Quantitative Methods for Systematic Portfolio Construction: *Applications of Spectral Theory, State-Space Models, and Adaptive Signal Processing*

Francisco A. Ibanez

A thesis presented for the degree of
Doctor of Philosophy

Faculty of Finance
Bayes Business School
City St George's, University of London
December 11, 2025

DECLARATION

I hereby declare that this thesis is my original work and has not been submitted in any form for the award of any degree or diploma at any other university or institution. All sources used in this thesis have been properly acknowledged and cited.

[redacted]

Signature: ________________

Francisco A. Ibanez

London, December 11, 2025

ACKNOWLEDGEMENTS

The last four years have been truly transformative. They marked not only the culmination of my academic journey but also the beginning of a new chapter in my personal life as we started our growing family. Navigating these two profound experiences simultaneously has defined this period, making this achievement all the more meaningful to me.

Academically, I would like to express my sincere gratitude to my supervisors, Aneel Keswani and Giovanni Urga, for their guidance and support throughout this doctoral journey. I owe a special debt of gratitude to Giovanni, whose insistence on methodological rigor pushed me to refine my research and achieve a higher standard of scholarship. His constructive criticism and challenging questions taught me to approach problems with greater precision; invaluable lessons that I will carry forward in my career. I am also grateful to my examiners, Richard Payne and Ian Tonks, for their detailed review of this thesis. I genuinely enjoyed the challenging and stimulating discussion during the viva, which provided valuable insights for my future work. I also thank Professor Sotiris Tsolacos for efficiently chairing the defense.

Finally, on a personal note, I could not have undertaken this academic endeavor without my wife by my side, taking the lead in managing our home and growing family. Her unwavering support and patience throughout this entire process have been the foundation upon which this work was built.

DEDICATION

To my loving and beautiful wife, who encouraged me to embark on this journey and supported me every step of the way.

To my family, who inspire me to continuously strive for excellence.

To my mother, to whom I owe everything I am.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The field of investment management has experienced a profound transformation over the last half-century. The practice has evolved from one dominated by qualitative judgment and individual intuition to a discipline increasingly defined by systematic processes, quantitative analysis, and scientific rigor. This paradigm shift has been propelled by the exponential growth in computational power and the proliferation of financial data, giving rise to the domain of mathematical and quantitative finance. The central premise of this domain is that the application of sophisticated statistical models can be used to identify and systematically harvest market inefficiencies and risk premia.

However, this evolution has introduced a new set of formidable challenges. The sheer volume and complexity of modern financial data present significant obstacles, including the prevalence of noise, the curse of dimensionality, and non-stationarity in market dynamics. Researchers and practitioners must also contend with the ever-present danger of data mining and overfitting statistical models. In this complex landscape, classical portfolio theories, while foundational, often prove to be insufficient. Specifically, traditional diversification strategies can falter when faced with the low-rank nature of financial data, where the returns of thousands of securities are often driven by a small number of underlying common factors. Furthermore, static, single-state models of asset returns fail to capture the abrupt and persistent shifts in market behavior, known as "regimes," which are empirically observed throughout economic cycles. The relentless academic and industrial search for predictive signals has also created a *factor zoo*, a landscape of hundreds of potential return predictors. This proliferation makes the task of combining signals into a single, robust investment score a non-trivial challenge, fraught with the risks of signal redundancy and cancellation.

This thesis argues that overcoming these modern challenges requires a departure from conventional approaches that analyze securities and signals in isolation. It posits that more resilient, efficient, and scalable investment frameworks can be developed by focusing on the identification, modeling, and exploitation of the market's latent structures. These underlying structures, whether they manifest as the implicit risk factors driving a covariance matrix, the unobservable macroeconomic regimes dictating return distributions, or the fundamental predictive waveforms hidden within noisy data, offer a more parsimonious and robust foundation for building investment portfolios. To this end, this thesis presents three distinct essays that leverage innovative methodologies from fields such as unsupervised machine learning, probabilistic modeling, and electrical engineering to address these critical limitations in modern portfolio construction.

1.2 A brief history of quantitative portfolio management

The field of quantitative portfolio management has undergone a profound transformation over the past seventy years, evolving from the foundational principles of modern portfolio theory into a sophisticated discipline that leverages advanced statistical modeling, optimization techniques, and, more recently, machine learning. This evolution has been driven by a continuous interplay between academic theory and practical application, spurred by the increasing availability of data, exponential growth in computational power, and a deeper understanding of market dynamics. This review traces the chronological development of portfolio construction methodologies, examines the latest advances over the past few decades, and situates the three essays of this thesis within the context of specific, pressing challenges in the literature.

The genesis of quantitative portfolio management can be traced to the seminal work of Markowitz (1952), which introduced the mean-variance optimization (MVO) framework. Markowitz's pivotal contribution was to formalize the trade-off between risk and return, demonstrating that portfolio risk, defined as variance, depends not only on the risk of individual assets but critically on their covariance. This established the mathematical foundation for diversification, showing that investors could construct an *efficient frontier* of portfolios that offer the maximum expected return for a given level of risk. The development of MVO is widely regarded as the catalyst for modern financial economics and remains a cornerstone of portfolio management to this day.

Building on Markowitz's framework, the 1960s saw the development of the Capital Asset Pricing Model (CAPM) by Sharpe (1964) and Lintner (1965), which provided a theoretical model for pricing assets based on their systematic risk, or beta (β). However, the foundational assumptions of MVO and CAPM, that investor utility is a quadratic function of returns and that asset returns are normally distributed—were soon met with skepticism. Early studies began to question the efficacy of variance as a comprehensive risk measure, as it penalizes upside volatility equally with downside risk, which is inconsistent with investor psychology. This led to the exploration of alternative risk measures and utility functions throughout the 1970s and 1980s.

The theory of stochastic dominance was introduced to provide a more general framework for comparing uncertain prospects without relying on specific utility functions. Concurrently, researchers developed downside risk measures that focused specifically on the part of the return distribution that investors are most concerned with. Fishburn (1977) laid the groundwork for downside risk, leading to the development of measures like semi-variance (Markowitz et al., 1993) and lower partial moments. The introduction of Value-at-Risk (VaR) provided a simple, intuitive measure of potential loss, though its statistical properties were found to be undesirable for optimization (Jorion, 1997). A significant breakthrough came with the formalization of Conditional Value-at-Risk (CVaR) by Rockafellar and Uryasev (2002), which measures the expected loss beyond the VaR threshold. Crucially, CVaR was shown to be a coherent risk measure (Artzner et al., 1999) and could be optimized using efficient linear programming techniques, making it highly practical for portfolio construction. Other measures, such as the mean-absolute deviation (MAD) model (Konno and Yamazaki, 1991), were also developed to offer computationally simpler, linear-programming-based alternatives to the quadratic MVO problem.

While foundational theories provided the essential grammar for portfolio management, their practical implementation revealed significant challenges. The most critical issue with MVO is its *error-maximization* property; the optimizer's tendency to be hypersensitive to input parameters, where small errors in estimating expected returns and covariances lead to extreme and unstable portfolio weights (Michaud, 1989; Chopra and Ziemba, 1993). The last two to three decades of research have been largely defined by the search for solutions to this fundamental problem, leading to several major streams of innovation.

Robust and Bayesian Methods: One major stream has focused on improving the quality of the inputs. Bayesian techniques offer a natural framework for this, allowing investors to blend historical data with prior beliefs or theoretical models. The Bayes-Stein shrinkage estimator, first applied to portfolios by Jorion (1986), shrinks unstable sample means toward a more stable common mean, such as the mean of the global minimum-variance portfolio. The Black-Litterman model (Black and Litterman, 1992) became a widely adopted industry standard by providing a disciplined framework for combining an investor's subjective views with market equilibrium returns, resulting in more intuitive and diversified portfolios. More formally, robust optimization directly tackles parameter uncertainty by optimizing for a worst-case scenario within a defined uncertainty set for the input parameters, producing portfolios that are less sensitive to estimation errors (Xidonas et al., 2020). A parallel approach involves regularization, where a penalty term (e.g., the L_1-*norm*) is added to the optimization objective to enforce sparsity and constrain portfolio weights, leading to lower turnover and more stable out-of-sample performance (Brodie et al., 2009; DeMiguel et al., 2009).

Risk-Based Asset Allocation: A second major innovation has been the shift away from reliance on notoriously difficult-to-estimate expected returns. Risk-based allocation methods construct portfolios based solely on risk parameters. The most prominent of these is the Risk Parity (RP) or Equal Risk Contribution (ERC) portfolio, which weights assets such that each contributes equally to the total portfolio risk (Maillard et al., 2010). Gaining significant traction after the 2008 Global Financial Crisis for its perceived resilience, RP seeks to achieve "true diversification" in risk terms, rather than capital terms. This approach has been extended to factor risk parity, where risk is balanced across underlying risk factors rather than asset classes (Bhansali et al., 2012).

Machine Learning and Artificial Intelligence: The most recent paradigm shift has been driven by the advent of machine learning (ML) and artificial intelligence (AI). ML offers a powerful toolkit to address the limitations of traditional models by capturing complex, non-linear relationships in vast datasets. The application of ML in portfolio optimization can be broadly categorized into two stages, mirroring the classic Markowitz framework (Lee et al., 2023). First, in the parameter estimation stage, supervised learning models like neural networks and gradient-boosted trees are used to forecast returns and covariance matrices with greater accuracy than traditional econometric models (Gu et al., 2020). Unsupervised learning techniques like hierarchical clustering are used to identify data-driven asset classes, providing more robust inputs for asset allocation (Raffinot, 2017). Second, in the optimization stage, ML helps solve complex, large-scale, or non-convex problems that are intractable for traditional solvers, such as portfolios with cardinality constraints.

Furthermore, advanced ML techniques are beginning to merge the prediction and optimization stages into a unified framework. Decision-focused learning trains predictive models not to minimize prediction error (e.g., Mean Squared Error) but to directly minimize the downstream decision loss (or "regret") of the final portfolio (Butler and Kwon, 2022). End-to-end models, often using deep learning architectures, bypass parameter estimation entirely, learning a direct mapping from raw market data to optimal portfolio weights (Zhang et al., 2020). Reinforcement learning takes this a step further, training an agent to learn an optimal dynamic trading strategy through trial-and-error interaction with a market environment, making it naturally suited for multi-period optimization (Jiang et al., 2017). While still an emerging area, these methods hold the potential to revolutionize portfolio construction by creating truly adaptive and data-driven strategies.

1.3 Research gap

While the evolution to quantitative finance has been transformative, it is not without its challenges. The sheer volume and complexity of financial data, coupled with issues of noise, non-stationarity, and the risk of overfitting, create a landscape where classical portfolio theories often prove insufficient. This thesis identifies three fundamental problems in modern portfolio construction where conventional approaches falter, revealing significant research gaps that motivate the subsequent essays.

The Challenge of True Diversification. A central problem is that traditional diversification strategies may fail when confronted with the low-rank nature of financial data. These methodologies typically work at the security level (e.g., equal-weighting or risk parity) and often ignore the powerful underlying common factors that drive portfolio returns and risk. As a result, allocating capital homogeneously across many securities can be an ineffective diversification exercise if a single factor dominates their returns. While recent studies have explored using matrix factorizations to diversify across these implicit factors, a research gap exists between these theoretical "thought experiments" and practical, implementable methodologies. Few studies have addressed the outstanding challenges of tradability and stability, as these methods often produce portfolios with extreme long and short positions and suffer from instability that leads to excess trading.

The Challenge of Market Non-Stationarity. A second major limitation of existing models is their static nature. Static, single-state models of asset returns fail to capture the abrupt and persistent shifts in market behavior, or *regimes*, that are empirically observed during economic cycles.

Although the existence of market regimes is well-documented, a research gap persists in their practical application to large-scale portfolios. Applying regime-switching models to the portfolios of real-world asset managers, which can contain hundreds or thousands of stocks, has remained a significant computational and statistical hurdle. This highlights the need for a scalable solution that can effectively incorporate time-varying market dynamics into the construction of large portfolios.

The Challenge of Signal Aggregation. The relentless academic and industrial search for predictive signals has produced a "factor zoo" of hundreds of potential return predictors. This proliferation presents a non-trivial challenge in combining them into a single, robust investment score, a task fraught with the risk of signal redundancy and cancellation. The common industry practice of simple averaging is demonstrably suboptimal, as many of these signals are both noisy and correlated. This creates a research gap for a more sophisticated and robust framework for signal aggregation that can systematically account for the complex and time-varying interrelationships between factors to maximize a portfolio's overall signal-to-noise ratio.

Collectively, these problems demonstrate a need to depart from conventional approaches that analyze securities and signals in isolation. The research gaps identified in diversification, regime-awareness, and signal combination all point toward the necessity of developing frameworks that can model and exploit the market's underlying latent structures.

1.4 Significance of the research

This thesis seeks to advance the theory and practice of quantitative portfolio management by developing and validating three novel frameworks designed to address fundamental limitations in conventional investment models. The research provides a significant contribution to the literature by offering novel, robust, and practical solutions to core problems in portfolio management. Collectively, the work underscores a unified theme: that the path to superior investment performance lies in a deeper, more structural understanding of the forces that shape financial markets, enabled by the thoughtful application of advanced quantitative techniques. By focusing on modeling the underlying latent structures that govern market dynamics, this research demonstrates that it is possible to build portfolios that are more intelligently diversified, more adaptive to changing conditions, and more resilient to market turmoil.

The specific contributions are threefold. First, the thesis addresses the fundamental challenge of achieving true portfolio diversification by proposing a novel framework, the Diversified Spectral Portfolio (DSP). This approach moves beyond traditional heuristics to diversify directly across implicit, data-driven factors, yielding superior risk-adjusted performance, particularly in the highly correlated market environments where traditional diversification is most needed. Second, it tackles the problem of time-varying market dynamics by introducing a scalable solution that leverages the low-rank factor structure of equity returns. This allows for the construction of large-scale, regime-aware portfolios that can dynamically adapt their posture to changing market conditions, delivering improved returns and more effective diversification compared to static, regime-agnostic baselines. Third, it confronts the challenge of signal aggregation in the era of the "factor zoo" by pioneering an interdisciplinary approach that draws a powerful analogy between combining investment signals and techniques used in electrical engineering. By applying adaptive beamforming, this framework optimally combines signals to maximize the portfolio's signal-to-noise ratio, dramatically outperforming traditional methods while providing superior resilience and shallower drawdowns during periods of market stress.

A key contribution of this research is its demonstration of the immense value of interdisciplinary innovation. The solutions proposed are inspired by advances in unsupervised machine learning, probabilistic modeling, and signal processing, showing that sophisticated toolkits from other quantitative fields can be adapted to yield powerful financial applications. The practical implications of this research are therefore significant, offering quantitative asset managers and hedge funds concrete methodologies for improving diversification, adapting to market regimes, and building more potent multi-factor models.

1.5 Thesis outline

This thesis builds upon the recent advances described in the previous section by arguing that more resilient and efficient investment frameworks can be developed by focusing on the identification, modeling, and exploitation of the market's latent structures. Conventional approaches often analyze securities and signals in isolation, proving insufficient in a landscape characterized by powerful common factors, non-stationary dynamics, and a proliferation of noisy signals. These underlying structures, whether they manifest as the implicit risk factors driving a covariance matrix, the unobservable macroeconomic regimes dictating return distributions, or the fundamental predictive waveforms hidden within noisy characteristic data, offer a more parsimonious and robust foundation upon which to build investment portfolios. To this end, this thesis presents three distinct essays that leverage innovative methodologies, drawing inspiration from fields such as unsupervised machine learning, probabilistic modeling, and electrical engineering, to address critical limitations in modern portfolio construction.

The first essay, "Diversified Spectral Portfolios: An Unsupervised Learning Approach to Diversification," addresses the fundamental challenge of achieving true portfolio diversification. It moves beyond traditional security-level heuristics like equal-weighting and risk parity, which often ignore the powerful common factors that drive portfolio risk. The essay proposes a novel framework, the Diversified Spectral Portfolio (DSP), that uses the Singular Value Decomposition (SVD) to peer into a portfolio's "eigenspace" and diversify directly across the implicit, data-driven factors. To solve the practical issues of tradability and stability that plague such methods, it introduces an unsupervised clustering step that groups securities based on their factor exposures, reducing dimensionality and mitigating the risk of extreme, unstable allocations. Through extensive Monte Carlo simulations, the essay demonstrates that this structurally aware approach to diversification yields superior risk-adjusted performance, particularly in the highly correlated market environments where traditional diversification is most needed.

The second essay, "Incorporating Market Regimes into Large-Scale Stock Portfolios: A Hidden Markov Model Approach," tackles the problem of time-varying market dynamics. While the existence of distinct market regimes (e.g., bull and bear markets) is well-documented, applying regime-switching models to the large-scale portfolios of real-world asset managers has remained a significant computational and statistical hurdle. This essay introduces a scalable solution that leverages the low-rank factor structure of equity returns. Instead of modeling thousands of individual stocks, a Hidden Markov Model (HMM) is fitted to a parsimonious set of common risk factors to identify the prevailing market state. It then develops a novel "Regime-Weighted Least Squares" (RWLS) methodology to estimate regime-dependent factor loadings for every stock in the universe. This allows for the construction of large-scale, regime-aware portfolios that can dynamically adapt their posture to changing market conditions. The empirical results, based on simulations calibrated to historical U.S. market data, show that this regime-aware framework delivers improved returns and more effective diversification compared to a static, regime-agnostic baseline.

The third essay, "Optimal Investment Signal Combination in Systematic Equity Portfolios: A Beamforming Approach," confronts the challenge of signal aggregation in the era of the "factor zoo." With hundreds of potential factors, many of which are noisy and correlated, the common industry practice of simple averaging is demonstrably suboptimal. This essay pioneers an interdisciplinary approach, drawing a powerful analogy between combining investment signals and the techniques used in electrical engineering to process signals from an antenna array. It re-frames investment signals as noisy waveforms and proposes the use of adaptive beamforming, a sophisticated signal processing methodology, to optimally combine them. This framework dynamically filters, aligns, and weights each signal to maximize the overall portfolio's signal-to-noise ratio, systematically accounting for their complex and time-varying interrelationships. A rigorous historical backtest shows that this engineering-inspired approach dramatically outperforms a traditional baseline, delivering substantially higher returns while providing superior resilience and shallower drawdowns during periods of market stress.

Collectively, these three essays contribute to the literature by providing novel, robust, and practical solutions to fundamental problems in portfolio management. They underscore a unified theme: that the path to superior investment performance lies in a deeper, more structural understanding of the forces that shape financial markets, enabled by the thoughtful application of advanced quantitative techniques.

Chapter 2

Diversified Spectral Portfolios: *An Unsupervised Learning Approach to Diversification*[1]

2.1 Introduction

Traditional portfolio diversification methodologies focus on either having a balanced weight profile (aiming at an equal-weighted portfolio) or achieving a balanced risk contribution profile (aiming at the risk-parity portfolio), working at the security level and ignoring the interactions of the underlying factors driving the returns and covariance of the investable universe. To illustrate this, the variability of returns of different securities could be dominated by a single shared underlying factor, making the exercise of allocating homogeneously across them very ineffective from the diversification point of view. This becomes even more pressing when a portfolio manager has the daunting task of allocating across numerous securities which, given the low-rank

[1]A version of this chapter has been published as: Ibanez, F. A. (2023). Diversified Spectral Portfolios: An Unsupervised Learning Approach to Diversification. *The Journal of Financial Data Science*, 5(2), 67–83. https://doi.org/10.3905/jfds.2023.1.118.

nature of financial data, will be mostly driven by very few shared factors.

When it comes to factors in finance, Martellini and Milhau (2017) identify three commonly adopted definitions. First, factors can be thought of as profitable systematic strategies that earn a premium in the long run and are backed by recognized anomalies (e.g., size, value, and momentum, in equities). Alternatively, factors can refer to time-varying state variables that can determine the conditional expected return, volatility, and correlations of asset classes (e.g., dividend yield, forward rates). Finally, factors can point to common sources of risk that may or may not be rewarded, which can be explicit in case their values are observable or implicit in case they have to be extracted from returns. Something the first two definitions have in common is that they require the user to define, ex ante, a parsimonious set of economically and statistically relevant factors. However, there is little agreement on either what these factors are or how many they are, when we analyze datasets other than the well-studied US equities case. The previous has sparked recent interest in the third definition, giving birth to a series of studies that use matrix factorizations to identify the underlying implicit factors and diversify across them. However, few studies have explored the greatest outstanding challenges of relying solely on these decompositions: the actual tradability of the resulting portfolios due to their extreme positioning (i.e., extreme long and short positions). Additionally, the amount of noise that typically surrounds financial data makes its estimation somewhat unstable in the presence of outliers, which translates into different optimal portfolios in consecutive rebalances, resulting in excess trading. These challenges have created a gap between what has so far been a thought experiment and methodologies that can be implemented while managing real portfolios.

This study aims to fill this gap by presenting a diversification-focused portfolio construction methodology that takes advantage of the *singular value decomposition* (SVD) to identify implicit factors and addresses the challenges surrounding its stability and tradability. This chapter is organized as follows. First, we review the existing literature upon which we build our argument. Then, we introduce the diversification framework; how it can be used to understand and diversify the risk taken by a portfolio and the challenges it has to overcome before being widely implemented. After that, we describe the features and mathematical formulation of our portfolio modeling approach. Next, we present the experiment conducted to test the methodology and its results. Finally, we present our conclusions.

2.2 Related literature

The idea of using matrix decomposition to identify the implicit factors driving portfolio risk and use these factors to build diversified portfolios is relatively recent. The first attempt is documented by Partovi and Caputo (2004), who used the eigendecomposition to translate the portfolio selection problem from one that chooses correlated assets into one that chooses uncorrelated linear combinations of them. These combinations are achieved through the change of basis that rearranges the covariance matrix into its eigenvectors, which are orthogonal to each other by construction. Given that these vectors are linear combinations of investable securities, the authors label these implicit factors as *principal portfolio*, and argue that they are natural instruments for portfolio analysis when short sales are allowed, because they are free from correlations. Attilio Meucci (2009) built upon this idea and uses the eigenvectors and eigenvalues of the covariance matrix of assets within a portfolio to project the capital weights into the principal portfolio weights, which can then be used to attribute the risk of the portfolio to each of these principal portfolios. Furthermore, the author suggested measuring the degree of portfolio risk concentration, by calculating the exponential of Shannon's entropy in these implicit factor risk attributions, which he called *effective number of bets* (ENB); the higher this metric, the more diversified a portfolio is.

The use of entropy-related measures when measuring portfolio concentration is not novel. Rudin and Morgan (2006) proposed to calculate the center of mass of the eigenvalues of the correlation matrix of the investmentable set to find a quantitative way to measure the degree of diversification in an investment portfolio. Similarly, Bera and Park (2008) suggested using cross-entropy measures on the portfolio weights while optimizing, which would be equivalent to a shrinkage towards the equal-weighted portfolio. Lohre et al. (2014) formalized this framework and called it *diversified risk parity* (DRP). Their work made two major contributions. First, the authors noted that the DRP strategy is an inverse volatility strategy along the principal portfolios that can be computed analytically. Second, they explain that maximizing the ENB measure does not allow for a unique solution in the absence of constraints. This diversification framework has been operationalized and expanded across different data sets in Lohre et al. (2012) (US stocks), Bernardi et al. (2018) (commodities), and Dichtl et al. (2020) (multi-asset), concluding that it is capable of providing better risk-adjusted return and diversification than other constructions, such as equal-weighted and risk parity.

However, using eigenvectors as implicit factors and diversifying across them comes with caveats. The most prevalent is the issue of multiple solution candidates that maximize the ENB measure. A workaround was suggested by Deguest et al. (2022), who arbitrarily picked two candidates; the solution that provided the highest Sharpe ratio and the one that provided the minimum variance. Another approach to overcome this issue is to simply relegate the ENB from the objective function to a constraint, as demonstrated in Martellini and Milhau (2017), who empirically found that imposing a minimum ENB as a constraint while optimizing is an effective way to improve the diversification of implicit factor exposure. Another caveat of this framework is the lack of interpretability of the factors. Meucci et al. (2015) state that working with the principal component is suboptimal, as they are purely statistical entities that are not related to the investment process. The authors try to remedy this by introducing the *minimum-torsion bets*, which are the uncorrelated basis of the factors closest to the factors the portfolio manager is familiar with, significantly improving the interpretability of the results obtained from the analysis. With similar intent, Kamauchi and Yokouchi (2021) explores the use of Gram-Schmidt orthonormalization to extract uncorrelated and understandable risk sources. In an empirical exercise using equity and bond indices from 2000 to 2016, the authors showed that their results are easier to interpret than other competing methodologies while showing a similar volatility profile. Finally, another challenge of the framework is the inability of the eigendecomposition to account for higher distribution moments, which are relevant in the case of security return data. Lassance et al. (2022) attempts addressing this by using the independent component analysis (ICA) instead, which is the rotation of the principal components that are maximally independent.

2.3 Diversification in eigenspace

2.3.1 Singular value decomposition

An alternative to theoretically defining a set of factors on an ex ante basis would be to infer the factor structure from the securities themselves. We can resort to one of the many matrix decompositions in numerical linear algebra to do so. Within these factorizations, Stewart (1993) argues that the singular value decomposition (SVD) assumes a special role, as it is an ideal vehicle for discussion of the geometry (in our case, the factor structure) of the n-space, it is stable to *small* perturbations, and it can provide an optimal low-rank approximation to the original data set. The decomposition is particularly useful in finance, as it can be applied to non-squared matrices, such as data sets of $t \times n$ security returns. Zhang (2015) argues that the SVD has become particularly relevant due to recent developments in machine learning, data mining, and theoretical data science, making matrices a language of data science. For these reasons, we propose using the SVD as a completely data-driven way of identifying the relevant implicit factors of a given investable universe.

Let us assume that we are interested in analyzing the implicit factor structure of $\boldsymbol{Z} \in \mathbb{R}^{txn}$, which contains the aligned t historical returns of the n assets that are part of the investable universe (where $t >> n$). Aligning the data[2] is of the utmost importance for SVD. Brunton and Kutz (2019) explain that the SVD rank explodes when objects in columns translate (nonzero historical return), rotate, or scale (variance different from 1). The economy-sized singular value decomposition of $\boldsymbol{Z}$ is then given by

$$\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T \tag{2.1}$$

where $\boldsymbol{U}$ is a tall-and-skinny $t \times n$ matrix that contains the eigenvectors of $\boldsymbol{Z}\boldsymbol{Z}^T$, $\boldsymbol{S}$ is a diagonal $n \times n$ matrix with the singular values of $\boldsymbol{Z}$ in diagonal terms arranged in descending order, and $\boldsymbol{V}$ is $n \times n$ matrix with the eigenvectors of $\boldsymbol{Z}^T\boldsymbol{Z}$. The information related to the cross-sectional implicit factors of our returns data set is described along the columns of $\boldsymbol{V}$. The eigenvector of $\boldsymbol{V}$ associated with the first (largest) value in $\boldsymbol{S}$ represents the most prevalent factor dynamic in the data set, continuing in decreasing order, until the variability of the set is completely explained.

[2]Aligning the data involves centering (i.e., subtracting its mean) and scaling (i.e., dividing by its standard deviation)

2.3.2 Spectral risk contribution

The variance of a portfolio, when expressed in terms of its correlation matrix ($\boldsymbol{R}$) and capital weights (w), is given by

$$\sigma_p^2 = \boldsymbol{w}^T \mathcal{D} \boldsymbol{R} \mathcal{D} \boldsymbol{w} \tag{2.2}$$

where

$$\mathcal{D} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix}_{n \times n}$$

and σ_i is the volatility of security i, before aligning. Given that the $\boldsymbol{Z}$ matrix is aligned, $\boldsymbol{Z}^T \boldsymbol{Z}$ is proportional to $\boldsymbol{R}$, so can rewrite Equation (2.2) as

$$\sigma_p^2 = \boldsymbol{w}^T \mathcal{D} \mathbf{Z}^T \mathbf{Z} \mathcal{D} \boldsymbol{w} \, (t-1)^{-1} \tag{2.3}$$

replacing (2.1) in (2.3) we get

$$\sigma_p^2 \propto \boldsymbol{w}^T \mathcal{D} \boldsymbol{V} \boldsymbol{S}^T \boldsymbol{S} \boldsymbol{V}^T \mathcal{D} \boldsymbol{w} \tag{2.4}$$

The importance of expression (2.4) lies in defining portfolio risk as a linear combination of eigenvectors and values, allowing us to move from capital weights and risk framework to *spectral weights* and *spectral risk* framework.

$$\boldsymbol{w}^T \mathcal{D} \boldsymbol{V} \underbrace{\boldsymbol{S}^T \boldsymbol{S}}_{\text{Spectral risk}} \overbrace{\boldsymbol{V}^T \mathcal{D} \boldsymbol{w}}^{\text{Spectral weights}}$$

We can manipulate (2.4) to construct the vector $\boldsymbol{\varphi} \in \mathbb{R}^{n \times 1}$, which contains the percentage of spectral risk that can be mapped to each implicit factor,

$$\boldsymbol{\varphi} := diag\left(\boldsymbol{V}^T \mathcal{D} \boldsymbol{w}\right) \boldsymbol{S}^T \boldsymbol{S} \boldsymbol{V}^T \mathcal{D} \boldsymbol{w} \left(\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}(n-1)\right)^{-1} \qquad (2.5)$$

Under this framework, we expect that the more uniformly distributed the contribution to spectral risk $\boldsymbol{\varphi}$, the more diversified the portfolio.

As discussed in the literature review, different attempts have been made to find a sufficient statistic that measures the degree of homegenity in a vector like $\boldsymbol{\varphi}$, and maximize it to find the most diversified portfolio. This concept is better illustrated in Exhibit 2.1. In the left panel, we can observe the capital allocation ($\boldsymbol{w}$) of two different portfolios; an equal-weighted portfolio (EW) that distributes capital in equal amounts across 13 ETFs tracking different US equity sectors, and a second portfolio with heterogeneous long and short positions, which we call Maximum Entropy (ME) in this example. In the right panel, we can see the contribution to the spectral risk ($\boldsymbol{\varphi}$) coming from each of these portfolios, which is the projection of the capital allocation of the left panel into the implicit factor space. Although the EW portfolio appears to be very well diversified in capital terms, its spectral risk contribution comes almost entirely from the first eigenvector, which represents the most dominant risk in the analyzed set. This result is sensible, as all 13 ETFs follow distinct sectors of the same market, making the risk of the equity market the most dominant risk in the sample. On the other hand, we have the very disperse capital allocation of the ME portfolio, which is able to provide a very homogeneous spectral risk profile. There is no doubt that, under this framework, the ME offers much better diversification than the EW portfolio.

2.3.3 Challenges of the framework

The implementation of this diversification framework in a real-life portfolio setting presents several challenges that, if not overcome, make this framework no more than a thought exercise.

To introduce the first challenge, let us look again at Exhibit 2.1. The portfolio resulting from equalizing the spectral risk contribution relies on being able to take very extreme positioning (i.e., large long and short positions in very few assets, with small weights everywhere else). This is problematic, as having the ability to freely take short positions is rather rare in the industry, with many pension funds, mutual funds, and retail investors partially or totally restricted from shorting securities. In addition, becoming a borrower within a securities lending program can be costly. Third, relying on taking short positions on assets with a positive expected return can be detrimental to long-term portfolio performance.

The second challenge to overcome is the non-unique set of portfolios that yield maximum diversification in this framework. This happens because the signs of the eigenvectors are arbitrary and their only relevant aspect is the direction in which they point. In other words, if $\boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{Z}^T\boldsymbol{Z}$, $-\boldsymbol{v}_i$ is also an eigenvector of $\boldsymbol{Z}^T\boldsymbol{Z}$ and is associated with the same eigen-

Figure 2.1: **Spectral risk contribution of equal-weighted (EW) and maximum entropy (ME) portfolios**

This figure illustrates the core concept of spectral risk diversification by comparing two distinct portfolios constructed from 13 U.S. equity sector ETFs. The left panel displays the capital allocation (portfolio weights) for an Equal-Weighted (EW) portfolio and a Maximum Entropy (ME) portfolio, which features heterogeneous long and short positions. The right panel shows the corresponding contribution to portfolio risk from each underlying implicit factor (eigenvector). The figure demonstrates that while the EW portfolio appears diversified in terms of capital allocation, its risk is almost entirely concentrated in the first and most dominant eigenvector. In contrast, the ME portfolio, despite its dispersed capital weights, achieves a highly uniform distribution of risk across all eigenvectors, representing superior diversification under this framework.

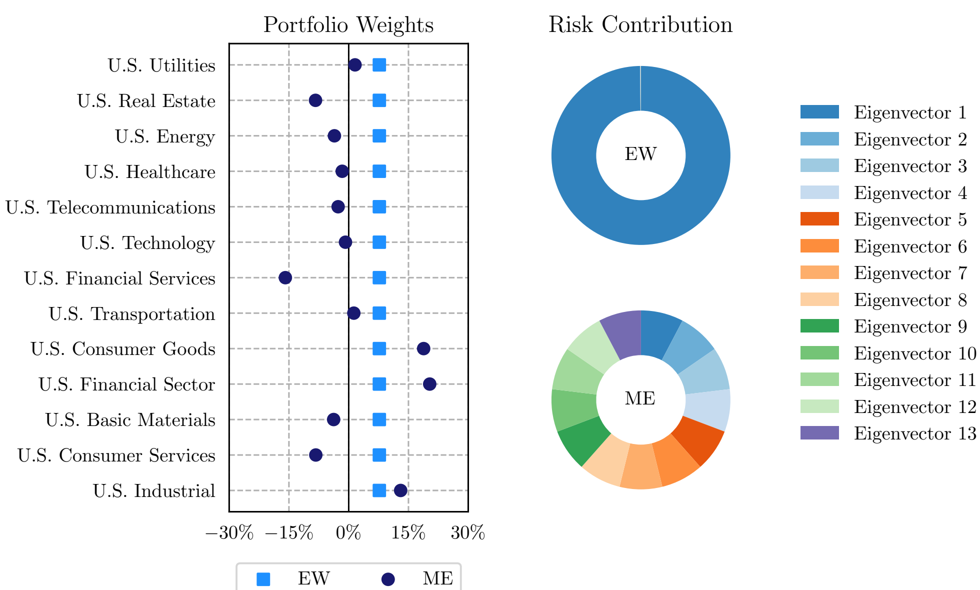
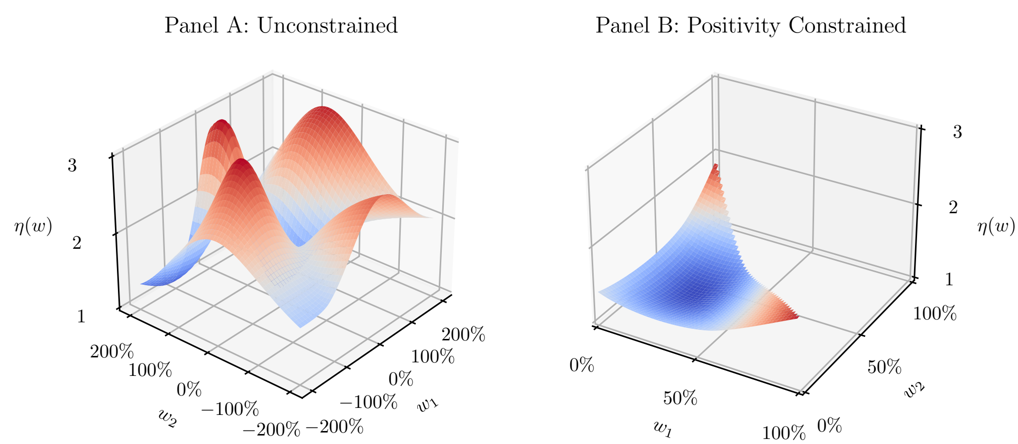


Figure 2.2: **Maximimum diversification with and without constraints**

This figure visualizes the solution space for maximizing the Effective Number of Bets (ENB) in a three-asset portfolio, illustrating the challenge of non-unique solutions. Panel A (Unconstrained) shows the ENB surface when long and short positions are permitted. The surface exhibits multiple peaks, indicating that several different portfolio weight combinations can achieve the maximum level of diversification, consistent with the findings of Lohre et al. (2014). Panel B (Positivity Constrained) shows the ENB surface when short selling is prohibited. The imposition of a positivity constraint reduces the number of local optima and leads to an overall lower maximum achievable ENB value, highlighting the trade-offs introduced by real-world investment constraints.



value. The resulting set of portfolios can be dramatically different in terms of capital allocations and risk-return characteristics. This is better illustrated in Panel A of Exhibit 2.2, which shows the resulting ENB coming from different combinations of three ETFs[3], allowing for long and short positions, and assuming a fully invested portfolio[4]. The first observation that we can make is the multiple optimal combination of weights that maximize the objective function, which is consistent with the observation of Lohre et al. (2014). We can also note that when a positivity constraint is in place, the number of local optima is reduced (in this case, to two), which is also highlighted by Deguest et al. (2022), and these optimal combinations of assets yield a much lower ENB value. To address this multiple-optimal issue, the authors opted for picking only two out of the 2^{n-1} candidate solutions: the one that produces the maximum Sharpe ratio within the set and the one that achieves the minimum variance. The authors also explore the alternative of setting another objective function, such as portfolio variance minimization, and setting a floor for the ENB measure as one of the optimization constraints. We believe that both approaches rely greatly on heuristics and arbitrary decisions, making the systematic implementation of the methodology difficult.

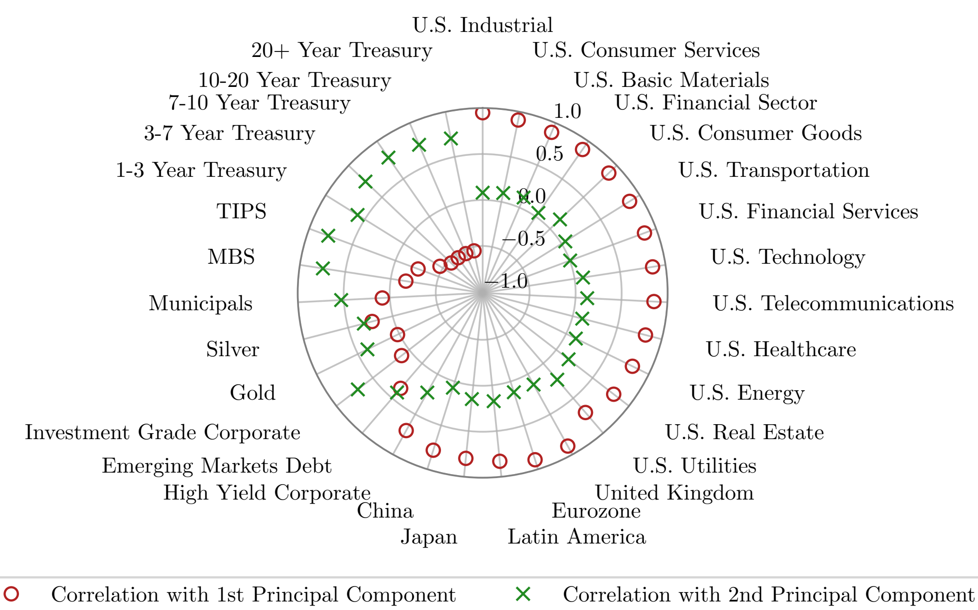
A third challenge comes in the form of collinearity. Given that it was already established that trying to achieve maximum diversification under this framework will possibly result in extreme portfolios and multiple solutions, one might want to incorporate constraints into the optimization problem, increasing its complexity. If this is the case, the optimizer will try to homogenize the contribution of spectral risk in (2.5) by loading more heavily on securities that have a high correlation with underweighted eigenvectors. If two (or more) securities are significantly correlated to the same eigenvector, the optimizer might see them as interchangeable. In practice, interchangeability could mean that, given small changes in the input data $\mathbf{Z}$, the op-

[3]U.S. Technology, U.S. Consumer Goods, U.S. Consumer Services

[4]$w_3 = 1 - w_1 - w_2$

Figure 2.3: **Correlation between assets and their first and second eigenvectors**

This polar plot illustrates the issue of collinearity by showing the correlation between 31 different ETFs and the first two implicit factors (eigenvectors) of the system. Each point on the plot represents an asset. Red circles indicate the correlation with the first principal component (eigenvector), while green 'x' markers show the correlation with the second. The plot reveals that multiple securities can be highly correlated with the same underlying factor. For example, "U.S. Industrial" and "U.S. Consumer Services" are almost perfectly correlated with the first eigenvector, while various Treasury ETFs are highly correlated with the second. This collinearity can make the securities interchangeable in an optimization, leading to high portfolio turnover.



timizer could allocate between these securities in different rebalances over time, significantly increasing the turnover of the portfolio and transaction costs associated with excess trading. The collinearity issue is illustrated in Exhibit 2.3, where the correlation between 31 ETFs and the first and second eigenvectors of this system is plotted. It can be easily observed that securities such as U.S. Industrial and U.S. Consumer Services are almost perfectly correlated to the first eigenvector (seemingly related to the equity risk premium), while the Treasury related ETFs are highly correlated to the second eigenvector (probably related to rates). We would expect the aforementioned securities to be the most susceptible to the interchangeability problem.

Unfortunately, the three challenges explored previously are not the only pitfalls of the framework. Other interesting discussion points include the sensitivity of the SVD results to outliers (common in financial data) and the noisiness and instability in the estimation of the eigenvectors associated with the smallest eigenvectors over time. Given that these points are related to the implementation of the SVD in financial data, rather than to the portfolio construction itself, they fall outside the scope of this study. Consequently, we will focus on addressing the three aforementioned challenges through the proposed portfolio construction methodology.

2.4 Building diversified spectral portfolios

2.4.1 Security clustering

In order to address the collinearity and extreme positioning issues, we propose bucketing the many securities of the investable universe into a few clusters, based on their correlation to the implicit factors. The expected benefits are twofold. First, it directly tackles the collinearity issue by collapsing securities that are susceptible to being interchanged into one, avoiding swapping allocations in consecutive rebalances, and reducing turnover. Second, it reduces the dimensionality of the problem, making the optimization problem more straightforward to solve in large investible universes (e.g., single-name equities).

To group investable securities in a manner that is consistent with the problem at hand, we need to estimate the correlation between these and their principal components (i.e., implicit factors). Given that $\boldsymbol{Z}$ is aligned (i.e., mean of 0 and standard deviation of 1), the correlation between each security and the standardized principal components $\boldsymbol{Z}\boldsymbol{V}\boldsymbol{S}^{-1}\sqrt{t-1}$ is given by

$$\frac{1}{t-1}\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{V}\boldsymbol{S}^{-1}\sqrt{t-1} \tag{2.6}$$

replacing (2.1) in (2.6) we get

$$\frac{1}{t-1}\boldsymbol{V}\boldsymbol{S}\boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T\boldsymbol{V}\boldsymbol{S}^{-1}\sqrt{t-1}$$

and given that $\boldsymbol{V}^T\boldsymbol{V} = \mathbf{I}$ and $\boldsymbol{U}^T\boldsymbol{U} = \mathbf{I}$ we finally obtain

$$\frac{1}{\sqrt{t-1}}\boldsymbol{V}\boldsymbol{S} \tag{2.7}$$

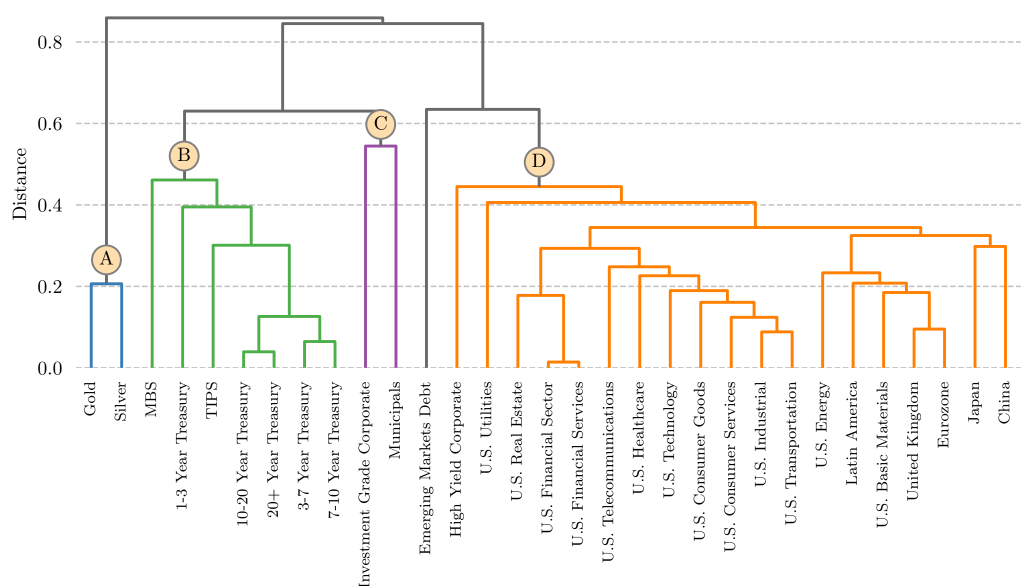
which is a $n \times n$ matrix that contains the correlation between each investable asset and the principal components. By measuring the cosine distance between each pair of row vectors in (2.6), we can identify securities with a similar correlation pattern with implicit factors and therefore susceptible to interchangeability during the optimization process. These cosine distances are stored in a distance matrix which will be used to group the securities based on them.

To perform the clustering, we propose hierarchical agglomerative clustering, which is a well-established technique in unsupervised machine learning. As Müllner (2011) explains, this technique starts by partitioning the data set into singleton nodes (one per security in the investable universe), and step by step merges the pair of the closest nodes into a new node, until there is only one final node left, which comprises all the data. The result of the algorithm is illustrated in Exhibit 2.4 as a graphical dendrogram. Using the distance matrix built on (2.7), the algorithm looks for the closest pair (i.e., the cosine distance between their respective vectors) of securities in the set, which in this example would be *U.S. Financial Sector* and *U.S. Financial Services*, and merges them into a new vector, which is then paired with its closest neighbor (in this case, *U.S. Real Estate*). The process will be repeated across the investible universe in a stepwise manner, moving up in the dendrogram until all the nodes are integrated and the tree is complete.

Although the methodology is able to hierarchically organize and relate the data set, the user has to determine which of the subtrees are actual clusters and which are just part of another cluster, as explained by Bar-Joseph et al. (2001). In practice, this is done by setting a distance threshold below which every connected (merged) node will be considered a cluster. For this reason, the final number of clusters to be found in the set is inversely related to the chosen distance threshold, and the most appropriate threshold will depend on the data set on which the cluster is performed. For example, in Exhibit 2.4, setting a distance threshold of 0.6 would give us four clusters; Precious metals, Treasury and asset-backed securities, corporate and municipal bonds, and global equities and high-yield bonds, along with one singleton ETF. Setting a higher distance threshold, such as 0.7, would result in the

Figure 2.4: **Agglomerative clustering of a sample investable universe**

This figure displays a dendrogram, the graphical output of a hierarchical agglomerative clustering algorithm applied to a universe of 31 ETFs. The algorithm groups securities based on the similarity of their correlation patterns with the underlying implicit factors, with the vertical axis representing the cosine distance between clusters. The dendrogram shows how individual assets (at the bottom) are progressively merged into larger clusters as the distance threshold increases. For instance, at a distance threshold of 0.6, the algorithm identifies four distinct clusters labeled A, B, C, and D, effectively bucketing assets with similar risk factor exposures to address the collinearity problem.



merging of clusters B, C, and D with the singleton; three clusters in total.

In the case of a given investable universe of size n, we can collapse it into k clusters, where $k << n$, using agglomerative clustering. Thus, we can then redesign the optimization problem, from one that chooses n security capital weights (w) to minimize the objective function, to one that chooses k cluster capital weights (θ). Therefore, we can express portfolio capital weights as the following linear transformation

$$\boldsymbol{w} = \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{C}}\boldsymbol{\theta} \tag{2.8}$$

where

$$\boldsymbol{\mathcal{W}} = \begin{bmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_n \end{bmatrix}_{n \times n}$$

is a $n \times n$ matrix that contains predetermined fixed allocations $(\omega_1, \omega_2, \dots, \omega_n)$ for each asset in the investable universe in the diagonal entries,

$$\boldsymbol{\mathcal{C}} = \begin{bmatrix} | & | & & | \\ \boldsymbol{c}_1 & \boldsymbol{c}_2 & \dots & \boldsymbol{c}_k \\ | & | & & | \end{bmatrix}_{n \times k}$$

is a $n \times k$ boolean matrix that indicates if the security i belongs to the cluster j with an entry of 1 and an entry of 0 otherwise, and finally $\boldsymbol{\theta}$ is a $k \times 1$ vector containing the cluster weights, which scales the fixed security allocations in $\boldsymbol{\mathcal{C}}$.

The linear transformation (2.8) effectively takes the weights of the given cluster in $\boldsymbol{\theta}$, linearly maps these values from the cluster space to the security space through $\boldsymbol{\mathcal{C}}$, and propagates these scaling factors through fixed security allocations $(\omega_1, \omega_2, \ldots, \omega_n)$. The question of how to define $\boldsymbol{\mathcal{W}}$ is discussed later in the study.

2.4.2 Cost function

In the spectral risk framework, diversifying a portfolio involves equalizing the vector of contribution to spectral risk $\boldsymbol{\varphi}$. We propose the following alternative way of diversifying (equalizing) the spectral risk contribution.

$$||\boldsymbol{S}\boldsymbol{V}^T\boldsymbol{\mathcal{D}}\boldsymbol{w} - \mathbf{1}_n||_2^2 \tag{2.9}$$

When not restricted, minimizing the expresion (2.9) has a well-known analytical solution. However, as Boyd and Vandenberghe (2004) noted, when linear inequality constraints (such as no short selling, in our case) are added, there is no longer a simple analytical solution.

We can rearrange the expression (2.9) into the quadratic program (QP) form and expand it to incorporate the additional model specifications. We can take advantage of (2.8) to reformulate this optimization problem, from one that chooses n security weights to one that chooses k cluster weights ($\boldsymbol{\theta}$). Finally, we can generalize for the case of different constraints imposed by a given investment policy, and we arrive at the following cost function that will yield our Diversified Spectral Portfolios (DSP):

$$\begin{aligned} \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \quad & \frac{1}{2}\boldsymbol{\theta}^T \boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{q}^T\boldsymbol{\theta} \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{\theta} \leq \boldsymbol{h} \\ & \mathbf{1}_n^T \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{C}}\boldsymbol{\theta} = 1 \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} \boldsymbol{P} &= \boldsymbol{\mathcal{C}}^T \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{D}}\boldsymbol{V}\boldsymbol{S}^T\boldsymbol{S}\boldsymbol{V}^T\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{C}} \\ \boldsymbol{q} &= -\boldsymbol{\mathcal{C}}^T \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{D}}\boldsymbol{V}\boldsymbol{S}\mathbf{1}_n \end{aligned}$$

and the matrices $\boldsymbol{G}$ and $\boldsymbol{h}$ help to generalize different capital weight constraints, originating from investment policy (internal) or market liquidity (external). For instance, in the case of lower and upper bounds for security allocation, these matrices can be accommodated in the following way:

$$\boldsymbol{G} = \begin{bmatrix} -\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{C}} \\ \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{C}} \end{bmatrix}_{2n\times k}, \quad \boldsymbol{h} = \begin{bmatrix} -\boldsymbol{l} \\ \boldsymbol{u} \end{bmatrix}_{2n\times 1}$$

Where $\boldsymbol{l}$ and $\boldsymbol{u}$ are $n \times 1$ vectors containing the weight floors and caps for each investable security, respectively. The matrices $\boldsymbol{G}$ and $\boldsymbol{h}$ can also be modified to accommodate weight constraints at the sector or asset class level, if needed.

We have not yet addressed the definition of predetermined fixed allocations $\boldsymbol{\mathcal{W}}$. The intention of defining the capital weight of the security in advance simplifies the optimization problem by reducing the number of decision variables from n securities to k clusters, where $k << n$. Given that the diagonal elements in $\boldsymbol{\mathcal{W}}$ must be set before trying to solve (2.10), it can be the result of a previous optimization step or it could be given by an optimization-free naive portfolio construction methodology. The optimization problem (2.10) would then scale up and down these weights by choosing different values for the vector $\boldsymbol{\theta}$. We suggest setting

$$\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{D}}^{-1}$$

for two reasons. First, it offers a more balanced initial allocation in risk terms; securities with higher volatility will get a smaller capital allocation, preventing them from dominating the overall risk of the portfolio. Second, we notice that $\boldsymbol{P}$ and $\boldsymbol{q}$ also get reduced to the following expressions:

$$\begin{aligned} \boldsymbol{P} &= \boldsymbol{\mathcal{C}}^T \boldsymbol{V} \boldsymbol{S}^T \boldsymbol{S} \boldsymbol{V}^T \boldsymbol{\mathcal{C}} \\ \boldsymbol{q} &= -\boldsymbol{\mathcal{C}}^T \boldsymbol{V} \boldsymbol{S} \boldsymbol{1}_n \end{aligned}$$

further simplifying the implementation of the DSP.

2.5 Simulation

2.5.1 Out-of-sample backtesting

A commonly accepted way to test the effectiveness and properties of the DSP methodology presented in this study would be to choose a specific set of securities, a particular historical sample, and perform a backtest. As Bailey et al. (2014) notes, a backtest is considered realistic when the in-sample performance is consistent with the out-of-sample performance. Furthermore, the authors argue that it is relatively simple to overfit and investment strategy so that it performs well in-sample, thanks to data snooping (observe the results, and modify the parameters and/or security universe to improve the results). Similar conclusions were drawn by López de Prado (2016) and López de Prado (2018), who outline a novel framework to test the efficacy of trading rules; describe the stochastic process that generates the return stream coming from a trading rule, find the optimal parameters using historical data, and perform multiple backtests using Monte Carlo simulations[5]. We borrowed this concept and extended it to simulate a large number of correlated securities and tested the effectiveness of our portfolio construction methodology using Monte Carlo simulations.

2.5.2 Benchmark constructions

The Diversified Spectral Portfolios (DSP) developed in the previous section will be contrasted with two other diversification-driven portfolio construction methodologies. The first benchmark will be the equally weighted portfolio. As mentioned above, this naive construction is capable of beating several optimization-driven methodologies, in risk-adjusted terms, due to the fact that it does not rely on noisy estimates of expected returns and the covariance matrix. The second benchmark will be the equal-risk contribution (ERC) portfolio, also called risk parity, which is achieved by the combination of weights that equalize the marginal risk contribution of each asset in the portfolio. We follow the specification detailed in Maillard et al. (2010) to construct this benchmark. With these two constructions as benchmarks, we compare the proposed methodology with equal allocation in both capital and risk terms.

[5] López de Prado (2018) dubbed the term "backtesting on synthetic data".

2.5.3 Experiment design

The objective of our experiment is to test and analyze the effectiveness of the DSP under a large number of dire market scenarios, without simply relying on historical returns data, for which we resort to Monte Carlo simulations. Each iteration of the simulation experiment is meant to represent one year of managing a portfolio of 50 securities. At the beginning of the iteration, a random 50×50 correlation matrix is generated and remains unchanged throughout the iteration. Using this matrix and the mathematical framework that will be discussed later, we generate daily return streams for 50 securities over a 3-year window (756 observations). Given that two of the three methodologies to be contrasted rely on the estimation of either the sample covariance matrix or the SVD, we use a rolling estimation window of 2 years (756 observations) to perform these calculations. Consequently, the first portfolio rebalancing takes place at the end of the second year, and the consequent rebalances are performed monthly (every 21 observation), resulting in one year of out-of-sample performance. On each rebalance day, we not only calculate the target weights for the DSP, but also for the two benchmark constructions, to make the results directly comparable.

The previous process will be repeated 10,000 times, giving us 10,000 years of out-of-sample data, which would not be possible just using historical stock market data. On top of that, the number of simulated observations will also be helpful in analyzing not only the effectiveness of our DSP, but also under which circumstances they perform the best.

2.5.4 Mathematical framework and parameter setting

Let us assume a market in which asset price dynamics can be described by the following stochastic differential equation (SDE), under the physical probability measure $\mathbb{P}$:

$$\frac{dS(t)}{S(t)} = \underbrace{\mu dt}_{\text{Drift}} + \underbrace{\sigma dW(t)}_{\text{Diffusion}} + \underbrace{dJ(t)}_{\text{Jump}} \tag{2.11}$$

where $S(t)$ is the price of the security modeled at time t, μ and σ are the unconditional mean and variance, respectively, of its return stream. $W(t)$ is a Wiener process and $J(t)$ is a compound Poisson process. The first two terms of (2.11) are helpful in describing a well-known geometric Brownian motion process, assuming normality in returns, while the third term in the equation gives the overall process the flexibility to also describe nonnormal price discontinuities observed in financial data.

To simulate a portfolio construction and study its features, we will need additional assets to include in the out-of-sample portfolios. Assuming a system composed of N securities with a correlation matrix $\boldsymbol{R}$, we can transform (2.11) into an SDE system that describes the dynamics of N correlated assets:

$$\dot{\boldsymbol{S}} = \boldsymbol{\mu} dt + \boldsymbol{D}\boldsymbol{L}\boldsymbol{W} + \boldsymbol{J} \tag{2.12}$$

where $\boldsymbol{D}$ is a diagonal matrix that has the standard deviation of the returns of the assets in the diagonal entries and zero everywhere else, $\boldsymbol{L}$ is the lower triangular matrix coming from the Cholesky decomposition of $\boldsymbol{R}$, and $\dot{\boldsymbol{S}}$, $\boldsymbol{W}$, $\boldsymbol{J}$ are the vectors of the security returns, Wiener processes and compound jump processes, respectively.

To make the securities comparable, we set the same values of expected returns and volatility for each of them at 5% and 10%, respectively. As a result of this, the difference will come from the realizations of $\boldsymbol{W}$ and $\boldsymbol{J}$. For the latter, we need to define two parameters; the frequency of the jumps and their intensity. Setting the frequency parameter at 0.15, we should expect to observe approximately three jumps in a given month. Regarding the size of the jump, these values will be drawn from a normal distribution with mean -5% and standard deviation 7.5%.

In addition to the above, we will also simulate different correlation structures at the beginning of each iteration. These correlation matrices will be generated randomly, following the numerically stable algorithm described in Davies and Higham (2000). To construct these matrices, it is necessary to specify a collection of eigenvalues (50 in our case), which will be drawn from an exponential distribution, simulating a system where the first few implicit factors are capable of explaining most of the risk; a common case in financial markets.

Finally, we have to set the distance threshold that controls how many clusters are found on each rebalance date. Although we recognize that refining and tailoring this number to the data set with which the investment manager is working should yield better results, we believe that finding a systematic and statistically sound way to choose this parameter falls outside the scope of this study. For this reason, we set the clustering threshold at 0.5, as it is the middle point of this cosine distance's range, which should provide enough clusters to address the collinearity issue and leave enough degrees of freedom in the system to allow the optimizer to work with.

2.5.5 Results

The Monte Carlo simulation exercise described in the previous section gives us 10,000 years of data for the Diversified Spectral Portfolios (DSP) and the other two chosen benchmarks; equal-weighted (EW), and equal risk contribution (ERC) portfolios. Exhibit 2.5 shows the Sharpe ratio distributions estimated with the simulation results. The cumulative distributions shown in

Table 2.1: **Two-sample one-sided Kolmogorov-Smirnov (KS) test**

This table reports the statistical results of a two-sample, one-sided Kolmogorov-Smirnov (KS) test for first-order stochastic dominance. The test compares the cumulative distribution of Sharpe ratios from the Diversified Spectral Portfolio (DSP) strategy against those of the Equal Risk Contribution (ERC) and Equal-Weighted (EW) strategies from the 10,000 Monte Carlo simulation trials. This provides strong statistical confirmation for the visual observation from Figure 2.5 that the DSP strategy stochastically dominates the two benchmarks.

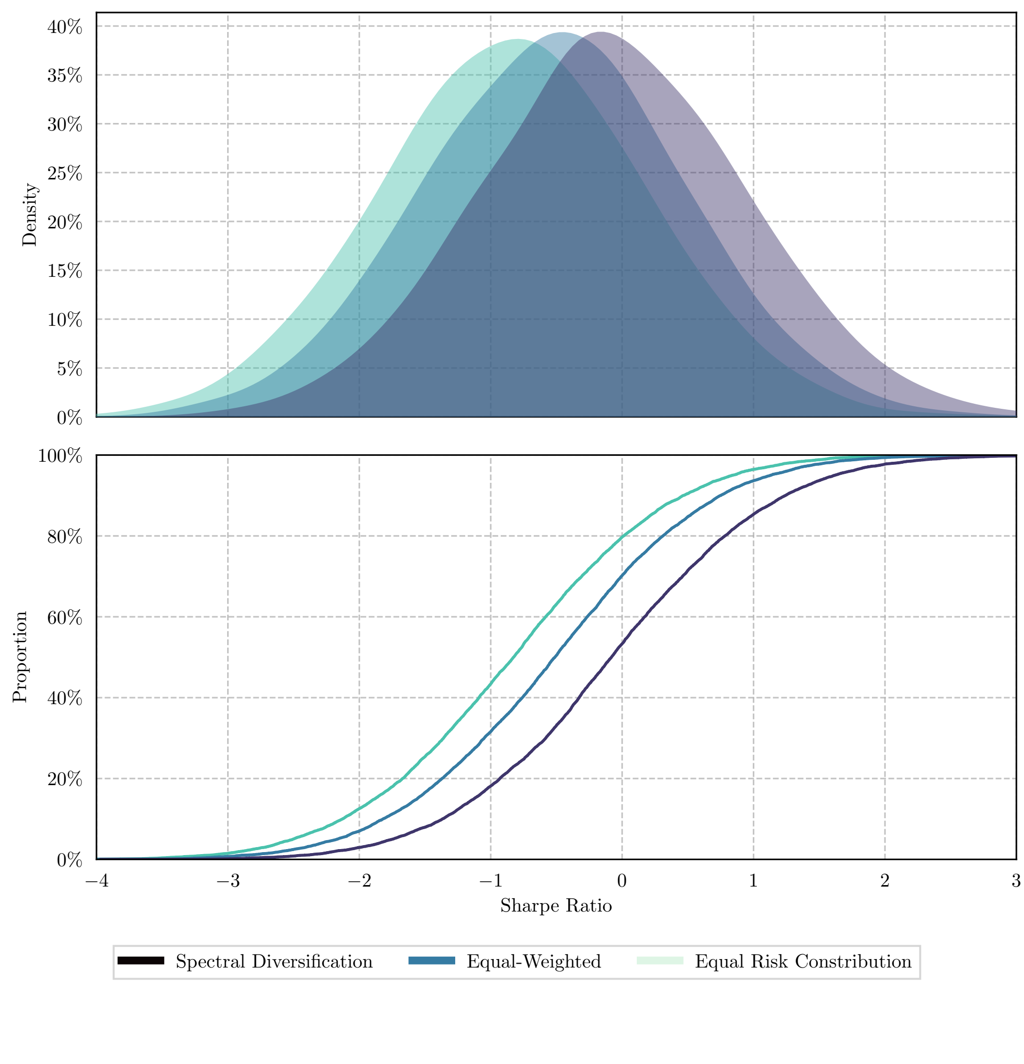
	$F_{DSP} > F_{ERC}$	$F_{DSP} > F_{ERC}$
KS Statistic	0.18	0.30
One-tailed p-value	0.00	0.00

the second panel hint at the first-order stochastic dominance of DSP over the two. As highlighted in Schmid and Trede (1996) the two-sample one-sided Kolmogorov-Smirnov test is among the most widely used tests for first-order stochastic dominance. We report the results of this test in Exhibit 2.1. The reported p-values seem to confirm our observation of stochastic dominance at a 5% confidence level.

The 10,000 trials were generated using an equal number of randomly generated correlation matrices. We can dissect these trails on the basis of the level of risk concentration of the manifested correlation structure. This is particularly interesting because one of the most commonly observed features of bear markets is a sudden spike in cross-asset correlations, which diminishes any diversification benefits that might be observed in bull markets. We choose to measure this risk concentration by calculating the Shannon entropy of the distribution of the eigenvalue matrix of the correlation matrix. Highly concentrated systems should have low entropy, since most of these ”probabilities” will be concentrated in one or few elements, making the systems less uncertain.

Figure 2.5: **Density of simulated Sharpe ratio (10,000 simulations)**

This figure presents the results of a 10,000-iteration Monte Carlo simulation comparing the out-of-sample performance of the proposed Diversified Spectral Portfolios (DSP) against two benchmarks: an Equal-Weighted (EW) portfolio and an Equal Risk Contribution (ERC) portfolio. The top panel shows the probability density functions of the annualized Sharpe ratios for the three strategies. The bottom panel displays the corresponding cumulative distribution functions (CDFs). The visual evidence from the CDFs suggests that the DSP strategy exhibits first-order stochastic dominance over both the EW and ERC benchmarks, implying a preferable risk-adjusted return profile across all outcomes.



The previous exercise can also be used to contrast the DSP with its two benchmarks in these challenging diversification scenarios. Exhibit 2.6 provides three different angles from which to view the results, using joint distributions to visualize the results. Within each joint distribution, a 45-degree line was added to facilitate comparison and serves as an "indifference" line. The first column of the exhibit depicts a similar picture as in Exhibit 2.5; DSP seems to be preferable to the other two constructions because both the center of the distribution and most of the joint probability mass fall in the lower triangular area, right to the indifference line. Interesting conclusions can be drawn from the least and most concentrated cases, which are plotted in the second and third columns, respectively. Here, we can see that the difference between these pairs of portfolio construction methodologies maximizes in the cases where the correlation structure is most concentrated, having most of the mass of the joint densities on the right of the indifference line. In the cases where the investable set exhibits the least correlation among the trials, the difference between the pairs is less noticeable, although still favorable for DSP.

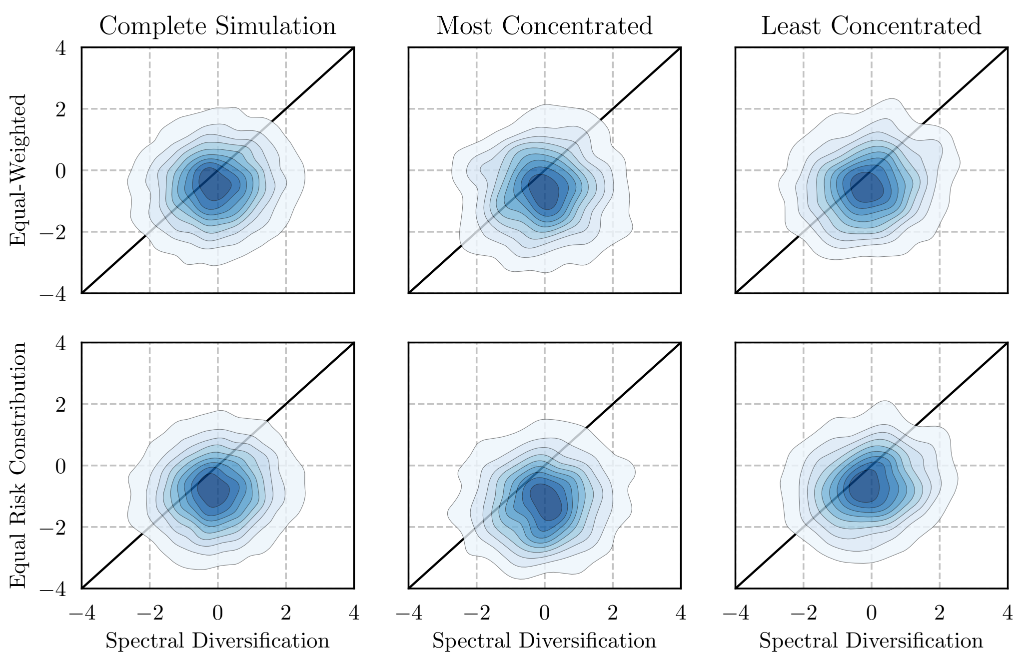
2.6 Conclusion

The question of how to diversify an investment portfolio has many candidate answers. We depart from conventional wisdom, and instead of diversifying a portfolio by leveraging the correlation matrix, we seek to balance out the risk attributed to its eigenvectors. Our Diversified Spectral Portfolio framework can be particularly useful when we are looking to construct diversified portfolios in a purely data-driven way or as a complementary approach to expand an already implemented portfolio optimization framework.

The out-of-sample results of our Monte Carlo simulation on a system of stochastic differential equations show that DSP is able to provide better risk-adjusted performance than competing, industry standard, portfolio di-

Figure 2.6: **Joint density of simulated Sharpe ratio (10,000 simulations)**

This figure provides a deeper analysis of the Monte Carlo simulation results by plotting the joint distributions of Sharpe ratios for the DSP strategy versus the two benchmarks under different market conditions. The plots are divided into three columns: the complete simulation, the scenarios with the most concentrated correlation structures (low entropy), and the scenarios with the least concentrated structures (high entropy). In each plot, the 45-degree line serves as an indifference line. The concentration of the probability mass to the right of this line indicates superior performance for DSP. The figure reveals that the outperformance of the DSP methodology is most pronounced in highly concentrated market environments, which are characteristic of bear markets and crises.



versification methodologies such as equal-weighted and risk parity. A closer look at the results of the simulation shows that this improved performance seems to come from better protection in the context of highly concentrated correlation structures, a characteristic of contractionary markets.

Although we are satisfied with the results from the out-of-sample methodology, there is still room for further developments. We can think of three main areas of improvement. First, the singular value decomposition, the backbone of the presented methodology, is sensitive to large non-normal outliers in the data, which are very common in financial data. Having a more robust way to estimate the eigenvectors and values of a set of stock returns would result in more reliable portfolios generated by our methodology. Second, given that the eigenvectors associated with the largest values are associated with the main dynamic driven by the data, the ones associated with the smallest values are usually driven by noise. Compressing the data by identifying and then dropping the eigenvectors that are potentially affected by this should increase the out-of-sample robustness of the methodology. Finally, one missing piece of our portfolio construction methodology is the expected return dimension. The introduction of the expected return associated with each eigenvector could improve the risk-adjusted performance of our spectral portfolios.

Chapter 3

Incorporating Market Regimes into Large-Scale Stock Portfolios: *A Hidden Markov Model Approach*

3.1 Introduction

Financial markets often exhibit abrupt changes in behavior that persist for extended periods, which often correspond to significant economic shifts, regulatory changes, or policy alterations, and secular changes (Ang and Timmermann, 2012). Empirical evidence strongly supports the presence of regimes in equity markets data, associating distinct return-generating distributions for different states. These regimes are closely tied to economic conditions with, for example, a high volatility and low growth regime being associated with deep recessions such as the Great Depression and the 2008 financial crisis (Yang, 2023). Uncertain environments like the aforementioned relatively favor demand-inelastic firms as non-essential spending declines, while economic expansions boost the earnings of *procyclical* companies via increased discretionary spending. Consequently, the equity market develops a regime-dependent behavior as investors attempt to anticipate the trajectory of the business cycle.

The impact of this phenomenon in the context of portfolio construction cannot be ignored, as these regime dynamics can explain and model the time-varying behavior of return volatility and correlations, together with other stylized facts of financial time series such as volatility clustering, gain/loss asymmetry, and excess kurtosis. Consequently, the incorporation of regime dynamics into asset allocation and portfolio construction has been of particular interest in the last couple of decades, and several studies have explored the benefits of implementing dynamic allocation strategies within a regime-aware framework over static approaches (see, e.g., Ang and Bekaert, 2002a; Guidolin and Timmermann, 2007; Guidolin and Timmermann, 2008; Kritzman et al., 2012; Sheikh and Sun, 2012; Nystrup et al., 2015; Mulvey and Liu, 2016; Reus and Mulvey, 2016; Nystrup et al., 2019; Costa and Kwon, 2019; Kelliher et al., 2022).

While the potential benefits of regime-aware investing are evident and well-documented, challenges remain in the practical implementation of such strategies. The estimation of regime-switching processes can be computationally demanding (Janczura and Weron, 2012) and prone to overfitting (Chen and Bunn, 2014), particularly in high-dimensional settings. This is the case of real-world equity portfolios, which can be significantly large, as sufficient diversification might be achieved only by combining hundreds of stocks (Statman, 2004), and modern systematic trading firms can potentially hold thousands of stocks in their books. In this context, the implementation of a regime-aware portfolio framework would be very challenging, and potentially unfeasible, due to the quadratic growth of parameters in the number of states (Bulla et al., 2011).

However, the low-rank nature of equity data offers a potential solution to this challenge by allowing the user to select a handful of statistically meaningful risk factors which can decently explain the variation in stock returns instead, to reduce the dimensionality and potentially the estimation error. Employing parsimonious factor models to simplify the complexity of large-scale portfolio optimization and risk modeling has been extensively studied in the academic literature (see, e.g., Perold, 1984; Jacobs et al., 2005; Deng et al., 2024) and has also become common practice in the financial industry. In this study, we leverage this approach to model the regime-dependent behavior of potentially hundreds of stocks.

The main contribution of this chapter is two-fold. First, we propose a regime-weighted least-squares method to estimate conditional regime-aware factor loadings based on the Hidden Markov Model (HMM) to obtain forward-looking factor loadings for a sizable universe of stocks. Second, we use these conditional factor loadings to construct large-scale stock portfolios that can be used to systematically manage investments in a regime-aware manner. Both our framework and its empirical application results expand on the current state of the literature by providing a way of implementing regime modeling in the context of a large-scale portfolio, and allowing its implementation in a real-world setting. We expect these results to be particularly interesting for quantitative asset managers and stock-focused hedge funds, who manage portfolios with hundreds or even thousands of stocks. The remainder of the chapter is organized as follows. In Section 2, we briefly analyze the related literature and highlight our contributions to it. In Section 3, we outline our regime-weighted least-squares methodology and how it can be leveraged to construct large-scale stock portfolios. In Section 4, we report the results of Monte Carlo simulations to show the out-of-sample performance of our approach, and contrast its performance against a regime-agnostic benchmark. Section 5 concludes.

3.2 Related literature

The literature on regime-switching models in finance is vast, with applications ranging from interest rate modeling (see, e.g., Ang and Bekaert 2002b, Bansal and Zhou 2002, Dai et al. 2007) and monetary policy and macro analysis (see, e.g., Owyang and Ramey 2004, Hamilton 2005, Sims and Zha 2006, Hamilton 2010 Baele et al. 2015) to option pricing (see, e.g., Chan 2014, Siu 2014). The incorporation of regime dynamics into asset allocation and portfolio construction has been of particular interest during the last couple of decades, expanding the related literature in multiple directions. Literature relevant to our methodology sits in the intersection of studies on the regime-dependent behavior of equity factors and applications of HMM in portfolio construction.

Relevant studies that have implemented HMMs to manage a portfolio include Bulla et al. (2011), who implement an HMM with t-distributions to model regime-switching in asset allocation, finding two regimes; a high-variance regime of relatively short duration, and a low-variance regime. Their strategy involves full investment in an index or risk-free asset based on predicted states. Their dataset consists of daily returns of five major international broad indices for over 20 years, starting in January 1976. The authors show that their strategy can lead to improved Sharpe ratios, mainly driven by lower overall portfolio risk. Along similar lines, Bae et al. (2014) fits a multivariate Gaussian HMM on three-dimensional input data; the S&P500, the 10-year US Government bond, and the GSCI Commodity Index to uncover four distinct market regimes: two extremely positive market conditions for the equity market, a transition period, and market crashes. Their results show that their framework outperforms other benchmarks but does especially outstanding during crash periods by avoiding risk during left-tail events. The authors use daily frequency observations from January 1980 to June 2012 and a stochastic programming framework to dynamically optimize portfolios.

Among studies that have analyzed the regime behavior of traditional equity factors we can find Guidolin and Timmermann (2008), who use a regime-switching model to analyze the joint distribution of returns on market, SMB, and HML portfolios, originally introduced in Fama and French (1993), and four regimes; *Bear* state, *Bull* state with low volatility, *Bull* state with positive returns, and a *Volatile* state with high returns. The authors consider monthly returns on US stock portfolios from December 1927 to December 2005, including six equity portfolios formed by intersecting two size portfolios and three book-to-market portfolios. Accounting for these regimes in portfolio strategy leads to better performance compared to single-state models. Separately, Costa and Kwon (2020) develop a regime-dependent portfolio framework, based on a univariate Gaussian HMM, that consistently outperforms its nominal and robust counterparts by achieving lower volatility and higher returns after 15 years of out-of-sample experiments. Using monthly market excess return data from January 1973 to June 2018, and the Bayesian information criterion (BIC), they identify two market regimes: *bullish* and *bearish*[1]. They use the 24 most recent observations corresponding to the most recent dominant market regime to estimate the factor loading of the Fama-French three-factor model parameters for 30 assets (sector portfolios). These parameters are then used to estimate the asset-level mean and covariance matrix, so they can be used in traditional mean-variance optimization (MVO) and minimum variance optimization. The authors argue that their framework allows for the construction of large, realistic portfolios at no additional computational cost during optimization.

There has also been a growing trend in employing non-parametric alternatives to HMMs (see, e.g., Nystrup et al., 2020; Zheng et al., 2021; Shu et al., 2024). The most relevant to our work is Aydınhan et al. (2024) who departs from traditional HMMs with discrete state sequences to a continuous

[1]The authors also note that, despite the two-regime model having the lowest BIC, the three- and four-regime models are not poor choices given their comparable BIC.

jump model that follows a Markov process on the probability simplex. The authors argue that a probabilistic approach provides more informative and valuable insights than labeling each time period with the most likely regime. In their simulations and an empirical exercise using daily data for the Nasdaq Composite index, the proposed continuous jump model outperforms both the HMM and discrete jump models.

This study builds upon the aforementioned literature that has implemented HMMs to incorporate regime-dependent estimates to manage investment portfolios. In particular, we present a regime-dependent portfolio framework that extends the work of Costa and Kwon (2020). In a similar fashion, we fit an HMM on a parsimonious set of factors rather than the whole investable set of assets, and use regression estimates to project back these estimates into asset space. However, instead of relying on a univariate Gaussian HMM, we expand it into a multivariate one by including other relevant factors besides the market risk premium while identifying equity market regimes. We also use weighted least squares to estimate regime-aware factor loadings by using the whole available history, rather than just the most recent observations. We believe the combination of these allows us to come up with richer regime-dependent estimates for stocks' expected returns and covariance matrix estimations that can be used in the optimization of a sizable portfolio. Additionally, we take the probabilistic approach of Aydınhan et al. (2024) to estimate continuous regime-dependent estimates, rather than just discretionarily labeling each observation with the most likely regime in order to capture the uncertainty associated with regime assignment. We leverage this probabilistic approach not only for regime identification, but also for the estimation of forward-looking factor loadings used in the stocks' expected returns and covariance matrix estimates.

3.3 Methodology

Although the advantages of considering market regimes in investment strategies are well-documented, the sheer complexity and high-dimensional nature of actual stock portfolios, particularly those managed through quantitative methods, could potentially make it unfeasible to directly apply regime-switching models on such vast datasets. A promising avenue for addressing these challenges lies in leveraging the low-rank structure inherent in stock return data. Instead of fitting a regime-switching model on thousands of stocks, we can instead focus on a parsimonious set of statistically significant common risk factors that substantially explain stock return variation; this way significantly reducing the dimensionality and potentially bringing down the estimation error. The application of factor models to simplify portfolio optimization and risk management is a well-established practice in both academic research (e.g., Perold, 1984; Jacobs et al., 2005; Deng et al., 2024) and the investment industry. Consequently, while the direct implementation of regime-switching models at the individual stock level may be impractical, fitting these models to a reduced set of factors offers a potentially viable alternative.

We begin by defining the relationship between a stock's excess return and a finite set of common risk factors, so the regime framework can be incorporated into this relationship. Taking a simple conditional asset pricing model framework, such as the one in Smith and Timmermann (2021) , we can linearly decompose the excess return x of any particular stock as follows:

$$x_{t+1} = \sum_{k=1}^{K} b_{k,t} \lambda_{k,t+1} + \epsilon_{t+1} \tag{3.1}$$

where $b_{k,t}$ and $\lambda_{k,t}$ are respectively the time-varying factor loading of the stock and the associated conditional risk premium, corresponding to factor

k in the model, and

$$\epsilon_{t+1} = \sum_{k=1}^{K} b_{k,t} \left(z_{k,t+1} - E_t[z_{k,t+1}] \right) + u_{t+1}$$

is an unpredictable component of the stock's excess return and u_{t+1} is a stock-specific idiosyncratic shock with zero mean. As pointed out by the authors, the return predictability in this context can arise either from the conditional factor loadings or from time-varying risk premia.

3.3.1 Regime-conditional risk premia

Let us assume that the common risk premiums that drive stock returns are governed by an unobservable stochastic process $\boldsymbol{s}$ which can assume a finite collection of states $\{s_1, s_2, \ldots, s_M\}$ of length M. Within each of these states, factor returns are drawn from a distinct state-dependent joint Gaussian distribution with an unconditional mean of the excess returns $\lambda_s \in \mathbb{R}^{K \times 1} := E[\boldsymbol{z} \mid s]$, and covariance matrix $\Omega_s \in \mathbb{R}^{K \times K} := Cov[\boldsymbol{z} \mid s]$.

Rather than relying on a discrete identification of the dominant regime at each point in time to dictate the marginal joint distribution to be used, we assume a probability simplex instead, which linearly combines the conditional joint distribution corresponding to each regime. This approach is also sensible from the empirical point of view as the marginal distribution of financial assets might not be properly described by a single probability distribution but rather by a combination of densities. In fact, modeling stock returns as a mixture of Gaussian components can describe the so-called stylized facts of financial time series (Rydén et al., 1998). Representing and modeling an overall population distribution as a combination of M densities is known as a *mixture model*. In this context, the regime-dependent behavior of the common factor risk premia can be better described as an m-component

Gaussian mixture model,

$$\hat{f}(z_t|\Psi) = \sum_{s=1}^{M} \gamma_{s,t}\phi(\lambda|\lambda_s, \Omega_s) \tag{3.2}$$

where $\hat{f}$ is the estimated distribution of λ_t; ϕ is the multivariate Gaussian probability density function. Ψ is the set of unknown parameters (γ_t, $\lambda_1, ...,$ λ_M, $\Omega_1, ..., \Omega_M$), where $\gamma_t := (\gamma_{1,t}, \gamma_{2,t}, \ldots, \gamma_{M-1,t})$ is the mixing distribution that stores the weight assigned to each state's density function, which satisfies $\forall \gamma_{s,t} \geq 0$ and $\sum_{s=1}^{M} \gamma_{s,t} = 1$. In the context of a regime-switching framework, each element in λ_s and Ω_s is directly associated with a different state in $\boldsymbol{s}$. Similarly, the mixing distribution γ_s takes the form of a simplex that contains the probabilities of observing each regime in $\boldsymbol{s}$ at time t. Something worth mentioning is that in the context of (3.2), despite the identification of λ and Ω being static irrespective of the time period, the time variation of the factor's risk premium λ_t comes primarily from the conditional mixing distribution γ_t, which continuously combines the unconditional joint densities differently on each observation t.

It is only sensible to expect that the probabilities associated with each state of $\boldsymbol{s} \in \{s_1, s_2, \ldots, s_M\}$, γ_t, will change in time reflecting the current stock market environment. One of the simplest approaches commonly taken to model these dynamics is to assume that the hidden process $\boldsymbol{s}$ evolves over time in an equally spaced sequence, forming a homogeneous first-order discrete-time Markov chain (DTMC). A process is said to be a DTMC if it satisfies the so-called *Markovian property*:

$$P(\boldsymbol{s}_{t+1} \mid \boldsymbol{s}_t, \boldsymbol{s}_{t-1}, \ldots, \boldsymbol{s}_0) = P(\boldsymbol{s}_{t+1} \mid \boldsymbol{s}_t)$$

where the future state of the underlying stochastic process (i.e., the market regimes in our case) depends only on its current state. The collection of the probabilities of transitioning from and to every state in $\boldsymbol{s}$ is organized and stored in the so-called one-step*transition matrix* $\Pi \in \mathbb{R}^{M \times M}$,

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \dots \pi_{1M} \\ \pi_{21} & \pi_{22} \dots \pi_{2M} \\ \vdots & \ddots \quad \vdots \\ \pi_{M1} & \pi_{M2} \dots \pi_{MM} \end{pmatrix}_{M \times M}$$

where $P(s_{t+1} = j \mid s_t = i) = \pi_{ij}$.

One sensible approach to extracting the underlying state-dependent densities of the common risk factor, while modeling the dynamics of the hidden process s through time as a DTMC, is the Gaussian variant of the well-received HMM. Formally, HMMs are double-stochastic discrete processes within the family of generative machine learning algorithms, which can be considered an extension of mixture models along the temporal axis (Bouguila et al., 2022). A Gaussian HMM of M hidden states can be completely determined by the parameter set $\{\Pi, \gamma^{(0)}, \Gamma, \boldsymbol{\lambda}, \boldsymbol{\Omega}\}$. The parameters Π, $\gamma^{(0)} \in \mathbb{R}^{M \times 1}$ are the previously introduced one-step-ahead transition matrix, and the set of initial state probabilities, respectively. $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_M\}$ and $\boldsymbol{\Omega} = \{\Omega_1, \dots, \Omega_M\}$ are the collection of state-dependent vectors of means and covariance matrices that describe the regime-dependent distributions of the common risk factors in each of M unobservable market regimes, where $\lambda_s \in \mathbb{R}^{K \times 1}$ and $\Omega_s \in \mathbb{R}^{K \times K}$ are the joint distribution means and covariances of the K risk factors in regime s. Finally, Γ is the mixing matrix, which contains the posterior probabilities of the observable variables being emitted by state $s \in \{s_1, \dots, s_M\}$, along the rows, conditional on the information available at time t, $\mathcal{F}_t$, and it is also true that Γ is row-stochastic (i.e., $\sum_{s=1}^{M} \gamma_{s,t} = 1$).

We proceed to fit a multivariate Gaussian HMM on the set of common risk factors. Let $\boldsymbol{\mathcal{Z}} \in \mathbb{R}^{T \times K} := (z_1, \dots z_K)$, where z_k be the sequences of

T historical returns of the K common risk factors. The priors for both the collection of initial state probabilities $\gamma^{(0)} \in \mathbb{R}^{M\times 1}$ and the transition matrix Π can be drawn from a Dirichlet distribution, favoring a transition matrix where most of the probability mass is concentrated on a small number of state transitions, while the priors for the state-dependent density parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\Omega}$ can be determined through, for example, k-means clustering. Once the parameters are initialized, an iterative Expectation-Maximization (EM) algorithm is used to arrive at the most likely initial state probabilities, transition matrix, and state-dependent density parameters. The posterior probability matrix Γ is computed through the Forward-Backward algorithm. With respect to the state-dependent covariance matrices Ω_s, we restrict the non-diagonal elements to be zero, as risk factors tend to be uncorrelated. On top of being rooted in economic foundations, this assumption will also reduce the number of parameters to be estimated by HMM, potentially reducing the estimation error and simplifying the math and computation.

We can now leverage the regime-dependent parameters estimated by the HMM to arrive at our time-varying risk premia. Let $\gamma_t \in \mathbb{R}^{1\times M} := (\gamma_{1,t}, \gamma_{2,t}, \ldots, \gamma_{M,t})$ be the filtered state probabilities at time t, conditioned on all the information available as of t (i.e., the last row of Γ) and $\hat{\gamma}_{t+1} \in \mathbb{R}^{1\times M} := (\hat{\gamma}_{1,t+1}, \hat{\gamma}_{1,t+1}, \ldots, \hat{\gamma}_{M,t+1})$ be the one-period-ahead forecasted state probabilities. The n-step-ahead state probabilities can be defined as $\hat{\gamma}_{t+n} = \gamma_t \Pi^n$, given that the homogeneous Markov chain satisfies the Chapman–Kolmogorov equations. With this in hand, our forecast for the one-step-ahead time-varying factor loadings will be given by

$$\begin{aligned} E_t[\lambda_{t+1}|\mathcal{F}_t] &= \sum_{s=1}^{M} \hat{\gamma}_{s,t+1}\lambda_s \\ Var_t[\lambda_{t+1}|\mathcal{F}_t] &= \sum_{s=1}^{M} \hat{\gamma}_{s,t+1}\left(\Omega_s + \lambda_s\lambda_s^T\right) - \hat{\gamma}_{s,t+1}\lambda_s^2 \end{aligned} \tag{3.3}$$

3.3.2 Regime-conditional factor loadings

As already pointed out in Equation (3.1), the return predictability is also a product of the conditional factor loadings. Nested within the same framework, given $\mathcal{F}_t$ let the time-varying factor loadings of stock i to factor k be defined as

$$b_{ik,t} = \sum_{s=1}^{M} \gamma_{s,t} b_{ik,s} \tag{3.4}$$

where $b_{ik,s}$ is the state-dependent factor loading of stock i associated with factor k, and $\gamma_{s,t}$ is the filtered probability corresponding to state s, previously defined. It can be seen in (3.4) that, within each regime, the state-dependent factor loadings are constant over time and their conditionality is solely driven by the dynamics of the probability simplex $\gamma_{s,t}$ which linearly combines these M sets of state-conditional factor loadings. We propose modeling this behavior by leveraging the matrix of posterior probabilities Γ estimated while fitting the HMM on $\mathcal{Z}$, to weight the historical observations of the common risk factors to estimate the state-specific factor loadings, and aggregate them accordingly to obtain the conditional factor loadings at the stock level. To estimate the regime-dependent factor loadings $\theta_s \in (\alpha_s, b_{1s}, b_{2s}, \ldots, b_{Ks})$, consider the following weighted least squares problem:

$$\underset{\theta_s}{\operatorname{argmin}} \quad \sum_{t=-T}^{0} \gamma_{s,t} \left(x_t - \left(\alpha_s + \sum_{k=1}^{K} b_{k,s} z_{k,t} \right) \right)^2 \tag{3.5}$$

where x_t is the excess return of a particular stock at time t, $b_s = (b_{1,s}, b_{2,s}, \ldots, b_{k,s})^T$ is the state-conditional (but time-invariant) factor loadings associated with each of the K common risk factors, and α_s is the state-conditional intercept of the regression when the dominant market regime is s. Repeating the (3.5) for each $s \in \{s_1, \ldots, s_M\}$ we obtain the collection of state-dependent factor loadings across the M unobservable regimes:

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_M) = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_m \\ b_{1,1} & b_{1,2} & \ldots & b_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{K,1} & b_{K,2} & \ldots & b_{K,M} \end{pmatrix}_{K \times M} \tag{3.6}$$

The result of this process represents a regime-weighted least squares (RWLS) estimator, which serves as the foundation of our methodology for obtaining regime-aware means and covariances of large-scale stock portfolios.

3.3.3 Building regime-aware stock portfolios

Consider an investable universe constituted by N stocks (where N is significantly large) whose returns can be explained by a set of K observable common risk factors (where $K << N$). We also assume that the dynamics of these K factors can be characterized by a finite set of M unobservable market regimes that evolve in time following a DTMC and can be modeled through the lens of a multivariate Gaussian HMM. The most accessible approach to construct stock portfolios when Gaussian returns are assumed, is to implement a traditional MVO or a minimum variance portfolio, as in Costa and Kwon (2020). However, we opt for moving away from these traditional approaches, as we believe that collapsing the M state-dependent joint distributions into one Gaussian to perform the optimization exercise would discard rich information provided by the multiple and distinct densities. For this reason, we instead follow Luxenberg and Boyd (2024), who introduced a portfolio construction approach that assumes asset returns that follow a Gaussian mixture distribution which is able to capture the information yielded by different densities belonging to different market states. Despite the apparent benefits of such an approach, implementing this framework requires estimating the expected return and covariance matrix of the N stocks in the universe in each of the

M regimes which, as previously stated, might result in significantly challenging and potentially infeasible if the investable universe is significantly large. Leveraging the regime-dependent densities and factor loadings developed in this section to obtain state-dependent factor risk premia and covariance estimates and project them back into stock space.

Equation (3.1) can be expanded to describe the behavior of the whole investable universe of N stocks through the regime-dependent behavior of the K common risk factors. Let $\mathcal{X}_{t+1} \in \mathbb{R}^{N\times 1} := (x_{1,t+1}, x_{2,t+1}, \ldots, x_{N,t+1})^T$ be a vector of excess returns over the next period for each of the N stocks in the investable universe. Then,

$$\mathcal{X}_{t+1} = \sum_{s=1}^{M} \hat{\gamma}_{s,t+1}^T \left(\Theta_s \lambda_s + \epsilon_s\right) \tag{3.7}$$

where

$$\Theta_s = \begin{pmatrix} \theta_{1,s} \\ \theta_{2,s} \\ \vdots \\ \theta_{N,s} \end{pmatrix}_{N\times K+1}$$

and $\theta_{i,s} = (\alpha_{i,s}, b_{1i,s}, \ldots, b_{ki,s}) \in \mathbb{R}^{1\times K+1}$ is the vector of state-dependent intercept and factor loadings corresponding to stock i in the investable universe when the market is in regime s. In order to understand the expected behavior of the investable universe in each of the regimes, let us assume that at time t we have complete certainty that the returns of the common risk factor are drawn from the multivariate Gaussian density corresponding to regime s. In this case, the expected return of our investable set of stocks, and its state-conditional corresponding covariance matrix will be given by:

$$\begin{aligned} E_t[\mathcal{X}_{t+1}|\mathcal{F}_t, s] &= \mu_{s,t+1} = \Theta_s \lambda_s \\ Var_t[\mathcal{X}_{t+1}|\mathcal{F}_t, s] &= \Sigma_{s,t+1} = \Theta_s^T \Omega_s \Theta_s + \eta_s \end{aligned} \tag{3.8}$$

where λ_s and Ω_s are the state-dependent factor risk premia and factor covariance matrix, given that the market is in regime s, and

$$\eta_s = \begin{pmatrix} \sigma_{1,s}^2 & 0 & \dots & 0 \\ 0 & \sigma_{2,s}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{N,s}^2 \end{pmatrix}_{N \times N}$$

and

$$\sigma_{i,s}^2 = \frac{1}{T} \left(x_i - \left(\alpha_{i,s} + \sum_{k=1}^{K} b_{ik,s} \lambda_{k,s} \right) \right)^2 \tag{3.9}$$

is the idiosyncratic variance corresponding to each stock i in the investable universe, yielded by the RWLS.

In the light we have state-dependent estimates of the stock-level vector of means and covariance matrix, $\mu_{s,t+1}$, and $\Sigma_{s,t+1}$, we can follow Luxenberg and Boyd (2024) and choose the portfolio w that maximizes the expected exponential utility function:

$$U_\varphi(w^T \mathcal{X}) = 1 - \exp\left(-\varphi w^T \mathcal{X}\right) \tag{3.10}$$

where $w^T \mathcal{X}$ is the portfolio return, and φ is a constant risk aversion parameter. The authors argue that $E_t[U_\varphi]$ can be maximized by minimizing the cumulant generating function:

$$\underset{w}{\operatorname{argmin}} \quad \log\left(\sum_{s=1}^{M}\exp\left(\log\hat{\gamma}_{s,t+1} - \varphi\mu_{s,t+1}^{T}w + \frac{\varphi^{2}}{2}w^{T}\Sigma_{s,t+1}w\right)\right) \tag{3.11}$$

where $\hat{\gamma}_{s,t+1}$ is the one-period-ahead forecasted state probabilities of landing in regime s, $\mu_{s,t+1}$ and $\Sigma_{s,t+1}$ state-dependent estimates for the expected return and covariance matrix of these stocks in regime s, given $\mathcal{F}_t$. Finally, the authors mention that Equation (3.11) is not only convex, but also conveniently collapses to the traditional MVO when the number of states is equal to one.

In the next section, we empirically assess the validity and practical benefits of this proposed framework; we turn next to an extensive Monte Carlo simulation exercise. Section 4 describes the design and results of this empirical analysis, demonstrating how effectively the RWLS methodology captures regime dynamics and improves portfolio performance relative to traditional approaches.

3.4 Experiments

This empirical section provides quantitative evidence, derived entirely from Monte Carlo simulations, supporting the advantage of implementing the proposed RWLS methodology for portfolio construction over traditional single-regime approaches. Our analysis is structured as follows: First, we delineate the characteristics of the dataset employed to calibrate the simulation parameters. Second, leveraging this calibration data, we determine the optimal number of distinct market regimes and characterize the joint distribution of the corresponding regime-dependent factor loadings. Third, we generate synthetic security and factor returns through extensive Monte Carlo simulation and conduct a parameter recovery exercise to ascertain the efficacy of retrieving the underlying, known regime-dependent factor loadings from the simulated data streams. Finally, utilizing this controlled, synthetic environment, we rigorously compare the performance metrics of large simulated equity portfolios constructed via the RWLS framework against benchmark portfolios generated using the conventional single-regime methodology, thereby isolating the contribution of explicitly modeling market regimes.

3.4.1 Data

In this chapter, we make use of a dataset comprising common stocks traded in the United States equity markets. The selection of the US market is motivated by its significant size, the large number of listed companies, high liquidity, and extensive data availability. We source monthly adjusted prices for common stocks listed on the NYSE, AMEX, and NASDAQ from the Center for Research in Security Prices (CRSP) database. The sample period spans from July 1963 to December 2024. Additionally, we collect the number of shares outstanding for each stock within our universe to compute daily market capitalization.

As the set of observable processes for our regime-switching model, we adopt the five-factor model proposed by Fama and French (2015). This model extends the seminal three-factor framework of Fama and French (1993), which includes the market excess return ($Mkt - Rf$), size (SMB), and value (HML) factors by incorporating profitability (RMW) and investment (CMA) factors. Furthermore, given its well-documented empirical significance in asset pricing, we augment this set with the momentum factor (UMD). Monthly return data for these six factors, along with the risk-free rate necessary for calculating stock excess returns, were obtained from Kenneth French's data library, covering the period from July 1963 (the earliest common availability for all factors) to December 2024.

The initial data set comprises 3,806 stocks actively traded as of December 2024, having valid adjusted returns and market capitalization data for that date. We apply sequential filtering criteria to refine this sample. Namely, first we exclude stocks with a market capitalization below $300 million to mitigate potential biases associated with micro-cap stocks, such as low liquidity, price distortions, and limited market representativeness. This step reduces the sample to 2,378 stocks. Second, we impose a requirement that stocks must possess at least 20 years of continuous monthly adjusted return data as of December 2024. This criterion ensures the inclusion of multiple economic cycles, including significant recessions and periods of market stress, which is crucial for robust market regime identification. While a more stringent historical data requirement might enhance regime characterization, it would substantially reduce the sample size, potentially compromising the representativeness of the findings. Applying this second filter yields a final sample of 1,132 stocks. Notably, this final sample collectively represents over 90% of the total US stock market capitalization as of the end of the sample period.

3.4.2 Market regime calibration and factor loadings identification

The initial step involves determining the optimal number of market regimes, represented as hidden states within a HMM, that most effectively captures the dynamics inherent in the six-factor model dataset. To achieve this, we employ a grid search procedure, evaluating integer values for the number of states (M) ranging from 1 to 5. Specifying more than 5 regimes is generally considered excessive and lacks substantial support in the existing literature, and introduces significant estimation challenges due to the increased number of parameters relative to the available time-series observations.

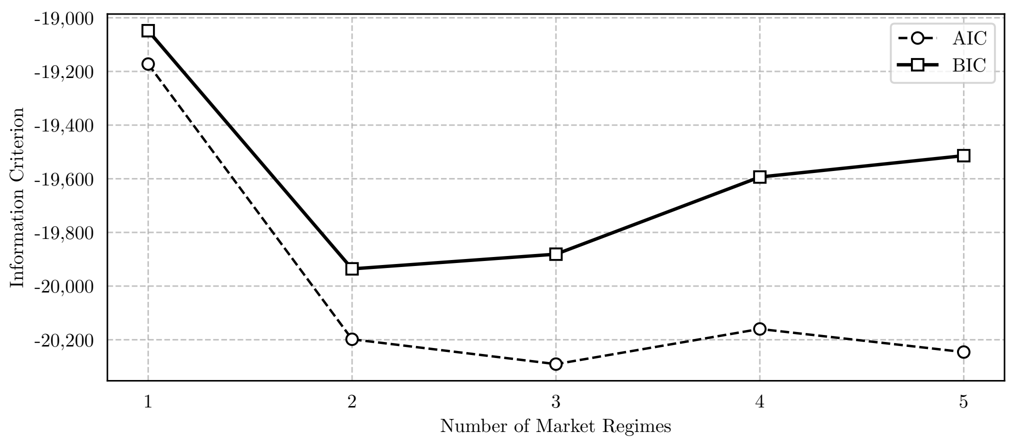
We iteratively fit a multivariate Gaussian HMM to the complete time series of the six factors (sourced from Kenneth French's data library, spanning July 1963 to December 2024), varying the number of hidden states $\boldsymbol{s} \in \{1, 2, \ldots, 5\}$, in each iteration. For each fitted model corresponding to a specific number regime, we compute the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) to assess model suitability. As illustrated in Figure 3.1, analysis of these criteria reveals that the AIC reaches its minimum value when $M = 3$ states are specified, while the BIC is minimized at $M = 2$ states. The divergence between these criteria stems from the BIC imposing a more stringent penalty for model complexity (i.e., the number of parameters).

Based on these findings, and prioritizing model parsimony as indicated by the BIC, we select $M = 2$ as the optimal number of hidden states to characterize the dynamics within our factor sample. This choice aligns with a significant body of financial literature that identifies distinct market regimes, often characterized as periods of high volatility/negative returns versus low volatility/positive returns, influencing asset return behavior. This determined number of states ($M = 2$) remains fixed throughout the subsequent recursive out-of-sample forecasting experiments detailed later in this study.

While not essential for implementing our framework, an economic interpretation of the two regimes may be of interest. One approach is to examine the resulting regime-dependent distributions of each of the six factors, as shown in Table 3.1. We can observe that the $Mkt - Rf$ factor is characterized by a strong positive return, accompanied by low market volatility. On top of the previous, we also see that in this regime the HML factor offers relatively low risk premia, hinting at a negligible difference between the performance of "value" and "growth stocks". Something similar is observed in the CMA factor, where companies that are more "conservative" when it comes to investments offer similar "aggressive" companies. When it comes to the UMD factor, we observe a relatively high Sharpe ratio, as this factor is expected to perform the best in tranquil and persistent environments, as

Figure 3.1: **Determining the number of market regimes**

We plot the information criteria resulting from a search grid designed to find the optimal number of hidden states to use in our exercise. Fitting an HMM on any dataset requires a pre-stablished number of densities to be identified from the set of observable (emmitted) variables as input. Using the training sample window, and the five factor dataset, we iteratively fit an HMM with multivariate Gaussian emissions to a grid of possible integer values ranging from 1 to 5 regimes, and register the Bayesian information criterion (BIC) and the Akaike information criterion (AIC) obtained by using that particular number of regimes. The results suggest that an optimal number of regimes to use for our dataset sits between 2 and 3 hidden states, which seems to be in line with the financial literature.



the momentum signal has a relatively long lookback (12 months, skipping the most recent one). These are characteristic features of positive trending market states, commonly known as *bull regimes*. Conversely, *Market Regime 2* represents the opposite extreme. This regime is characterized by a strong negative return and high volatility in the $Mkt - Rf$ factor. Furthermore, value stocks and companies with conservative and robust profiles outperform their counterparts, as they offer some protection during this turbulent regime and outperform weaker counterparts. The UMD factor exhibits a stark contrast to its performance during Regime 1, which is expected due to "momentum crashes" experienced in highly volatile and negative equity markets, documented by Daniel and Moskowitz (2016). Because of recently described behaviors, we label this negative trending market state as a *bear regime*. This interpretation is further supported by the Markov transition matrix that governs the evolution of the hidden process and the transition between these two regimes over time, illustrated in Panel B of Table 3.1. The two states are highly persistent; Regime 1 has a probability of 95.2% of staying in the same market regime within a month horizon, while Regime 2 has 78.9% of doing the same. The excess persistence of Regime 1 over Regime 2 confirms the intuition built in the previous analysis, as bull regimes tend to be stickier than bear regimes, and by consequence last longer, which is confirmed by empirical data.

3.4.3 Factor loading recovery exercise

A critical component of our empirical validation involves a parameter recovery exercise designed to assess the capability of the RWLS methodology to accurately identify known parameters within a controlled simulation environment. Leveraging the HMM previously calibrated on the extensive historical market dataset, we generate a synthetic dataset comprising 1,000 independent simulation paths. Each path spans 600 observations, representing 50 years of monthly frequency data. For each path, we simulate

Table 3.1: **Regime-conditional distributions**

The table below presents the estimated regime-dependent distributions of six asset pricing factors under a two-state HMM, selected as optimal based on the Bayesian Information Criterion (BIC). Panel A reports the conditional risk premia (λ) and their corresponding standard deviations (σ). Regime 1 is characterized by positive market excess returns and lower volatility, consistent with a "bull market" interpretation, while Regime 2 reflects a high-volatility, negative-return environment, indicative of a "bear market" regime. Panel B displays the estimated one-step-ahead transition probabilities between regimes. Both regimes exhibit high persistence, with Regime 1 showing a notably higher self-transition probability (95.2%) compared to Regime 2 (78.9%), aligning with empirical evidence that bull markets tend to be more enduring than bear markets.

Panel A: Regime-Conditional Factor Distribution				
	Regime 1		Regime 2	
	λ	σ	λ	σ
Mkr-Rf	9.57%	12.68%	-4.35%	24.42%
SMB	1.57%	8.63%	5.90%	17.06%
HML	1.90%	7.79%	9.64%	18.09%
RMW	3.32%	5.22%	3.85%	14.70%
CMA	1.34%	5.68%	10.75%	12.14%
UMD	10.22%	9.64%	-5.45%	27.03%

Panel B: Transition Matrix		
	To	
From	Regime 1	Regime 2
Regime 1	95.2%	4.8%
Regime 2	21.1%	78.9%

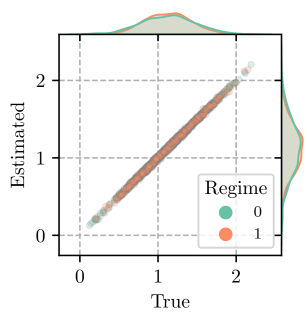
sequences of regime-smoothed probabilities corresponding to the two market regimes identified in the prior section, ensuring consistency with the transition dynamics empirically observed in the six-factor dataset. Concurrently, we simulate the factor returns for the identical 600-observation period. To establish the "ground truth" parameters for this exercise, we draw specific vectors of regime-dependent factor loadings for each simulated security from the empirical joint distribution characterized earlier. These known, true loadings are then combined with the simulated factor returns to generate synthetic individual stock return series across the full 600-observation horizon for each path. Subsequently, utilizing only the simulated factor returns and the derived synthetic stock returns as inputs mimicking the data available in a practical estimation setting, we apply the RWLS procedure to infer the latent regime states and estimate the corresponding stock-level regime-weighted factor loadings inherent in the simulated data. Our findings indicate that, contingent upon the assumption that the data generating process aligns with the HMM specification, the RWLS methodology successfully retrieves the underlying regime-weighted factor loadings. The accuracy of this retrieval process is visually substantiated in Figure 3.2, which depicts the close alignment between the estimated loadings and their known ground truth counterparts across the simulations.

3.4.4 Portfolio simulation

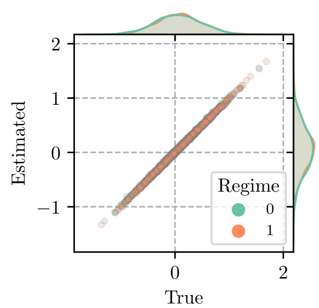
Evaluating the empirical efficacy of any investment framework necessitates a rigorous simulation environment that closely mimics real-world conditions. As discussed in Bailey et al. (2014), a backtest is only considered meaningful when in-sample performance is consistent with out-of-sample behavior, cautioning against the ease of overfitting due to data snooping and parameter tweaking. López de Prado (2016, 2018) formalizes this concern by introducing a Monte Carlo simulation-based methodology, commonly referred to as "backtesting on synthetic data", where strategy robustness is assessed across

Figure 3.2: **Monte Carlo simulation results assessing parameter recovery accuracy**

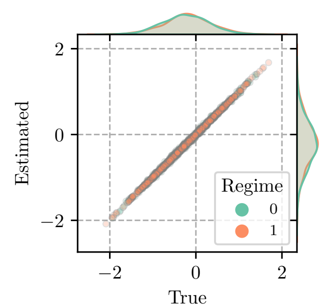
Subplots show estimated versus true values for the parameters associated with the six factors: (a) Mkt-Rf, (b) SMB, (c) HML, (d) RMW, (e) CMA, and (f) UMD. Data points are generated under two distinct regimes (Regime 0 and Regime 1). The close alignment to the identity line and congruent marginal kernel density estimates validate the estimation procedure's effectiveness for all factors across both regimes.



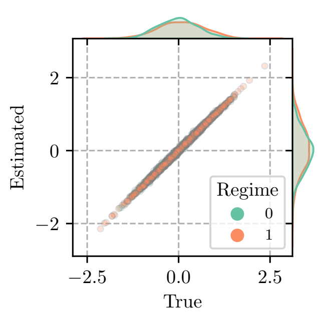
(a) Mkt-Rf



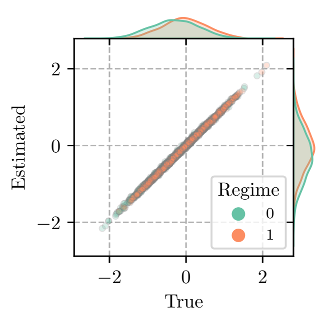
(b) SMB



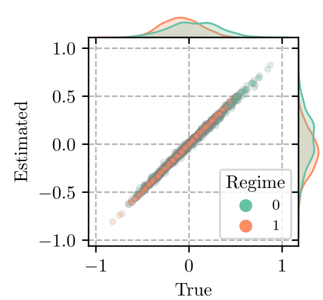
(c) HML



(d) RMW



(e) CMA



(f) UMD

numerous return paths generated from known stochastic processes. More recently, Ibanez (2023) applies a similar simulation framework using stochastic differential equations to examine the resilience of diversified portfolios under severe market conditions. Motivated by these studies, we implement a controlled Monte Carlo simulation exercise to benchmark the performance of our proposed RWLS framework.

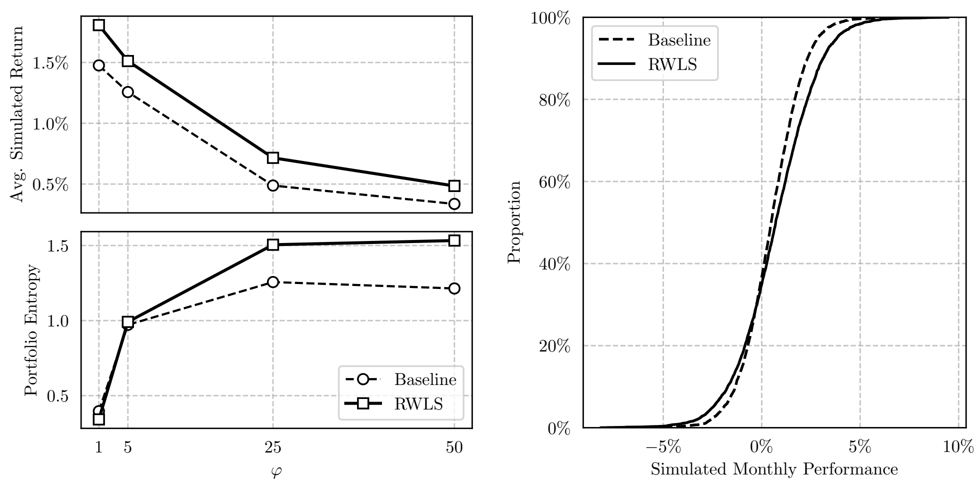
Assuming that equity-market behavior can be represented by a multivariate Gaussian HMM, we employ the model fitted in the factor-loading recovery exercise to carry out 5,000 Monte Carlo simulations to test the efficacy of our model while managing a large portfolio of 500 stocks. For each trial, the HMM first draws a pair of successive hidden-state realizations following the calibrated one-step transition matrix. The first state identifies the regime prevailing at the rebalancing date (t), whereas the second captures the regime anticipated to govern returns over the following one-month holding period ($t+1$). Conditional on this regime sequence, the HMM simulates contemporaneous paths for the six systematic risk factors, ensuring that the draws respect the regime-specific multivariate distribution inferred from historical data. In parallel to the factor simulation, we sample vectors of regime-dependent factor loadings for each of the 500 constituent securities from the joint distributions recovered in the earlier estimation exercise. This step is repeated for every replication, so that each Monte Carlo path is paired with an internally consistent cross-section of 3,000 factor loadings (500 stocks by 6 factors), all drawn from the empirically estimated regime-conditional distributions, just as in the previous section. Combining these loadings with the simulated factor returns corresponding to $t+1$ yields synthetic, regime-aware one-month excess-return vectors for the full cross-section of 500 hypothetical stocks. The resulting data-generating process provides a controlled environment in which to benchmark the proposed RWLS framework against a single-regime baseline, thereby isolating the incremental benefit of explicitly modeling regime dynamics when constructing large-scale equity portfolios.

In each of these trials, we construct a regime-aware portfolio by solving (3.11). We estimate the one-period-ahead forecasted state probabilities of landing in regime s, $\hat{\gamma}_{s,t+1}$, by premultiplying the regime transition matrix by the first hidden-state realization; the prevailing market regime at the rebalancing date. Instead of choosing only one constant risk aversion parameter, we define a grid of possible values between 1 and 50, to have a comprehensive view of the resulting portfolios. As a baseline comparison, we define a regime-agnostic framework by ignoring the regime dynamics and assuming unconditional factor loadings and factor distributions. To obtain these unconditional estimates, we collapse the pairs of regime-aware loadings and risk premia distribution parameters using the limiting distribution of the Markov chain (i.e., $\lim_{n\to\infty} \Pi^n$). As previously mentioned, solving the problem (3.11) for the case of a single regime is equivalent to solving a traditional mean-variance optimization problem. In the case of both the RWLS framework and the baseline comparison, we impose a long-only constraint (i.e., $\forall w \geq 0$).

Figure 3.3 presents the results of a Monte Carlo simulation exercise designed to assess the out-of-sample performance of the proposed RWLS portfolio construction methodology relative to a regime-agnostic baseline. The simulation comprises 5,000 trials of a portfolio containing 500 stocks. Panel (a) demonstrates that the RWLS framework achieves higher average simulated returns across all levels of risk aversion considered, with the performance differential being most pronounced at lower levels of the constant risk aversion parameter φ. In addition to improved returns, the RWLS methodology consistently produces portfolios with higher entropy, reflecting more diversified allocations under the Shannon entropy measure. Panel (b) illustrates the empirical cumulative distribution functions of monthly portfolio returns for both methods at a fixed risk aversion level of $\varphi = 25$. The RWLS distribution stochastically dominates the baseline, and a one-tailed Kolmogorov–Smirnov test confirms that this improvement is statistically significant at the 1% level.

Figure 3.3: **Monte Carlo Simulated Portfolio Performance**

The figure below summarizes the results of a Monte Carlo simulation study evaluating portfolio performance under the proposed Regime-Weighted Least Squares (RWLS) framework relative to a regime-agnostic baseline. The analysis is based on 5,000 simulated trials of a portfolio composed of 500 stocks. Panel (a) reports average simulated returns (top) and portfolio entropy (bottom) across varying levels of the risk aversion parameter (φ). The RWLS approach consistently outperforms the baseline in terms of expected return, while also yielding portfolios with higher entropy, indicating greater diversification as measured by Shannon entropy of the portfolio weights. Panel (b) displays the cumulative distribution function (CDF) of simulated monthly portfolio returns for both methods under a fixed risk aversion level ($\varphi = 25$). The empirical CDF for RWLS stochastically dominates that of the baseline, and a one-tailed Kolmogorov–Smirnov test confirms with 99% confidence that the RWLS methodology significantly improves performance.



(a) Simulated Portfolio Performance

(b) Performance CDF ($\varphi = 25$)

These results also hold for every value tested for the risk aversion parameter.

These findings suggest that incorporating regime information into the portfolio optimization process yields more robust and diversified outcomes, particularly in risk-averse settings.

3.5 Conclusions

This study introduces a novel RWLS methodology designed for estimating conditional means and covariance matrices in large-scale stock portfolios through the integration of market regime dynamics using a HMM. The proposed approach addresses significant limitations inherent in traditional mean-variance optimization, particularly those associated with dimensionality and estimation complexity in constructing extensive regime-aware equity portfolios.

Our empirical results, derived from rigorous out-of-sample forecasting experiments through comprehensive Monte Carlo simulations, demonstrate that portfolios constructed using the RWLS framework consistently outperform traditional regime-agnostic strategies. Specifically, our approach achieves better risk-adjusted returns and enhanced diversification, reflected in higher portfolio weights entropy across varying levels of investor risk aversion.

By effectively capturing regime-dependent behaviors of systematic risk factors, our framework significantly contributes to practical asset management, enabling quantitative asset managers and hedge funds to systematically adapt their portfolio strategies to prevailing market conditions. Furthermore, our results underscore the importance of explicitly modeling regime dynamics, which leads to materially improved investment performance, particularly during periods characterized by market stress and volatility.

While our findings provide compelling evidence in favor of incorporating regime-awareness into portfolio management, several avenues remain open for future research. Particularly promising is the integration of endogenous regime-switching models that respond dynamically to market innovations. Further exploration of regime characteristics across different economic cycles and markets could enhance the robustness and applicability of the framework, expanding its relevance across diverse investment contexts. This ongoing research promises further valuable insights and improvements in the strategic management of large-scale investment portfolios. This is part of an ongoing research agenda.

Chapter 4

Optimal Investment Signal Combination in Systematic Equity Portfolios: A Beamforming Approach

4.1 Introduction

The investment management industry, specially in the case of equity-focused funds, has experienced a significant transformation since the turn of the millennium. Previously, investment decisions relied heavily upon individual experience, skill, and intuition. In recent years, however, these subjective methods have increasingly been supplanted by rigorous, scientific, and replicable processes characterized by the systematic design and implementation of rules-based investment strategies. These contemporary strategies primarily rely on empirical evidence rather than traditional seasoned judgment.

The widespread adoption of systematic investment methodologies is noteworthy. Harvey et al. (2017) estimated that, by 2014, approximately 26% of total assets under management (AUM) were already systematically managed. It is reasonable to infer that this proportion has grown substantially over the past decade, given the significant expansion observed among hedge funds; the primary users of systematic strategies. Indeed, hedge fund AUM has dramatically increased from roughly $200 billion around the year 2000 (Sullivan, 2020) to over $5 trillion nowadays. This rapid expansion highlights the growing preference for systematic approaches within the industry, underscoring their effectiveness, scalability, and appeal to institutional investors.

In quantitative investment funds, researchers identify, distill, and curate certain stock characteristics (e.g., book-to-price, market capitalization, etc.) that are statistically proven to be positively correlated with the future excess return of the stock. Such characteristics, also called investment signals, typically come either from peer-reviewed journals or in-house research. Once these individual signals are identified, a common practice in the systematic equity trading industry is to aggregate them into one single score (Novy-Marx, 2015), and based on this score a portfolio is constructed where stocks with a high aggregate signal are bought (or overweighted with respect to a given benchmark) while stocks scoring relatively low on the aggregate signal are sold (or underweighted).

The academic literature in finance has witnessed a proliferation of factors purported to explain the cross-section of expected stock returns, a phenomenon often described as a "factor zoo" (Cochrane, 2011). The sheer volume of discoveries is extensive, with Harvey et al. (2016) documenting over 300 significant return predictors. This figure has continued to grow since, encompassing over 450 anomalies in comprehensive replications like that of Hou et al. (2020). This rapid explosion of factors has led to significant concerns about data mining; the implications of this intense search for significant characteristics, together with the well-known publication bias effect, make it likely that many of these findings are spurious. Indeed, Harvey et al. (2016) suggests that most claimed research findings in financial economics are likely false after adjusting for multiple testing. Underscoring the vastness of the search space, Yan and Zheng (2017) constructed a universe of over 18,000 potential fundamental signals from financial statements, finding that many still exhibited predictive power even after accounting for overfitting. A crucial issue complicating this landscape is that these numerous factors are not independent; many are correlated with one another, a point emphasized by Arnott et al. (2019). However, this correlation does not necessarily imply redundancy. Kozak et al. (2020) argues that there is not enough redundancy among the dozens of known anomalies to allow for a parsimonious, characteristics-sparse factor model to adequately summarize the cross-section of returns, challenging the sufficiency of popular three-, four-, or five-factor models.

Given this crowded and correlated factor landscape, the challenge for portfolio construction becomes paramount. The dubious statistical significance of many signals, highlighted by the high t-statistic hurdles proposed by Harvey et al. (2016) and the widespread replication failures documented by Hou et al. (2020), means that many factors have overstated historical performance that is unlikely to persist. Furthermore, even for factors that appear robust, their return streams are often characterized by significant noise and non-normality, including fat tails and larger-than-expected drawdowns (Arnott et al., 2019). This combination of issues, correlation that undermines naive diversification, questionable predictive power, and adverse tail risk, makes the method by which signals are combined critically important. A simple approach of selecting a few historically successful factors is likely to be suboptimal. Instead, as suggested by Kozak et al. (2020), robust methods such as imposing economically motivated priors that shrink the contributions of noisy signals are required to distill the true predictive information from the vast and complex cross-section of factors. The focus thus shifts from discovering individual factors to developing a robust stochastic discount factor that can effectively combine information from a large set of predictors.

While the optimal method for combining signals is an ongoing debate among academics and financial engineers, the challenge of aggregating noisy signals has long been a central focus of other disciplines, notably signal processing and electrical engineering. Feng and Palomar (2016) argue that while financial engineering and electrical engineering may appear to be unrelated fields, they are built upon the same mathematical foundations. The core of both disciplines relies on the statistical modeling and analysis of systems to make predictions and optimize strategies. Financial engineering focuses on the statistical analysis of numerical time series to model the behavior of financial markets, which allows for the systematic optimization of investment strategies. In a similar vein, electrical engineering, particularly in areas like wireless communications, employs statistical signal processing to model communication channels to optimize transmission strategies. This fundamental parallel suggests that optimizing an investment strategy in a financial market is conceptually equivalent to optimizing a signal transmitted from an antenna. The recognition of this shared mathematical framework opens up the possibility for both fields to benefit from methodologies that were often developed independently (Feng and Palomar, 2016). For instance, the Autoregressive Moving Average (ARMA) model, a popular tool for modeling financial time series (Tsay, 2010), is also a foundational rational or pole-zero model in signal processing (Manolakis et al., 2005). In the realm of risk management, robust covariance matrix estimation is critical. The financial practice of "shrinking" a sample covariance matrix (Ledoit and Wolf, 2004) is mathematically identical to the "diagonal loading" technique used for robust adaptive beamforming in signal processing for decades (Abramovich, 1981, Carlson, 1988, Cox et al., 1987). Furthermore, the problem of designing a minimum variance portfolio (Markowitz, 1952) is mathematically identical to the design of a filter or beamformer in signal processing (Monzingo et al., 2011; Zhang et al., 2013).

This chapter bridges an interdisciplinary gap by applying a proven methodology from signal processing to address this central problem in quantitative finance. We argue that the challenge of combining noisy investment signals is mathematically equivalent to the problem of combining signals from an antenna array in wireless communications (as in Feng and Palomar, 2016). We implement adaptive beamforming, a methodology used by electrical engineers for decades, as a robust framework for optimally combining investment signals. This approach moves beyond static weighting schemes to a dynamic system that maximizes the portfolio's signal-to-noise ratio. The remainder of this chapter is organized as follows: Section 3.2 reviews the academic literature on combining investment signals and the challenges posed by the "factor zoo". Section 3.3 introduces our proposed methodology, establishing the theoretical link between financial engineering and signal processing and detailing the multi-stage beamforming framework for signal combination. Section 3.4 presents a comprehensive empirical backtest of this framework, where we apply the methodology to a set of seven distinct value and momentum signals using historical U.S. equity data. We compare its performance against a baseline strategy that reflects common industry practice. Finally, Section 3.5 concludes and discusses the implications of our findings for systematic portfolio managers.

4.2 Related literature

In empirical asset pricing, factors are systematic drivers that explain the co-movement and expected returns of securities. This concept, rooted in the Arbitrage Pricing Theory of Ross (1976), posits that the cross-section of stock returns is driven by securities' covariances with these underlying factors. In practice, factor portfolios are constructed not by directly measuring these abstract betas, but by sorting stocks based on observable firm characteristics. The most widely accepted factors, which have demonstrated persistent abnormal returns over long periods, include size (market capitalization), value (e.g., book-to-market), momentum (past returns), profitability (e.g., return on equity), and low volatility/low beta.

The persistent outperformance associated with these characteristics is often referred to as investment *anomalies*, as they cannot be explained by traditional general equilibrium-based models such as the Capital Asset Pricing Model (CAPM), and has provided considerable evidence against the traditional efficient market hypothesis. For example, value stocks are defined as those with high fundamentals-to-price ratios, and the extensive literature on the value anomaly shows that these stocks have historically tended to outperform glamour stocks, which exhibit low fundamental-to-price ratios. These findings led to the development of multifactor asset pricing models, most notably the Fama and French three-factor model (Fama and French, 1993) and the later five-factor model (Fama and French, 2015). These models have become standard benchmarks for evaluating investment performance. However, even these influential models face scrutiny. Blitz et al. (2018) identify several concerns with the five-factor model, including its retention of the disputed CAPM relation between beta and return, its inability to explain the momentum premium, and the questionable robustness of its new factors.

Over recent decades, there has been a dramatic proliferation in the number of factors and anomalies identified in asset pricing literature, with discoveries accelerating sharply to approximately eighteen new factors per year recently, up from just one per year in earlier periods (Harvey et al., 2016). This rapid growth, driven by enhanced data availability, computational capacity, and extensive data mining, poses significant statistical and practical challenges. Firstly, replication issues have become acute, as evidenced by Hou et al. (2020), who found that around 65% of 452 tested anomalies failed rigorous statistical replication procedures, with the failure rate rising to approximately 82% after correcting for multiple testing. Moreover, Arnott et al.

(2019) emphasized practical pitfalls such as factors losing their profitability due to crowding, data mining-induced overfitting, and the complexity of managing nonnormal distributions and time-varying correlations, especially in stressed market conditions. Additionally, Yan and Zheng (2017) documented pervasive data mining within fundamental signals, underscoring the difficulty in distinguishing genuine predictors from random chance. Addressing this complexity, Kozak et al. (2020) advocated methodologies like shrinkage techniques and principal component analysis, aiming to effectively reduce the dimensionality of this growing *factor zoo* (Cochrane, 2011). Consequently, the exponential growth in discovered factors necessitates increasingly rigorous empirical validation and innovative analytical methods to ensure robust and economically meaningful investment strategies.

While individual factors have historically earned return premiums, they have all experienced periods of underperformance, and crucially, these periods have not always occurred at the same time. This cyclicality creates an opportunity for diversification. By combining multiple factors, investors can create more stable portfolios and potentially achieve more consistent returns. Furthermore, a simple single-factor approach can lead to significant and often undesirable exposures to other factors. A portfolio targeting one factor in isolation may inadvertently hold securities with low or even negative scores on other important factors. For example, Blitz and Vidojevic (2019) shows that generic factor strategies often invest a substantial portion of their portfolio (around 20% or more) in stocks with negative implied market-relative returns due to poor characteristics on other factors. Small-cap stocks, for instance, are often "junky" and have poor quality characteristics that detract from their returns. A multifactor approach helps mitigate this dilution, as it accounts for the cross-sectional interactions of factors. Combining factors can also have other benefits due to the inadvertent interaction between characteristics. Asness et al. (2018) argues that the well-documented weakness of the size premium is an artifact of the strong negative correlation between firm size and quality. The authors show that by controlling for a firm's "junk" characteristics, a significant and robust size premium that is stable over time and across international markets is resurrected.

The inherent low correlation between these factors and investment anomalies offers interesting avenues to construct efficient stock portfolios. Combining multiple factors provides considerable diversification benefits, reducing vulnerability to factor-specific underperformance (Bender and Wang, 2016). Such diversification allows multifactor portfolios to offer higher risk-adjusted returns compared to single-factor strategies. Firoozye et al. (2023) highlights the value of combining signals by illustrating how Canonical Correlation Analysis (CCA) helps identify optimal combinations of assets and predictive signals, improving performance and offering theoretical grounding for practical implementation. Furthermore, Pätäri et al. (2018) argues that multicriteria decision-making methods effectively aggregate various value and momentum indicators, enhancing portfolio stability and performance.

The literature on combining investment signals and factors into multifactor portfolios highlights two primary methodological approaches: portfolio blending (top-down) and signal blending (bottom-up), each with distinct theoretical foundations and practical implications. Clarke et al. (2016) assert that portfolios built directly from individual securities (signal blending) generally exhibit greater mean-variance efficiency compared to portfolios constructed from factor subportfolios (portfolio blending). They demonstrate empirically that signal blending captures a larger portion of the potential improvement in the Sharpe ratio relative to the market portfolio, attributed primarily to broader latitude in leveraging cross-sectional variation in factor exposures. Conversely, Ghayur et al. (2018) suggests a nuanced view, finding that portfolio blending provides superior risk-adjusted returns at low-to-moderate levels of tracking error, whereas signal blending outperforms at higher tracking error levels. This finding implies practical considerations for investors depending on their risk tolerance and investment constraints. Leippold and Rueegg (2018), however, present a skeptical perspective, suggesting that prior findings favoring integrated (signal blending) approaches over mixed (portfolio blending) approaches might be statistical artifacts. Their rigorous analysis indicates negligible empirical differences between these methods once robust testing frameworks are employed. Lester (2019) provides theoretical support for signal blending, demonstrating analytically that bottom-up multifactor portfolios can significantly outperform top-down constructions, especially when factors are orthogonal or exhibit low correlations. This theoretical underpinning is supported by empirical results indicating higher expected returns and improved information ratios from bottom-up portfolios. Amenc et al. (2017, 2018) critically evaluate both methodologies, emphasizing the merits of simplicity, transparency, and ease of attribution in top-down approaches. They highlight that bottom-up portfolios, despite potentially higher factor exposures, often incur higher implementation costs due to greater turnover and concentration risks. Nevertheless, they acknowledge that a carefully executed top-down method that accounts for cross-factor interactions can deliver comparable efficiency. Finally, Bender and Wang (2016) strongly advocates for bottom-up multifactor portfolio construction, arguing that it captures beneficial interactions at the security level that top-down methods might overlook. Their empirical results support the bottom-up approach, indicating meaningful enhancements in portfolio efficiency arising from security-level interactions of factors.

One of the ongoing debates when it comes to combining signals is the case of negatively correlated factors. Value and momentum are two prominent anomalies in financial markets, each individually providing robust return premiums. However, combining them effectively in portfolio strategies presents distinct challenges due to their consistently observed negative correlation. Early foundational work by Asness (1997) highlights this critical interaction, demonstrating that value strategies perform strongest among stocks with weak recent performance (low momentum), whereas momentum strategies excel particularly among expensive (low-value) stocks. This fundamental interaction requires careful consideration when these strategies are combined. Further research supports this negative correlation and underscores its implications. Asness et al. (2015) reiterates that integrating these two factors into a single portfolio can significantly enhance overall risk-adjusted returns by harnessing their diversification benefits. However, Grobys and Huhta-Halkola (2019) emphasize that the negative correlation between value and momentum appears predominantly driven by growth stocks, making strategic integration nuanced. They find notably similar returns between winner stocks that are value or growth stocks, implying that the negative correlation is not uniform across all market segments. The combination methodologies themselves introduce further complexity. Fisher et al. (2016) and Cooper and Jiao (2024) delve into the operational details, demonstrating that simultaneous integration of value and momentum signals at the security level, rather than combining separate portfolios afterward, significantly reduces transaction costs and leverages unfavorable signals more efficiently. Particularly, Fisher et al. (2016) find that directly incorporating momentum and value signals in a single portfolio not only enhances cost efficiency but better utilizes available information, highlighting the advantage of a more integrated approach over independent strategies. Addressing risk, Barroso and Santa-Clara (2015) highlight the inherent volatility and crash risk associated with momentum strategies. They suggest that volatility-targeted (risk-managed) momentum strategies significantly mitigate momentum's severe downturns, thus potentially stabilizing the combination of momentum with value, which typically exhibits lower volatility. Li (2018) further provides theoretical grounding by proposing a unified, investment-based framework explaining how distinct risks underlie momentum and value strategies. This research indicates that momentum captures short-term productivity shocks, whereas value strategies reflect longer-term risks, thus justifying their negative correlation and advising nuanced integration. Moreover, Pani and Fabozzi (2021) introduce an innovative approach, incorporating trends within fundamental metrics, which enhances the value component when combined with momentum, again underscoring the importance of nuanced methodological details. Additionally, Wang and Kochard (2012) propose employing a z-score approach for blending momentum and value, emphasizing the need for precise methodologies to navigate their negative interaction effectively. Effectively combining negatively correlated stock characteristics such as value and momentum strategies is intricate, demanding careful consideration of their negative correlation, volatility management, methodological integration, and theoretical underpinnings. Successful integration requires sophisticated techniques, risk controls, and a nuanced understanding of how these factors interplay across different market environments and stock characteristics.

The literature above illustrates substantial debate and complexity in integrating multiple investment signals, particularly when signals exhibit negative correlation, such as value and momentum. Research demonstrates theoretical and empirical advantages in the bottom-up signal blending approach, especially due to enhanced factor exposures and reduced transaction costs. Nevertheless, achieving efficient integration is challenging, particularly with negatively correlated factors, due to potential dilution effects and increased volatility. In this study, we extend the existing literature by exploring efficient bottom-up signal blending methodologies through the application of techniques commonly employed in the signal processing field of electrical engineering. Our objective is to optimize the combination of investment signals, managing their intrinsic correlations and volatility dynamically, thus enabling robust portfolio construction even in the presence of strongly negatively correlated factors.

4.3 A signal processing framework for investment signals

4.3.1 From financial engineering to electrical engineering: establishing the analogy

While the challenges inherent in systematic investing, such as identifying predictive signals, managing risk, and constructing optimal portfolios, are often considered unique to finance, the underlying mathematical problems are not. Other quantitative disciplines, notably signal processing within electrical engineering, have long focused on a conceptually identical task: extracting a desired signal from a noisy, multi-source environment. As Feng and Palomar (2016) argues, financial engineering and signal processing, while appearing to be unrelated fields, are built upon the same mathematical foundations, allowing for a powerful cross-pollination of ideas and techniques.

The core of both disciplines relies on the statistical modeling of complex systems to make predictions and optimize strategies. This fundamental parallel is not merely abstract; it manifests in the direct equivalence of specific, foundational models used in both fields. The parallels highlighted by Feng and Palomar (2016) are shown in Table 4.1, which notably include:

- **Time-series modeling:** The Autoregressive Moving Average (ARMA) models, which are standard tools for modeling financial time series (Tsay, 2010), are mathematically identical to the *pole-zero* models that form the basis of filter design in signal processing (Manolakis et al., 2005).

- **Robust risk management:** In finance, robust covariance matrix estimation is critical for portfolio construction. The common practice of *shrinking* a sample covariance matrix toward a more structured target to improve its stability, popularized by Ledoit and Wolf (2004), is a technique known as *diagonal loading* in signal processing, where it has been used for decades to achieve robust adaptive beamforming (Carlson, 1988; Cox et al., 1987).

- **Portfolio optimization:** Most directly, the problem of designing a minimum-variance portfolio as formulated by Markowitz (1952) is mathematically equivalent to the design of an optimal filter or beamformer in signal processing (Monzingo et al., 2011). Both seek to find an optimal set of weights to combine multiple input streams to minimize variance (noise) for a given level of signal strength (return).

This shared mathematical framework allows us to re-conceptualize the central task of a quantitative portfolio manager. In this view, optimizing an investment strategy is analogous to optimizing a signal received by an array of antennas. Individual investment factors (e.g., Value, Momentum) act as the separate antennas, each receiving a noisy version of the desired information (alpha). The portfolio construction process, therefore, becomes a *beamformer*, an algorithm designed to intelligently combine these inputs to amplify the true signal while canceling out the uncorrelated noise and interference. Recognizing this shared foundation opens the door for finance to benefit from robust methodologies that have been developed and refined independently over decades in engineering.

4.3.2 Investment signals as noisy data streams

In electrical engineering, a signal is defined as a function of one or more independent variables that conveys information about a physical phenomenon, encompassing everything from speech and audio to biomedical, seismic, and radar signals. These phenomena typically manifest as variations over time or space and can be represented in various mathematical forms, including discrete, continuous, or digital signals (Sundararajan, 2023, Anand, 2022).

Table 4.1: **Financial engineering and signal processing** In Feng and Palomar (2016), the authors show that both fields focus on extracting signals from noisy data. In finance, this involves detecting trends in unstable market data. This similarity enables interdisciplinary applications, where signal processing techniques such as beamforming, filter design, and random matrix theory can be applied to financial challenges like risk management, statistical arbitrage, and market impact modeling, potentially enhancing the accuracy and efficiency of financial strategies.

Connections between financial engineering and signal processing		
	Financial Engineering	**Signal Processing**
Modeling	Autoregressive moving average (ARMA)	Rational/Pole-zero model
Robust covariance matrix estimation	Shrinkage sample covariance matrix estimator	Diagonal loading in beamforming
Asymptotic analysis	Large-dimensional general asymptotics	Random matrix theory
Optimization	Mean-variance portfolio optimization	Filter/Beamforming design
Sparsity	Index tracking	Sparse signal recovery

However, signals encountered in practical applications are often contaminated by noise, unwanted disturbances that obscure or alter the intended information. Noise can manifest as random variations or structured interference, often reducing the effectiveness of communication and measurement systems. Sources of noise are diverse and can include thermal fluctuations in electronic components (Johnson-Nyquist noise), electromagnetic interference from external devices, environmental conditions such as atmospheric disturbances, mechanical vibrations, and inherent physical limitations of measurement systems such as sensor inaccuracies or quantization errors (Tan and Jiang, 2019). Understanding and mitigating these noise sources is crucial to enhancing the fidelity and reliability of signals across various engineering and technological applications.

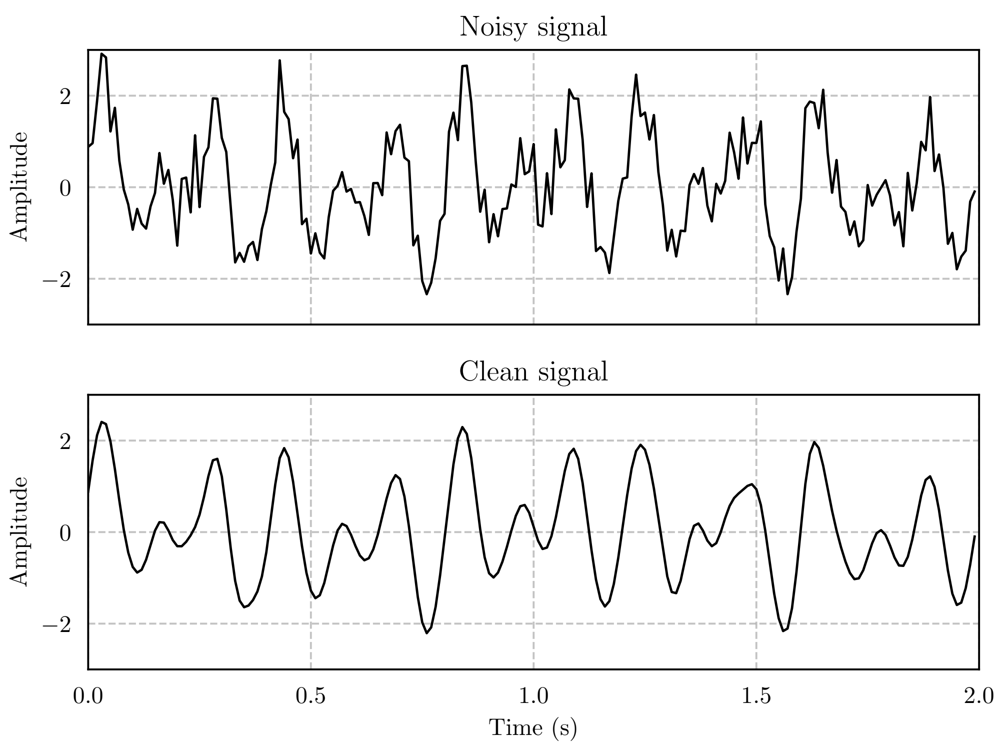
Consequently, the quality of a signal processing system is indicated by the Signal-to-Interference-and-Noise Ratio (SINR); defined as the ratio between the power of the desired signal and the power of background noise (Sundararajan, 2023), which is usually measured in decibels (dB). A higher SINR implies a clearer and more distinguishable signal, directly correlating with improved performance in communication systems, signal detection, and accurate information retrieval. This makes SINR an essential parameter in optimizing signal processing techniques for noise reduction, ensuring the reliability and efficiency of various electronic and digital systems.

A visual representation of this core challenge is provided in Figure 4.1. The top panel illustrates a raw, noisy signal, analogous to the return stream of a single investment factor where the underlying predictive information is obscured by high-frequency volatility and random market fluctuations. The bottom panel shows the same signal after a filtering process has been applied, stripping away the noise to reveal the cleaner, more fundamental waveform. This act of isolating a true signal from a contaminated data stream is the fundamental objective of signal processing and serves as the conceptual basis for our portfolio construction methodology.

In the landscape of financial engineering, an investment signal, often referred to as an alpha signal, is an empirical regularity in which a firm's observable characteristics are statistically correlated with its future excess stock returns. These signals represent market anomalies that are not explained by classical asset pricing models like the Capital Asset Pricing Model (CAPM). The vast majority of these signals operate cross-sectionally, a crucial concept meaning that the absolute value of a characteristic is less important than its rank or value relative to other companies in the investable universe. For

Figure 4.1: **Illustration of a noisy signal and its filtered counterpart**

This figure visually represents the core challenge in signal processing. The top panel shows a "noisy signal," where a clean, underlying waveform is contaminated by random, high-frequency disturbances. This is analogous to a raw investment factor's return series, which contains both a predictive "signal" (alpha) and significant market "noise." The bottom panel displays the "clean signal" after a filtering process has been applied to remove the unwanted noise, revealing the core informational waveform.



instance, in the case of the value anomaly, a company with a book-to-market ratio of 0.5 might be considered a value stock in a growth-dominated market environment, yet it could be classified as a growth stock during a period of deep economic distress where average valuations are much lower. This relativity is precisely why quantitative practitioners typically de-mean and re-scale raw characteristic data through a process known as z-scoring. This transformation converts the often-meaningless absolute value into a standardized, relative score, which serves the dual purpose of identifying a stock's standing within its peer group and creating a common unit of measurement, allowing for the direct comparison and subsequent combination of heterogeneous signals, such as value and momentum. However, this common practice is a linear and static transformation that implicitly assumes each normalized signal contributes equally and independently to a final composite score, an assumption that fails to capture the complex, dynamic interplay between factors.

The academic literature is replete with hundreds of potential investment signals, creating a "factor zoo" that presents both opportunities and challenges for investors (Cochrane, 2011). The sheer volume of discoveries has led to valid concerns about data mining, with many findings failing to replicate out of sample or requiring an exceptionally high statistical hurdle (e.g., a t-statistic greater than 3.0) to be considered significant (Harvey et al., 2016). While a comprehensive cataloging of this factor zoo is beyond the scope of this chapter, we can classify most robust signals into two intuitive categories based on their underlying economic rationale. The first category is "convergency" signals, which identify mispricings that are expected to correct over time, pointing to a company's market price moving closer to an estimate of its intrinsic value. Factors such as book-to-market, earnings-to-price, and dividend yield fall into this category, as they are based on the premise that market prices will eventually revert to a mean dictated by fundamental value. The second category is "divergency" signals, which capture trends or behavioral biases that are expected to persist, pointing to a stock's price continuing to move away from a prior reference point. Price momentum is the canonical example, betting on the continuation of a recent trend often attributed to investor underreaction to information. Other signals, like asset growth, can also be viewed as divergencies, as they often capture extrapolative expectations from investors. This taxonomy is useful because these two types of signals are often driven by different economic and behavioral forces, leading them to have distinct risk profiles and, frequently, a negative correlation.

A critical property of these signals is that a company's exposure to them is dynamic and changes over time. Figure 4.2 illustrates this for Walmart Inc. from 1985 to 2025. The company's book-to-market ratio fluctuates through long cycles; it entered a period of deep relative undervaluation in the early 2000s before its market capitalization surged, driving the ratio down into "growth" territory for much of the subsequent decade. A similar pattern is observed for its price momentum, which experiences periods of strong, persistent price appreciation followed by corrections. This cyclicality is not random noise but a structured, periodic behavior, directly analogous to the sinusoidal waveforms that represent acoustic or electromagnetic signals, as depicted in Figure 4.1. Given the documented association of these signals with future returns, a stock's performance is expected to be strongest when its signal is at a peak. However, the critical challenge for portfolio construction is that these signal "waveforms" are often out of phase. As Figure 4.2 clearly shows, the peaks for Walmart's value and momentum signals often occur at different times, a firm-level example of the well-documented negative correlation between these two factors. A naive, static combination method, such as simple averaging, would see these powerful but opposing signals effectively neutralize each other during extended periods. Therefore, any robust portfolio construction methodology must move beyond simple signal aggregation and instead employ a framework that can explicitly model and adapt to these complex temporal and cross-serial correlations. It must behave less like a simple mixer and more like a sophisticated signal processor, capable of understanding when signals are reinforcing and when they are interfering. This is precisely the problem that adaptive beamforming is designed to solve.

4.3.3 The beamforming approach to signal combination

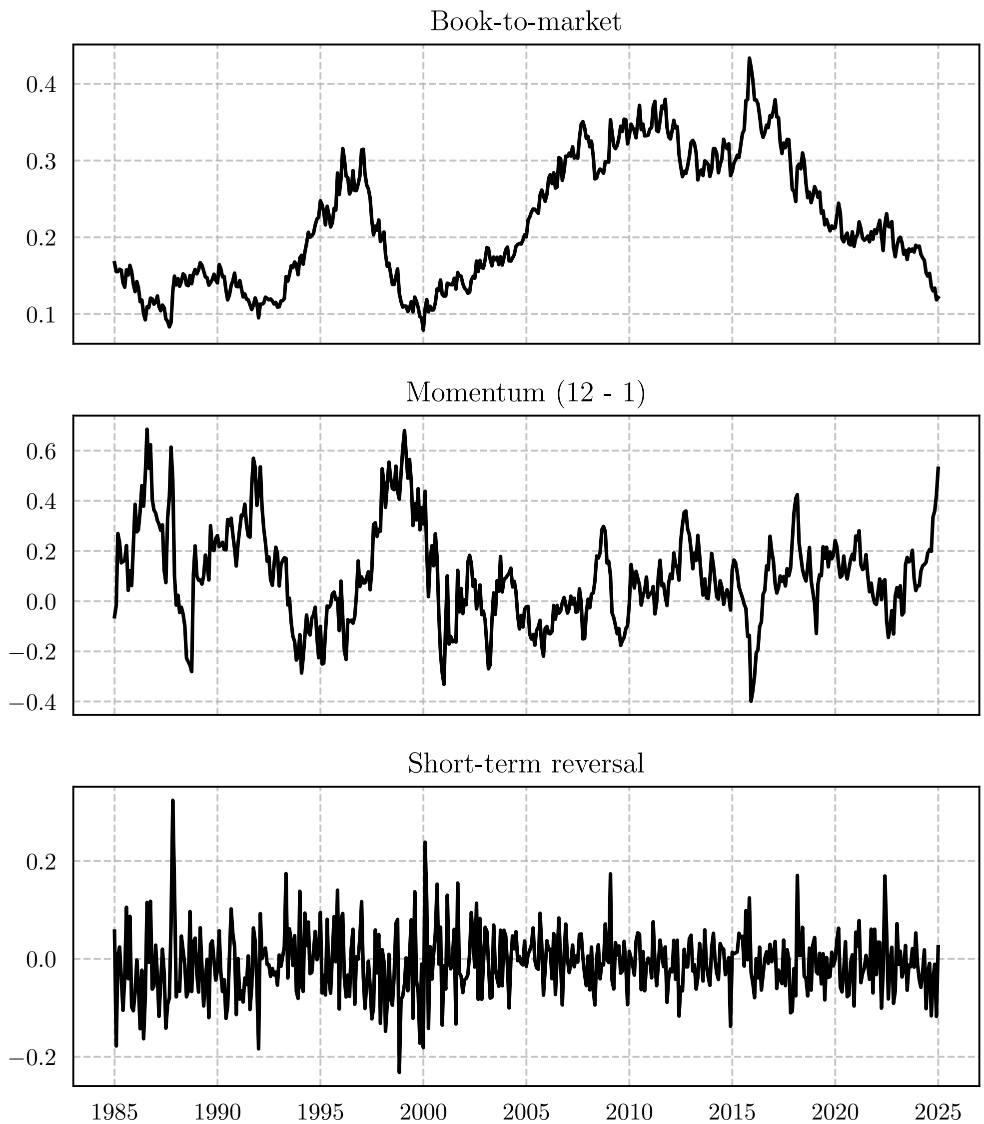
In both communications and radar systems, a phased array is a collection of individual antennas, often called elements, that work in concert. These elements, typically simple omnidirectional antennas, are arranged in a specific geometry, such as a line or a grid. By electronically coordinating the signals received by each element, the array can achieve performance far superior to that of a single antenna. In financial engineering, we can conceptualize a collection of individual investment signals (e.g., Value, Momentum, Quality) as a financial "array". Each signal acts as a sensor, providing a unique, albeit noisy, perspective on the market.

A beamformer is the signal processing operation that transforms the multiple inputs from an antenna array into a single, optimized output. It functions as a spatial filter, enhancing signals from a desired direction while suppressing unwanted interference and noise from all other directions. This is achieved by applying a set of complex weights, or coefficients, to the signal from each element before they are summed together. By precisely manipulating these weights, the beamformer can "steer" a highly sensitive beam toward a target signal, effectively increasing the SINR of the desired information. The process of forming and directing these sensitivity patterns gives the technique its name: *beamforming*.

Central to beamforming is the concept of the steering vector. This vector mathematically describes how a signal from a specific direction arrives at each element of the array with a unique time and phase delay relative to the

Figure 4.2: **The dynamic nature of value and momentum signals for Walmart Inc.**

The chart displays the time-varying, z-scored book-to-market (value) and 12-1 month momentum signals for Walmart Inc. from 1985 to 2024. The signals exhibit cyclical, sine-wave-like behavior and are often out of phase, illustrating the challenge of static signal combination.



others. For a simple uniform linear array (ULA), the steering vector, $\mathbf{s}$, can be expressed as a function of the signal's angle of arrival, θ, and the distance between elements, d:

$$\mathbf{s}(\theta) = \begin{bmatrix} 1 \\ e^{-2j\pi d \sin(\theta)} \\ e^{-2j\pi d(2)\sin(\theta)} \\ \vdots \\ e^{-2j\pi d(N_r-1)\sin(\theta)} \end{bmatrix}$$

This vector provides the precise phase adjustments needed to align the signals from a desired direction, causing them to add together constructively. In essence, the steering vector is the blueprint for pointing the array's focus.

While conventional beamformers use the steering vector to simply maximize gain in a known direction, adaptive beamforming techniques adjust the weights based on the statistical properties of the received signals themselves. A powerful and widely used adaptive method is the *Minimum Variance Distortionless Response (MVDR)*, also known as the Capon Beamformer. The objective of MVDR is to maintain a perfect, distortionless response (a gain of one) in the direction of the desired signal while simultaneously minimizing the total output power, or variance, from all other sources. Since the desired signal's gain is fixed, minimizing the total output power is equivalent to minimizing the power of all interfering signals and noise. This makes MVDR a "statistically optimal" beamformer for maximizing the SINR. The optimal weight vector for the MVDR beamformer is calculated as:

$$\omega_{\text{mvdr}} = \frac{\mathbf{R}^{-1}\mathbf{s}}{\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}}$$

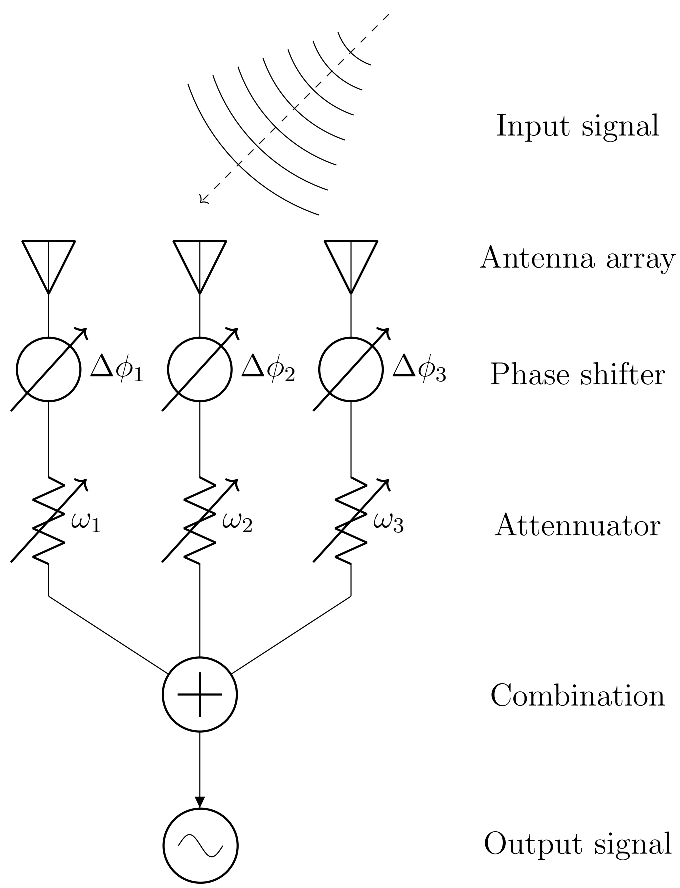
Here, $\mathbf{s}$ is the steering vector for the desired signal, and $\mathbf{R}$ is the spatial covariance matrix of the received signals. This matrix captures the correlation structure between the signals received at each antenna element. The crucial step is the inversion of this covariance matrix ($\mathbf{R}^{-1}$), which allows the beamformer to systematically place deep nulls in the directions of interfering signals.

Figure 4.3 provides a detailed block diagram of a multi-channel adaptive beamformer, illustrating the sequential processing steps required to transform multiple noisy inputs into a single, optimized output. In this basic illustration, the process begins where an array of 3 antennas receives the incoming wavefronts, capturing 3 distinct, spatially-sampled versions of the same input signal. The core of the adaptive beamforming occurs in the next two stages, where a complex weight, comprising a phase shift ($\Delta\phi_k$) and an amplitude adjustment (ω_k), is applied to each signal path. The phase shift is calculated to align the desired signal components from all channels, while the amplitude weight is determined by an adaptive algorithm to strategically suppress sources of interference. Finally, all individually processed signals are coherently summed at a combination point, resulting in an output signal where the desired components have been constructively reinforced and the interference and noise have been destructively canceled, yielding a significantly higher SINR.

Translating this to financial engineering, while we do not have physical antennas, we have an array of investment signals. The physical "interference" an engineer seeks to nullify becomes the statistical correlation between these investment signals. A strong value signal, for example, can be "interfered

Figure 4.3: **Block Diagram of a Basic Beamformer and Its Financial Analogy**

This figure illustrates the core components of a beamformer. An incoming signal is received by an array of antennas (the individual investment signals). Each signal is then passed through an attenuator that applies a specific weight (ω_i) and a phase shifter that applies a delay ($\Delta\phi_i$). In our financial framework, this combined weighting and phasing process is equivalent to the optimal capital allocation determined by the portfolio construction algorithm. These individually processed signals are then summed together to produce a single, optimized output signal with a higher signal-to-noise ratio (the final multifactor portfolio with an enhanced Information Ratio).



with" by a simultaneous, negatively correlated momentum signal. Instead of measuring signal power in watts, we can measure the risk premia associated with each signal. The MVDR framework thus provides a rigorous methodology for constructing a composite portfolio. It uses the covariance matrix of the factor returns to understand their interrelationships (the "interference") and a steering vector (representing the desired alpha source) to enhance the target premium while systematically canceling out the detrimental effects of correlated factor exposures.

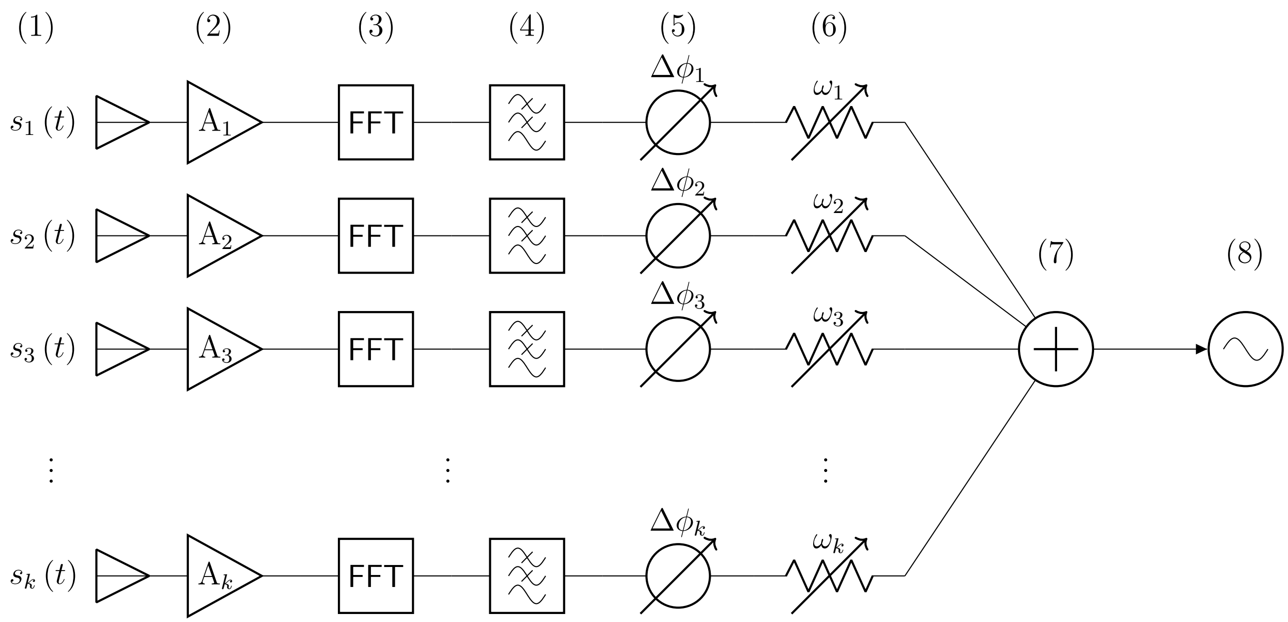
4.3.4 A Beamforming Approach to Investment Signals Aggregation

Our proposed methodology translates the principles of adaptive beamforming from signal processing into a rigorous, multi-stage framework for constructing systematic equity portfolios. The process, illustrated in Figure 4.4, is designed to systematically receive, amplify, filter, and optimally combine multiple noisy investment signals into a final, robust portfolio. Each stage has a direct counterpart in the engineering world, allowing us to leverage a sophisticated and proven toolkit to address the unique challenges of financial data.

The process begins at each monthly rebalance date with the reception of raw investment signals for all stocks in the investable universe (1). Given the cross-sectional and relative nature of these signals, we first perform an amplification and normalization step to make them comparable and robust (2). Each raw signal is cross-sectionally transformed using a non-parametric quantile transformation, which maps the data to a standard normal distribution. This process, unlike simple z-scoring, is robust to outliers and effectively handles non-linear relationships in the data. In addition to this cross-sectional normalization, we also standardize each stock's signal time series in the longitudinal dimension. We implement a Kalman filter to estimate the conditional mean and volatility of each stock's signal exposure over time. This step ensures that the signal is standardized relative to its own history, mitigating the impact of structural breaks or time-varying volatility in a firm's characteristics. The output of this stage is a set of amplified, robustly normalized signals ready for further processing.

Figure 4.4: **Block Diagram of the Proposed Investment Process**

The diagram shows the signal flow for k investment signals in our proposed methodology. Each signal is received, amplified, filtered, and then optimally phase-shifted and weighted before being combined to produce a final, enhanced output that informs portfolio construction.



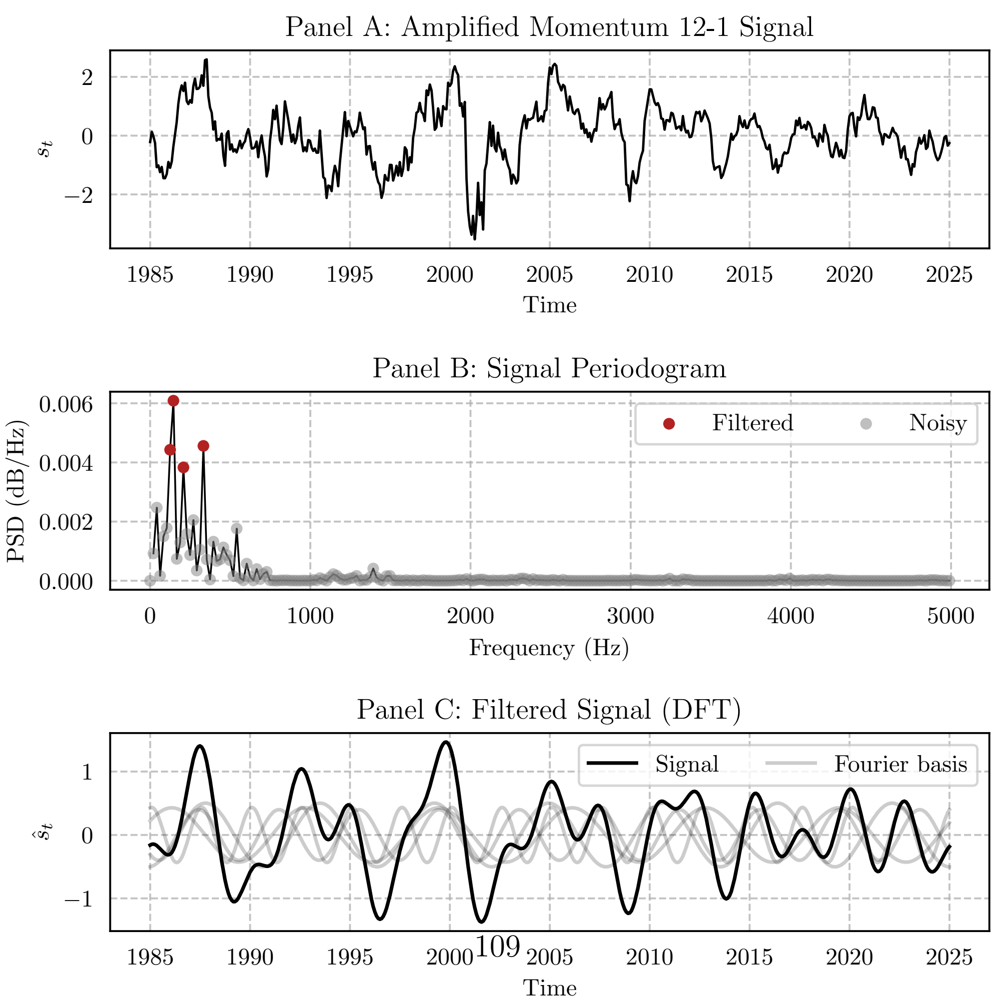
With the amplified signals in hand, the next stage is to separate the underlying predictive trend from high-frequency noise. We move the signal from the time domain into the frequency domain using a Fast Fourier Transform (FFT) (3). The FFT decomposes the time series of each signal into a sum of sinusoidal waves of different frequencies and amplitudes. This allows us to analyze the signal's power spectrum, as shown in the periodogram in Figure 4.5, which illustrates this procedure using the 12-1 month momentum signal for Walmart Inc. as a case study. The volatile time-domain signal (Panel A) is first decomposed into its constituent frequencies using a Discrete Fourier Transform. In the frequency domain (Panel B), we identify the low-frequency components that represent the persistent, underlying trend of the signal, while discarding the higher-frequency components, which are treated as noise. These selected, powerful frequencies are then used to reconstruct the final, filtered signal (Panel C), which is a smoothed and more robust representation of the original investment characteristic.

This crucial filtering step ensures that our beamforming algorithm is fed with cleaner, more reliable inputs, thereby improving the stability and efficacy of the final portfolio construction. In the frequency domain, we apply a k-means clustering method to isolate the most relevant frequencies. We posit that the persistent, predictive component of an investment signal resides in its lower-frequency components, which represent secular trends. The higher-frequency components, conversely, are treated as transient market noise and are discarded by setting their amplitudes to zero. By taking the inverse FFT of only the selected low-frequency components, we transform the signal back into the time domain, resulting in a smoothed, "cleaned" representation that is a more robust proxy for the true underlying economic anomaly.

One of the most critical steps in beamforming is the alignment of signals, or phasing (5). Just as an engineer must adjust the phase of signals

Figure 4.5: **Frequency Domain Filtering of the Momentum Signal**

This exhibit illustrates the process of filtering an investment signal using a Discrete Fourier Transform (DFT), with the 12-1 month momentum signal for Walmart Inc. as an example. Panel A displays the raw, amplified momentum signal in the time domain, exhibiting significant volatility. Panel B shows the signal's periodogram, which decomposes the time-series into its constituent frequencies and their corresponding power. The strongest, low-frequency components, identified as the true underlying "signal" (red dots), are selected for reconstruction, while the weaker, higher-frequency components are discarded as "noise". Panel C displays the final "filtered signal" (black line), reconstructed by summing the sinusoidal waveforms (the Fourier basis, shown in gray) corresponding only to the selected components. This process effectively isolates the persistent trend from short-term noise, creating a cleaner input for the beamforming combination algorithm.



from different antennas to ensure they add constructively, we must align the cyclical peaks of our different investment signals. This step is of the utmost relevance for combining factors with a persistent negative correlation, such as value and momentum. By analyzing the cyclical properties of our cleaned signals, we apply a dynamic phase shift to align their peak exposures. This ensures that when we combine a value signal that is peaking with a momentum signal that may be in a trough, the combination algorithm can correctly interpret their opposing states rather than allowing them to simply cancel each other out. This alignment is crucial for building a composite signal that benefits from the diversification of negatively correlated factors instead of being diluted by it.

With the signals filtered and aligned, we proceed to the final weighting and combination stage. In beamforming, the attenuator weights (ω_k) (6) are calculated to maximize the SINR. To do this in our financial context, we require robust, forward-looking estimates of both the "signal power" (expected risk premia) and the "interference" (the dynamic covariance of the signals). First, we estimate the time series of the risk premia for each of our seven signals by running monthly Fama-MacBeth regressions across the entire CRSP/Compustat universe. This provides a historical record of the reward for each factor. We then fit a Vector Autoregressive (VAR(1)) model to this system of seven risk premia time series. The one-step-ahead forecast from this VAR model serves as our dynamic estimate of the expected signal power for each factor.

Next, we model the interference using a Dynamic Conditional Correlation (DCC) GARCH model on the residuals from the VAR(1) system. To ensure the stability of this process, particularly given the high dimensionality, we first compute the unconditional covariance matrix of the residuals using the robust shrinkage estimator of Ledoit and Wolf (2004). This technique produces a well-conditioned covariance matrix by optimally combining the sample covariance matrix with a highly structured target (the identity matrix). This approach is conceptually identical to the "diagonal loading" technique used in signal processing to stabilize covariance estimates for antenna arrays. We then transform this static, shrunk covariance matrix into a dynamic one by replacing the unconditional volatility estimates on its diagonal with the conditional volatility forecasts from individual GARCH(1,1) models fitted to each residual series. With these dynamic estimates of the expected risk premia (signal) and the signal covariance matrix (interference), we compute the optimal attenuator vector, ω_k, that maximizes the SINR. This vector provides the dynamic weights used to combine the seven investment signals into a final composite score at the combination step (7). The resulting "output signal" (8) is a single, robust score for each stock. This aggregate score then serves as the basis for portfolio construction. Stocks with a positive composite score are assigned to the long leg of the portfolio, while those with a negative score constitute the short leg. Within each leg, individual stock weights are set to be proportional to the absolute value of their composite scores. Finally, the weights in both the long and short legs are scaled to sum to 100%, creating a self-financed, dollar-neutral portfolio with a total leverage of 2.

4.3.5 Beamforming in the Context of Modern Factor Aggregation Methods

The proliferation of return-predictive signals, famously dubbed the "factor zoo" (Cochrane, 2011), has shifted the central challenge in empirical asset pricing from factor discovery to factor aggregation. With hundreds of documented anomalies, many of which fail rigorous replication or are highly correlated, the task for a systematic manager is no longer simply to find new signals but to robustly combine existing ones into a single, effective portfolio. In response, the academic literature has proposed several sophisticated statistical and machine-learning frameworks to "tame the factor zoo." Situating our beamforming approach within this ongoing debate highlights its novelty and motivates the turn to electrical engineering for solutions.

Contemporary methods for managing the factor zoo can be broadly categorized into three groups. The first approach employs shrinkage and economic priors to discipline high-dimensional models. Kozak et al. (2020) argue that while the stochastic discount factor (SDF) is driven by a few latent sources of risk, these sources have dense loadings across a large number of firm characteristics. Consequently, a parsimonious model with only a few factors is misspecified. Their solution is to build a reduced-form SDF using a large cross-section of characteristic-managed portfolios and then apply Bayesian shrinkage, which is guided by an economically motivated prior on the maximum achievable Sharpe ratio. This prevents the model from placing extreme weights on noisy, in-sample correlations, leading to more robust out-of-sample performance.

A second approach focuses on high-dimensional model selection to identify the most potent and non-redundant return predictors. Feng et al. (2020) introduce a new testing methodology that provides a formidable statistical hurdle for any newly proposed factor. Their framework accommodates time-varying factor loadings and builds a benchmark model from a large set of existing factors, against which the marginal contribution of a new factor is tested. This helps distinguish true predictive power from statistical noise. In a similar vein, Freyberger et al. (2020) employ a non-parametric methodology using adaptive group LASSO. Their approach avoids the constraints of linear factor models by allowing for complex, non-linear relationships and interactions between firm characteristics, selecting the most relevant predictors without pre-specifying a functional form.

The third approach centers on factor timing. Rather than focusing on which factors to include, Neuhierl et al. (2023) we investigate when to invest in them. They document significant time-series predictability in the premia of well-known anomaly factors, driven by macroeconomic conditions and factor-specific attributes like valuation spreads. Their findings suggest that a dynamic strategy that times factor exposures can dramatically outperform a static combination, highlighting the importance of capturing the time-varying nature of expected returns.

Our proposed beamforming framework offers a novel alternative that differs from these methods in its conceptual foundation and operational mechanics. Unlike model selection techniques (Feng et al., 2020; Freyberger et al., 2020), beamforming is fundamentally an aggregation, not a selection, methodology. It does not seek to identify a sparse subset of "best" factors. Instead, it assumes a set of potentially valuable signals is given and focuses on optimally combining all of them, dynamically adjusting their weights to maximize the clarity of the final composite signal.

Furthermore, beamforming is disciplined by an engineering prior, not an economic one. Whereas Kozak et al. (2020) use priors on the plausible magnitude of the Sharpe ratio to regularize their model, beamforming's objective is to maximize the Signal-to-Interference-and-Noise Ratio (SINR). This concept, rooted in physics and information theory, provides a powerful and orthogonal motivation. It explicitly reframes the statistical cross-correlation between investment signals as "interference", an unwanted disturbance to be actively nulled out. This is a distinct conceptualization from standard portfolio optimization, where covariance is a component of risk to be managed rather than an interference pattern to be canceled. The inversion of the signal covariance matrix in the MVDR beamformer is specifically designed to achieve this cancellation, attenuating the contributions of interfering signals.

Finally, beamforming provides a more holistic form of dynamic adaptation than factor timing models. While Neuhierl et al. (2023) focus on the predictability of the first moment (expected returns), our framework, through its use of VAR(1)-DCC-GARCH modeling, simultaneously forecasts the entire signal environment: time-varying risk premia (the "signals"), volatilities ("noise"), and correlations ("interference"). The optimal combination weights are derived from this complete, forward-looking view of the signals' joint distribution. This ability to adapt to the full, time-varying covariance structure is a key differentiator and a primary motivation for looking to signal processing. By reframing the factor aggregation challenge as a problem of information processing, beamforming provides a robust, data-driven, and highly adaptive solution that directly confronts the core task of isolating alpha from a noisy, crowded, and interference-prone environment.

4.4 Empirical Exercise

4.4.1 Sourcing and Data Preparation

We work with a US-centric dataset, due to the size of this market, the number of companies listed, its high liquidity, and data availability. The historical performance of our investable universe is made up of US-based common stocks and is sourced from the monthly return files of the Center for Research in Security Prices (CRSP) database. We select all securities identified as common stock (share codes 10 or 11) listed on the NYSE, AMEX, and NASDAQ (exchange codes 1, 2, 3, 31, 32, and 33). For each of these stocks, we collect daily returns starting in July 1963, and ending in December 2024. We adjust these returns in the case of stock delisting in the same way as described in Shumway (1997), also explained in Bali et al. (2016). From these returns, we subtract the risk-free rate from Kenneth French's website to compute the excess returns. In addition, for each of the stocks in the universe, we also collect time series of daily closing prices (field PRC), number of shares outstanding (field SHROUT), to calculate each stock's market capitalization. Finally, we obtain the Global Industry Classification Standard (GICS) codes from COMPUSTAT, and link them to our return dataset.

The number of stocks used in the dataset used for the iterative backtesting ranges from 3,608 to 5,071 stocks, averaging 4,122 during the sample. For the purpose of the empirical exercise to be described below, we subset the complete dataset based on data availability in order to arrive at our investable universe. We define the investable universe on a given date as every stock in the dataset with valid returns, market capitalization, and industry classification entries on that particular day.

4.4.2 Investment signals

For each one of these companies, on each observation, we compute a set of investment signals to test the proposed signal combination framework. Given the wide body of knowledge supporting the existence and robustness of value and momentum anomalies, we select the following set of investment signals:

- **Book-to-market:** A valuation ratio that compares a company's book value of equity to its market value of equity (market capitalization). It is calculated as

$$BM = \frac{\text{Book Value of Equity}}{\text{Market Capitalization}}$$

  A high book-to-market ratio is often associated with value stocks, which have historically tended to outperform growth stocks (those with low ratios). This is commonly known as the value anomaly.

- **Cash-flow-to-price:** A valuation ratio that measures a company's cash-flow relative to its stock price (market capitalization). It's calculated as

$$CP = \frac{\text{Operating Cash Flow}}{\text{Market Capitalization}}$$

Like the book-to-market ratio, a high cash-flow-to-price ratio can indicate that a stock is undervalued.

- **Sales-to-price:** A valuation ratio that compares a company's total sales (revenue) to its market capitalization, calculated as

$$SP = \frac{\text{Total Annual Sales}}{\text{Market Capitalization}}$$

It is the reciprocal of the more common Price-to-Sales ratio. A high sales-to-price ratio might suggest a company is undervalued relative to its sales generation, and it's particularly useful for valuing companies that are not yet profitable.

- **Earnings-to-price:** This ratio, also known as the earnings yield, compares a company's earnings per share (EPS) to its market price per share. It is the reciprocal of the Price-to-Earnings (P/E) ratio:

$$EP = \frac{\text{Income Bef Extraord Items - Depreciation \& Amortization}}{\text{Market Capitalization}}$$

A high earnings yield suggests that an investor is getting more earnings for each dollar invested and may signal an undervalued stock.

- **12-minus-1 price momentum:** A momentum indicator that captures an asset's performance over the past year, but crucially excludes the most recent month. It is calculated as the cumulative return from 12 months ago to one month ago:

$$MOM_{12-1} = \frac{P_{t-1}}{P_{t-12}} - 1$$

  where P_t is the price at the end of month t. The most recent month is excluded to avoid the short-term reversal effect, where very recent top performers tend to underperform in the immediate future. The momentum anomaly is the tendenzillow cy for assets with high past returns to continue performing well.

- **6-minus-1 price momentum:** A shorter-term version of the price momentum signal. It measures the asset's cumulative return over the past six months, again excluding the most recent month. The calculation is

$$MOM_{6-1} = \frac{P_{t-1}}{P_{t-12}} - 1$$

  This signal also aims to capture the medium-term persistence in stock returns.

- **Residual price momentum:** A more sophisticated measure of momentum that isolates the portion of a stock's return that is not explained by common systematic risk factors (like market, size, and value). It is the firm-specific excess return, or alpha, obtained from a factor model regression (e.g., the Fama-French three-factor model) over a given look-back period. The goal is to capture momentum that is unique to the company itself, rather than momentum driven by broad market or factor trends. For the sake of this exercise use the CAPM to residualize the monthly excess returns.

The fundamental data used in the computation of the value-related signals is sourced from the Compustat database, while the excess returns used in the technical signals are sourced from CRSP, and the risk-free and market returns were obtained from Ken French's database.

4.4.3 Backtesting

At the end of each month, we deem as investable any stock that has valid values for each of these seven investment signals. In an effort to mitigate the possible distortive influence of exceedingly small stocks, which are categorized under the micro-cap classification, we exclude from consideration those companies that fall within the lowest one percentile of the overall market capitalization spectrum. This leaves us with an investable universe with an average of 2,230 stocks during the validation sample window, ranging from 1,671 to 2,943 companies, covering at least 83% of the total capitalization of the US stock market at each point in time, and averaging nearly 90% of market capitalization coverage during the backtesting period.

To assess the efficacy of our proposed beamforming approach, we establish a baseline comparison portfolio that reflects a common methodology for signal combination in the investment management industry. At each monthly rebalance, we first cross-sectionally normalize each individual investment signal into a z-score to ensure comparability across different metrics. Following a widely adopted heuristic, these normalized scores are then combined into a single composite signal for each stock using a simple average. This aggregate score serves as the basis for portfolio construction; stocks with a positive composite score are assigned to the long leg of the portfolio, while those with a negative score constitute the short leg. Within each leg, the individual stock weights are set to be proportional to the absolute value of their respective composite signal scores. Finally, the weights in both the long and short legs are scaled to sum to 100%, creating a self-financed, dollar-neutral portfolio with a total leverage of 2.

In contrast to the heuristic baseline, our proposed methodology employs a sophisticated and dynamic framework derived from signal processing to optimally combine investment signals. At each monthly rebalance date, we observe the seven normalized investment signals for all stocks in the investable universe, with the objective of constructing an optimal steering vector that intelligently aggregates these signals into a single composite score. To estimate the predictive "signal power" of each of the seven "antennas," we adopt a robust estimation procedure. Leveraging the full historical CRSP and Compustat database, we first run monthly Fama-MacBeth regressions to derive the time series of the seven factor risk premia. We then model the dynamics of this system by fitting a Vector Autoregressive model of order one, VAR(1), whose one-step-ahead forecast serves as our sophisticated mean model for expected signal intensity. To model the "interference" among the signals, we construct a dynamic conditional covariance matrix. We begin by taking the residuals from the VAR(1) system and applying the robust shrinkage estimator of Ledoit and Wolf (2004) to the unconditional covariance matrix. To capture time-varying volatility, we then fit a GARCH(1,1) model to each of the seven residual series, integrating these conditional volatility forecasts to produce a fully dynamic and robust estimate of the signal covariance structure. With these rigorous estimates for both expected signal and interference, we construct the optimal steering vector by maximizing the Signal-to-Interference-and-Noise Ratio (SINR). This resulting vector provides the dynamic weights used to combine the seven investment signals into a final composite score. This approach systematically accounts not only for the predictive power of each individual signal but also for the complex, time-varying correlations and interference among them, resulting in a far more robust and efficient aggregation than is possible with simple averaging. The weights are then computed in a similar fashion as in the baseline approach.

4.4.4 Results

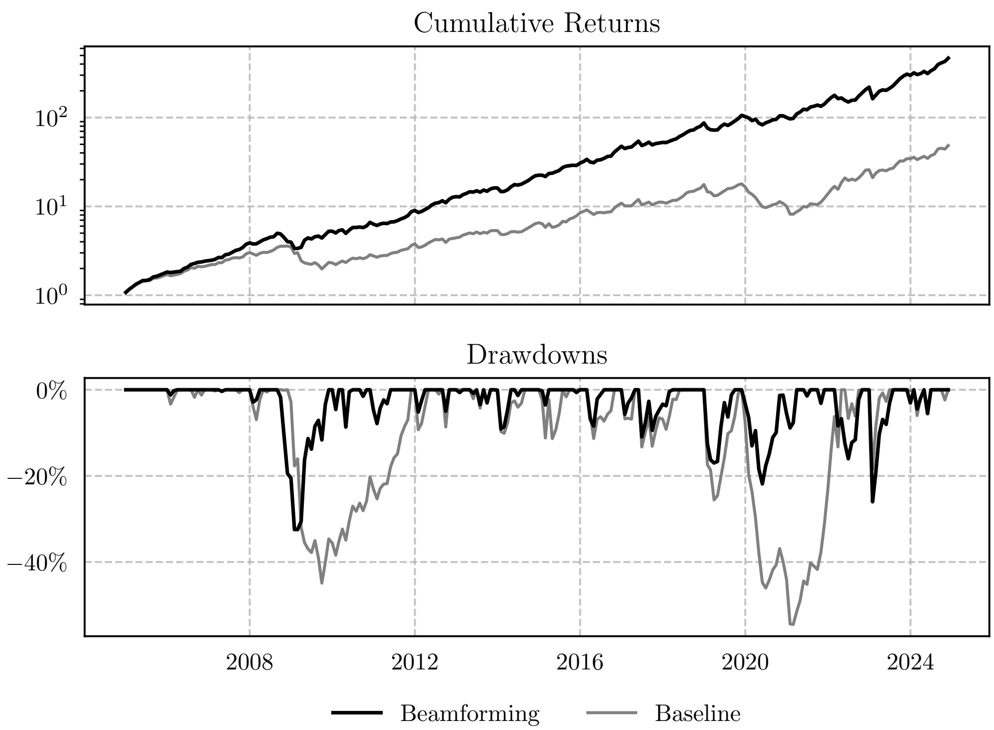
The empirical results reveal a stark contrast between the adaptive beamforming methodology and the traditional baseline approach, with the former demonstrating superior performance across nearly every dimension of return and risk. A visual inspection of the strategies' long-term performance in Figure 4.6 provides a clear narrative. The top panel, which plots cumulative returns on a logarithmic scale, shows that the beamforming portfolio consistently compounds wealth at a significantly higher rate than the baseline. While both strategies begin at a similar point, the beamforming approach begins to separate itself decisively after the 2008 Global Financial Crisis and continues to widen its lead over the subsequent decade, indicating a more efficient and robust engine for capital growth.

However, the most compelling visual evidence lies in the analysis of drawdowns, shown in the bottom panel of Figure 4.6. The beamforming methodology exhibits a structurally more resilient risk profile. During periods of significant market stress, such as the 2008 crisis and the COVID-19 crash of 2020, the baseline portfolio suffers from profoundly deep and prolonged drawdowns. In contrast, the beamformer's drawdowns are consistently shallower and its recoveries are markedly swifter. This superior downside protection suggests that the adaptive nature of the signal combination is not merely enhancing returns during favorable periods but, more critically, is providing a robust defense during market turmoil.

The quantitative metrics detailed in Table 4.2 provide rigorous validation for these visual observations. The beamforming strategy delivers a substan-

Figure 4.6: **Cumulative Returns and Drawdowns of Beamforming vs. Baseline Strategies**

The chart displays the cumulative returns (log scale) and historical drawdowns for the proposed Beamformer strategy and the traditional Baseline strategy from 2004 to 2024. The Beamformer demonstrates both superior long-term growth and significantly more contained drawdowns during periods of market stress.



tially higher Annualized Return of 35.97% compared to the baseline's 24.14%. Remarkably, this significant outperformance is achieved with a lower Annualized Volatility (18.44% vs. 19.93%). This combination of higher return and lower risk leads to a dramatically improved risk-adjusted performance, as captured by the Sharpe Ratio, which stands at 1.95 for the beamformer versus 1.21 for the baseline. The outperformance on a risk-adjusted basis is even more pronounced when focusing solely on downside risk; the Sortino Ratio of 2.33 is nearly double that of the baseline's 1.39. The drawdown data quantifies the resilience seen in the chart, with the beamformer's Max Drawdown limited to -32.47% compared to the baseline's severe -53.54%. Furthermore, the Max Drawdown Duration is less than half that of the baseline (15 months vs. 34 months), confirming its ability to recover capital far more rapidly.

A deeper analysis of the risk profile reveals that the beamforming approach offers superior protection against extreme events. At both the 5% and 1% levels, the Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) are consistently lower for the beamforming portfolio, indicating a more controlled tail risk profile. The only notable trade-off for this substantially improved performance and risk profile is a higher One-Way Portfolio Turnover (879.72% vs. 514.76%). This is an expected outcome of the methodology's adaptive nature; the higher turnover reflects the model's dynamic response to the changing statistical properties and interrelationships of the investment signals. While this implies a greater consideration for transaction costs in a live implementation, the overwhelming improvements across all other metrics suggest that the benefits derived from this sophisticated, adaptive signal combination process far outweigh its implementation costs. In essence, the beamforming methodology successfully translates a more rigorous signal processing framework into tangible portfolio benefits, delivering not just higher returns, but a structurally more robust and resilient investment strategy.

Table 4.2: **A comparison of performance and risk metrics for two trading strategies**

The table provides a detailed statistical comparison of the Baseline and Beamforming strategies. The Beamforming approach shows superior metrics across returns, risk-adjusted returns (Sharpe, Sortino, Calmar), and tail risk measures (VaR, CVaR), albeit with higher portfolio turnover.

Metric	**Baseline**	**Beamforming**
Annualized Return	24.14%	35.97%
Annualized Volatility	19.93%	18.44%
Sharpe Ratio	1.21	1.95
Sortino Ratio	1.39	2.33
Calmar Ratio	0.45	1.11
Downside Volatility	17.38%	15.43%
Max Drawdown	-53.54%	-32.47%
Max Drawdown Duration (Months)	34	15
Hit Ratio	71.67%	76.67%
Skewness	-1.15	-1.09
Kurtosis (Excess)	3.64	3.63
VaR (5%)	-8.03%	-6.81%
CVaR (5%)	-13.57%	-11.12%
VaR (1%)	-14.26%	-11.91%
CVaR (1%)	-20.57%	-17.85%
One-Way Portfolio Turnover	514.76%	879.72%

4.5 Conclusions

The primary objective of this chapter was to introduce and validate a novel framework for combining investment signals by drawing upon the well-established principles of adaptive beamforming from signal processing. We began by establishing the strong theoretical parallels between the two fields, arguing that the challenge of extracting alpha from noisy, correlated financial factors is mathematically analogous to an antenna array isolating a desired signal from a field of interference. By re-framing investment signals as noisy waveforms with distinct properties, we demonstrated how techniques like Fourier analysis can be used to filter and prepare these signals for a more robust combination.

Our proposed methodology replaces the heuristic and often suboptimal industry practice of simple signal averaging with a dynamic, adaptive system. The beamforming approach systematically models both the predictive power of each individual signal and, critically, the complex, time-varying covariance structure between them. This allows the model to intelligently amplify strong, independent signals while mitigating the detrimental effects of noise and interference, such as the persistent negative correlation between value and momentum.

The empirical backtest provided compelling evidence for the efficacy of this approach. In a direct comparison, the beamforming portfolio delivered substantially higher absolute and risk-adjusted returns than a baseline strategy representing traditional industry methods. More importantly, it exhibited a structurally more resilient risk profile, with significantly shallower drawdowns and faster recovery times during periods of market stress. This demonstrates that the benefits of the methodology are not confined to return enhancement but also extend to robust risk management. While the adaptive nature of the beamforming strategy results in higher portfolio turnover, the dramatic improvements in performance and risk metrics suggest that the rewards of this sophisticated signal combination process far outweigh the potential implementation costs. In essence, by bridging the gap between financial engineering and signal processing, we have outlined a superior framework for translating the theoretical promise of alpha signals into a tangible and resilient investment strategy.

Chapter 5

Conclusion

This thesis has advanced the theory and practice of quantitative portfolio management by developing and validating three novel frameworks, each designed to address a fundamental limitation in conventional investment models. The central argument connecting these distinct essays is that more robust, scalable, and high-performing investment strategies can be constructed by moving beyond the analysis of individual securities and signals to instead model the underlying latent structures that govern market dynamics. This work has demonstrated that by identifying and exploiting these structures, whether they are the implicit factors of a covariance matrix, unobservable market regimes, or the predictive essence of noisy signals, we can build portfolios that are more intelligently diversified, adaptive to changing conditions, and resilient to market turmoil.

5.1 Findings and Contributions

The research presented herein critically synthesizes findings across three distinct but related fronts, situating each within the mainstream portfolio management literature that motivated the research.

The first essay, "Diversified Spectral Portfolios", confronted the challenge of achieving true diversification in a market where returns are often driven by a few dominant, implicit factors. It moved beyond traditional heuristics like equal-weighting, which can be misleading, and established that effective diversification must occur not at the level of capital weights but in the eigenspace of underlying risk factors. By employing unsupervised machine learning to cluster securities based on their factor exposures, we developed a practical and scalable methodology that overcomes the tradability and stability issues plaguing purely statistical portfolios. The resulting Diversified Spectral Portfolios demonstrated superior risk-adjusted performance, particularly in the highly correlated environments where traditional methods falter. A key limitation of this work, however, is its reliance on historical covariance data, which can be susceptible to estimation error and may not fully capture future risk dynamics, especially during unforeseen market shocks. Furthermore, the framework is agnostic to expected returns, focusing solely on risk-based diversification.

The second essay, "Incorporating Market Regimes", addressed the critical issue of non-stationarity in financial markets. It introduced a scalable framework for regime-aware investing capable of handling large-scale equity portfolios. By applying a Hidden Markov Model to a parsimonious set of risk factors rather than thousands of individual stocks, it became possible to construct portfolios that dynamically adjust their posture to prevailing market conditions. This approach proved highly effective, with the resulting regime-aware portfolios consistently outperforming their static, regime-agnostic counterparts. The primary limitation here lies in the ex-post nature of regime identification; while the model adapts to shifts, it inherently lags in recognizing a new regime until sufficient data has accumulated. This can expose the portfolio to losses in the early stages of an abrupt market transition. Additionally, the model is constrained to the number of pre-defined states, which may not capture the full spectrum of market behaviors.

Finally, the third essay, "Optimal Investment Signal Combination", tackled the "factor zoo" by drawing a novel and powerful analogy to electrical engineering. We reframed the problem of combining multiple, often noisy and correlated, investment signals as one of optimal signal processing. By treating signals as data streams and applying adaptive beamforming techniques, we developed a method to systematically account for their time-varying correlations and interference. The result was a strategy with vastly superior performance and downside protection compared to traditional signal-averaging techniques. This engineering-inspired approach, however, comes with its own limitations. The adaptive nature of the beamforming algorithm leads to higher portfolio turnover, which would incur greater transaction costs in a real-world implementation. The model's efficacy also depends on the stability of the relationships between signals; a structural break in these relationships could temporarily degrade the model's performance.

5.2 Implications and Future Research

A common thread woven through all three essays is the immense value of interdisciplinary innovation. The solutions proposed were inspired by advances in unsupervised machine learning, probabilistic modeling, and signal processing, demonstrating that many of finance's most challenging problems have conceptual analogs in other quantitative fields. The practical implications of this research are significant, offering asset managers concrete methodologies for improving diversification, adapting to market regimes, and building more potent multi-factor models.

From a broader policy perspective, these findings hold relevance for systemic risk and financial stability. The spectral diversification and regime-switching frameworks offer tools that could help regulators and large institutional investors better understand and manage portfolio concentrations that might contribute to systemic vulnerability during market crises. By promoting a deeper understanding of latent risk factors and market states, these models can encourage more robust risk management practices, potentially enhancing overall market efficiency and stability.

Despite the compelling results, this research represents a step, not a final destination. Its limitations open several promising avenues for future inquiry. The frameworks presented, while powerful, could be extended and integrated. For instance, the Diversified Spectral Portfolio methodology could be enhanced by incorporating forward-looking risk estimates or by using more robust estimators for the covariance matrix that are less sensitive to outliers. The regime-switching framework could be expanded to explore endogenous models where regime transitions are influenced by macroeconomic variables, providing a more forward-looking signal. The beamforming approach could be evolved to explicitly incorporate transaction cost optimization within its objective function.

A particularly exciting direction lies in integrating these frameworks. One could construct a portfolio that is both spectrally diversified and regime-aware, where asset clusters are allocated dynamically based on the identified market state. Similarly, the beamforming signal combination technique could itself be made regime-dependent, employing different optimal weights during bull and bear markets. Such a synthesis could lead to a truly holistic investment model, one that is structurally sound, dynamically adaptive, and optimally constructed.

Finally, it must be acknowledged that the applications in this thesis are largely U.S.-centric. Broadening the scope of this research to global markets would be a crucial next step. Testing these frameworks across different market structures, regulatory environments, and economic cycles would significantly strengthen both the generality and the relevance of the contributions.

In closing, the increasing complexity of modern financial markets demands a continuous evolution in our tools and techniques. The research in this thesis contributes to this evolution by advancing a unified perspective focused on uncovering and leveraging the market's hidden structures. By continuing to look beyond the surface of security prices and embracing a more profound, structurally aware, and interdisciplinary approach, the field of quantitative finance is poised to develop the next generation of truly intelligent investment strategies.

Bibliography

Yu I. Abramovich. Controlled method for adaptive optimization of filters using the criterion of maximum SNR. *Radio Engineering and Electronic Physics*, 26(3):87–95, 1981.

Noël Amenc, Frédéric Ducoulombier, Mikheil Esakia, Felix Goltz, and Sivagaminathan Sivasubramanian. Accounting for Cross-Factor Interactions in Multifactor Portfolios without Sacrificing Diversification and Risk Control. *The Journal of Portfolio Management*, 43(5):99–114, 3 2017. ISSN 0095-4918. doi: 10.3905/jpm.2017.43.5.099.

Noël Amenc, Felix Goltz, and Sivagaminathan Sivasubramanian. Multifactor Index Construction: A Skeptical Appraisal of Bottom-Up Approaches. *The Journal of Index Investing*, 9(1):6–17, 5 2018. ISSN 2154-7238. doi: 10.3905/jii.2018.9.1.006.

R. Anand. *Digital Signal Processing.* De Gruyter, 6 2022. ISBN 9781683928010. doi: 10.1515/9781683928010.

Andrew Ang and Geert Bekaert. International Asset Allocation With Regime Shifts. *Review of Financial Studies*, 15(4):1137–1187, 7 2002a. ISSN 0893-9454. doi: 10.1093/rfs/15.4.1137.

Andrew Ang and Geert Bekaert. Regime Switches in Interest Rates. *Journal* *of Business & Economic Statistics*, 20(2):163–182, 4 2002b. ISSN 0735-0015. doi: 10.1198/073500102317351930.

Andrew Ang and Allan Timmermann. Regime Changes and Financial Markets. *Annual Review of Financial Economics*, 4(1):313–337, 10 2012. ISSN 1941-1367. doi: 10.1146/annurev-financial-110311-101808.

Rob Arnott, Campbell R Harvey, Vitali Kalesnik, and Juhani Linnainmaa. Alice's Adventures in Factorland: Three Blunders That Plague Factor Investing. Technical report, 2019.

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228, 7 1999. ISSN 0960-1627. doi: 10.1111/1467-9965.00068.

Clifford Asness, Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz, and Lasse H. Pedersen. Size matters, if you control your junk. *Journal of Financial Economics*, 129(3):479–509, 9 2018. ISSN 0304405X. doi: 10.1016/j.jfineco.2018.05.006.

Clifford S. Asness. The Interaction of Value and Momentum Strategies. *Financial Analysts Journal*, 53(2):29–36, 3 1997. ISSN 0015-198X. doi: 10.2469/faj.v53.n2.2069.

Clifford S. Asness, Antti Ilmanen, Ronen Israel, and Tobias J. Moskowitz. Investing with style. *Journal Of Investment Management*, 13(1):27–63, 2015.

Attilio Meucci. Managing Diversification. *Risk*, pages 74–79, 5 2009.

Afşar Onat Aydınhan, Petter N. Kolm, John M. Mulvey, and Yizhan Shu. Identifying patterns in financial markets: extending the statistical jump model for regime identification. *Annals of Operations Research*, pages 1–37, 5 2024. ISSN 0254-5330. doi: 10.1007/s10479-024-06035-z.

Geum Il Bae, Woo Chang Kim, and John M. Mulvey. Dynamic asset allocation for varied financial markets under regime switching framework. *European Journal of Operational Research*, 234(2):450–458, 4 2014. ISSN 03772217. doi: 10.1016/j.ejor.2013.03.032.

Lieven Baele, Geert Bekaert, Seonghoon Cho, Koen Inghelbrecht, and Antonio Moreno. Macroeconomic Regimes. *Journal of Monetary Economics*, 70:51–71, 3 2015. ISSN 03043932. doi: 10.1016/j.jmoneco.2014.09.003.

David H. Bailey, Jonathan M. Borwein, Marcos López de Prado, and Qiji Jim Zhu. Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance. *Notices of the American Mathematical Society*, 61(5):458–471, 5 2014. ISSN 0002-9920. doi: 10.1090/noti1105.

Turan G. Bali, Robert F. Engle, and Scott Murray. *Empirical Asset Pricing: The Cross Section of Stock Returns*. Wiley, Hoboken, NJ, 2016.

Ravi Bansal and Hao Zhou. Term Structure of Interest Rates with Regime Shifts. *The Journal of Finance*, 57(5):1997–2043, 10 2002. ISSN 0022-1082. doi: 10.1111/0022-1082.00487.

Ziv Bar-Joseph, David K Gifford, and Tommi S Jaakkola. Fast Optimal Leaf Ordering for Hierarchical Clustering. *Bioinformatics*, 17:S22–S29, 2001.

Pedro Barroso and Pedro Santa-Clara. Momentum has its moments. *Journal of Financial Economics*, 116(1):111–120, 4 2015. ISSN 0304405X. doi: 10.1016/j.jfineco.2014.11.010.

Jennifer Bender and Taie Wang. Can the Whole Be More Than the Sum of the Parts? Bottom-Up versus Top-Down Multifactor Portfolio Construction. *The Journal of Portfolio Management*, 42(5):39–50, 7 2016. ISSN 0095-4918. doi: 10.3905/jpm.2016.42.5.039.

Anil Bera and Sung Y Park. Optimal Portfolio Diversification Using the Maximum Entropy Principle. *Econometric Reviews*, 27(4-6):484–512, 2008.

Simone Bernardi, Markus Leippold, and Harald Lohre. Maximum diversification strategies along commodity risk factors. *European Financial Management*, 24(1):53–78, 8 2018. ISSN 13547798. doi: 10.1111/eufm.12122.

Vineer Bhansali, Josh Davis, Graham Rennison, Jason Hsu, and Feifei Li. The Risk in Risk Parity: A Factor-Based Analysis of Asset-Based Risk Parity. *The Journal of Investing*, 21(3):102–110, 8 2012. ISSN 1068-0896. doi: 10.3905/joi.2012.21.3.102.

Fischer Black and Robert Litterman. Global Portfolio Optimization. *Financial Analysts Journal*, 48(5):28–43, 9 1992. ISSN 0015-198X. doi: 10.2469/faj.v48.n5.28.

David Blitz and Milan Vidojevic. The Characteristics of Factor Investing. *The Journal of Portfolio Management*, 45(3):69–86, 2 2019. ISSN 0095-4918. doi: 10.3905/jpm.2019.45.3.069.

David Blitz, Matthias X. Hanauer, Milan Vidojevic, and Pim van Vliet. Five Concerns with the Five-Factor Model. *The Journal of Portfolio Management*, 44(4):71–78, 3 2018. ISSN 0095-4918. doi: 10.3905/jpm.2018.44.4.071.

Nizar Bouguila, Wentao Fan, and Manar Amayri. *Hidden Markov Models and Applications*. Springer International Publishing, Cham, 2022. ISBN 978-3-030-99141-8. doi: 10.1007/978-3-030-99142-5.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004. ISBN 9780521833783. doi: DOI: 10.1017/CBO9780511804441.

Joshua Brodie, Ingrid Daubechies, Christine De Mol, Domenico Giannone, and Ignace Loris. Sparse and stable Markowitz portfolios. *Proceedings* *of the National Academy of Sciences*, 106(30):12267–12272, 7 2009. ISSN 0027-8424. doi: 10.1073/pnas.0904287106.

Steven L Brunton and J Nathan Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019. ISBN 9781108380690. doi: 10.1017/9781108380690.

Jan Bulla, Sascha Mergner, Ingo Bulla, André Sesboüé, and Christophe Chesneau. Markov-Switching Asset Allocation: Do Profitable Strategies Exist? *Journal of Asset Management*, 12(5):310–321, 11 2011. ISSN 1470-8272. doi: 10.1057/jam.2010.27.

Andrew Butler and Roy H. Kwon. Integrating prediction in mean-variance portfolio optimization. 11 2022.

B.D. Carlson. Covariance matrix estimation errors and diagonal loading in adaptive arrays. *IEEE Transactions on Aerospace and Electronic Systems*, 24(4):397–401, 7 1988. ISSN 00189251. doi: 10.1109/7.7181.

Leunglung Chan. An Exact Formula for Pricing American Exchange Options with Regime Switching. In R. Mamon and R. Elliott, editors, *Hidden Markov Models in Finance*, volume 209, pages 211–226. Springer, Boston, MA., 2014. doi: 10.1007/978-1-4899-7442-6_9.

Dipeng Chen and Derek Bunn. The Forecasting Performance of a Finite Mixture Regime-Switching Model for Daily Electricity Prices. *Journal of Forecasting*, 33(5):364–375, 8 2014. ISSN 0277-6693. doi: 10.1002/for.2297.

Vijay Kumar. Chopra and William T. Ziemba. The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice. *The Journal of Portfolio Management*, 19(2):6–11, 1 1993. ISSN 0095-4918. doi: 10.3905/jpm.1993.409440.

Roger Clarke, Harindra de Silva, and Steven Thorley. Fundamentals of Efficient Factor Investing (corrected May 2017). *Financial Analysts Journal*, 72(6):9–26, 11 2016. ISSN 0015-198X. doi: 10.2469/faj.v72.n6.3.

John H. Cochrane. Presidential Address: Discount Rates. *The Journal of Finance*, 66(4):1047–1108, 8 2011. ISSN 0022-1082. doi: 10.1111/j.1540-6261.2011.01671.x.

Ricky Cooper and Zixuan Jiao. Combining Value, Profitability, and Momentum: The Details Matter. *The Journal of Portfolio Management*, 51(3): 73–85, 12 2024. ISSN 0095-4918. doi: 10.3905/jpm.2024.1.647.

Giorgio Costa and Roy Kwon. Risk parity portfolio optimization under a Markov regime-switching framework. *Quantitative Finance*, 19(3):453–471, 3 2019. ISSN 1469-7688. doi: 10.1080/14697688.2018.1486036.

Giorgio Costa and Roy Kwon. A regime-switching factor model for mean–variance optimization. *The Journal of Risk*, 4 2020. ISSN 14651211. doi: 10.21314/JOR.2020.432.

H. Cox, R. Zeskind, and M. Owen. Robust adaptive beamforming. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(10):1365–1376, 10 1987. ISSN 0096-3518. doi: 10.1109/TASSP.1987.1165054.

Qiang Dai, Kenneth J. Singleton, and Wei Yang. Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields. *Review of Financial Studies*, 20(5):1669–1706, 9 2007. ISSN 0893-9454. doi: 10.1093/rfs/hhm021.

Kent Daniel and Tobias J. Moskowitz. Momentum crashes. *Journal of Financial Economics*, 122(2):221–247, 11 2016. ISSN 0304405X. doi: 10.1016/j.jfineco.2015.12.002.

Philip I Davies and Nicholas J Higham. Numerically Stable Generation of Correlation Matrices and Their Factors. *Bit Numerical Mathematics*, 40 (4):640–651, 2000. ISSN 00063835. doi: 10.1023/A:1022384216930.

Romain Deguest, Lionel Martellini, and Attilio Meucci. Risk Parity and Beyond—From Asset Allocation to Risk Allocation Decisions. *The Journal of Portfolio Management*, 48(4):108, 2 2022. doi: 10.3905/jpm.2022.1.340.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5):1915–1953, 2009. ISSN 08939454, 14657368.

Weichuan Deng, Paweł Polak, Abolfazl Safikhani, and Ronakdilip Shah. A Unified Framework for Fast Large-Scale Portfolio Optimization. *Data Science in Science*, 3(1):1–21, 12 2024. ISSN 2694-1899. doi: 10.1080/26941899.2023.2295539.

Hubert Dichtl, Wolfgang Drobetz, Harald Lohre, and Carsten Rother. Active Factor Completion Strategies. *The Journal of Portfolio Management*, 47 (2):9–37, 8 2020. ISSN 0095-4918. doi: 10.3905/jpm.2020.1.193.

Eugene F. Fama and Kenneth R. French. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33(1): 3–56, 2 1993. ISSN 0304405X. doi: 10.1016/0304-405X(93)90023-5.

Eugene F. Fama and Kenneth R. French. A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, 116(1):1–22, 4 2015. ISSN 0304405X. doi: 10.1016/j.jfineco.2014.10.010.

Guanhao Feng, Stefano Giglio, and Dacheng Xiu. Taming the Factor Zoo: A Test of New Factors. *Journal of Finance*, 75(3):1327–1370, 6 2020. ISSN 15406261. doi: 10.1111/jofi.12883.

Yiyong Feng and Daniel P. Palomar. A Signal Processing Perspective of Financial Engineering. *Foundations and Trends in Signal Processing*, 9 (1-2):1–231, 2016. ISSN 1932-8346. doi: 10.1561/2000000072.

Nikan Firoozye, Vincent Tan, and Stefan Zohren. Canonical portfolios: Optimal asset and signal combination. *Journal of Banking & Finance*, 154: 106952, 9 2023. ISSN 03784266. doi: 10.1016/j.jbankfin.2023.106952.

Peter C Fishburn. Mean-Risk Analysis with Risk Associated with Below-Target Returns. 67(2):116–126, 1977.

Gregg Fisher, Ronnie Shah, and Sheridan Titman. Combining value and momentum. *Journal Of Investment Management*, 14(2):33–48, 2016.

Joachim Freyberger, Andreas Neuhierl, and Michael Weber. Dissecting Characteristics Nonparametrically. *The Review of Financial Studies*, 33(5): 2326–2377, 5 2020. ISSN 0893-9454. doi: 10.1093/rfs/hhz123.

Khalid Ghayur, Ronan Heaney, and Stephen Platt. Constructing Long-Only Multifactor Strategies: Portfolio Blending vs. Signal Blending. *Financial Analysts Journal*, 74(3):70–85, 7 2018. ISSN 0015-198X. doi: 10.2469/faj.v74.n3.5.

Shihao Gu, Bryan Kelly, and Dacheng Xiu. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5):2223–2273, 5 2020. ISSN 0893-9454. doi: 10.1093/rfs/hhaa009.

M. Guidolin and A. Timmermann. Size and Value Anomalies Under Regime Shifts. *Journal of Financial Econometrics*, 6(1):1–48, 12 2008. ISSN 1479-8409. doi: 10.1093/jjfinec/nbm021.

Massimo Guidolin and Allan Timmermann. Asset Allocation Under Multivariate Regime Switching. *Journal of Economic Dynamics and Control*, 31 (11):3503–3544, 11 2007. ISSN 01651889. doi: 10.1016/j.jedc.2006.12.004.

James D. Hamilton. What's Real About the Business Cycle? *Review, Federal Reserve Bank of St. Louis*, 87(4):435–452, 7 2005. doi: 10.3386/w11161.

James D. Hamilton. Regime Switching Models. In *Macroeconometrics and Time Series Analysis*, pages 202–209. Palgrave Macmillan UK, London, 2010. doi: 10.1057/9780230280830_23.

Campbell R. Harvey, Yan Liu, and Heqing Zhu. and the Cross-Section of Expected Returns. *Review of Financial Studies*, 29(1):5–68, 1 2016. ISSN 14657368. doi: 10.1093/rfs/hhv059.

Campbell R. Harvey, Sandy Rattray, Andrew Sinclair, and Otto Van Hemert. Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance. *The Journal of Portfolio Management*, 43(4):55–69, 7 2017. ISSN 0095-4918. doi: 10.3905/jpm.2017.43.4.055.

Kewei Hou, Chen Xue, and Lu Zhang. Replicating Anomalies. *The Review of Financial Studies*, 33(5):2019–2133, 5 2020. ISSN 0893-9454. doi: 10.1093/rfs/hhy131.

Francisco A. Ibanez. Diversified Spectral Portfolios: An Unsupervised Learning Approach to Diversification. *The Journal of Financial Data Science*, 5(2):67–83, 4 2023. ISSN 2640-3943. doi: 10.3905/jfds.2023.1.118.

Bruce I. Jacobs, Kenneth N. Levy, and Harry M. Markowitz. Portfolio Optimization with Factors, Scenarios, and Realistic Short Positions. *Operations Research*, 53(4):586–599, 8 2005. ISSN 0030-364X. doi: 10.1287/opre.1050.0212.

Joanna Janczura and Rafał Weron. Efficient estimation of Markov regime-switching models: An application to electricity spot prices. *AStA Advances in Statistical Analysis*, 96(3):385–407, 7 2012. ISSN 1863-8171. doi: 10.1007/s10182-011-0181-2.

Zhengyao Jiang, Dixing Xu, and Jinjun Liang. A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem. 7 2017.

Philippe Jorion. Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis*, 21(3):279, 9 1986. ISSN 00221090. doi: 10.2307/2331042.

Philippe. Jorion. *Value at risk : the new benchmark for controlling market risk*. Irwin Professional Pub., 1997. ISBN 0786308486.

Kensuke Kamauchi and Daisuke Yokouchi. A Method for Risk Parity/Budgeting Portfolio Based on Gram-Schmidt Orthogonalization. *Hitotsubashi Journal of Commerce and Management*, 54(1):15–27, 2021.

Chris Kelliher, Avishek Hazrachoudhury, and Bill Irving. A Novel Approach to Risk Parity: Diversification across Risk Factors and Market Regimes. *The Journal of Portfolio Management*, 48(4):73–90, 2 2022. ISSN 0095-4918. doi: 10.3905/jpm.2022.1.327.

Hiroshi Konno and Hiroaki Yamazaki. Mean-Absolute Deviation Portfolio Optimization Model and Its Applications to Tokyo Stock Market. *Management Science*, 37(5):519–531, 5 1991. ISSN 0025-1909. doi: 10.1287/mnsc.37.5.519.

Serhiy Kozak, Stefan Nagel, and Shrihari Santosh. Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292, 2 2020. ISSN 0304405X. doi: 10.1016/j.jfineco.2019.06.008.

Mark Kritzman, Sébastien Page, and David Turkington. Regime Shifts: Implications for Dynamic Strategies (corrected). *Financial Analysts Journal*, 68(3):22–39, 5 2012. ISSN 0015-198X. doi: 10.2469/faj.v68.n3.3.

Nathan Lassance, Victor DeMiguel, and Frédéric Vrins. Optimal Portfolio Diversification via Independent Component Analysis. *Operations Research*, 70(1):55–72, 8 2022. ISSN 0030-364X. doi: 10.1287/opre.2021.2140.

Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2): 365–411, 2 2004. ISSN 0047259X. doi: 10.1016/S0047-259X(03)00096-4.

Yongjae Lee, John R. J. Thompson, Jang Ho Kim, Woo Chang Kim, and Francesco A. Fabozzi. An Overview of Machine Learning for Asset Management. *The Journal of Portfolio Management*, 49(9):31–63, 8 2023. ISSN 0095-4918. doi: 10.3905/jpm.2023.1.526.

Markus Leippold and Roger Rueegg. The mixed vs the integrated approach to style investing: Much ado about nothing? *European Financial Management*, 24(5):829–855, 11 2018. ISSN 1354-7798. doi: 10.1111/eufm.12139.

Ashley Lester. On the Theory and Practice of Multifactor Portfolios. *The Journal of Portfolio Management*, 45(3):87–100, 2 2019. ISSN 0095-4918. doi: 10.3905/jpm.2019.45.3.087.

Jun Li. Explaining Momentum and Value Simultaneously. *Management Science*, 64(9):4239–4260, 9 2018. ISSN 0025-1909. doi: 10.1287/mnsc.2017.2735.

John Lintner. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1):13, 2 1965. ISSN 00346535. doi: 10.2307/1924119.

Harald Lohre, Ulrich Neugebauer, and Carsten Zimmer. Diversified Risk Parity Strategies for Equity Portfolio Selection. *The Journal of Investing*, 21(3):111–128, 8 2012. ISSN 1068-0896. doi: 10.3905/joi.2012.21.3.111.

Harald Lohre, Heiko Opfer, and Gabor Orszag. Diversifying Risk Parity. *Risk*, 16(5):53–79, 2014.

Marcos López de Prado. Building Diversified Portfolios that Outperform Out of Sample. *The Journal of Portfolio Management*, 42(4):59–69, 5 2016. ISSN 0095-4918. doi: 10.3905/jpm.2016.42.4.059.

Marcos López de Prado. *Advances in Financial Machine Learning*. John Wiley & Sons, Hoboken, NJ, 1 edition, 2018.

Eric Luxenberg and Stephen Boyd. Portfolio construction with Gaussian mixture returns and exponential utility via convex optimization. *Optimization and Engineering*, 25(1):555–574, 3 2024. ISSN 1389-4420. doi: 10.1007/s11081-023-09814-y.

Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche. The Properties of Equally Weighted Risk Contribution Portfolios. *The Journal of Portfolio Management*, 36(4):60–70, 8 2010. ISSN 0095-4918. doi: 10.3905/jpm.2010.36.4.060.

Dimitris. Manolakis, Vinay. Ingle, and Stephen. Kogon. *Statistical and Adaptive Signal Processing : Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*. Artech House, 2005. ISBN 0071166602.

Harry Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77, 3 1952. ISSN 00221082. doi: 10.2307/2975974.

Harry Markowitz, Peter Todd, Ganlin Xu, and Yuji Yamane. Computation of mean-semivariance efficient sets by the Critical Line Algorithm. *Annals of Operations Research*, 45(1):307–317, 12 1993. ISSN 0254-5330. doi: 10.1007/BF02282055.

Lionel Martellini and Vincent Milhau. Proverbial Baskets Are Uncorrelated Risk Factors! A Factor-Based Framework for Measuring and Managing Diversification in Multi-Asset Investment Solutions. *The Journal of Portfolio Management*, 44(2):8–22, 8 2017. ISSN 0095-4918. doi: 10.3905/jpm.2018.44.2.008.

Attilio Meucci, Alberto Santangelo, and Romain Deguest. Risk budgeting and diversification based on optimized uncorrelated factors. *Risk*, 11:70–75, 2015.

Richard O Michaud. The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, 45(1):31–42, 1989. ISSN 0015198X.

Robert A. Monzingo, Randy L. Haupt, and Thomas W. Miller. *Introduction to Adaptive Arrays*. Institution of Engineering and Technology, 1 2011. ISBN 9781891121579. doi: 10.1049/SBEW046E.

Daniel Müllner. Modern Hierarchical, Agglomerative Clustering Algorithms. *arXiv preprint arXiv:1109.2378*, 2011.

John M. Mulvey and Han Liu. Identifying Economic Regimes: Reducing Downside Risks for University Endowments and Foundations. *The Journal of Portfolio Management*, 43(1):100–108, 10 2016. ISSN 0095-4918. doi: 10.3905/jpm.2016.43.1.100.

Andreas Neuhierl, Otto Randl, Christoph Reschenhofer, and Josef Zechner. Timing the Factor Zoo. *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4376898.

Robert Novy-Marx. Backtesting Strategies Based on Multiple Signals. Technical report, National Bureau of Economic Research, Cambridge, MA, 7 2015.

Peter Nystrup, Bo William Hansen, Henrik Madsen, and Erik Lindström. Regime-Based Versus Static Asset Allocation: Letting the Data Speak. *The Journal of Portfolio Management*, 42(1):103–109, 10 2015. ISSN 0095-4918. doi: 10.3905/jpm.2015.42.1.103.

Peter Nystrup, Stephen Boyd, Erik Lindström, and Henrik Madsen. Multi-period Portfolio Selection With Drawdown Control. *Annals of Operations Research*, 282(1):245–271, 11 2019. ISSN 0254-5330. doi: 10.1007/s10479-018-2947-3.

Peter Nystrup, Erik Lindström, and Henrik Madsen. Learning hidden Markov models with persistent states by penalizing jumps. *Expert Systems with Applications*, 150:113307, 7 2020. ISSN 09574174. doi: 10.1016/j.eswa.2020.113307.

Michael T. Owyang and Garey Ramey. Regime Switching and Monetary Policy Measurement. *Journal of Monetary Economics*, 51(8):1577–1597, 11 2004. ISSN 03043932. doi: 10.1016/j.jmoneco.2004.01.004.

Bijon Pani and Frank J. Fabozzi. Finding Value Using Momentum. *The Journal of Portfolio Management*, 48(2):264–283, 12 2021. ISSN 0095-4918. doi: 10.3905/jpm.2021.1.272.

M Hossein Partovi and Michael Caputo. Principal portfolios: Recasting the efficient frontier. *Economics Bulletin*, 7(3):1–10, 2004.

Eero Pätäri, Ville Karell, Pasi Luukka, and Julian S Yeomans. Comparison of the multicriteria decision-making methods for equity portfolio selection: The U.S. evidence. *European Journal of Operational Research*, 265(2): 655–672, 3 2018. ISSN 03772217. doi: 10.1016/j.ejor.2017.08.001.

Andre F. Perold. Large-Scale Portfolio Optimization. *Management Science*, 30(10):1143–1160, 10 1984. ISSN 0025-1909. doi: 10.1287/mnsc.30.10.1143.

Thomas Raffinot. Hierarchical Clustering-Based Asset Allocation. *The Journal of Portfolio Management*, 44(2):89–99, 12 2017. ISSN 0095-4918. doi: 10.3905/jpm.2018.44.2.089.

Lorenzo Reus and John M. Mulvey. Dynamic allocations for currency futures under switching regimes signals. *European Journal of Operational Research*, 253(1):85–93, 8 2016. ISSN 03772217. doi: 10.1016/j.ejor.2016.02.024.

R.Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, 7 2002. ISSN 03784266. doi: 10.1016/S0378-4266(02)00271-6.

Stephen A Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360, 12 1976. ISSN 00220531. doi: 10.1016/0022-0531(76)90046-6.

Alexander M Rudin and jonathan S Morgan. A Portfolio Diversification Index. *The Journal of Portfolio Management*, 32(2):81–89, 2006. ISSN 0095-4918. doi: 10.3905/jpm.2006.611807.

Tobias Rydén, Timo Teräsvirta, and Stefan Åsbrink. Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics*, 13(3):217–244, 5 1998. ISSN 0883-7252. doi: 10.1002/(SICI)1099-1255(199805/06)13:3¡217::AID-JAE476¿3.0.CO;2-V.

Friedrich Schmid and Mark Trede. Testing for First-Order Stochastic Dominance: A New Distribution-Free Test. *The Statistician*, 45(3):371, 1996. ISSN 00390526. doi: 10.2307/2988473.

William F. Sharpe. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3):425, 9 1964. ISSN 00221082. doi: 10.2307/2977928.

Abdullah Z. Sheikh and Jianxiong Sun. Regime Change: Implications of Macroeconomic Shifts on Asset Class and Portfolio Performance. *The Journal of Investing*, 21(3):36–54, 8 2012. ISSN 1068-0896. doi: 10.3905/joi.2012.21.3.036.

Yizhan Shu, Chenyu Yu, and John M. Mulvey. Downside risk reduction using regime-switching signals: a statistical jump model approach. *Journal of Asset Management*, 25(5):493–507, 9 2024. ISSN 1470-8272. doi: 10.1057/s41260-024-00376-x.

Tyler Shumway. The Delisting Bias in CRSP Data. *The Journal of Finance*, 52(1):327–340, 3 1997.

Christopher A Sims and Tao Zha. Were There Regime Switches in U.S. Monetary Policy? *American Economic Review*, 96(1):54–81, 2 2006. ISSN 0002-8282. doi: 10.1257/000282806776157678.

Tak Kuen Siu. A Hidden Markov-Modulated Jump Diffusion Model for European Option Pricing. In R. Mamon and R. Elliott, editors, *Hidden Markov Models in Finance*, volume 209, pages 185–209. Springer, Boston, MA., 2014. doi: 10.1007/978-1-4899-7442-6_8.

Simon C Smith and Allan Timmermann. Break Risk. *The Review of Financial Studies*, 34(4):2045–2100, 3 2021. ISSN 0893-9454. doi: 10.1093/rfs/hhaa084.

Meir Statman. The Diversification Puzzle. *Financial Analysts Journal*, 60 (4):44–53, 7 2004. ISSN 0015-198X. doi: 10.2469/faj.v60.n4.2636.

G W Stewart. On the Early History of the Singular Value Decomposition. *SIAM Review*, 35(4):551–566, 8 1993. ISSN 0036-1445. doi: 10.1137/1035134.

Rodney N. Sullivan. Hedge Fund Alpha: Cycle or Sunset? *The Journal of Alternative Investments*, 23(3):55–79, 12 2020. ISSN 1520-3255. doi: 10.3905/jai.2020.1.118.

D. Sundararajan. *Signals and Systems*. Springer Nature Switzerland, Cham, 2023. ISBN 978-3-031-19376-7. doi: 10.1007/978-3-031-19377-4.

Lizhe Tan and Jean Jiang. *Digital Signal Processing*. Academic Press, 3 edition, 2019. ISBN 9780128150719. doi: 10.1016/C2017-0-02319-4.

Ruey S. Tsay. *Analysis of Financial Time Series*. Wiley, 8 2010. ISBN 9780470414354. doi: 10.1002/9780470644560.

Peng Wang and Larry Kochard. Using a Z-score Approach to Combine Value andMomentum in Tactical Asset Allocation. *The Journal of Wealth Management*, 15(1):52–71, 4 2012. ISSN 1534-7524. doi: 10.3905/jwm.2012.15.1.052.

Panos Xidonas, Ralph Steuer, and Christis Hassapis. Robust portfolio optimization: a categorized bibliographic review. *Annals of Operations Research*, 292(1):533–552, 9 2020. ISSN 0254-5330. doi: 10.1007/s10479-020-03630-8.

Xuemin (Sterling) Yan and Lingling Zheng. Fundamental Analysis and the Cross-Section of Stock Returns: A Data-Mining Approach. *The Review of Financial Studies*, 30(4):1382–1423, 4 2017. ISSN 0893-9454. doi: 10.1093/rfs/hhx001.

Wei Yang. Business Cycles, Regime Shifts, and Return Predictability. *Journal of Financial and Quantitative Analysis*, 58(7):3058–3084, 11 2023. ISSN 0022-1090. doi: 10.1017/S0022109023000261.

Mengyi Zhang, Francisco Rubio, and Daniel P. Palomar. Improved Calibration of High-Dimensional Precision Matrices. *IEEE Transactions on Signal Processing*, 61(6):1509–1519, 3 2013. ISSN 1053-587X. doi: 10.1109/TSP.2012.2236321.

Zhihua Zhang. The Singular Value Decomposition, Applications and Beyond. 8 2015.

Zihao Zhang, Stefan Zohren, and Stephen Roberts. Deep Learning for Portfolio Optimization. *The Journal of Financial Data Science*, 2(4):8–20, 10 2020. ISSN 2640-3943. doi: 10.3905/jfds.2020.1.042.

Kai Zheng, Yuying Li, and Weidong Xu. Regime switching model estimation: spectral clustering hidden Markov model. *Annals of Operations Research*, 303(1-2):297–319, 8 2021. ISSN 0254-5330. doi: 10.1007/s10479-019-03140-2.