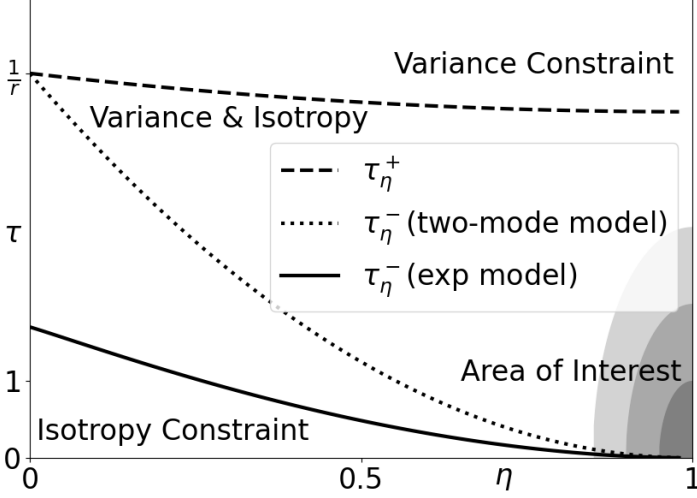


Figure 1: η - τ Region Diagram.

To better understand the boundary where the variance constraint disappears, we focus on the two-mode case that can easily be solved explicitly. We express the solution as:

$$\begin{cases} \theta_{i \leq m}^2 &= \eta + \sqrt{2\tau}x_1 \\ \theta_{i > m}^2 &= \eta + \sqrt{2\tau}x_2 \end{cases}$$

We have the following:

$$\begin{cases} mx_1^2 + (n-m)x_2^2 &= n \\ c_{\min}(\eta + \sqrt{2\tau}x_1)x_1^2 &= c_{\max}(\eta + \sqrt{2\tau}x_2)x_2^2 \end{cases}$$

The limit case with $c_{\min} = 0$ and $\tilde{\psi} = \frac{m}{n}$ is enlightening. We observe that:

$$\text{case } c_{\max} > c_{\min} = 0 \quad \begin{cases} \theta_{i \leq m} &= \sqrt{\eta + \sqrt{2\tau} \frac{n}{m}} \\ \theta_{i > m} &= \sqrt{\eta} \end{cases}$$

and the variance is:

$$\text{case } c_{\max} > c_{\min} = 0 \quad \frac{1}{n} \sum \theta_k^2 = \eta + \sqrt{2\tau} \frac{m}{n}$$

In this simple scenario (i.e. $c_{\min} = 0$ and $\tilde{\psi} = \frac{m}{n}$), we note that:

- When η is chosen as $\eta = 1 - \sqrt{2\tau} \frac{m}{n} = 1 - \sqrt{2\tau} \tilde{\psi}$, the variance

ends up around σ^2 , as desired.

- As τ decreases towards 0, we converge towards full isotropy with $\theta_i \rightarrow \sqrt{\eta}$.

- Conversely, as $\tau \rightarrow \frac{n}{2m} = \frac{1}{2\tilde{\psi}} \gg 1$ while $\eta = 1 - \sqrt{2\tau} \frac{m}{n} = 1 - \sqrt{2\tau} \tilde{\psi} \rightarrow 0$, we end up with zero exposure on the lower mode and $\sqrt{\frac{n}{m}} = \sqrt{\tilde{\psi}} > 1$ on the higher mode.

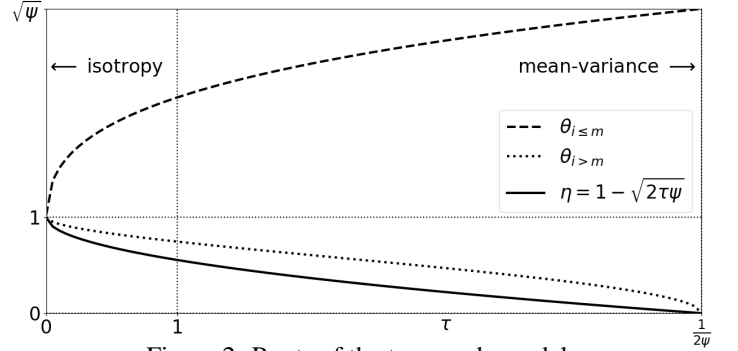


Figure 2: Roots of the two-mode model.

With the variance saturated through the choice of $\eta = 1 - \sqrt{2\tau} \tilde{\psi}$, the resulting Sharpe of the strategy is:

$$\text{Sharpe}(\theta^\tau, c) = \frac{1}{n} \sum \theta_k^\tau c_k = \frac{m}{n} \theta_{\max}^\tau c_{\max} = \sqrt{\eta + \sqrt{\frac{2\tau}{\tilde{\psi}}} \frac{m}{n} c_{\max}}$$

where $\frac{m}{n} c_{\max}$ is the value of the (in-sample) isotropic Sharpe.

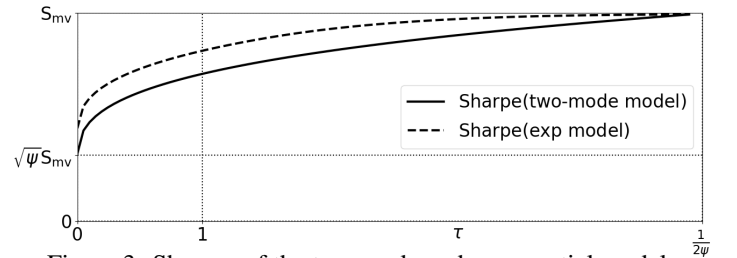


Figure 3: Sharpes of the two-mode and exponential models.

We quickly see that as $\tau \rightarrow \frac{1}{2\tilde{\psi}}$, the (in-sample) Sharpe converges towards its mean-variance value (as expected!):

$$\text{Sharpe}(\tau \rightarrow \frac{1}{2\tilde{\psi}}) = \sqrt{\frac{m}{n} c_{\max}} = \sqrt{\frac{1}{\tilde{\psi}} \frac{m}{n} c_{\max}}$$

4.5 Take-Aways

- By decomposing a general portfolio allocation $\mathbf{w} = \mathbf{L}^\top \mathbf{s}$ as $\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{T} \mathbf{\Xi}^{-\frac{1}{2}}$, fixing σ as $\sigma = \text{Tr}(\mathbf{\Omega} \mathbf{L} \mathbf{E} \mathbf{L}^\top)$, one can easily measure (and potentially penalize) the portfolio isotropy thanks to Eq. 48. With variance fixed, the isotropy metric is a function of the participation ratio of the eigenspectrum of \mathbf{T} .
- By enforcing exactly isotropy, “pure” isotropic allocations (i.e. isotropic-mean portfolios) could deviate significantly from the original mean-variance allocation. In order to retain some amount of control, we augment the mean-variance framework by adding an isotropy constraint, thereby offering an adjustable trade-off between return maximization, variance minimization, and isotropic control.
- The isotropy constraint acts as a geometric regularizer, in the orthonormal SVD basis of the normalized predictability matrix $\tilde{\mathbf{\Pi}}$. As a function of its inverse participation ratio (when variance saturates on its constraint), it prevents loading up on too few concentrated modes.
- The approach defines the solution as a modulation of the shape of the eigenspectrum $\tilde{\Psi}$ of the normalized predictability matrix $\tilde{\mathbf{\Pi}}$. The resulting solution Θ offers a trade-off between pure isotropy (with flat allocation) and mean-variance (proportional to $\tilde{\Psi}$).
- The parameters η and τ controlling the amount of isotropy can be fine-tuned. The region $1 - \sqrt{2\tau\tilde{\Psi}} \leq \eta \leq 1$ with $\tau \approx 1$ defines an area of interest where both constraints co-exist: setting $\eta = \tau = 1$ is generally a sensible choice. The general solution solves n -cubic equations coupled through both constraints.

Isotropic Mean-Variance Framework		Solution			
$\mathbf{w}_e = \mathbf{L}_\star^\top \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{T}_\star \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \sum_{k=1}^N \theta_k \tilde{\mathbf{w}}_k$					
$\mathbf{T} = \tilde{\mathbf{B}} \mathbf{\Theta}_n \tilde{\mathbf{U}}_n^\top$ where $\tilde{\mathbf{\Pi}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^\top$					
$\mathbf{T}_\star = \arg_T \max \frac{1}{\sqrt{n}} \text{Tr} \left(\mathbf{T}^\top \tilde{\mathbf{\Pi}} \right) - \frac{\gamma}{2n} \text{Tr} \left(\mathbf{T} \mathbf{T}^\top \right) - \frac{\lambda}{4n} \mathbf{T} \mathbf{T}^\top - \eta \mathbb{I}_n _{\mathbb{F}}^2$					
Constraints	{	variance:	$\frac{1}{n} \text{Tr} \left(\mathbf{T} \mathbf{T}^\top \right) \leq \sigma^2$		
		isotropy:	$\frac{1}{n} \mathbf{T} \mathbf{T}^\top - \eta \mathbb{I} _{\mathbb{F}}^2 \leq 2\tau$		
$\sqrt{n} \tilde{\Psi}_i = \gamma \theta_i + \lambda \theta_i (\theta_i^2 - \eta) = (\gamma - \eta \lambda) \theta_i + \lambda \theta_i^3$					
Canonical Portfolios: $\tilde{\mathbf{w}}_k = \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_k \tilde{\mathbf{U}}_k^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$					

	mean-variance	full-isotropic	general case
condition	$\tau \rightarrow \infty$	$\tau \rightarrow 0$	$\tau \leq 1 \ \eta \leq 1$
θ_i	$\sqrt{\frac{n}{\sum \tilde{\Psi}_i^2}} \tilde{\Psi}_i$	$\sqrt{\eta}$	θ_i
PnL	$\sigma \sqrt{\sum \tilde{\Psi}_i^2}$	$\sigma \sqrt{\frac{\eta}{n}} \sum \tilde{\Psi}_i$	$\sigma \frac{1}{\sqrt{n}} \sum \theta_i \tilde{\Psi}_i$
Risk	σ^2	$\sigma^2 \eta$	$\frac{\sigma^2}{n} \sum \theta_i^2$
Sharpe	$\sqrt{\sum \tilde{\Psi}_i^2}$	$\frac{1}{\sqrt{n}} \sum \tilde{\Psi}_i$	$\frac{\sum \theta_i \tilde{\Psi}_i}{\sum \theta_i^2}$

- The approach emphasizes the importance of canonical portfolios as essential building blocks. However, it also highlights the limitations of the approach:
 - The described framework depends critically on the estimation and stability of both return and signal covariances $\mathbf{\Omega}$ and $\mathbf{\Xi}$ (through the whitening operators $\mathbf{\Omega}^{-\frac{1}{2}}$ and $\mathbf{\Xi}^{-\frac{1}{2}}$). When the requirement is met, the approach would be able to manage some level of uncertainty present in the prediction matrix, e.g. $\mathbf{\Pi}$ or equivalently $\tilde{\mathbf{\Pi}}$. But those are strong assumptions, certainly not met in practice, particularly for the signal covariance $\mathbf{\Xi}$.
 - The uncertainty is only tackled by modifying the shape of the eigencurve, nothing more. Analysis of the eigenspectrum of the normalized predictability matrix $\tilde{\mathbf{\Pi}}$ is therefore of critical importance.
- Although we explicitly choose to work from the perspective of the Riccati basis \mathbf{b}_i and \mathbf{u}_i , focusing on the predictability matrix $\mathbf{\Pi}_{bu}$, any other isotropic basis could have been used (the approach is intrinsic). For example, one might prefer to work with Cholesky decomposition $\mathbf{\Omega} = \mathbf{L}_\Omega \mathbf{L}_\Omega^\top$ and $\mathbf{\Xi} = \mathbf{L}_\Xi \mathbf{L}_\Xi^\top$ (see Eq. 14), due to numerical stability.

5 Application: Sector Trend-Following

5.1 Setup

We consider a simple sector model where the returns \mathbf{r}_t of a set of n similar assets are driven by white noises ϵ_t and some stochastic autocorrelated trends μ_t slowly mean-reverting around zero as defined in [10, 11]:

$$\begin{cases} \mathbf{r}_t &= \beta \mu_t + \epsilon_t \\ \mu_t &= q \mu_{t-\delta t} + \xi_t \end{cases} \quad (64)$$

The stochastic variables ϵ_t and ξ_t are supposed to be independent and identically distributed through time, with zero mean and correlation structures:

$$E[\epsilon_t \epsilon_u^\top] = \delta_{t-u} \mathbf{\Omega}_\epsilon \text{ and } E[\xi_t \xi_u^\top] = \delta_{t-u} \mathbf{\Omega}_\xi$$

In a sector model (e.g. pharma, banking, futures on Equity, futures on bonds), the trend innovations ξ_t should be heavily influenced by common factors, whereas the noise components ϵ_t should reflect idiosyncratic shocks, likely to be less “synchronized”. We do expect the overall level of correlation to be higher for $\mathbf{\Omega}_\xi$ than for $\mathbf{\Omega}_\epsilon$.

The parameter β (identical for all assets for simplicity reason) scales the trend and is of the order of the signal-to-noise ratio, i.e. $\beta \ll 1$. The parameter q (also chosen constant across assets in our naive sector model) captures the speed at which the trend mean-reverts. It is a critical market parameter that shapes the dynamics of returns:

$$\mathbf{r}_t = \epsilon_t + \beta \sum_{k \geq 0} q^k \xi_{t-k\delta t}$$

Most often, the value of $1 - q \ll 1$ models a slow frequency, typically several months with $1 - q \approx 1\%$ or less.

We consider a standard trend-following strategy where the trading signals are computed as exponential moving averages:

$$s_t = \sqrt{1 - p^2} \sum_{k > 0} p^{k-1} \mathbf{r}_{t-k\delta t}$$

The strategy parameter p should be chosen to approximately match the mean-reversion speed q . Unfortunately, the market parameter q is not known exactly [10] (and prone to sudden temporary changes). During a crisis, negative returns $\mathbf{r}_t \approx -\sigma_\epsilon \mathbb{1}$ (positively auto-correlated in time and positively correlated in space) would pile up (with higher volatility $\sigma_\epsilon > 1$), while most model assumptions would break (e.g. increase volatility and fat tails, sudden decrease of the q parameter, strong auto-correlation of the stochastic variables, ...).

We denote $\beta = \beta_0 \sqrt{1 - q^2}$. Following [10], we use an order of magnitude around $\beta_0 \approx 0.1$, while $p \approx q \approx 0.99$; we also use $n = 10$ in our numerical simulations.

We can compute the different matrices:

$$\begin{cases} \mathbf{\Omega} &= \mathbf{\Omega}_\epsilon + \beta_0^2 \mathbf{\Omega}_\xi \\ \mathbf{\Pi} &= \frac{q\sqrt{1-p^2}}{1-pq} \beta_0^2 \mathbf{\Omega}_\xi \\ \mathbf{\Xi} &= \mathbf{\Omega}_\epsilon + \frac{1+pq}{1-pq} \beta_0^2 \mathbf{\Omega}_\xi \end{cases}$$

Despite its simplicity, this model is a sensible reflection of reality (that is ignoring the non-Gaussian nature of financial distributions, the presence of fat tails, asymmetry, and so on). Combined with trend-following signals, we obtain a convincing allocation problem faced by portfolio managers (e.g. CTAs). Unfortunately, even in this simple model, the dynamics is complex and hard to solve. As noticed in [11], unexpected properties appear.

The “alpha”, e.g the predictive information needed to trade successfully future returns, is embedded in the cross-correlation matrix $\mathbf{\Pi}$, that is implicitly within the covariance matrix $\mathbf{\Omega}_\xi$. Unfortunately, $\mathbf{\Omega}_\xi$ is much harder to accurately estimate than $\mathbf{\Omega}_\epsilon$. Hidden in a sea of noise, it is also typically less stable.

The single-asset case is straight-forward. The theoretical in-sample (annualized) Sharpe ratio of a trend-following strategy applied to a single asset is:

Single-Asset Trend-Following

$$S_1 = \sqrt{252} \frac{q\sqrt{1-p^2}}{\sqrt{Q^2 + 2Q + R}} \approx 0.78 \quad (65)$$

where $Q = (1 - pq)/\beta_0^2 \approx 1.99$ and $R = 1 + q^2 - 2p^2 q^2 \approx 0.058$ (using the same notations as in [11]). The value ≈ 0.78 is obviously unrealistic and massively inflated. In practice, the expected Sharpe ratio of a single-asset trend-following system would be barely positive, around $0.1 - 0.2$.

From the symmetry of $\mathbf{\Pi}$, we know that the solution \mathbf{L} is also symmetrical. We can develop the following equalities:

$$\begin{aligned} \mathbf{\Xi L \Omega L^\top} &= \mathbf{\Omega}_\epsilon \mathbf{L \Omega}_\epsilon \mathbf{L} + \frac{2\beta_0^2}{1-pq} \mathbf{\Omega}_\epsilon \mathbf{L \Omega}_\xi \mathbf{L} + \frac{1+pq}{1-pq} \beta_0^4 \mathbf{\Omega}_\xi \mathbf{L \Omega}_\xi \mathbf{L} \\ \mathbf{\Pi L \Pi L} &= \frac{q^2(1-p^2)}{(1-pq)^2} \beta_0^4 \mathbf{\Omega}_\xi \mathbf{L \Omega}_\xi \mathbf{L} \end{aligned}$$

The second variance term $\text{Tr}(\mathbf{\Pi L \Pi L})$ is typically much smaller than the first one $\text{Tr}(\mathbf{\Xi L \Omega L^\top})$. Its inclusion rarely changes signifi-

cantly the solution, while complicating significantly the methodology (e.g. the complexity is obvious in [11]).

We can derive the expected PnL and total variance, obtaining the same expressions as in [11]:

$$E[\mathbf{w}^\top \mathbf{r}] = \text{Tr}(\mathbf{L}\mathbf{\Pi}) = \frac{q\sqrt{1-p^2}}{Q} \text{Tr}(\mathbf{L}\mathbf{\Omega}_\xi) \quad (66)$$

$$\text{Var}[\mathbf{w}^\top \mathbf{r}] = \text{Tr}\left(\mathbf{\Omega}_\epsilon \mathbf{L} \mathbf{\Omega}_\epsilon \mathbf{L} + \frac{2}{Q} \mathbf{\Omega}_\epsilon \mathbf{L} \mathbf{\Omega}_\xi \mathbf{L} + \frac{R}{Q^2} \mathbf{\Omega}_\xi \mathbf{L} \mathbf{\Omega}_\xi \mathbf{L}\right) \quad (67)$$

The first-order condition writes:

$$\frac{1}{\gamma} \mathbf{\Pi} = \mathbf{\Omega}_\epsilon \mathbf{L} \mathbf{\Omega}_\epsilon + \frac{2}{Q} \mathbf{\Omega}_\epsilon \mathbf{L} \mathbf{\Omega}_\xi + \frac{R}{Q^2} \mathbf{\Omega}_\xi \mathbf{L} \mathbf{\Omega}_\xi \quad (68)$$

where γ is the leverage coefficient (e.g. Lagrange multiplier of the variance constraint).

The last term is often negligible and the exact mean-variance solution barely differs from our departing closed-form solution:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{\text{Tr}(\tilde{\mathbf{\Pi}}^\top \tilde{\mathbf{\Pi}})}} \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1} \quad (69)$$

We discuss those differences in a simplifying scenario below.

5.2 Simplifying Assumption: Uniformity

To simplify further, we model the two covariance matrices as:

$$\mathbf{\Omega}_\epsilon = (1 - \rho_\epsilon) \mathbb{I} + \rho_\epsilon \mathbb{J} \text{ and } \mathbf{\Omega}_\xi = (1 - \rho_\xi) \mathbb{I} + \rho_\xi \mathbb{J} \quad (70)$$

with $\mathbb{J} = \mathbb{1}\mathbb{1}^\top$, the matrix full of ones. That is we are assuming that all return and signal correlations are equal to $\rho_\epsilon \geq -\frac{1}{n-1}$ and $\rho_\xi \geq -\frac{1}{n-1}$ respectively (note that the variances are also the same for the assets and signals). This scenario is explored in details in [11], which we refer for another perspective. Our results corroborate their findings.

In a typical sector, such as equity stocks (e.g. pharma or banking), equity futures, or bond futures, the correlation of the trend innovations ρ_ξ is generally higher than the correlation of the idiosyncratic noise ρ_ϵ , as trends capture systematic, sector-wide movements, while noise reflects asset-specific fluctuations.

Although it is difficult to provide some accurate order of magnitudes, it is sensible to estimate ρ_ξ around 0.6 to 0.9, reflecting strong sector-wide correlations in trends, versus ρ_ϵ around 0.1 to 0.4, reflecting lower residual correlations in idiosyncratic noise, with potential increases during crises.

5.2.1 Attractive Properties

Thanks to the assumption of Eq. 70, all matrices end up of the form $a\mathbb{I} + b\mathbb{J}$ and commute (since $\mathbb{J}^2 = n\mathbb{J}$). This has several implications. At a high level, we already know that the optimal solution \mathbf{w} will be of the form $\mathbf{w} = (a_w \mathbb{I} + b_w \mathbb{J}) \mathbf{s}$. As such, it can only have two expositions, the idiosyncratic signals $a_w \mathbf{s}$ and the market mode through $nb_w \bar{\mathbf{s}} \mathbb{1}$ where $\bar{\mathbf{s}} = \frac{1}{n} \mathbb{1}^\top \mathbf{s}$.

There are only two eigenspaces identical for all operators. Choosing an allocation (e.g. mean-variance or isotropic-mean) only amounts to modulating the weights allocated to both modes. Besides, one can easily solve the exact first-order condition of Eq. 68 by working independently on each eigenmode (and where the Lagrange γ coefficient links the modes together through the variance constraint).

Practically, we have the following:

$$\begin{aligned} \mathbf{\Omega} : \quad a_\Omega &= 1 - \rho_\epsilon + \beta_0^2(1 - \rho_\xi) & b_\Omega &= \rho_\epsilon + \beta_0^2 \rho_\xi \\ \mathbf{\Pi} : \quad a_\Pi &= \frac{q\sqrt{1-p^2}}{1-pq} \beta_0^2(1 - \rho_\xi) & b_\Pi &= \frac{q\sqrt{1-p^2}}{1-pq} \beta_0^2 \rho_\xi \\ \mathbf{\Xi} : \quad a_\Xi &= 1 - \rho_\epsilon + \frac{1+pq}{1-pq} \beta_0^2(1 - \rho_\xi) & b_\Xi &= \rho_\epsilon + \frac{1+pq}{1-pq} \beta_0^2 \rho_\xi \end{aligned}$$

For a matrix of the form $a\mathbb{I} + b\mathbb{J}$ (with $a > 0$ and $a + nb > 0$ to ensure positive definiteness), we have some basic properties :

$$\begin{aligned} (a\mathbb{I} + b\mathbb{J})^{-1} &= \frac{1}{a} \left(\mathbb{I} - \frac{b}{a + nb} \mathbb{J} \right) \\ (a\mathbb{I} + b\mathbb{J})^{-\frac{1}{2}} &= \frac{1}{\sqrt{a}} \left(\mathbb{I} + \frac{\sqrt{a} - \sqrt{a + nb}}{n\sqrt{a + nb}} \mathbb{J} \right) \end{aligned}$$

One can also compute the eigenvalues and singular values without any difficulty and the whole problem is solvable in closed-form. The symmetric matrix $a\mathbb{I} + b\mathbb{J}$ is diagonalizable, with two eigenvalues: one eigenvalue $\lambda_{a\mathbb{I} + b\mathbb{J}}^1 = a + nb$ with multiplicity 1 and eigenvector $\frac{1}{\sqrt{n}} \mathbb{1}$ and a second one $\lambda_{a\mathbb{I} + b\mathbb{J}}^2 = a$ with multiplicity $n - 1$ in the space orthogonal to $\mathbb{1}$ (e.g. the basis can be computed using the Gram-Schmidt process). Depending on the coefficient b , we could have λ^1 smaller than λ^2 . We also have:

$$\text{Tr}(a\mathbb{I} + b\mathbb{J}) = n(a + b) = \lambda_{a\mathbb{I} + b\mathbb{J}}^1 + (n - 1)\lambda_{a\mathbb{I} + b\mathbb{J}}^2$$

Since all matrices are of the form $a\mathbb{I} + b\mathbb{J}$, they commute and have the same basis of eigenvectors. This greatly simplifies the computations and analysis, yet demonstrates incidentally the limitations of the approach. As long as the covariances $\mathbf{\Omega}_\epsilon$ and $\mathbf{\Omega}_\xi$ maintains the simplified structure of Eq. 70, no complex dynamics can emerge as the eigenspaces remaining stable.

5.2.2 Portfolio Allocation Form

From the symmetry of the assets and signals, we know that any optimal solution will be of the form:

$$\begin{aligned} \mathbf{L} &= \mathbf{L}^\top = a_w \mathbb{I} \mathbb{I} + b_w \mathbb{J} \\ \mathbf{w} &= (a_w \mathbb{I} \mathbb{I} + b_w \mathbb{J}) \mathbf{s} \\ &= a_w \mathbf{s} + n b_w \bar{s} \mathbb{I} \\ &= \lambda_w^2 \mathbf{s} + (\lambda_w^1 - \lambda_w^2) \bar{s} \mathbb{I} \end{aligned}$$

where $\bar{s} = \frac{1}{n} \mathbb{I}^\top \mathbf{s}$. The ratio $x_w = \frac{n b_w}{a_w}$ captures a position trade-off between being exposed to the average signal factor \bar{s} and idiosyncratic signals \mathbf{s} (note that the eigenvalues λ_w^1 and λ_w^2 are for the operator \mathbf{L} ; they are not the ones of $\tilde{\mathbf{\Pi}}$, nor the weights of canonical portfolios as encoded in $\frac{\sigma}{\sqrt{n}} \mathbf{\Theta}$).

The coefficient b_w is referred as the lead-lag term in [11]. It provides an exposition to the average signal factor \bar{s} (the exposure is multiplied by n). The ratio x_w measures the deviation from a conventional trading where cross-asset allocations term are ignored:

$$\text{lead-lag ratio: } x_w = n \frac{b_w}{a_w} = \frac{\lambda_w^1}{\lambda_w^2} - 1$$

In a “conventional trading” trend-following strategy, where the lead-lag term is ignored, setting $b_w = 0$ and $\mathbf{L} = a_w \mathbb{I} \mathbb{I}$, we can easily compute the theoretical in-sample (annualized) Sharpe ratio S_n^{tf} as:

Conventional Trend-Following

$$S_n^{tf}(\rho_\epsilon, \rho_\xi) = \frac{\mathbf{L}^{tf} \propto \mathbb{I} \mathbb{I}}{\sqrt{252} \sqrt{n} q \sqrt{1-p^2}} \quad (71)$$

$$\sqrt{Q^2 + 2Q + R + (n-1)(Q^2 \rho_\epsilon^2 + 2Q \rho_\epsilon \rho_\xi + R \rho_\xi^2)}$$

We find as expected that $S_n^{tf}(\rho_\epsilon = 0, \rho_\xi = 0) = \sqrt{n} S_1$: in the uncorrelated case, the Sharpe ratio scales in \sqrt{n} . However, the benefit of diversification appears to diminish in the presence of correlations. To illustrate this point, we plot below the ratio $\frac{1}{\sqrt{n}} S_n^{tf}(\rho_\epsilon, \rho_\xi)$ as a function of ρ_ϵ for different value of ρ_ξ .

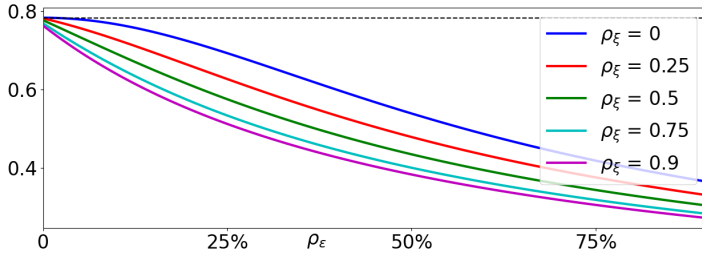


Figure 4: Annualized Sharpe ratio per asset computed as $\frac{1}{\sqrt{n}} S_n^{tf}(\rho_\epsilon, \rho_\xi)$ as a function of ρ_ϵ .

Surprisingly, the inclusion of the lead-lag term through the proper optimization of the functional allows one to compensate the loss. This property is explained clearly in [11].

To illustrate this point, let's consider the case where $\rho_\epsilon = \rho_\xi$. All the matrices of interest (e.g. $\mathbf{\Omega}$, $\mathbf{\Xi}$, $\mathbf{\Pi}$) are proportional to $\mathbf{\Omega}_\epsilon = \mathbf{\Omega}_\xi$, and the mean-variance solution takes the form:

$$\mathbf{\Omega}_\epsilon = \mathbf{\Omega}_\xi \implies \mathbf{L} \propto \mathbf{\Omega}_\epsilon^{-1}$$

From there, we can quickly compute the annualized Sharpe ratio and the lead-lag ratio:

Optimal Allocation when $\rho_\epsilon = \rho_\xi = \rho$

$$\mathbf{\Omega}_\epsilon = \mathbf{\Omega}_\xi \implies \mathbf{L}^{opt} \propto \mathbf{\Omega}_\epsilon^{-1} \quad (72)$$

$$x_w^{opt} = -\frac{n\rho}{1+(n-1)\rho}$$

$$S_n^{opt} = \frac{\sqrt{252} \sqrt{n} q \sqrt{1-p^2}}{\sqrt{Q^2 + 2Q + R}} = \sqrt{n} S_1$$

When $\rho_\epsilon = \rho_\xi = \rho$, the annualized Sharpe ratio does not depend on the correlation level ρ and equals $\sqrt{n} S_1$. The addition of a lead-lag term compensates exactly the drop observed in the conventional trend-following allocation.

5.2.3 Eigen-Equations

The strategies we consider are as follows:

1. Conventional Trend-Following

In a standard trend-following strategy, the positions are usually directly proportional to the signals, while the lead-lag term is ignored:

$$\mathbf{L} \propto \mathbb{I} \mathbb{I}$$

2. Isotropic-Mean

The framework is described in Eq. 39. Because all matrices commute, the isotropic-mean allocation takes the following form:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi}^{-\frac{1}{2}}$$

3. Closed-Form Mean-Variance

Neglecting the second variance term leads to the simple closed-form solution of Eq. 69 that we have used throughout this work:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{\text{Tr}(\tilde{\mathbf{\Pi}}^\top \tilde{\mathbf{\Pi}})}} \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1}$$

4. Exact Mean-Variance Allocation

The exact mean-variance solution can be obtained by directly solving Eq. 68.

5. Isotropy-Regularized Mean-Variance Framework

We also investigate the impact of adding an isotropy constraint using the framework described in Section 4 (with parameters $\tau = 1$ and $\eta = 1$).

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}} \mathbf{\Theta} \tilde{\mathbf{B}}^\top \mathbf{\Xi}^{-\frac{1}{2}}$$

where $\tilde{\mathbf{B}}$ is the orthonormal matrix of eigenvectors (e.g. $\tilde{\mathbf{B}}_1 = \frac{1}{\sqrt{n}} \mathbf{1}$) and the coefficients θ_1 and θ_2 verify 2 cubic equations (see Eq. 55).

We have the following equations for the eigenvalues of the corresponding operators:

$$\begin{aligned} \lambda_w^i(\text{trend-following}) &= \frac{\sigma}{\sqrt{\lambda_\Omega^1 \lambda_\Xi^1 + (\lambda_\Pi^1)^2 + (n-1)(\lambda_\Omega^2 \lambda_\Xi^2 + (\lambda_\Pi^2)^2)}} \\ \lambda_w^i(\text{iso-mean Eq. 39}) &= \frac{\sigma}{\sqrt{n \lambda_\Omega^1 \lambda_\Xi^1}} \\ \lambda_w^i(\text{mean-variance Eq. 69}) &= \frac{\sigma}{\sqrt{\sum \tilde{\Psi}_k^2}} \frac{\lambda_\Pi^i}{\lambda_\Omega^i \lambda_\Xi^i} \\ &= \frac{\sigma}{\sqrt{n \lambda_\Omega^i \lambda_\Xi^i}} \frac{\tilde{\Psi}_i}{\sqrt{\frac{1}{n} \sum \tilde{\Psi}_k^2}} \quad (73) \\ \lambda_w^i(\text{exact mean-var Eq. 68}) &\propto \frac{\sigma}{(\lambda_\epsilon^i)^2 + \frac{2}{1-pq} \beta_0^2 \lambda_\epsilon^i \lambda_\xi^i + \frac{1+q^2-2p^2q^2}{(1-pq)^2} \beta_0^4 (\lambda_\xi^i)^2} \\ \lambda_w^i(\text{mean-var-iso}) &= \frac{\sigma}{\sqrt{n \lambda_\Omega^i \lambda_\Xi^i}} \theta_i \end{aligned}$$

where $\tilde{\Psi}_i = (\lambda_\Omega^i \lambda_\Xi^i)^{-\frac{1}{2}} \lambda_\Pi^i$ and $\sum \tilde{\Psi}_k^2 = \tilde{\Psi}_1^2 + (n-1) \tilde{\Psi}_2^2$.

5.2.4 Validating the Closed-Form Solution Eq. 69

Before analyzing those strategies, we first validate the closed-form solution of Eq. 69. We verify that the second variance term can be neglected. Figure 5-left displays the ratio $\text{Tr}(\mathbf{\Pi} \mathbf{L} \mathbf{\Pi} \mathbf{L}) / \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top)$ when the allocation is determined as the closed-form mean-variance solution of Eq. 69. For most values of ρ_ϵ and ρ_ξ , the ratio remains below 2%. In the worst cases, corresponding to parameters where $|\rho_\epsilon - \rho_\xi| \gg 0$, the ratio barely exceeds a couple of %. The resulting Sharpe inflation due to the approximation is negligible.

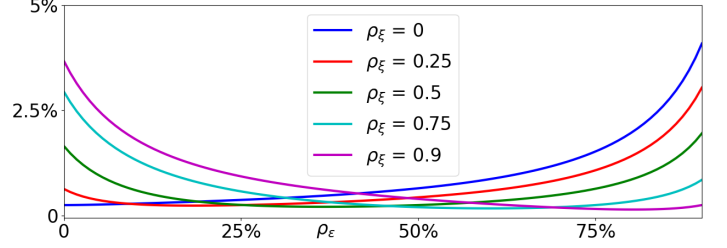


Figure 5: Ratio $\text{Tr}(\mathbf{\Pi} \mathbf{L} \mathbf{\Pi} \mathbf{L}) / \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^\top)$ as a function of ρ_ϵ for different values of ρ_ξ ($n = 10$) for the closed-form solution $\mathbf{L}^\top \propto \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1}$ of Eq. 69.

We also compare the exact solution Eq. 68 with the closed-form approach Eq. 69 used throughout this work. We display in Figure 6 the ratio λ_w^i (mean-variance Eq. 69) over λ_w^i (solution Eq. 68) for a range of parameters. We observe that both eigenmodes rarely diverge by more than 2%. The impact on the theoretical in-sample Sharpe ratio is also minimal (not displayed here).

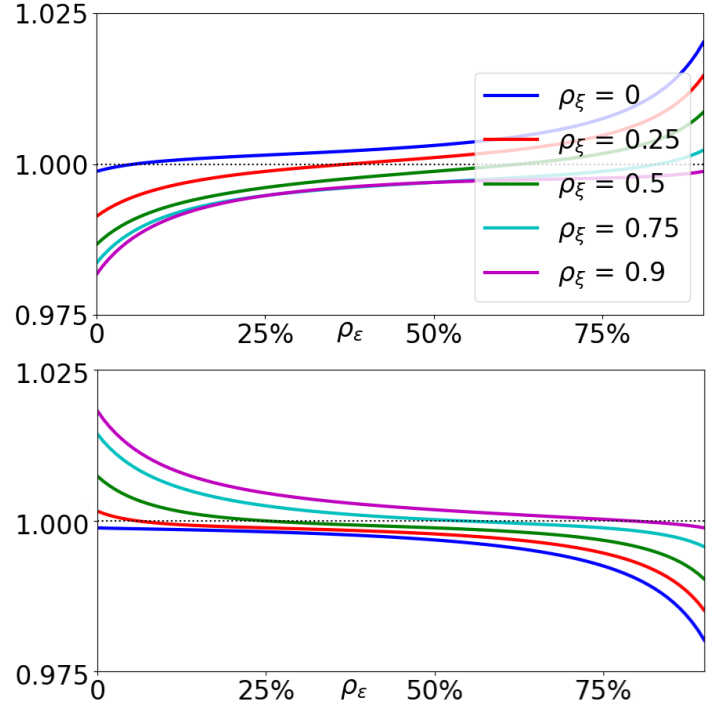


Figure 6: Ratios of eigenmodes (Top = first eigenmode; Bottom = second eigenmode) for the exact model Eq. 68 and the closed-form approach Eq. 69.

5.2.5 Lead-Lag Correction

Without lead-lag term, the presence of positive correlations (in returns and/or signals) deteriorates significantly the expected in-sample Sharpe ratio. The expected diversification benefit expected from using a large numbers of assets and signals is muted as correlations increase.

However, as we saw above, the introduction of a lead-lag term can drastically change the situation and help recover (and even improve) the diversification effect, expected to be in the magnitude of $\sqrt{n} S_1$. The lead-lag ratio x_w is the amount of exposure into \bar{s} per unit of exposure in s :

$$w = a_w (s + x_w \bar{s} \mathbb{1})$$

For instance, in the case of equal correlations, i.e. $\rho_e = \rho_\xi = \rho$, the optimal lead-lag ratio is equal to:

$$x^{opt} = -\frac{n\rho}{1 + (n-1)\rho} \quad (74)$$

with a resulting (in-sample) Sharpe ratio exactly equal to $\sqrt{n} S_1$. The exposure per asset to \bar{s} is always negative, with x^{opt} converging towards -1 as n increases (and $\rho > 0$).

In the general case where $\rho_e \neq \rho_\xi$, a proper accounting of the correlations through an optimized lead-lag term can improve even further the in-sample Sharpe ratio, reaching (theoretical and non-realistic) values greater than $\sqrt{n} S_1$. This is a strong result, which was noticed and emphasized in [11].

The Isotropic-Mean Case

In the case of the isotropic-mean allocation of Eq. 39, the lead-lag ratio takes the following form:

$$\begin{aligned} x^{ep} &= \left((1 + n \frac{b_\Omega}{a_\Omega}) (1 + n \frac{b_\Xi}{a_\Xi}) \right)^{-\frac{1}{2}} - 1 \leq 0 \\ &= \frac{1 - \rho_\xi}{1 + (n-1)\rho_\xi} \sqrt{\frac{(\frac{1-\rho_e}{1-\rho_\xi} + \beta_0^2)(\frac{1-\rho_e}{1-\rho_\xi} + \frac{1+pq}{1-pq}\beta_0^2)}{(\frac{1-(n-1)\rho_e}{1-(n-1)\rho_\xi} + \beta_0^2)(\frac{1-(n-1)\rho_e}{1-(n-1)\rho_\xi} + \frac{1+pq}{1-pq}\beta_0^2)}} - 1 \end{aligned}$$

The non-diagonal coefficient is always negative and quickly becomes significant (e.g. when $|\rho_e - \rho_\xi| > 15\%$ or as soon as $\rho_\xi > 50\%$).

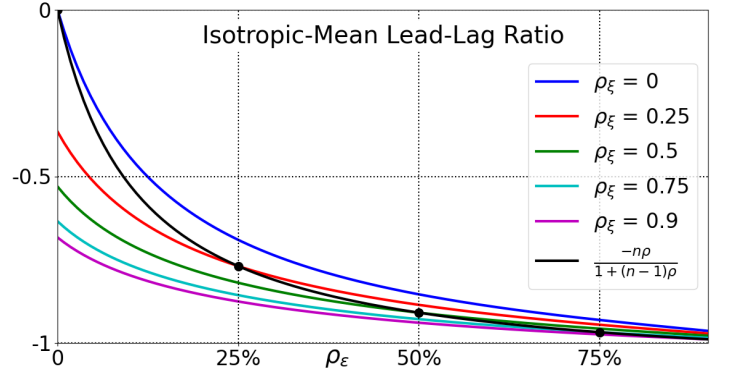


Figure 7: Lead-lag ratio for the Eigenrisk Allocation ($n = 10$).

The position of each asset S^i is a linear combination of its own signal s^i and a negative contribution of the average signal \bar{s} . This negative contribution serves as a hedge and tends to diminish the over-reliance on the individual signals.

As expected when $\rho_e = \rho_\xi$, we end up with the same allocation as in Eq. 74 with a corresponding Sharpe equal to $\sqrt{n} S_1$. We observe that when $\rho_e \geq \rho_\xi$, the annualized Sharpe ratio per asset is even higher (see Figure 8). Yet, in the most likely scenario where $\rho_e > \rho_\xi$, the Sharpe per asset remains lower than S_1 . The advocated hedging is costly in-sample, but could prevent some painful situation as we discuss below in Section 5.2.6.

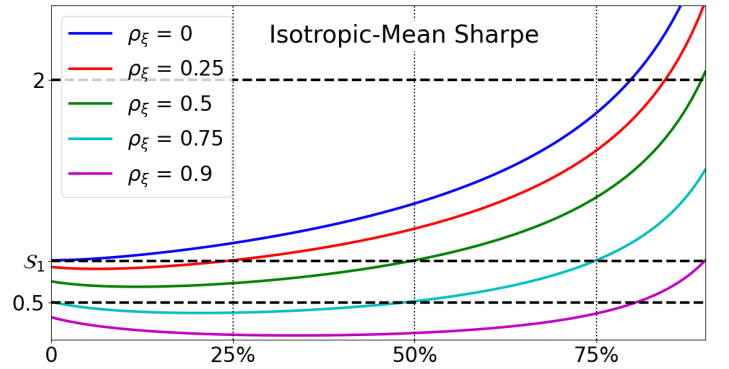


Figure 8: Annualized Sharpe per asset for the Isotropic-Mean Portfolio ($n = 10$).

The Case of Mean-Variance

The case of the mean-variance framework is even more interesting, as the sign of the lead-lag term depends on the choice of parameters (e.g. $\rho_e > 0$, $\rho_\xi > 0$, but also n). The lead-lag ratio takes the following form:

$$x^{opt} = \frac{1 - \rho_\xi}{1 + (n-1)\rho_\xi} \frac{(\frac{1-\rho_e}{1-\rho_\xi} + \beta_0^2)(\frac{1-\rho_e}{1-\rho_\xi} + \frac{1+pq}{1-pq}\beta_0^2)}{(\frac{1-(n-1)\rho_e}{1-(n-1)\rho_\xi} + \beta_0^2)(\frac{1-(n-1)\rho_e}{1-(n-1)\rho_\xi} + \frac{1+pq}{1-pq}\beta_0^2)} - 1$$

The optimal lead-lag ratio can turn positive when the noise correlation ρ_e is much smaller than the innovation correlation ρ_ξ , e.g. when $\rho_\xi \approx 0$ and $\rho_e > 0$. In this scenario, the optimization of the mean-variance functional leads to some positions that reinforce the individual signal views s with a positive exposure to \bar{s} , thereby increasing the level of risk associated with the strategy. This could prove dangerous in the case of a sudden market crash (see below in Section 5.2.6).

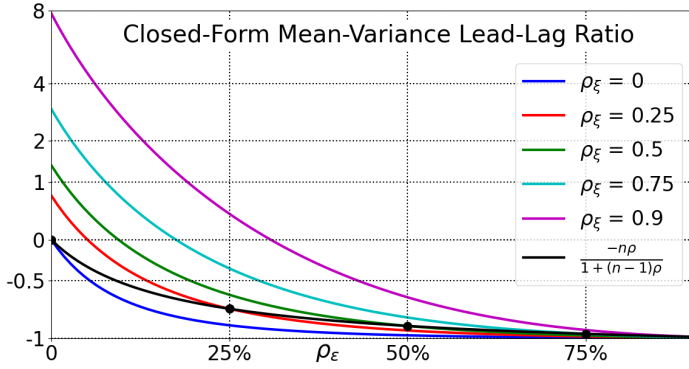


Figure 9: Lead-lag ratio for the Mean-Variance Solution ($n = 10$).

We display the Sharpe ratio of the mean-variance solution in Figure 10 for $n = 10$. Many parameter configurations shows an annualized Sharpe per asset higher than S_1 . This is particularly the case when the noise correlation is much higher than the innovation correlation, i.e. $\rho_e > \rho_\xi$.

However, the case that concerns us more in practice where $\rho_e < \rho_\xi$ is less advantageous. Taking into account the proper correlations obviously leads to an improvement over the Sharpe $S_n^{tf}(\rho_e, \rho_\xi)$ of a conventional strategy (see Eq. 71), yet the gain appears more marginal.

In the realistic region where $\rho_\xi \approx 0.75$ and $\rho_e \approx 0.25 - 0.5$, both isotropic-mean and mean-variance annualized Sharpe ratio per asset hover around 0.5 with a negative lead-lag ratio lower than -0.5 .

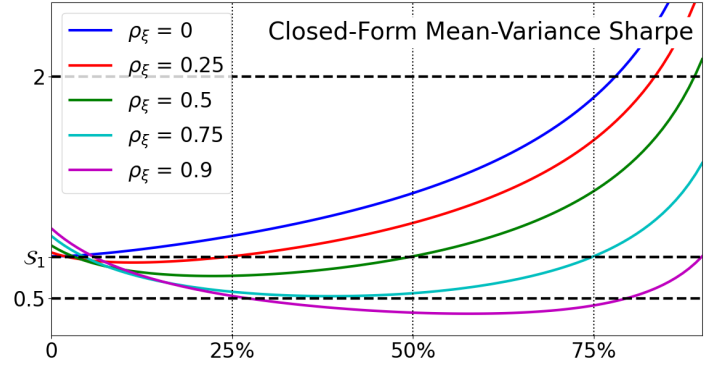


Figure 10: Annualized Sharpe per asset for the Mean-Variance Solution ($n = 10$).

The region where $\rho_e \approx 0$ and $\rho_\xi \gg \rho_e$ merits some comments. As ρ_e tends towards 0, the Sharpe ratio per asset converges around S_1 . There even appears to be an improvement of the Sharpe per asset over S_1 , but it remains minimal. However, as displayed in Figure 9, this corresponds to lead-lag ratios significantly positive (increasing as ρ_ξ increases), which boosts the Sharpe by reinforcing the individual signal s with a positive exposure to \bar{s} . This is a well-known fallacy of the mean-variance framework, which creates at times large, unreasonable, and risky positions just for the sake of maximization.

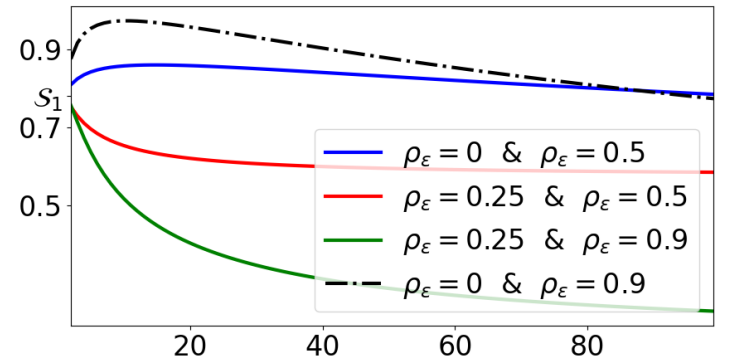


Figure 11: Evolution of the annualized Sharpe per asset as a function of n for different parameters ρ_e and ρ_ξ .

Going further, the region where $0 \leq \rho_e \ll \rho_\xi$ exhibits some non-intuitive properties, as they also depend on the number n of assets. As noted in [11], there is a non-monotonous behavior, with an inflexion point around 50 assets. For a large number of assets (e.g. higher than 100), the lead-lag ratio and the Sharpe ratio per asset would start to decrease. This is visible in Figure 11.

Mean-Variance versus Eigenrisk: Differences

Both isotropic-mean and mean-variance allocations are defined by their exposures to two (orthogonal) eigenspaces. For each mode, the exposure ratio is captured by $\frac{\tilde{\Psi}_i}{\sqrt{\frac{1}{n} \sum \tilde{\Psi}_k^2}}$ (see Eq. 73).

This ratio is linked to the notion of effective rank and participation ratio as defined in Eq. 63. A low effective rank/participation ratio would happen when the imbalance $\tilde{\Psi}_1 \gg \tilde{\Psi}_2$ is large, typically in the region with strongly correlated stochastic trends and low noise correlation (see Figure 14 further below). This is where the lead-lag difference is at its maximum. The mean-variance solution exploits meaningful differences to optimize the Sharpe ratio pushing the allocation into dangerous territory.

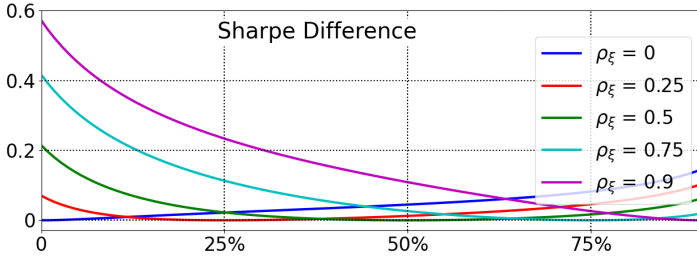


Figure 12: Difference of Sharpe ratios between the mean-variance solution Eq. 69 and the isotropic-mean approach Eq. 39 (note that it cancels exactly when $\rho_\epsilon = \rho_\xi$).

At the opposite, isotropic-mean allocation enforces a negative lead-lag term and acts as a hedge. The impact on the in-sample Sharpe ratio could be large depending on current parameters (see Figure 12). This is particularly the case in the region when $\rho_\xi \gg \rho_\epsilon$, as up to 0.6 point of Sharpe could be lost between both approaches (see Figure 12).

However, those are in-sample measures. In case of a sudden crash, this difference might not become so important anymore, as the benefit of a negative lead-lag ratio might really make the difference.

5.2.6 Impact of Market Crash

During a market crash, things go haywire. Our set of assumptions and the whole model would not make much sense anymore. We assume that the returns would all suddenly drop $\mathbf{r} \approx -\sigma_\epsilon \mathbf{1}$. As a consequence, the realized PnL would take the following form:

$$\mathbf{w}^\top \mathbf{r} = \lambda_w^2 \mathbf{s}^\top \mathbf{r} + (\lambda_w^1 - \lambda_w^2) \bar{\mathbf{s}}^\top \mathbf{r} \approx -n\sigma_\epsilon \lambda_w^1 \bar{\mathbf{s}}$$

Depending on the sign of $\bar{\mathbf{s}}$ right before the crash, it is certainly possible to envision an unexpected PnL jump to the upside. This would be the case if there is a progressive crisis build-up, reflected

in a progressive downtrend of the market leading to negative signals before the crash. However, sudden, unanticipated, and negative news would typically work against the general macro-environment, and would likely cause large losses rather than large gains.

The first eigenvalue λ_w^1 (with multiplicity 1) is the one modulating the risk during the crash. We have:

$$\lambda_w^1(\text{mean-variance}) \geq \lambda_w^1(\text{iso-mean})$$

if and only if $\tilde{\Psi}_1 \geq \tilde{\Psi}_2$, or equivalently:

$$1 + n \frac{b_\Pi}{a_\Pi} \geq \sqrt{(1 + n \frac{b_\Omega}{a_\Omega})(1 + n \frac{b_\Xi}{a_\Xi})}$$

Now, at first-order in β_0^2 , assuming a large number of assets n and strictly positive correlations ρ_ϵ and ρ_ξ , this condition is equivalent to:

$$\rho_\xi \geq \rho_\epsilon > 0$$

which seems a reasonable assumption in the case in a standard sector model. The isotropic-mean allocation would have a smaller exposure to the principal market mode $-\sigma_\epsilon \mathbf{1}$.

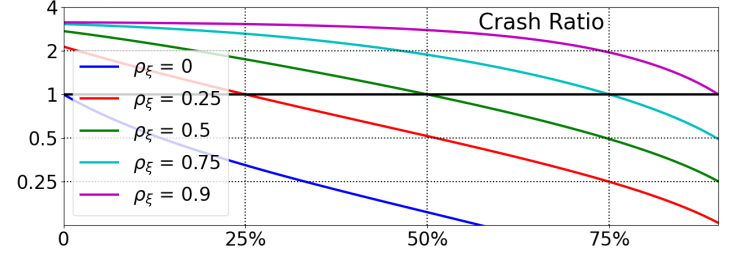


Figure 13: Log2-ratio $\lambda_w^1(\text{mean-variance})/\lambda_w^1(\text{iso-mean})$ of first eigenmodes of mean-variance over isotropic-mean.

The crash ratio, defined as $\lambda_w^1(\text{mean-variance})/\lambda_w^1(\text{iso-mean}) - 1$, takes the form:

$$\frac{\tilde{\Psi}_1}{\sqrt{\frac{1}{n} \sum \tilde{\Psi}_k^2}} = \frac{1 + n \frac{b_\Pi}{a_\Pi}}{\sqrt{\frac{1}{n} (1 + n \frac{b_\Pi}{a_\Pi})^2 + \frac{n-1}{n} (1 + n \frac{b_\Omega}{a_\Omega})(1 + n \frac{b_\Xi}{a_\Xi})}}$$

Using some sensible approximations (e.g. $0 < \rho_\epsilon \leq \rho_\xi < 1$, $p \approx q$ so that $\frac{1+pq}{1-pq} \beta_0^2 \gg 1$), an order of magnitude can be computed as:

$$\frac{\tilde{\Psi}_1}{\sqrt{\frac{1}{n} \sum \tilde{\Psi}_k^2}} \approx \frac{b_\Pi}{a_\Pi} \sqrt{\frac{a_\Omega}{b_\Omega} \frac{a_\Xi}{b_\Xi}} \approx \sqrt{\frac{\rho_\xi}{1 - \rho_\xi} \frac{1 - \rho_\epsilon}{\rho_\epsilon}} \quad (75)$$

The crash ratio is easily around 150% as soon as $\rho_\xi - \rho_\epsilon > 25\%$. Figure 13 displays the ratio of the first eigenmode between the closed-form solution and the isotropic-mean allocation.

5.2.7 Isotropy-Regularized Mean-Variance as a Safeguard

We now investigate how our isotropy-regularized mean-variance framework would naturally safeguard against perilous regions. We simply set $\tau = \eta = 1$ and solve our isometric mean-variance functional.

To start, we plot the isotropy metric of the exact mean-variance closed-form solution. Figure 14 displays $\frac{1}{\tilde{\psi}} - 1$ (in y-log-coordinate) where $\tilde{\psi}$ is the participation ratio of the normalized predictability matrix $\tilde{\mathbf{\Pi}}$. As we can observe, the departure from isotropy is maximal in the region of interest $0 \approx \rho_\epsilon \ll \rho_\xi$. The isotropy penalization would kick-in in that risky region and naturally avoid zones that are more isotropic to start with.

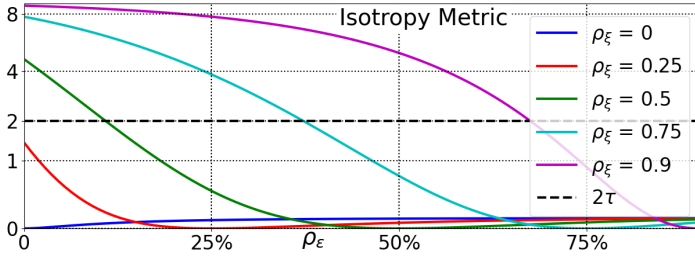


Figure 14: Isotropy metric of the mean-variance solution.

By capping the isotropy metric at 2τ , we prevent absurdly large lead-lag ratios through the over-optimization of the Sharpe ratio. Figure 15 displays the resulting lead-lag ratio. We also observe in Figure 14 some clear inflection points where the isotropy metric reaches its cap 2τ . Interestingly, the resulting lead-lag ratios do not exceed 1, even in the worst cases.

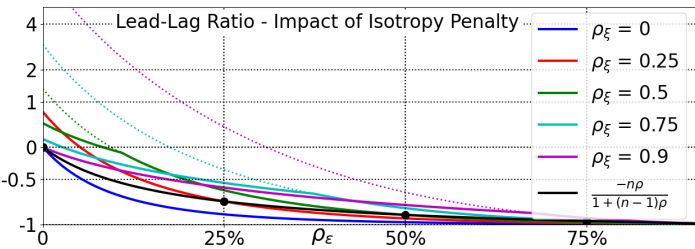


Figure 15: Lead-lag ratio of the Isotropy-Regularized Mean-Variance Solution ($n = 10, \eta = \tau = 1$). The mean-variance ratios are indicated by dotted lines.

We also investigate how the crash ratio, which we now define as $\lambda_w^1(\text{mean-variance})/\lambda_w^1(\text{iso-reg-mean-var}) - 1$, evolves as a function of ρ_ϵ and ρ_ξ . In the riskiest region, the exposure to the first eigenmode is decreased by more than 30% (see Figure 16).

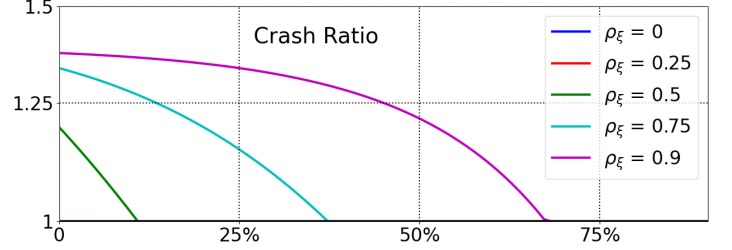


Figure 16: Log2-ratio $\lambda_w^1(\text{mean-variance})/\lambda_w^1(\text{iso-reg-mean-var})$ of first eigenmodes of mean-variance over our isotropy-regularized mean-variance approach.

The cost in Sharpe ratio is rather small, as displayed in Figure 17. Even in the problematic region, the mean-variance Sharpe ratio is only marginally better, thereby confirming that the excessive magnitude of the lead-lag ratio is a dangerous by-product of the untamed mean-variance optimization. The isotropic penalty works by preventing corner solution, and avoid naturally unbalanced risky allocations.

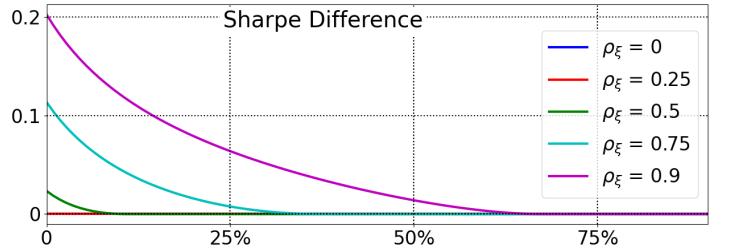


Figure 17: Difference of Sharpe ratios between the mean-variance solution Eq. 69 and the isotropic-mean approach Eq. 39 (note that it cancels exactly when $\rho_\epsilon = \rho_\xi$).

6 Final Words

It seems important to start with a **disclaimer** stating clearly that the described approach is overly simplistic and unrealistic. Markets show fat tails, sudden regime breaks, and execution frictions (costs, slippage, impact), while signals and covariances are rarely joint Gaussians. Those were all ignored for the sake of simplicity, tractability, and clarity.

Besides, we focused on a specific narrow problem, that is how to mitigate the over-reliance on untrustworthy signals within a mean-variance framework. It does so by introducing isotropy as a safeguard, but relies on a critical assumption that $\mathbf{\Omega}$ and $\mathbf{\Xi}$ are well-estimated, accurate, and stable through time, a property that is rarely met in practice.

On that point, we also did not discuss the estimation of parameters. Even if a model as straight-forward as the sector trend-following model presented in Section 5 had the good behavior of being valid, the estimation of the unknown parameters, especially the covariances, e.g. ρ_ϵ and ρ_ξ , would be a challenging and critical task.

In light of those strong limitations, it should be clear that the conclusions should be taken with a large pinch of salt and that additional considerations are required in practice.

However, Basis Immunity (BI) introduces some original ideas and sheds light on a few noteworthy issues:

- By careful analysis of the concept of isotropy, we designed a sound framework where the uncertainty of the signals is mitigated through the concept of isotropy. We properly defined isotropic bases and showed the importance of canonical portfolios as building blocks.
- A general portfolio allocation that takes the form (without loss of generality) $\mathbf{w} = \mathbf{L}^\top \mathbf{s} = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-1} \mathbf{M}^\top \mathbf{s}$ with $\mathbf{M}^\top \in \mathbb{R}^{n \times m}$ can be expressed as:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \left(\mathbb{R}_{\hat{\mathbf{b}}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} \right) \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \quad (76)$$

where $\mathbb{R}_{\hat{\mathbf{b}}}$ and $\mathbb{R}_{\hat{\mathbf{u}}}$ are two rotations, corresponding to the isotropic bases $\{\hat{\mathbf{b}}\}$ and $\{\hat{\mathbf{u}}\}$ respectively. The linear operator:

$$\mathbf{T} = \mathbb{R}_{\hat{\mathbf{b}}}^\top \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{u}}} \in \mathbb{R}^{n \times m}$$

facilitates the computation of a few important expressions:

return	$E[\mathbf{w}^\top \mathbf{r}] = \frac{\sigma}{\sqrt{n}} \text{Tr}(\mathbf{T}^\top \mathbf{\Pi}_{\hat{\mathbf{b}} \hat{\mathbf{u}}})$
variance	$\text{Var}[\mathbf{w}^\top \mathbf{r}] = \frac{\sigma^2}{n} \text{Tr}(\mathbf{T} \mathbf{T}^\top)$
anisotropy	$\frac{1}{n} \ \mathbf{T} \mathbf{T}^\top - \eta_T \mathbb{I}_n\ _F^2, \eta_T = \frac{\text{Tr}(\mathbf{T} \mathbf{T}^\top)}{n} \text{ or } \eta_T = \text{cst}$

where $\mathbf{\Pi}_{\hat{\mathbf{b}} \hat{\mathbf{u}}} = \mathbb{R}_{\hat{\mathbf{b}}}^\top \tilde{\mathbf{\Pi}} \mathbb{R}_{\hat{\mathbf{u}}}$ and $\tilde{\mathbf{\Pi}} = \mathbf{\Pi}_{\mathbf{b} \mathbf{u}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}}$ is the normalized predictability matrix.

These measures are intrinsic and do not depend on a specific choice of isotropic basis (because the trace and Frobenius norm are invariant under simultaneous rotations of $\mathbb{R}_{\hat{\mathbf{b}}}$ and $\mathbb{R}_{\hat{\mathbf{u}}}$).

In Eq. 76, fixing the volatility σ at $\sigma^2 = \text{Tr}(\mathbf{\Omega} \mathbf{L} \mathbf{\Xi} \mathbf{L}^\top)$ shows that the isotropy metric is inversely proportional to the participation ratio of $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}}$ (computed from its eigenspectrum).

In the mean-variance framework, a closed-form solution can be (approximately) expressed as ($\mathbf{M}^\top = \mathbf{\Pi} \mathbf{\Xi}^{-1}$):

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{\text{Tr}(\tilde{\mathbf{\Pi}} \tilde{\mathbf{\Pi}}^\top)}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbb{R}_{\hat{\mathbf{b}}} \left(\mathbb{R}_{\hat{\mathbf{b}}}^\top \tilde{\mathbf{\Pi}} \mathbb{R}_{\hat{\mathbf{u}}} \right) \mathbb{R}_{\hat{\mathbf{u}}}^\top \mathbf{\Xi}^{-\frac{1}{2}} \quad (77)$$

- Enforcing full isotropy (with variance at σ^2) while preserving some directional information encoded within the matrix \mathbf{M}^\top can be achieved by identifying the two operators $\mathbb{R}_{\hat{\mathbf{b}}}$ and $\mathbb{R}_{\hat{\mathbf{u}}}$ so that \mathbf{T} , which is of rank n , becomes as close as possible (in the sense of the Frobenius norm) to the linear operator $[\mathbb{I}_n, \mathbf{0}_{n, m-n}]$ and then replacing \mathbf{T} by $[\mathbb{I}_n, \mathbf{0}_{n, m-n}]$ in Eq. 76, i.e. keeping only the first n right singular vectors.

This is easily achieved through the singular value decomposition of the matrix:

$$\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^\top \mathbf{\Xi}^{\frac{1}{2}} = \dot{\mathbf{B}} \dot{\mathbf{\Psi}} \dot{\mathbf{U}}^\top = \dot{\mathbf{B}} \dot{\mathbf{\Psi}}_n \dot{\mathbf{U}}_n^\top$$

leading to:

$$\mathbb{R}_{\hat{\mathbf{b}}} = \dot{\mathbf{B}}, \mathbb{R}_{\hat{\mathbf{u}}} = \dot{\mathbf{U}}, \mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \dot{\mathbf{B}} \dot{\mathbf{U}}_n^\top \mathbf{\Xi}^{-\frac{1}{2}}$$

Some noteworthy comments:

- The solution can be decomposed into a set of n orthogonal portfolios $\mathbf{\Omega}^{-\frac{1}{2}} \dot{\mathbf{B}}_i \dot{\mathbf{U}}_i^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}$, equally weighted.
- There are $m - n$ signal basis vectors that span a linear space with no contribution - those are uninformative for the n -returns. Those could form a basis for statistical arbitrage on signal residuals, analogous to idiosyncratic risk in factor models. This will be explored in future work.

- When $m = n$ and $\mathbf{M}^\top = \mathbb{I}_n$ the allocation is the same as the one proposed in [22] and takes the form:

$$\mathbf{L}^\top = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Xi} \mathbf{\Omega}^{-\frac{1}{2}} \right)^{-\frac{1}{2}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{s},$$

It is slightly different from the ERP approach of [2] (except when $\mathbf{\Omega}$ and $\mathbf{\Xi}$ commute).

- Agnostic Risk Parity [2] (ARP) is a special case of ERP, where the signal covariance $\mathbf{\Xi}$ is chosen as $\mathbf{\Xi} \propto \varphi \mathbf{\Omega} + (1 - \varphi) \mathbb{I}$. In this scenario BI=ARP.
- **Isotropic-Mean Allocation**

In the case of the mean-variance approach $\mathbf{M}^\top = \mathbf{\Pi} \mathbf{\Xi}^{-1}$, the orthogonal portfolios $\tilde{\mathbf{w}}_k$ are constructed from the singular vectors $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{U}}$ of the normalized predictability matrix $\tilde{\mathbf{\Pi}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}}$, also known as canonical portfolios [8]:

$$\tilde{\mathbf{w}}_k = \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_k \tilde{\mathbf{U}}_k^\top \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}$$

- Full isotropy could potentially deform significantly the theoretical closed-form solution of Eq. 77. In order to retain some amount of control, we augment the mean-variance framework with a tunable isotropy penalty, thereby offering an adjustable trade-off between return maximization, variance minimization, and isotropic control:

$$\arg_T \max \frac{1}{\sqrt{n}} \text{Tr} \left(\mathbf{T}^\top \tilde{\mathbf{\Pi}} \right) - \frac{\gamma}{2n} \text{Tr} \left(\mathbf{T} \mathbf{T}^\top \right) - \frac{\lambda}{4n} \|\mathbf{T} \mathbf{T}^\top - \eta \mathbb{I}_n\|_{\mathbb{F}}^2$$

Canonical portfolios $\tilde{\mathbf{w}}_k$ emerge naturally as the core building blocks:

$$\mathbf{w} = \frac{\sigma}{\sqrt{n}} \sum_{k=1}^n \theta_k \tilde{\mathbf{w}}_k$$

where the parameters θ_i solve n coupled cubic equations:

$$\sqrt{n} \tilde{\Psi}_i = (\gamma - \eta \lambda) \theta_i + \lambda \theta_i^3$$

where $\tilde{\Psi}_i$ are the singular values of $\tilde{\mathbf{\Pi}}$.

This creates a smooth trade-off between isotropic-mean portfolios and mean-variance allocations. Pure isotropy flattens allocations ($\theta_i = \sqrt{\eta}$), while mean-variance scales them by eigenvalue strength ($\theta_i \propto \tilde{\Psi}_i$).

The parameters η and τ controlling the amount of isotropy can be fine-tuned (generally, setting $\tau = \eta = 1$ appears a sensible choice).

- Although the general solution and the decomposition into canonical portfolios do not depend on the specific choice of isotropic bases, one could employ alternative ones, such as those designed for enhanced stability (e.g. Cholesky or others).
- We showed an existing link with the principal portfolio approach [13]. Principal portfolios are not purely intrinsic and depend on the choice of basis (modulo an invariance to rotations).

Principal portfolios emerge naturally as canonical portfolios when the triple norm is expressed between isotropic bases, e.g. $\{\hat{\mathbf{b}}\}$ and $\{\hat{\mathbf{u}}\}$. Therefore, similar techniques of principal beta portfolios and principal alpha portfolios could be applied (see [13]), and will be explored in further work.

- As an application, we reviewed the sector trend-following model introduced in [11] (mixing stochastic trends with noise). We recovered the same expressions²² and reached similar conclusions. Some features of the models are counter-intuitive and could generate some risky allocations.

Depending on parameters, such as the number of assets n , but particularly the correlations ρ_ϵ (noise) and ρ_ξ (trend), the optimal cross-asset position (referred to as lead-lag term) could either be negative (e.g. when $\rho_\epsilon \ll \rho_\xi$) or turn significantly positive (e.g. when $\rho_\epsilon \gg \rho_\xi$). This is particularly the case in the realistic scenario where stochastic trends are significantly correlated.

The isotropy constraint would certainly help in this case. Isotropic-mean allocation always possess a negative lead-lag term, acting as a hedging component with a negative exposure to the average signal \bar{s} . Depending on the market regime and unknown model parameters, such allocations would be less impacted by sudden regime changes as a market crash.

Our isotropy-regularized mean-variance (IRMV) approach naturally tames the propensity of the mean-variance framework to amplify imbalances as captured by the participation ratio of the normalized predictability matrix, thereby preventing undiversified corner solutions.

While fragile to estimation error and regime shifts, our framework reframes signal uncertainty as a *measurable geometric defect* and mitigates it via *canonical, isotropic structure*. The isotropy-regularized mean-variance portfolios interpolates between full isotropic portfolios (i.e. isotropic-mean portfolios) and mean-variance allocations.

²²We note that the closed-form mean-variance solution of Eq. 77 has the advantage of greatly simplifying some of the calculations.

7 Summary

Extending Mean-Variance	
$\mathbf{\Omega} = E[\mathbf{r}\mathbf{r}^T]$ asset covariance $\mathbf{\Xi} = E[\mathbf{s}\mathbf{s}^T]$ signal covariance $\mathbf{\Pi} = E[\mathbf{r}\mathbf{s}^T]$ prediction/cross-covariance	$\{\mathbf{e}_i\}$ natural basis $\{\mathbf{b}_i\}$ Riccati basis $\mathbf{b}_i = \mathbf{\Omega}^{-\frac{1}{2}}\mathbf{e}_i$ $\{\mathbf{u}_i\}$ Riccati basis $\mathbf{u}_i = \mathbf{\Xi}^{-\frac{1}{2}}\mathbf{e}_i$
regression-based $\mathbf{r} = \beta\mathbf{s} + \epsilon$ $\beta = E[\mathbf{r}\mathbf{s}^T] E[\mathbf{s}\mathbf{s}^T]^{-1} = \mathbf{\Pi}\mathbf{\Xi}^{-1}$ $E[\mathbf{r} \mathbf{s}] = \beta\mathbf{s} = \mathbf{\Pi}\mathbf{\Xi}^{-1}\mathbf{s}$ $\mathbf{w}_\star = \arg_{\mathbf{w}} \max \mathbf{w}^T E[\mathbf{r} \mathbf{s}] - \frac{\gamma}{2} \mathbf{w}^T \mathbf{\Omega} \mathbf{w}$ $\mathbf{w}_\star = \frac{1}{\gamma} \mathbf{\Omega}^{-1} E[\mathbf{r} \mathbf{s}] = \frac{1}{\gamma} \mathbf{\Omega}^{-1} \mathbf{\Pi} \mathbf{\Xi}^{-1} \mathbf{s}$	general mean-variance $\mathbf{w} = \mathbf{L}^T \mathbf{s}$ $E[\mathbf{w}^T \mathbf{r}] = \text{Tr}(\mathbf{L} \mathbf{\Pi})$ $\text{Var}[\mathbf{w}^T \mathbf{r}] = \text{Tr}(\mathbf{\Xi} \mathbf{L} \mathbf{\Omega} \mathbf{L}^T) + \text{Tr}(\mathbf{\Pi} \mathbf{L} \mathbf{H} \mathbf{L}^T)$ $\mathbf{L}_\star = \arg_{\mathbf{L}} \max E[\mathbf{s}^T \mathbf{L} \mathbf{r}] - \frac{\gamma}{2} \text{Var}[\mathbf{s}^T \mathbf{L} \mathbf{r}]$ $\mathbf{L}_\star = \frac{1}{\gamma} \mathbf{\Xi}^{-1} \mathbf{\Pi}^T \mathbf{\Omega}^{-1}$
$\mathbf{w}_e = \frac{1}{\gamma} \mathbf{\Omega}^{-\frac{1}{2}} \underbrace{\left(\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}} \right)}_{\text{in } \{\mathbf{b}_i\}} \underbrace{\mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e}_{\text{in } \{\mathbf{u}_i^*\}} \leftarrow \{\mathbf{u}_i^*\} \text{ in } \{\mathbf{u}_i^*\}$	
Canonical Portfolios [$m \geq n$] [8] $\mathbf{w}_e = \mathbf{L}_\star^T \mathbf{s}_e = \frac{1}{\gamma} \sum_{k=1}^n \tilde{\Psi}_k \tilde{\mathbf{w}}_k$ $\mathbf{L}_\star = \arg_{\mathbf{L}} \max E[\mathbf{s}^T \mathbf{L} \mathbf{r}] - \frac{\gamma}{2} \text{Var}[\mathbf{s}^T \mathbf{L} \mathbf{r}]$ $\mathbf{L}_\star = \frac{1}{\gamma} \mathbf{\Xi}^{-1} \mathbf{\Pi}^T \mathbf{\Omega}^{-1}$ $\tilde{\mathbf{\Pi}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^T$ canonical portfolios $\tilde{\mathbf{w}}_k = \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_k \tilde{\mathbf{U}}_k^T \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$	Principal Portfolios [$m = n$] [13] $\mathbf{w}_e = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{w}_b = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{L}_{bb}^T \mathbf{s}_b = \frac{1}{\gamma} \sum_{k=1}^n \ddot{\mathbf{w}}_k$ $\mathbf{L}_{bb} = \arg_{\mathbf{L}} \max_{\ \mathbf{L}\ \leq 1} E[\mathbf{s}_b^T \mathbf{L} \mathbf{r}_b]$ $\mathbf{L}_{bb} = \frac{1}{\gamma} (\mathbf{\Pi}_{bb}^T \mathbf{\Pi}_{bb})^{-\frac{1}{2}} \mathbf{\Pi}_{bb}^T$ with $\mathbf{\Pi}_{bb} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Omega}^{-\frac{1}{2}}$ $\mathbf{L}_\star = \ddot{\mathbf{U}} \ddot{\mathbf{B}}^T = \sum_k \ddot{\mathbf{U}}_k \ddot{\mathbf{B}}_k^T$ $\ddot{\mathbf{w}}_k = \mathbf{\Omega}^{-\frac{1}{2}} \ddot{\mathbf{B}}_k \ddot{\mathbf{U}}_k^T \mathbf{s}_b = \mathbf{\Omega}^{-\frac{1}{2}} \ddot{\mathbf{B}}_k \ddot{\mathbf{U}}_k^T \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{s}_e$
Isotropic-Mean [$m \geq n$] Eq. 39 $\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}} \tilde{\mathbf{U}}_n^T \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \sum_{k=1}^N \tilde{\mathbf{w}}_k$ $\tilde{\mathbf{\Pi}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^T$ $\tilde{\mathbf{w}}_k = \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_k \tilde{\mathbf{U}}_k^T \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$ Case when $E[\mathbf{r} \mathcal{F}] \propto \mathbf{M}^T \mathbf{s}$ $\mathbf{w}_e = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \dot{\mathbf{B}} \dot{\mathbf{U}}_n^T \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$ with $\mathbf{\Omega}^{-\frac{1}{2}} \mathbf{M}^T \mathbf{\Xi}^{+\frac{1}{2}} = \dot{\mathbf{B}} \dot{\Psi} \dot{\mathbf{U}}^T$	Isotropy-Regularized Mean-Variance [$m \geq n$] $\mathbf{w}_e = \mathbf{L}_\star^T \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{T}_\star \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e = \frac{\sigma}{\sqrt{n}} \sum_{k=1}^N \theta_k \tilde{\mathbf{w}}_k$ $\mathbf{T} = \tilde{\mathbf{B}} \mathbf{\Theta} \tilde{\mathbf{U}}^T$ where $\tilde{\mathbf{\Pi}} = \mathbf{\Omega}^{-\frac{1}{2}} \mathbf{\Pi} \mathbf{\Xi}^{-\frac{1}{2}} = \tilde{\mathbf{B}} \tilde{\Psi} \tilde{\mathbf{U}}^T$ $\mathbf{T}_\star = \arg_{\mathbf{T}} \max \frac{1}{\sqrt{n}} \text{Tr}(\mathbf{T}^T \tilde{\mathbf{\Pi}}) - \frac{\gamma}{2n} \text{Tr}(\mathbf{T} \mathbf{T}^T) - \frac{\lambda}{4n} \ \mathbf{T} \mathbf{T}^T - \eta \mathbf{I}\ _{\mathbb{F}}^2$ $\sqrt{n} \tilde{\Psi}_i = \gamma \theta_i + \lambda \theta_i (\theta_i^2 - \eta) = (\gamma - \eta \lambda) \theta_i + \lambda \theta_i^3$ $\tilde{\mathbf{w}}_k = \mathbf{\Omega}^{-\frac{1}{2}} \tilde{\mathbf{B}}_k \tilde{\mathbf{U}}_k^T \mathbf{\Xi}^{-\frac{1}{2}} \mathbf{s}_e$

References

- [1] N. Baltas. Optimising cross-asset carry. *Factor Investing*, 5 2017.
- [2] R. Benichou, Y. Lempérière, E. Sérié, J. Kockelkoren, P. Seager, J.-P. Bouchaud, and M. Potters. Agnostic risk parity: Taming known and unknown-unknowns. *arXiv:1610.08818*, Oct. 2016.
- [3] J.-P. Bouchaud and M. Potters. Financial applications of random matrix theory: a short review. *arXiv*, 2009.
- [4] J.-P. Bouchaud, M. Potters, and L. Laloux. Financial applications of random matrix theory: Old laces and new pieces. *arXiv*, 2005.
- [5] B. Bruder and T. Roncalli. Managing risk exposures using the risk budgeting approach. *Working Paper*, 2012.
- [6] J. Bun, J.-P. Bouchaud, and M. Potters. Cleaning correlation matrices. *Risk*, Apr. 2016.
- [7] Y. Choueifaty and Y. Coignard. Toward maximum diversification. *The Journal of Portfolio Management*, 35(1):50–51, 2008.
- [8] N. Firoozye, V. Tan, and S. Zohren. Canonical portfolios: Optimal asset and signal combination. *arXiv*, 2023.
- [9] G. H. Golub and C. F. V. Loan. *Matrix Computations - 4th Edition*. Johns Hopkins University Press, 2013.
- [10] D. S. Grebenkov and J. Serror. Following a trend with an exponential moving average: Analytical results for a gaussian model. *Physica A: Statistical Mechanics and its Applications*, 394:288–303, Jan 2014.
- [11] D. S. Grebenkov and J. Serror. Optimal allocation of trend following strategies. *Physica A: Statistical Mechanics and its Applications*, 433:107–125, Sep 2015.
- [12] N. Higham. Matrix nearness problems and applications. *Oxford University Press*, pages 1–27, 1989.
- [13] B. T. Kelly, S. Malamud, and L. H. Pedersen. Principal portfolios. *Swiss Finance Institute Research Paper*, pages 20–67, 2020.
- [14] R. S. J. Koijen, T. J. Moskowitz, L. Pedersen, and E. B. Vrugt. Carry. *Fama-Miller Working Paper*, 11 2016.
- [15] O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.
- [16] H. M. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, Mar. 1952.
- [17] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, revised edition, 1991.
- [18] A. Meucci. *Risk and Asset Allocation*. Springer, 2007.
- [19] A. Meucci, A. Santangelo, and R. Deguest. Risk budgeting and diversification based on optimized uncorrelated factors. *arxiv*, 2015.
- [20] G. A. Paleologo. *The Elements of Quantitative Investing*. Wiley, 2025.
- [21] P. H. Schonemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 1966.
- [22] F. Segonne. *Quant Basis 101*. Amazon, 2024.
- [23] G. V. Stevens. On the inverse of the covariance matrix in portfolio analysis. *The Journal of Finance*, 53(5):1821–1827, 1998.