PORTFOLIO OPTIMIZATION WHEN ASSET RETURNS HAVE THE GAUSSIAN MIXTURE DISTRIBUTION

IAN BUCKLEY, DAVID SAUNDERS, AND LUIS SECO

Abstract. Portfolios of assets with Gaussian mixture (GM) distributed returns are optimised in the static setting to find portfolio weights and efficient frontiers using a selection of objective functions, including the Sharpe ratio, the probability of outperforming a target return, Hodges' modified Sharpe ratio, expected shortfall and semi-variance. The probability of outperformance is seen as the natural generalisation to the GM setting of the Sharpe ratio objective in the mean-variance setting. However, both objectives suffer mild pathologies that ultimately lead us to favour alternatives. We associate component Gaussian distributions in a bi-Gaussian mixture with tranquil and distressed market regimes. Just as in the mean-variance case with the Sharpe ratio objective, our algorithm for solving the (non-linear) problem relies on our ability to restrict our search to the efficient frontier, a subspace of the low-dimensional portfolio coordinate space. Sufficient conditions on the portfolio coordinate and non-linear objective functions to guarantee that optima be efficient are that they be concave and increasing, respectively. Because our coordinates are linear or quadratic functions of portfolio weights, then points on the efficient frontier are solutions to linear-quadratic problems. Thus, justified by the Karush-Kuhn-Tucker theorem, the non-linear programming problem in the portfolio weight space is reduced to the non-linear optimisation of a three-parameter family of quadratic programs. Because for Gaussian returns an objective of the form of an expected utility function is only decreasing in variance if the utility is concave, we favour expected shortfall, semi-variance, expected exponential utility and Hodges ratio objectives for use in the GM setting. Probability of shortfall can be made to be increasing in variance only by imposing an inconvenient restriction on the target value.

Contents

Date: January 29, 2004.

Key words and phrases. Mixture of normals distribution, Gaussian mixture distribution, portfolio optimization, market distress, hedge fund portfolio, Sharpe ratio, lower partial moment, efficient frontier, Hodges' modified Sharpe ratio, exponential utility, correlation switching, regime switching, asset allocation, commodity trading advisor, probability of shortfall, distress sensitivities, fund of funds.

We would like the thank the organizers and participants of the conference for the opportunity to present this material and Gustavo Comezaña and Ben Djerroud for many helpful discussions.

One of us (IB) would like to thank Mr. Gerry Salkin and Prof. Nicos Christofides of the Centre for Quantitative Finance, Imperial College for financial support and the opportunity to perform sabbatical work in Toronto, and his colleagues at RiskLab, University of Toronto, for making him welcome there.

1. Introduction

In this paper we generalize Markowitz mean-variance portfolio theory [25], the cornerstone of single-period investment theory, to describe portfolios of assets whose returns are described by the (finite) *Gaussian mixture* (GM) (alternatively *mixture of normals*) distribution. We adopt novel objective functions, such as the probability of outperformance, which give rise to non-linear problems (NLPs), and from lessons learned from studying the mean-variance case with Sharpe ratio objective, we are able to construct a general framework in which to maximise these efficiently. A theorem provides conditions under which objectives will necessarily have efficient extrema and provides the key to an algorithm.

Whilst the assets in the universe could be of the conventional variety, such as equities or bonds, our primary goal is to develop a framework for the management of portfolios of funds, such as hedge funds or *commodity trading advisors* (CTAs).

1.1. Objectives for a novel distribution.

The standard assumption that asset returns have the multivariate Gaussian distribution is reasonable and permits the construction of tractable and useful models. More elaborate models frequently aim to capture richer, non-Gaussian features of asset returns, such as the skewed and leptokurtotic nature of distributions of individual assets and the correlation breakdown phenomenon associated with the dependence structure of multiple assets. This motivates our adoption here, as in [7], of the GM distribution, which is capable of handling such features – in particular lobed iso-probability contours – in a tractable way.

Apart from the distributional assumption, the second fundamental difference between the GM approach and the mean-variance approach[1] is that we are obliged to adopt alternative objective functions. Otherwise, whatever distribution assumptions we make, our model will still be of the standard mean-variance type and recommend the same optimal portfolios.

We introduce the notion of a *portfolio coordinate*, which is a function mapping from the *portfolio weight space* into the *portfolio coordinate space*. The set of feasible portfolios and the *efficient frontier* are subsets of the portfolio weight space, whilst the *investment opportunity set* is a subset of the portfolio coordinate space. Objectives are dependent on portfolio weights through their dependence on portfolio coordinates.

This paper can be seen as a generalisation of the procedure used in the mean-variance approach for maximising the portfolio Sharpe ratio [28]. Because in the Gaussian mean-variance setting the Sharpe ratio and probability of outperformance objectives are equivalent in that the latter reduces to an increasing function of the former, we take a particular interest in these two objectives. The probability of outperformance plays a role in the GM approach analogous to the Sharpe ratio in the mean-variance approach; the former objective is a generalisation of the latter one. However, in due course we will discover that the probability of outperformance is mildly pathological (in fact through features inherited from the Sharpe ratio) and we will cast our net wider to propose other, more robust objectives to incorporate into the GM approach.

[1] The mean-variance approach assumes *any* distribution and a mean-variance objective: one dependent on the portfolio return distribution only through its first two moments.

We give definitions for these two non-linear objectives to give a flavour of the challenges we face, at this stage without introducing portfolio coordinates.

Sharpe ratio: in the classical mean-variance setting, takes the form:

$$F_k^{\mathrm{SR}}(\boldsymbol{\theta}) = \frac{\boldsymbol{\mu}'.\boldsymbol{\theta} - k}{\sqrt{\boldsymbol{\theta}'.\boldsymbol{V}.\boldsymbol{\theta}}}$$

where $\boldsymbol{\theta}$ is the vector of asset weights in the portfolio, $\boldsymbol{\mu}$ and $\boldsymbol{V}$ are the vector of means and matrix of covariances for the asset returns and k is the portfolio return target.

Probability of outperformance: in a two-regime GM setting is given by:

$$F_k^{\mathrm{PO}}(\boldsymbol{\theta}) = (1-w)\ \Phi\left(\alpha_{\mathrm{T}}(\boldsymbol{\theta},k)\right) + w\ \Phi\left(\alpha_{\mathrm{D}}(\boldsymbol{\theta},k)\right) \tag{1.1.1}$$

where

$$\alpha_i(\boldsymbol{\theta},k) = \frac{\boldsymbol{\mu}_i'.\boldsymbol{\theta} - k}{\sqrt{\boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}}}$$

is the *regime Sharpe ratio* for regime i in terms of the vectors of means $\boldsymbol{\mu}_i$ and matrices of covariances $\boldsymbol{V}_i$ for the asset returns, with $i \in \{\mathrm{T},\mathrm{D}\}$, in the *tranquil* and *distressed* market regimes, w is the GM distribution regime mixing parameter, k is the portfolio return target and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Whilst neither of these objectives, nor any of the others investigated, give rise to convex problems in portfolio weight space (Section 3.4.3), it is gratifying that we are nevertheless able to employ the mainstay of convex optimisation theory, the Karush-Kuhn-Tucker (KKT) theorem, to prove our main result (Section 3.5).

An important class of objectives contains those that are the expected value of the utility of the return. In the Gaussian or GM setting, these are necessarily increasing (increasing and decreasing in the return and variance, respectively) when the utility function is concave. It is only for such objectives that we are always entitled to use the search-space dimension reduction algorithm. From this perspective target shortfall, target semi-variance, and expected exponential and quadratic utility functions, as expectations of concave utility functions, are increasing objectives and therefore desirable objectives. Conversely, probability of outperformance can only be used with caution because it is only increasing for a sufficiently small, possibly negative, target parameter.

1.2. **An algorithm.** When we maximise the Sharpe ratio in the mean-variance setting, we first solve a family of linear-quadratic programs (LQPs) to find the efficient frontier. Having done so, we conduct our non-linear search for the portfolio with maximum Sharpe ratio only along the efficient frontier, which is a 1-dimensional curve in the 2-dimensional portfolio coordinate space consisting of the first two moments. Without this short-cut, we would have had to search for the maximum to an NLP in the portfolio weight space, which is of dimensionality equal to the number of assets.

By analogy with the above, in the GM setting we adopt the same strategy of reducing the dimensionality of the problem by conducting our search within the efficient subspace of the portfolio coordinate space, rather than within the portfolio weight space. As in the mean-variance setting, the portfolio coordinate space is typically of lower dimension than the portfolio weight space. To this end we consider sufficient conditions for the GM case and in general under which the solutions to the NLPs will be Pareto efficient and hence correspond to solutions to associated *portfolio coordinate linear* problems.

A theorem states that in any model in which the following functions have the given properties:

- portfolio coordinates - *convex* (as a function of portfolio weights)
- objective - *increasing* (as a function of portfolio coordinates)

each optimum will be efficient and hence the optimum of an associated linear problem. That is to say, the optimum of an increasing objective is necessarily the optimum of a linear objective in portfolio coordinate space.

The benefit of this is that if the portfolio coordinates are themselves linear or quadratic functions of the weight space, as we can choose them to be in the GM case if we use the regime means and variances for this purpose, then the portfolio coordinate linear problems will be *linear-quadratic problems* (LQPs) in portfolio weight space.

For non-linear objectives, by performing the search inside the subspace of portfolios that are efficient with respect to a particular basis of portfolio coordinates, the dimensionality of the NLP can be reduced to improve the efficiency and speed of the solution algorithm. The NLP in the $(m-1)$-dimensional budget-constrained portfolio weight space is reduced to the non-linear optimization of a $p-1$ parameter family of LQPs, where p and m are the numbers of portfolio coordinates and assets, respectively. In a two regime GM model ($n = 2$), taking two moments per regime, $p - 1 = 2n - 1 = 3$.

1.3. **Plan for paper.** The plan for the paper is as follows. We are introduced to the GM distribution in Section 2 and Appendix A where we find definitions of GM distributed random variables, estimation issues, and some identities concerning moments and the distribution of linear combinations of GM variables. We review the literature on mixture distributions and the related topic of mixture processes. In Section 3 we assume GM distributed asset returns, take a number of non-linear objectives including the probability of outperformance, and test models based on these ingredients. Here we describe our key theoretical result: in a general framework for the construction of static optimisation problems we give conditions on portfolio coordinates and objectives that are sufficient to ensure that optimal portfolios are efficient and can be obtained as global solutions to associated LQPs. In Section 4 numerical results, in particular the dependence of portfolio asset weights and objectives against the target return and mixing parameters, are presented graphically. Section 5 contains our conclusions.

2. Gaussian mixture distribution

2.1. **Limitations of the mean-variance approach.** As [29] explain,

> the need for models that go beyond the Gaussian paradigm is vividly felt by practitioners, regulatory agencies and is also advocated in the academic literature.

The standard assumption that asset returns have the multivariate Gaussian distribution is a reasonable first approximation to reality and gives rise to tractable theories. Many theories forming the foundations of mathematical finance implicitly adopt this conjecture, including Markowitz portfolio theory[2] [25], [30], and the CAPM and APT equity pricing models. The assumption in Black-Scholes-Merton option pricing theory that stock prices follow geometric Brownian motion is the natural generalisation of this idea to the continuous-time dynamic setting.

However, it is well-known that for assets, both in the conventional sense of equities and bonds, but also in a broader sense, for example in the form of country or sector-based equity or bond indices, and alternative investments such as hedge funds and CTAs, the situation is more complex [15]. The purpose of the generalization described in this paper is to address two well-known deficiencies of the use of the multivariate Gaussian distribution with constant parameters to model asset returns:

- The skewed (asymmetric around the mean) and leptokurtotic (more kurtotic or 'fat-tailed' than a Gaussian distribution) nature of marginal probability density functions (pdfs)
- The *asymmetric correlation* (or *correlation breakdown* or *perfect storm* or *contagion*) phenomenon, which describes the tendency for the volatilities of and correlations between asset returns to be dependent on the prevailing direction of the market. A widely held view is that volatilities and correlations are larger in a bear market than a bull market, particularly in an international context. Important investigations of this phenomenon include [13], [23] and [8].

The first point refers to the univariate distributions for returns to individual assets considered in isolation; the second describes effects exhibited by multiple assets considered together. Whilst both phenomena can be comprehensively modelled only in a dynamic setting, even in the static setting it is possible to build models that capture at least elements of both effects simply by adopting non-Gaussian distributions.

2.2. **Choosing a distribution.** The Gaussian mixture distribution is selected from the range of parametric alternatives to the Gaussian distribution for its tractability: calculations using it often closely resemble those using the Gaussian distribution. Other alternatives to the Gaussian tend to suffer from one of two evils: either they are too restrictive in the variety of pdf shapes that can be achieved, or not restrictive enough, in the sense that they have too many degrees of freedom for calibration to be feasible.

Alternative distributions to the multivariate Gaussian for asset returns include the generalised hyperbolic [12] and the multivariate stable [27]. To calibrate these models, the former requires only a finite number of parameters to be found, whereas for the latter, it is a considerable feat to specify a *spectral measure*. Many of the distributions that are special cases of the generalised hyperbolic, such as the multivariate t, normal inverse Gaussian, and Gaussian distributions, are elliptic, or close to being so. A benefit of using elliptic multivariate pdfs, i.e., those with

[2]In fact, Markowitz merely assumes that investors are concerned with the mean and the variance of the portfolio return, i.e., that they possess quadratic utility functions, and makes no explicit assumptions about the distribution beyond the finiteness of its first two moments. However, mean-variance optimal portfolios maximise the expected value of any concave, increasing utility function when asset returns are governed by an elliptic distribution, such as the multivariate Gaussian. In this sense the mean-variance approach implicitly assumes a Gaussian (or other elliptic) distribution.

ellipsoidal iso-probability surfaces, such as the Gaussian and multivariate t, is that the correlation matrix fully and parsimoniously describes the dependence structure of the constituent risks. Conversely, a general non-elliptic distribution requires cross moments of all orders to fully describe the dependence structure[3].

Whilst such distributions are tractable, symmetric elliptical pdfs for asset returns are not what we observe in real financial markets. The multivariate GM distribution is appealing in that whilst it is highly non-elliptic, its pdf and dependence structure are fully and conveniently specified by the (multiple) mean return vectors, (multiple) covariance matrices and relative weights of the constituent Gaussian components; i.e., a finite number of parameters. Techniques such as principal component analysis exist for reducing the dimensionality of the covariance matrix in each regime. Similar tools for more exotic distributions are still in their infancy.

Almost uniquely amongst parametric distributions and unlike the generalised hyperbolic family of distributions, the GM distribution can model protuberances on the probability iso-density contours. Mixture distributions have the appeal that by adding together a sufficient number of component distributions, any multivariate distribution may be approximated to arbitrary accuracy.

2.3. **Mixture literature.**

2.3.1. *Random variables.* The GM distribution is seen quite often in the field of finance, mostly in its univariate guise for the estimation of Value at Risk (VaR). [20] develop a model for estimating VaR in which the user is free to choose any probability distribution for the daily changes in each market variable and employ the univariate mixture of normals distribution as an example. In the same field, [34] assumes probability distributions for each of the parameters describing the mixture of normals and uses a Bayesian updating scheme; and [32] uses a quasi-Bayesian maximum likelihood estimation procedure. The current RiskMetrics™ methodology uses GM with a mixture of two normal distributions. More recently GM models have been used [22] to model futures markets and for portfolio risk management and by [14] for credit risk. [33] develops an efficient analytical Monte Carlo method for generating changes in asset prices using a multivariate mixture of normal distributions with arbitrary covariance matrix.

The *finite* GM distribution is closely related to the semi-parametric *normal variance-mean mixture* distributions of [14] and [6], which combine normal distributions continuously. This family of distributions contains the generalised hyperbolic distribution, and therefore the multivariate t, hyperbolic, normal inverse Gaussian and Gaussian distributions, as special cases.

2.3.2. *Stochastic processes.* Although this paper is solely concerned with the static case, we do maintain an interest in the dynamic case, i.e., the multiple period discrete or continuous-time process setting, because we wish to motivate the use of the GM distribution and we prefer to construct static models that extend naturally to the dynamic case.

There is a growing body of work in which exotic (asset return) stochastic processes have been constructed by mixing simpler ones. Processes for asset returns may be constructed from unconditional or conditional distributions. As an example of the latter case, by mixing autoregressive processes such as ARCH and GARCH,

---

[3]At least for distributions for which these are finite.

processes can be constructed that can account for both the heteroscedastic and leptokurtic nature of financial time series. [31] describe computational tools for the calculation of VaR and other more sophisticated risk measures such as shortfall, Max-VaR, conditional VaR and conditional risk measures that aim to take account of the heteroskedastic structure of time series. More recent papers on this topic include [26] and [16].

GM distributions can arise naturally as the level of certain stochastic processes at a point in time, conditional on the level at an earlier time e.g., Markov (regime) switching models and jump processes. Regime switching models describe processes in which parameters of a continuous diffusion process may change discontinuously according to the realized stochastic path through an associated Markov chain. Recent applications of regime switching asset return processes include: in the field of Merton-style option pricing theory, [11], and in portfolio management [9] (CAPM) and [1], [2] (international diversification).

2.4. **Definition.**

Definition 2.4.1. If U, is a discrete random variable taking on values $i = 1, 2, \dots n$ with $\mathbb{P}[U = i] = p_i$ and the random variable $\boldsymbol{X}$ is conditionally multivariate normal on the outcome of U

$$\boldsymbol{X}|(U = i) \sim N_m(\boldsymbol{\mu}_i, \boldsymbol{V}_i),$$

where $\boldsymbol{V}_i$ is a symmetric positive definite $m \times m$ matrix and $\boldsymbol{\mu}_i$ is an m-vector, then $\boldsymbol{X}$ has the Gaussian mixture distribution.

2.5. **Estimation.** When calibrating a model, a disadvantage of using the GM distribution is that the log-likelihood function does not have a global minimum. A resolution to this problem explored in [17] is to use a modified log-likelihood function. Because of the use of the GM distribution and other mixture distributions in image processing, clustering, and unsupervised learning a host of estimation techniques have been developed for it [10]. When using the GM distribution to model asset returns, [33] employs the EM algorithm.

2.6. **Identities.** In Appendix A a further definition of the GM distribution and various identities involving its moments and the behaviour of linear combinations of GM variables are presented.

Key observations include:

- The overall variance and covariance of a GM distribution depends not only on the regime variances and covariances, but also on the difference between the regime means (Equations A.2.4 and A.2.6)
- The portfolio return is itself (univariate) GM distributed if asset returns are (multivariate) GM distributed. (Equation A.3.1.) In Figure 1 we see that, within each regime, the mechanism for diversification carries over intact from the mean-variance case. Each curve is the pdf for the return to a two-asset portfolio.

2.7. **Asset return regimes.** In the GM framework for portfolio management, the n Gaussian contributions to the GM distribution are associated with n *regimes*. Regime i is allotted a weight w_i, with $\sum_{i=1}^{n} w_i = 1$. We shall refer to w_i as the *probability* of or *mixing parameter* for regime i. Each regime has an m-vector of means $\boldsymbol{\mu}_i$ and an $m \times m$ covariance matrix $\boldsymbol{V}_i$, where m is the number of assets in the universe.

In all the numerical examples, we take $n = 2$, and refer to the two regimes as *tranquil* and *distressed*. The tranquil (distressed) regimes are typically associated with low (high) levels of both variance and correlation.

2.8. **Visualization.** In Figure 2, in four contour plots of pdfs, we see how two bivariate Gaussian distributions (top row) can be added to yield a GM distribution (bottom, left). Note the potential for highly non-elliptic iso-probability contours when the GM distribution is used. For comparison a bivariate normal with the same overall means, variances and covariances is included (bottom, right). If the GM distribution given were used to describe the returns to two assets, the lobe pointing down and to the left of the figure would describe the propensity of the two assets to decline sharply together, which is an essential feature of the asymmetric correlation phenomenon. The Gaussian distribution, with its elliptic contours, is clearly unable to capture this feature.

These examples illustrate the two asset case. With three assets the contours for the Gaussian distribution are three-dimensional ellipsoids (rather than ellipses in two dimensions), and for the GM distribution the contours are complicated lobed surfaces embedded in a three-dimensional space. In the m-asset case with two regimes, one exhibiting low and one high correlation and with similar regime means, a typical iso-probability surface in the m-dimensional asset return space resembles a (short, fat ellipsoid) ball pierced by a (long, thin ellipsoid) stick: a useful image is of a cherry on a cocktail stick.

3. Portfolio optimization

3.1. **Portfolio coordinates.** All objectives are expressed in terms of statistics of the distribution, called *portfolio coordinates*, which map points in the m-dimensional space of portfolio asset weights into the p-dimensional space of portfolio coordinates. In the parlance of multiple objective optimisation theory, the portfolio weight and portfolio coordinate spaces are described as *design* and *criterion* spaces, respectively. We will choose our sign convention such that large values for all coordinates are undesirable: we prefer smaller coordinate values in preference to larger ones.

Definition 3.1.1. A function $f_i : \mathbb{R}^m \to \mathbb{R}$, $i = 1, \ldots, p$ is called a *portfolio coordinate.* A collection of p portfolio coordinates $\boldsymbol{f} = (f_1, \ldots, f_p) : \mathbb{R}^m \to \mathbb{R}^p$ maps the *portfolio weight space*, $\mathbb{R}^m$ into the *portfolio coordinate space*, $\mathbb{R}^p$.

Figure 3 shows the portfolio weight (design) space $\mathbb{R}^m$, portfolio coordinate (criterion) space $\mathbb{R}^p$ and the portfolio coordinate function $\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^p$ and objectives with coordinate space $\hat{F} : \mathbb{R}^p \to \mathbb{R}$ and weight space $F : \mathbb{R}^m \to \mathbb{R}$ domains.

3.1.1. *Mean variance portfolio coordinates.* The classical mean-variance theory uses a two-dimensional portfolio coordinate space

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{\theta}) = (\bar{L}(\boldsymbol{\theta}), Q(\boldsymbol{\theta})) \in \mathbb{R}^2,$$

where the negative mean (linear) and variance (convex[4]) of the portfolio return are:

$$\bar{L}(\boldsymbol{\theta}) = -L(\boldsymbol{\theta}) = -\boldsymbol{\mu}.\boldsymbol{\theta}$$
$$Q(\boldsymbol{\theta}) = \boldsymbol{\theta}'.\boldsymbol{V}.\boldsymbol{\theta},$$

where $\boldsymbol{\mu}$ is the m-vector of asset means, $\boldsymbol{V}$ is the $m \times m$ matrix of asset covariances and $\boldsymbol{\theta}$ is the m-vector of asset portfolio weights. We take the portfolio mean with negative sign so that our objectives to be minimised (maximised) will be increasing (decreasing) in all portfolio coordinates. Whilst this is our preferred sign convention, we shall not always abide by it, particularly on the axes of certain figures.

3.1.2. *Gaussian mixture portfolio coordinates.* In the GM setting $p = 2n$, where n is the number of regimes.

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{\theta}) = (\bar{L}_1(\boldsymbol{\theta}), \ldots, \bar{L}_n(\boldsymbol{\theta}), Q_1(\boldsymbol{\theta}), \ldots, Q_n(\boldsymbol{\theta})) \in \mathbb{R}^{2n}.$$

The regime specific linear negative first and quadratic second moment functions for regime $i = 1, \ldots, n$ are

$$\bar{L}_i(\boldsymbol{\theta}) = -L_i(\boldsymbol{\theta}) = -\boldsymbol{\mu}_i'.\boldsymbol{\theta}$$
$$Q_i(\boldsymbol{\theta}) = \boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}.$$

In terms of these, the Sharpe ratio for regime i becomes:

$$\alpha_i(\boldsymbol{\theta}, k) = \frac{L_i(\boldsymbol{\theta}) - k}{\sqrt{Q_i(\boldsymbol{\theta})}}.$$

For the n regime case, this $2n$-dimensional portfolio coordinate space of regime means and variances is highly appropriate for calculational purposes, because in this space the efficient frontier is the solution to a family of relatively much easier to solve LQPs.

In the two regime case of our numerical studies, $p = 2n = 4$ and $i \in \{\mathrm{T}, \mathrm{D}\}$ (tranquil or distressed):

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{\theta}) = (\bar{L}_{\mathrm{T}}(\boldsymbol{\theta}), \bar{L}_{\mathrm{D}}(\boldsymbol{\theta}), Q_{\mathrm{T}}(\boldsymbol{\theta}), Q_{\mathrm{D}}(\boldsymbol{\theta})) \in \mathbb{R}^4.$$

3.1.3. *Portfolio coordinates reconciled.* To compare objectives from the mean-variance and GM approaches, we need to express coordinates from the former in terms of those from the latter:

$$\bar{L}(\boldsymbol{\theta}) = \sum_{i=1}^{n} w_i \bar{L}_i(\boldsymbol{\theta})$$
$$Q(\boldsymbol{\theta}) = \sum_{i=1}^{n} w_i Q_i(\boldsymbol{\theta}) + \sum_{i,j<i}^{n} w_i w_j (\bar{L}_i(\boldsymbol{\theta}) - \bar{L}_j(\boldsymbol{\theta}))^2,$$

where we have used Equation A.2.6 to express the overall asset return covariance matrix $\boldsymbol{V}$, which gives only partial information about the dependence structure in the GM setting, in terms of the parameters of the GM distribution.

[4]That portfolios of non-perfectly correlated assets have lower variance than the weighted average of the asset variances is the basis of diversification. The quadratic portfolio variance function $Q(\boldsymbol{\theta}) = \boldsymbol{\theta}'.\boldsymbol{V}.\boldsymbol{\theta}$ in terms of the asset weights $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{V}$, satisfies the identity

$$Q(\lambda\boldsymbol{\theta}_A + (1-\lambda)\boldsymbol{\theta}_B) - (\lambda Q(\boldsymbol{\theta}_A) + (1-\lambda)Q(\boldsymbol{\theta}_B)) = -\lambda(1-\lambda)Q(\boldsymbol{\theta}_A - \boldsymbol{\theta}_B) \leq 0,$$

for all $\lambda \in [0, 1]$.

3.1.4. *Dimension reduction legitimate?* At first sight being able to reduce the dimensionality of the problem in this way looks too good to be true. Given that we are projecting from a space of higher dimension into one of lower dimension, the legitimacy of the method relies fundamentally on the fact that that optimal points in portfolio coordinate space are associated with unique points in the portfolio weight space. Whilst this is not true in general, it is true for efficient boundary portfolios in the portfolio coordinate space. Consequently, it is a matter of considerable importance that we find out which objectives have efficient extrema.

3.2. **Feasible portfolios and the efficient frontier.**

3.2.1. *Feasible portfolios and the investment opportunity set. Feasible* portfolios in portfolio weight space are those that satisfy budget and short-selling constraints. The *investment opportunity set* contains all attainable points in the portfolio coordinate space corresponding to the feasible portfolios. See Figures 3 and 4.

Definition 3.2.1. The set of *feasible* portfolio weights is:

$$C = \left\{ \boldsymbol{\theta} \in \mathbb{R}^m | \boldsymbol{\theta}'.\mathbf{1} = 1, \theta_i \geq 0, i = 1, \ldots, m \right\}.$$

The image in the portfolio coordinate space of C under $\boldsymbol{f}$, the set of *attainable* portfolio coordinates, $D = \boldsymbol{f}(C)$, is the *investment opportunity set* (IOS).

Typical shapes of investment opportunity sets can be seen in the following figures:

- Figure 4 is a schematic diagram to show the investment opportunity set in three-dimensional portfolio coordinate space for four assets. With four assets, the simplex has the topology of a tetrahedron (see inset box). The front face and solid body of the investment opportunity set are the images of parts of the interior of the simplex, under the convex portfolio coordinate map. The rear boundary of the investment opportunity set is composed of the images of the faces and edges of the tetrahedral simplex. With $p = 3$, the efficient frontier is two-dimensional. In this case, return target k-dependent objectives favour portfolios along a curve. The loci of optimal portfolios as k is varied for two different objectives are indicated by curves in the efficient frontier. In this figure, we are using the preferred sign convention for portfolio coordinates: all are to be minimised.
- Figure 5 shows the investment opportunity set for a typical three-asset example of the standard mean-variance approach. The axes are the portfolio variance and the mean. The investment opportunity set is overlaid on a contour plot of the Sharpe ratio objective. Desirable portfolios are in the upper, left-hand corner. In this figure we use the standard convention in this context of having the positive mean on the y-axis, whereas our preferred sign convention is to use the *negative* mean instead.

Further examples of investment opportunity sets will be given in Section 4.4.

3.2.2. *Efficiency.* A portfolio is said to be *efficient* (or *Pareto-optimal*) if all other portfolios have a higher value for at least one of the portfolio coordinates, or else have the same value for all coordinates. The following definition is standard in the field of multiple objective optimisation theory:

Definition 3.2.2. With respect to a set of portfolio coordinates, a portfolio, $\boldsymbol{\theta}^* \in C$, is said to be (globally) *efficient* if and only if there is no $\boldsymbol{\theta} \in C$ such that

$f_i(\boldsymbol{\theta}) \leq f_i(\boldsymbol{\theta}^*)$ for all $i \in 1, 2, ...p$, with at least one strict inequality. The set of all such points is the *efficient frontier*.

Other terms for efficient points are *non-dominated* or *non-inferior* points. We shall apply the same terminology to images of efficient portfolios and efficient frontiers in the portfolio coordinate space.

Remark 3.2.3. Whilst (negative) regime Sharpe ratios could in principle be used to define an n-dimensional space of portfolio coordinates, they are not ideal for this purpose because a) they are not convex functions, b) the efficient frontier in this space can only be found by solving NLPs and c) only the probability of outperformance objective can be expressed in terms of these.

A natural question to ask is whether it is possible to define an efficient frontier in an appropriate space of portfolio coordinates for the novel objectives as we did for the Sharpe ratio problem. The answer is in the affirmative. Echoing the Sharpe ratio case, the appropriate efficient frontier to use in the GM approach is found as the solution to a family of LQPs. We will prove that the extrema of all objectives that are increasing in convex coordinates lie on it.

3.3. **Objective functions - definitions.** To get results from the GM approach that differ from those from the mean-variance approach, we adopt objective functions that probe aspects of the distribution over and above the first two moments. In particular, all of the novel objectives that we have tested are functions of the first two moments of the regime-specific Gaussian components of the GM distribution. Only one objective, *Gaussian mixture mean-variance*, the linear combination of the portfolio coordinates, is a linear-quadratic function of the asset weights θ_i, and hence particularly easy to optimise. The rest are non-linear functions of the asset weights with multiple local minima.

We shall consider two objectives based on the overall first and second moments of the distribution:

- mean variance (MV)
- out-performance Sharpe ratio (SR)

and a number of 'GM-aware' objectives, which are functions of the first two moments in each regime:

- Gaussian mixture mean-variance (GMMV)
- probability of shortfall, (PS) is the zeroth-order lower partial moment (LPM_0); the probability of outperformance (PO) is one minus the probability of shortfall
- target shortfall (TS) is the first-order lower partial moment (LPM_1),
- target semi-variance (TSV) is the second-order lower partial moment: (LPM_2),
- Expected exponential utility (EEU)
- Hodges' modified Sharpe ratio (HR).

Many objectives (SR, PS, PO, TS, TSV, all LPMs and UPMs) are dependent on a target portfolio return parameter, k. The objective EEU depends on a parameter γ.

We proceed by defining a number of objective functions, all with tidy, closed-form expressions in terms of the portfolio coordinates:

3.3.1. *Mean variance.*

$$F_{\upsilon}^{\mathrm{MV}}(\boldsymbol{\theta}) = L(\boldsymbol{\theta}) - \upsilon\, Q(\boldsymbol{\theta}) \tag{3.3.1}$$

where $\upsilon > 0$ is the risk aversion parameter.

3.3.2. *Outperformance Sharpe ratio.*

$$F_k^{\mathrm{SR}}(\boldsymbol{\theta}) =: \alpha(\boldsymbol{\theta}, k) := \frac{L(\boldsymbol{\theta}) - k}{\sqrt{Q(\boldsymbol{\theta})}}. \tag{3.3.2}$$

In the numerical study we change the sign and minimise this objective, for consistency with the objectives against which it is compared.

3.3.3. *Gaussian mixture mean variance.* This objective is a linear combination of the regime mean and variance portfolio coordinates for the GM distribution:

$$F_{\boldsymbol{a},\boldsymbol{b}}^{\mathrm{GMMV}}(\boldsymbol{\theta}) = \sum_{i=1}^{n} a_i L_i(\boldsymbol{\theta}) - b_i Q_i(\boldsymbol{\theta}) \tag{3.3.3}$$

where $a_i, b_i > 0$ are real coefficients.

3.3.4. *Probability of outperformance and shortfall.* In the GM setting of this paper the probability of outperformance (PO) objective is the probability that a univariate GM random variable, with Gaussian component means L_i and variances Q_i, exceeds the threshold k. From the expression for the univariate GM cumulative distribution function (cdf) Eqn. A.1.2, in Appendix A, it can be shown that:

Proposition 3.3.1. *The probability that the portfolio return outperforms the target k is:*

$$F_k^{\mathrm{PO}}(\boldsymbol{\theta}) = \sum_{i=1}^{n} w_i \Phi\left(\frac{L_i(\boldsymbol{\theta}) - k}{\sqrt{Q_i(\boldsymbol{\theta})}}\right). \tag{3.3.4}$$

The objective depends on the portfolio coordinates only through the regime Sharpe ratios $\alpha_i(\boldsymbol{\theta}, k)$. The *probability of shortfall* (PS) objective, the probability that the portfolio return falls short of the target, is one minus the probability of outperformance. Obviously we aim to maximise the probability of outperformance and minimise the probability of shortfall.

Remark 3.3.2. In the mean-variance setting the expression for the probability of outperformance in terms of the Sharpe ratio is simply $\Phi\left(\alpha(\boldsymbol{\theta}, k)\right)$. Because the cumulative distribution function is increasing, the probability of outperformance and Sharpe ratio objectives are maximised by the same portfolio. In that sense these two objectives are equivalent. Because in the special case in which there is only a single regime $n = 1$ the GM approach reduces to the mean-variance approach, and therefore the probability of outperformance objective reduces to the Sharpe ratio objective, we see the probability of outperformance objective is a natural generalisation to the GM approach of the Sharpe ratio objective of the mean-variance theory.

3.3.5. *Expected exponential utility.* In the GM setting, the expected value of the exponential utility function $u_\gamma(w) = -e^{-\gamma w}, \gamma \in \mathbb{R}^+$. gives the objective:

$$F_\gamma^{\mathrm{EEU}}(\boldsymbol{\theta}) = -\sum_{i=1}^{n} w_i \exp\left[-\left(\gamma\, L_i(\boldsymbol{\theta}) - \frac{1}{2}\,\gamma^2\, Q_i(\boldsymbol{\theta})\right)\right],$$

where we have used the identity that for the random variable $X \sim N(\mu, \sigma^2)$, $\mathbb{E}[u_\gamma(X)] = -\exp[-(\gamma\,\mu - \gamma^2\,\sigma^2/2)]$.

In the mean-variance setting the EEU iso-objective contours are straight lines. In the GM setting the EEU iso-objective hyper-surfaces are not planes, but within the ith regime, the intersection between the surfaces and the (L_i, Q_i) plane with all other portfolio coordinates fixed, are straight lines.

3.3.6. *Hodges ratio.* To address the paradoxes inherent in using the Sharpe ratio [28] as a measure for ranking the desirability of payoff distributions, Hodges [19] introduces an intuitive measure: the *modified Sharpe ratio* or *Hodges ratio* (HR), based on the exponential utility function given above. In addition to the risky opportunity, whose desirability the approach gauges, it is assumed that the investor also has access to a risk-free cash investment. She divides her wealth between the two assets with weight ξ in the risky asset.

Because the Hodges ratio uses a utility function with constant absolute risk aversion (CARA), the composition of the optimal portfolio is independent of the coefficient of risk tolerance γ and so without loss of generality this can be taken to be one. The Hodges ratio is a generalization of the Sharpe ratio, reducing to it for normally distributed returns. However, unlike Sharpe ratio, the Hodges ratio is compatible with stochastic dominance: an investment opportunity that outperforms another in every state of the world necessarily has a higher Hodges ratio. This is a property of a coherent risk measure [4] that the Sharpe ratio lacks. An apparent pathology[5] of the Hodges ratio objective [24], has a simple remedy.

In the GM setting, using the identity that for the random variable $X \sim N(\mu, \sigma^2)$, $\mathbb{E}[u_\gamma(\xi X)] = -\exp[-(\gamma\,\xi\,\mu - \gamma^2\,\xi^2\,\sigma^2/2)]$, and setting $\gamma = 1$, we obtain a closed-form expression for the objective:

$$\begin{aligned} F^{\mathrm{HR}}(\boldsymbol{\theta}) &= \max_{\xi} F^{\mathrm{HR}}(\boldsymbol{\theta}, \xi) \\ &= \max_{\xi} \left\{ -\sum_{i=1}^{n} w_i \exp\left[-\left(\xi\, L_i(\boldsymbol{\theta}) - \frac{1}{2}\,\xi^2\, Q_i(\boldsymbol{\theta})\right)\right]\right\}. \end{aligned}$$

3.4. **Objective functions - discussion.**

3.4.1. *Linear and increasing objectives.* Two categories of objectives of particular importance are those that are *portfolio coordinate linear* and those that are *portfolio coordinate increasing.* Whilst portfolio coordinate increasing problems are typically (hard to solve) NLPs, in the special case in which portfolio coordinates are linear or quadratic, such as in the GM setting, portfolio coordinate linear problems are (relatively easier to solve) LQPs in portfolio weight space. It is for portfolio coordinate increasing objectives that the efficient frontier will retain the same, useful role as a hunting-ground for extrema that it has in the mean-variance setting with

[5]Madan and McPhail point out that the Hodges ratio sometimes perversely considers distributions with large negative skew to be desirable investments, because the approach effectively shorts the risky asset. This problem is resolved by scaling the utility function by the sign of ξ. In this study, optimal $\xi^* > 0$ for all sensible parameter values, so this refinement was unnecessary.

Sharpe ratio objective. This is true for the GM case, or in general, for any model using convex portfolio coordinates. We use the KKT theorem to show that increasing objective problem solutions, which are efficient, solve corresponding linear problems.

Definition 3.4.1. A *portfolio coordinate linear* objective $F_{\boldsymbol{\eta}}(\boldsymbol{\theta}) = \hat{F}_{\boldsymbol{\eta}}(\boldsymbol{f}(\boldsymbol{\theta}))$ is one that is linear in all portfolio coordinates: $\hat{F}_{\boldsymbol{\eta}}(\boldsymbol{x}) = \boldsymbol{\eta}'.\boldsymbol{x}$, for $\boldsymbol{\eta} \in \mathbb{R}^p$.

Definition 3.4.2. A *portfolio coordinate increasing* objective $F(\boldsymbol{\theta}) = \hat{F}(\boldsymbol{f}(\boldsymbol{\theta}))$ is one that is increasing in all portfolio coordinates. I.e., the gradient of $\hat{F}(\boldsymbol{x})$ is in the positive orthant: $\boldsymbol{\nabla}\hat{F}(\boldsymbol{x}) \geq \mathbf{0}$.

When it is possible to do so without causing ambiguity about their domain we will omit the 'portfolio coordinate' prefix and talk simply of *linear* and *increasing* objectives.

3.4.2. *Optimisation problem.* All objectives are optimised over the portfolio weight space $\boldsymbol{\theta} \in \mathbb{R}^m$ subject to linear budget equality and short-selling inequality constraints:

$$\begin{aligned} \max/\min_{\boldsymbol{\theta}} \quad & F(\boldsymbol{\theta}) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \boldsymbol{\theta} \geq \mathbf{0}, \end{aligned} \tag{3.4.1}$$

apart from the Hodges ratio, which is optimised over both the portfolio weight space $\boldsymbol{\theta} \in \mathbb{R}^m$ and the total weight in risky assets, ξ:

$$\begin{aligned} \max_{\boldsymbol{\theta},\xi} \quad & F^{\text{HR}}(\boldsymbol{\theta}, \xi) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \boldsymbol{\theta} \geq \mathbf{0}. \end{aligned} \tag{3.4.2}$$

3.4.3. *Two objectives with non-convex problems.* Before attempting to solve the optimisation problems of extremising the chosen objectives, it is instructive to ask whether they are convex problems[6] in the portfolio weight space. We demonstrate how the convexity of a problem is established using as examples the Sharpe ratio and probability of outperformance cases, starting with the former. We consider the convexity properties of the following functions in the Sharpe ratio problem:

Equality and inequality constraints: (budget and short-selling, respectively) are linear in the weights, and therefore convex.

Objective: Because we are trying to maximise the Sharpe ratio objective the problem will be convex if the objective is concave[7]. This is always true in coordinate space and true in weight space whenever $\boldsymbol{\mu}'.\boldsymbol{\theta} - k \geq 0$.

[6]A convex problem is that of minimising a convex function (or equivalently, maximising a concave function) over a convex set. A convex function is one with Hessian determinant greater than zero, over the domain of the function. If $g(\cdot)$ is convex and $f(\cdot)$ is convex and increasing, then $f(g(\cdot))$ is convex.

[7]The function $\hat{F}_k^{\text{SR}}(l, q) = (l - k)/\sqrt{q}$, $q > 0$ is concave in the space (l, q), with Hessian determinant $-\frac{1}{4q^3}$. The Sharpe ratio is concave in portfolio coordinate space and the portfolio coordinates in turn are convex functions in portfolio weight space. However, the Sharpe ratio is concave in portfolio weight space only if it is decreasing in the variance (in the denominator), which is only true when the expected excess return relative to the target (in the numerator) is positive.

Conversely, when the target exceeds the expected value for the portfolio return, the maximum Sharpe ratio portfolio may be inefficient.

Therefore, the common practice of restricting the search for Sharpe ratio optimal portfolios to only those along the efficient frontier when solving the Sharpe ratio problem is only valid when the target, k, is sufficiently small. Figure 5 shows the investment opportunity set overlaid on the contour plot for the Sharpe ratio objective, in the classical mean-variance setting, for the non-pathological case in which the expected return exceeds the target. The contours on the plot are the Sharpe ratio iso-objective curves.

The probability of outperformance objective, in common with the majority of objectives investigated, is convex in neither weight nor coordinate spaces. It is increasing[8] in coordinate space, and increasing in weight space when $L_i(\boldsymbol{\theta}) = \boldsymbol{\mu}_i'.\boldsymbol{\theta} > k$ for all i.

The fact that neither the probability of outperformance objective, nor any of the other exotic objectives are convex might appear to frustrate our use of convex optimisation theory and in particular the use of the KKT theorem to prove that optimal portfolios are efficient. As we explain, the KKT theorem is nevertheless the appropriate tool to bring to bear on any problem with convex portfolio coordinates.

3.4.4. *Expected utility objectives.* An important class of objectives consists of those that can be expressed as the expected value of a utility function for returns:

$$F(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}}[u(Z)] \tag{3.4.3}$$

where $\mathbb{E}_{\boldsymbol{\theta}}[u(Z)]$ is the $\boldsymbol{\theta}$-dependent expectation of a function $u : \mathbb{R} \to \mathbb{R}$ of the portfolio return Z.

It is well-known[9] that in a Gaussian setting, only utility functions that are concave for all values in return space (i.e., over the real line) are guaranteed to have expectations that are decreasing in the variance. It is for these objectives that we are entitled to use the search space dimension reduction algorithm. From this perspective *target shortfall* (TS), *target semi-variance* (TSV), expected (negative) exponential utility[10] and expected quadratic utility are desirable objectives.

Despite having the potential to exhibit the undesirable behaviour of favouring more variance over less, an objective that is the expectation of a non-concave utility function may nevertheless be used with caution if it is well-behaved (decreasing in regime variances) for a restricted range of parameter values. In particular the *probability of outperformance* objective, which is the expected value of a (non-concave) step utility function, is decreasing in the regime variances only for those regimes in which the regime expected return exceeds the return target. Consequently, with this objective we are obliged to set our portfolio return target k to be smaller than the smallest regime expected return. Given that it is quite possible for distressed regimes to have negative expected returns, the target might be required to be negative too.

[8]After changes of sign on the objective and the first moment portfolio coordinate argument.

[9]That $u(x)$ concave, X Gaussian implies $\frac{\partial \mathbb{E}[u(X)]}{\partial Q} \leq 0$ follows from the fact that $\frac{\partial \mathbb{E}[u(X)]}{\partial Q} = \frac{1}{2Q^2}\int_{-\infty}^{\infty} u(x)((x-L)^2 - Q)\phi_{L,\sqrt{Q}}(x)dx$ would be zero if $u(x)$ were linear. However, because the function is assumed to be concave and therefore the infimum of a family of linear functions, the integral satisfies the inequality.

[10]The case $u(x) = -\exp(-\gamma x)$, $\gamma > 0$, is particularly interesting because for this utility the iso-objective curves are straight lines with positive slope in mean-variance space.

3.4.5. *GM-aware.* We shall adopt the term *GM-aware* for those objectives that depend on the regime specific portfolio coordinates, i.e., first two moments of regime portfolio return, rather than on the overall mean and variance of the portfolio return. By this classification, PO, PS, TS, TSV etc. are GM-aware, whereas SR is not.

3.4.6. *Comparison with Higher Moment Models.* Because with two regimes the portfolio coordinate space has four dimensions, the GM approach in this case invites comparison with a *higher-moment model* in which the portfolio coordinates are the first four moments of the portfolio return [3], [21], [18].

In the higher moment model case an undesirable feature is that certain portfolio coordinates (e.g. the third moment) are neither concave nor convex functions in weight space. Furthermore, with such models usually the only way to express common objective functions as closed-form expressions in terms of the moments is approximately, as the first few terms of an expansion. The efficient frontier as the solution to a family of (relatively hard to solve) NLPs is hard to calculate, and knowledge of its form not particularly useful given that optimal portfolios for perfectly reasonable objective functions are not guaranteed to be found on it. The non-convexity of higher moments may lead to redundancy of the optimal portfolios in weight space. (I.e., the portfolio coordinate function is many-to-one). Whilst it is relatively easy to express preferences in terms of moments in vague terms: return and skew are good, variance is bad, etc., it is much harder to use such observations as design criteria for a sensible objective, presumably in an investor-specific way. As Harvey et al [18] observe,

> as the dimensionality of the efficient frontier increases, it becomes less obvious that an investor can easily interpret the geometry of the frontier and reasonably select a portfolio.

Conversely, in the GM model all of the portfolio coordinates are convex functions, in fact linear or quadratic functions, and most familiar objectives give rise to simple closed-form expressions in terms of these. These can be used straight-forwardly to establish the desirability of points on the efficient frontier. The efficient frontier can be found by solving (relatively easy to solve) LQPs, and optimal portfolios to increasing objectives are guaranteed to be found on it. Because portfolio coordinates are convex, (all convex and at least one strictly convex), points on the efficient frontier in coordinate space correspond to unique points in the weight space (I.e., the portfolio coordinate map f is one-to-one). In common with a four-moment higher moment model, the GM approach with two regimes uses a four-dimensional portfolio coordinate space and permits considerable freedom as to the third, fourth and higher moments that can be described, but manages to do so in such a way that the efficient frontier concept is as useful as for the mean-variance case.

3.5. **Theorem underpins algorithm.** Our goal is to find an efficient algorithm for solving the NLPs that arise from the use of novel objectives in the two regime case. We achieve this by reducing our NLP in the portfolio weight space to the non-linear optimization of a three-parameter family of quadratic programs. To do so, we need to show that portfolio coordinate increasing objective optimal portfolios are efficient or equivalently portfolio coordinate linear objective optimal.

The system of KKT conditions is a set of necessary and/or sufficient conditions for the optimality of a NLP. The observation made previously that our increasing objectives are typically not convex in weight space does not prevent us from using the KKT theorems, provided that the portfolio coordinates are convex. The critical observation is that although the increasing NLP is typically not convex, the linearized problem will be. Conveniently, the increasing and associated linear problems more or less share the same system of conditions. The systems can be made equivalent by appropriate choice of the parameters of the linear problem. The KKT conditions are

- *necessary* for the (non-convex) *increasing* problem and
- necessary and *sufficient* for the the (convex) *linear* problem.

That the conditions for the increasing problem are *necessary* is true irrespective of the convexity of the objective. All that the KKT theorem requires for the conditions to be necessary is that the constraints be of affine form, as they are in this case. Thus, the conditions are satisfied for the solution to the (non-convex) increasing problem. Conversely, the linear problem *is* (globally) convex, in which case the system is *sufficient* (as well as necessary) for the (global) optimality of the linear problem.

The steps of the proof are as follows.

- Given an optimal portfolio, i.e. an increasing problem solution, the (increasing problem) system of KKT conditions holds there necessarily.
- By clever choice of the linear problem the linear problem conditions can be made to correspond to the increasing problem conditions. Now, linear and increasing problem KKT conditions hold.
- That the linear problem conditions hold there is sufficient for the point to be a linear problem solution.

The increasing problem solution solves a linear problem.

3.5.1. *Theorem.* We are considering the portfolio problem on the simplex:

$$\begin{aligned} &\min_{\boldsymbol{\theta}} \hat{F}(\boldsymbol{f}(\boldsymbol{\theta})) \\ &\boldsymbol{\theta} \cdot \mathbf{1} = 1 \\ &\boldsymbol{\theta} \geq \mathbf{0} \end{aligned} \tag{3.5.1}$$

where $\hat{F}$ is the objective and $\boldsymbol{f} = (f_1, \ldots, f_p) : \mathbb{R}^m \to \mathbb{R}^p$ is the portfolio coordinate function.

We would like to show that there is a p-vector $\boldsymbol{\eta} \geq \mathbf{0}$ such that the portfolio coordinate linear problem

$$\begin{aligned} &\min_{\boldsymbol{\theta}} \boldsymbol{\eta}'.\boldsymbol{f}(\boldsymbol{\theta}) \\ &\boldsymbol{\theta} \cdot \mathbf{1} = 1 \\ &\boldsymbol{\theta} \geq \mathbf{0} \end{aligned} \tag{3.5.2}$$

has the same optimal solution.

The following result covers the smooth case, and can be generalized.

Theorem 1. *Suppose that the functions $\hat{F}$, $\boldsymbol{f}$ are smooth (on the feasible set), and let $\boldsymbol{\theta}^*$ be an optimal solution to (3.5.1). Suppose that the functions f_i are convex, and that $\boldsymbol{\nabla}\hat{F}(\boldsymbol{f}(\boldsymbol{\theta}^*)) \geq \mathbf{0}$. Then there exists a vector $\boldsymbol{\eta} \in \mathbb{R}^p_+$ such that $\boldsymbol{\theta}^*$ solves (3.5.2).*

Proof. Since the constraints in (3.5.1) are linear, a local optimum $\boldsymbol{\theta}^*$ must satisfy the KKT conditions (see [5], Prop. 3.4.1, page 292). Thus, there exist, $\lambda \in \mathbb{R}$, $\boldsymbol{\mu} \in \mathbb{R}^m_+$ such that:

$$\sum_{i=1}^{p} \frac{\partial \hat{F}}{\partial x_i}(f(\boldsymbol{\theta}^*))\boldsymbol{\nabla} f_i(\boldsymbol{\theta}^*) + \lambda \mathbf{1} + \boldsymbol{\mu} = 0 \tag{3.5.3}$$

$$\boldsymbol{\mu} \cdot \boldsymbol{\theta}^* = 0$$

$$\boldsymbol{\mu} \geq \mathbf{0}.$$

Letting $\eta_i = \frac{\partial \hat{F}}{\partial x_i}(\boldsymbol{f}(\boldsymbol{\theta}^*))$ we have

$$\sum_{i=1}^{p} \eta_i \boldsymbol{\nabla} f_i(\boldsymbol{\theta}^*) + \lambda \mathbf{1} + \boldsymbol{\mu} = 0 \tag{3.5.4}$$

$$\boldsymbol{\mu} \cdot \boldsymbol{\theta}^* = 0$$

$$\boldsymbol{\mu} \geq \mathbf{0},$$

which is the KKT system for the problem (3.5.2). The constraint on $\boldsymbol{\nabla}\hat{F}(\boldsymbol{f}(\boldsymbol{\theta}^*))$ ensures that (3.5.2) (with this choice of $\boldsymbol{\eta}$) is a convex programming problem with linear constraints. The KKT system above is therefore sufficient to guarantee that $\boldsymbol{\theta}^*$ is an optimal solution to (3.5.2) (see [5], Prop. 3.4.2, pages 295-297). □

We have learned that a portfolio coordinate increasing objective has an associated portfolio coordinate linear objective, such that they share the same extremum. The latter is necessarily increasing also.

3.5.2. *Algorithm.* In the GM setting, the non-linear problem (Equation 3.4.1) for a typical objective has dimension equal to the (typically large) number of assets in the universe, m, and can have multiple extrema, with as many as n of these, the number of regimes. It can be solved easily enough using standard NLP optimizers, but such tools are unable to exploit the LQP-like aspects of the NLP and as a consequence will struggle when presented with large problems, taking longer than necessary and being prone to error. Theorem 1 suggests a better way (in fact the time-honoured way of maximising the Sharpe ratio in the mean-variance case). The theorem states that for an increasing objective, the optimal portfolio will be embedded in the efficient frontier, the $(2n-1)$-dimensional space of solutions to a family of portfolio coordinate linear problems, which are ultimately solved as LQPs in weight space. Ultimately there is no escape from the need to solve an NLP, but doing so only over the efficient frontier in coordinate space in preference to over all feasible portfolios in the weight space, is a very worthwhile shortcut.

3.5.3. *Counterexample.* We have deduced that an increasing objective necessarily has a Pareto efficient extremum. Furthermore, essentially by definition, such an efficient point is the extremum of an (increasing) linear objective.

However, for this to be a genuinely interesting result, we require the existence of objectives that do *not* fall into the increasing class for which the extrema *cannot* be obtained as the global extremum of a linear objective. This is to rule out the possibility that the extrema of *all* objectives, irrespective of whether or not they are increasing, are global linear objective extrema. That is to say, we wish to prove the falsity of the assertion that *all optimal portfolios solve portfolio coordinate linear problems.*

Finding such a concrete counterexample to this (false) assertion is somewhat tricky in the two-regime GM case because five or more assets are necessary ($m-1 \geq p$) in order that the investment opportunity set and efficient frontier have their full complements of four ($p = 2n$) and three dimensions ($p-1$), respectively. A related, but simpler problem is to consider objectives defined in a bi-regime portfolio coordinate space of the variances only (i.e. consider an objective that is ambivalent to the levels of the regime first moments). In this case, $p = 2$, and we can search for a counterexample with three assets only, which will be easier to visualize. Because an objective that is ambivalent to regime returns is a valid example of a GM objective, the counterexample to the simplified problem is a legitimate counterexample to the general problem too.

Figure 6 gives a three asset counterexample. In the tranquil regime (x axis), all assets are perfectly correlated, whilst in the distressed regime (y axis), two assets are perfectly correlated and the third perfectly anti-correlated with this pair. The plot is in $(S_{\mathrm{T}}, S_{\mathrm{D}})$ space, where $S_i = \sqrt{Q_i}$. We choose to plot on S_i axes, rather than Q_i axes, because the loci of portfolios containing pairs of perfectly correlated or perfectly anti-correlated assets will be straight lines (with kinks for the perfectly anti-correlated case) on these axes, which simplifies the figure. The contours are for the objective $\Phi(\frac{L_{\mathrm{T}}-k}{S_{\mathrm{T}}}) + \Phi(\frac{L_{\mathrm{D}}-k}{S_{\mathrm{D}}})$ for fixed $L_{\mathrm{T}}, L_{\mathrm{D}}$, and $L_{\mathrm{D}} < k < L_{\mathrm{T}}$, with $L_{\mathrm{T}} - 0.4 = k = L_{\mathrm{D}} + 0.1$. This objective is the probability of outperformance objective; with these parameter values, it is not even locally increasing. Efficient portfolios are found along the bottom left boundary of the investment opportunity set. Because the objective is not increasing, the optimum is inside one of the concavities behind the efficient frontier, behind a line joining the left and upper assets. Therefore, the optimum cannot be obtained as the global solution to a linear problem.

4. Numerical results

4.1. **Overview.** We present a three asset example with five objectives:

- Sharpe ratio (SR)
- Probability of shortfall (PS) (zeroth order lower partial moment)
- Target shortfall (TS) (first order lower partial moment)
- Target semi-variance (TSV) (second order lower partial moment)

.

Our results include:

- Contour plots of two-dimensional sections through the three-dimensional pdfs for asset returns
- Contour plots of two-dimensional projections of the four-dimensional investment opportunity sets
- Optimal weights and objectives against the target return, k
- Optimal weights and objectives against the regime mixing parameter, w
- Weight distress sensitivities (rate of change of optimal weight with respect to mixing parameter, $\frac{d\theta_i}{dw}$) against the target return, k.

We require five or more assets for the investment opportunity set and efficient frontier to have their full quotas of four and three dimensions, respectively. The three assets of the numerical study are too few for the investment opportunity set to be 'solid', so the investment opportunity set is a two-dimensional manifold (embedded in a four-dimensional space) and the efficient frontier is two-dimensional sub-manifold of this. Fortunately, three assets is a sufficient number for the results to be interesting.

4.2. **Parameter values.** All results are presented graphically. Parameter values for mean vectors $\boldsymbol{\mu}_i$, volatility vectors $\boldsymbol{\sigma}_i$ and correlation matrices $\boldsymbol{\rho}_i$, $i \in \{\mathrm{T}, \mathrm{D}\}$ are:

$$\boldsymbol{\mu}_{\mathrm{T}} = \begin{pmatrix} 0.21 \\ 0.29 \\ 0.41 \end{pmatrix} \qquad \boldsymbol{\mu}_{\mathrm{D}} = \begin{pmatrix} 0.0475 \\ 0.0775 \\ 0.105 \end{pmatrix} \tag{4.2.1}$$

$$\boldsymbol{\sigma}_{\mathrm{T}} = \begin{pmatrix} 0.2 \\ 0.3 \\ 0.4 \end{pmatrix} \qquad \boldsymbol{\sigma}_{\mathrm{D}} = \begin{pmatrix} 0.4 \\ 0.608 \\ 0.812 \end{pmatrix} \tag{4.2.2}$$

$$\boldsymbol{\rho}_{\mathrm{T}} = \begin{pmatrix} 1. & 0. & 0. \\ 0. & 1. & 0. \\ 0. & 0. & 1. \end{pmatrix} \qquad \boldsymbol{\rho}_{\mathrm{D}} = \begin{pmatrix} 1. & 0.986 & 0.985 \\ 0.986 & 1. & 0.992 \\ 0.985 & 0.992 & 1. \end{pmatrix}. \tag{4.2.3}$$

4.3. **GM pdfs for asset returns.** Figure 7 shows two-dimensional sections (variables in suppressed dimensions set to zero) of the GM pdf as contour plots. The contours indicate iso-probability density curves.

4.4. **Investment opportunity set.** Some examples of investment opportunity sets were given in Section 3.2.1. Here we add to this list examples from the the current numerical example plotted on both portfolio coordinate and regime Sharpe ratio axes.

4.4.1. *Portfolio coordinate space.* Figure 8 shows the investment opportunity set and probability of outperformance objective optimal portfolio in the four-dimensional portfolio coordinate space of tranquil (T) and distressed (D) regime (negative) means and variances displayed as projections (side, top, front view and fourth dimensional analogue of these) into the two-dimensional planes obtained by taking the axes in pairs, for a three asset example. The objective parameters are $k = 0$ and $w = 0.5$.

Because we favour small values for all portfolio coordinates, efficient portfolios are those boundary points whose normals point into the negative orthant in the four-dimensional portfolio coordinate space. Following projection into two-dimensions, optimal and therefore efficient portfolios tend to be found to the bottom-left of each sub-figure. However, due to our inability to think in four dimensions it is sometimes hard to see how an optimal point can possibly be efficient when it appears to be decidedly inefficient when viewed on a picture of the projected investment opportunity set! A good example of this is the probability of outperformance optimal portfolio, which in the $(Q_{\mathrm{T}}, Q_{\mathrm{D}})$ subplot deceptively appears to be 'on top' of the investment opportunity set. Contrary to appearances, in four dimensions the optimal portfolio really is 'underneath' the investment opportunity set.

Similarly, on the same subplot it is hard to see how the optimal point is able to move from its current ($k = 0$, $w = 0.5$) position on the figure, to pure Asset 3 (blue) only via 'downward facing' efficient portfolios. Our studies indicate that this transition occurs as either the appetite for risk, as measured by the target return k, or alternatively, the probability of distress, w, is increased. Again, this apparent problem is merely an artifact of the poor representation of what is really happening in the four-dimensional portfolio coordinate space.

4.4.2. *Regime Sharpe ratio space.* Figure 9 shows the investment opportunity set and probability of outperformance objective optimal portfolio in two-dimensional regime Sharpe ratio space $(\alpha_{\mathrm{T}}(k), \alpha_{\mathrm{D}}(k))$, a three asset case with return target $k = 0$. The objective has mixing parameter $w = 0.5$.

Because the benefits of diversification are poor in the distressed regime, the composition of the optimal portfolio is primarily dictated by the desire to maximise the Sharpe ratio in the tranquil regime. However, if the probability of distress, w, is increased towards one, then the slopes of the contours become less negative and the maximum probability of outperformance point moves upwards and leftwards, eventually towards the top (blue) asset, which has the highest return in both regimes.

The parameter values we have used are for an extreme case in which correlations are zero in the tranquil regime and very close to one in the distressed regime. As a consequence, the 'tails' of the investment opportunity set point to the left with negligible vertical tendency. In general, with lower correlations during times of distress, the tails would point to the left and downwards. The efficient frontier is to the top-right of this diagram for all parameter values. (A similar plot with negative regime Sharpe ratios on the axes, in accordance with our preferred sign convention for portfolio coordinates, would have had the efficient frontier to the bottom-left.)

4.5. **Legend.** The line graphs to follow can be interpreted with the aid of the following key:

4.5.1. *Weights* $\theta_i(k, w)$*.* The assets are in order of mean return:

Asset 1: Red, solid
Asset 2: Green, dash-dot
Asset 3: Blue, dotted

4.5.2. *Objective values* $F_k^a(\boldsymbol{\theta}_k^b)$*.*

SR: Red, solid
PS: Yellow, dot-dash (short)
TS: Green, dot-dash (long)
TSV: Cyan, dash-dash (long-short)

4.6. **Weights and objectives against target return** k**.** For the following plots the regime mixing parameter is fixed at $w = 40\%$. The plots show the dependence of certain functions on the target level parameter k, which is a measure of appetite for risk and return. Agents with small values of k are risk averse.

For consistency we change the sign of the Sharpe ratio objective. This way, all of the objectives are to be minimized.

- Figure 10 shows the optimal portfolio asset weights $\boldsymbol{\theta}_i(k)$ for asset i.
- Figure 11 shows $F_k^a(\boldsymbol{\theta}_k^b)$: Objective a evaluated using the optimal weights for objective b. Within each plot for an objective a, each curve is for an objective b, where $a, b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$.
- Figure 12 shows $F_k^{\mathrm{PS}}(\boldsymbol{\theta}_k^b)$: probability of shortfall evaluated at the optimal weights for objective $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$. Obviously, the optimal weights for a given objective outperform those from all other objectives with respect to the given objective.

- Figure 13 shows $F_k^{\mathrm{PS}}(\boldsymbol{\theta}_k^b) - F_k^{\mathrm{PS}}(\boldsymbol{\theta}_k^{\mathrm{PS}})$: probability of shortfall penalty for using non-probability of shortfall objective b to obtain optimal weights $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$. The penalties in terms of increased probability of failure to meet the target, for using the 'wrong' objectives are in the order SR $>$ TSV $>$ TS $>$ PS $= 0$.

In the Sharpe ratio example, the mean-variance approach recommends a quick move from a diversified portfolio to a pure asset as k is increased. This is because an mean-variance approach investor overestimates the riskiness of a diversified portfolio, and so, even for relatively low values of the target, seeing no apparent risk benefit in diversifying, opts instead to maximise return by holding the asset with the highest return.

Conversely, an investor with a GM-aware objective, which acknowledges that the tranquil regime exists and will occur with some probability, has a greater tendency to favour a diversified portfolio, even for relatively large values of k. The GM investor is able to benefit from diversification effects in the tranquil regime at least, even if diversification in the distressed regime is a lost cause.

For large enough k, all objectives eventually eschew balanced portfolios in favour of risky single asset portfolios because high target levels correspond to a large appetite for risk and return.

The GM-aware objectives: PS, TS, TSV recommend similar portfolios as k changes and there is little deterioration in the objective if the optimal portfolio from one objective is evaluated using another. We conclude that for the GM approach the precise choice of objective is not critical, so long as it is GM-aware.

4.7. **Weights and objectives against regime mixing parameter w.** For the following plots the target return is fixed at $k = 10\%$ and the probability of distress, w is varied. Note that to the far left or right of each figure, $w = 0$ or $w = 1$ corresponding to pure tranquil or pure distressed regime. At these extremities the GM approach reduces to the mean-variance approach, although using the regime mean and variance parameters rather than the mean and variance of the distribution overall.

- Figure 14 shows the optimal portfolio weights $\boldsymbol{\theta}_i(w)$ for assets $i = 1, \ldots, m$. Assets with positive (negative) slope on this diagram, i.e., those with a tendency to have big (small) positions in the distressed regime, are good for speculating on the market entering a distressed (tranquil) state.
- Figure 15 shows $F_w^a(\boldsymbol{\theta}_w^b)$: Objective a evaluated using the optimal weights for objective b against mixing parameter w. Within each plot for an objective a, each curve is for an objective b, where $a, b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$.
- Figure 16 shows $F_w^{\mathrm{PS}}(\boldsymbol{\theta}_w^b)$: probability of shortfall evaluated at the optimal weights for objective $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$.
- Figure 17 shows $F_w^{\mathrm{PS}}(\boldsymbol{\theta}_w^b) - F_w^{\mathrm{PS}}(\boldsymbol{\theta}_w^{\mathrm{PS}})$: probability of shortfall penalty for using non-probability of shortfall objective b to obtain optimal weights $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$. The penalties in terms of increased probability of failure to meet the target, for using the 'wrong' objectives are in the order SR $>$ TSV $>$ TS $>$ PS $= 0$.

For zero probability of distress, $w = 0$, all objectives agree on a diversified optimal portfolio. Similarly, for unit probability of distress, $w = 1$, all objectives agree on a pure asset (Asset 3, blue) as the optimal portfolio. It is for the intermediate values of w, that the GM theory gives recommendations different to those of the mean-variance theory.

There is greater disparity between the portfolio weight recommendations for GM-aware objectives (PS, TS, TSV) as w changes than there was as k changed. However, amongst the GM-aware objectives there is only mild loss of performance as measured by one objective when it is supplied with optimal weights from a different objective. Evaluating the GM-unaware Sharpe ratio objective optimal weights with the probability of shortfall objective, we observe a more serious degradation of the probability that the portfolio will meet or exceed its target. As above, we conclude that for the GM approach the precise choice of objective is not critical, so long as it is GM-aware.

4.8. **Weight distress sensitivities against target return k.** Figure 18 shows the *weight distress sensitivities* $\frac{d\theta_i^*}{dw}$ against target k, for $w = 40\%$. These are not to be confused with *objective distress sensitivities*, which are the rate of change of the value of the optimal objective with respect to the regime mixing parameter: $\frac{dF(\theta_i^*)}{dw}$.

Assets with a positive value of $\frac{d\theta_i}{dw}$ are a good hedge against an increase in the probability of distress. These are the assets into which investors should place a greater fraction of their wealth if w increases. We observe that investors are not unanimous about which assets provide a good hedge against increased probability of distress, as this is dependent on the risk appetite of the investor. The contention concerns Asset 2 (green). For the GM-aware objectives, an increase in w prompts investors with a low (high) return target level to buy (sell) Asset 2 (green). However, investors of all risk tolerances agree that Asset 1, red and Asset 3, blue should be sold and bought, respectively, if w is increased.

5. Conclusions

We provide evidence that the GM approach, namely the assumption of a multivariate finite Gaussian mixture distribution for asset returns, used in conjunction with a suitable objective, such as the target shortfall, target semi-variance or expected exponential utility, will be useful for a large range of portfolio management applications, in addition to the fund of fund and CTA management role that motivated its development.

Our key finding is that when portfolio coordinate functions are *convex*, then objectives that are *increasing* in the portfolio coordinate space will give rise to optimal portfolios that are *efficient* and therefore solutions to associated *linear* problems in the portfolio coordinate space. This applies to the GM approach, but also applies equally well in general to any framework in which objectives depend on portfolio weights through their dependence on portfolio coordinate functions.

Furthermore, because the portfolio coordinates that we choose to use for the GM approach, namely the regime means and variances, are linear or quadratic functions in the portfolio weight space, the associated portfolio coordinate linear problems are LQPs rather than NLPs in weight space. Therefore we only need to solve LQPs to obtain the GM efficient frontier.

By analogy with the familiar procedure for maximising the Sharpe ratio in the mean-variance case, we have proposed an efficient algorithm for solving portfolio problems with convex coordinates and increasing objectives: the theorem tells us that these can be extremised by performing a restricted search over the (low-dimensional) efficient frontier, rather than the (high-dimensional) weight space.

When evaluating objectives in the important class of those that are the expected values of utility functions of return, our new-found preference for increasing objectives now leads us to favour those derived from *concave* utility functions. Consequently, target shortfall (expected loss below a target), target semi-variance (variance of loss below a target) and expected exponential and quadratic utility functions are highly suitable for the GM approach. The probability of outperformance, associated with a step utility function (i.e. not concave) can be used but only for sufficiently small values of the target parameter.

Knowledge of the theorem sheds light on existing portfolio management approaches. For example, because third and higher moments are not necessarily convex, the theorem explains a drawback of higher moment models. For these the efficient frontier in a coordinate space of the first three or four moments is both a) hard to find (consisting of solutions to NLPs) and b) once found not particularly useful. In particular, the efficient frontier cannot be used as the basis of an algorithm to extremise non-linear objective functions; instead all objectives must be maximised by brute force in the weight space.

The GM approach is barely harder to implement than the standard Gaussian mean-variance approach, with many features in common. We retain the following features:

- The dependence structure is encoded in (multiple) covariance matrices
- Use made of Gaussian distribution functions, and functionals of these such as moments and quantiles
- Benefits of diversification are due to a similar mechanism (within regimes)
- The efficient frontier is obtained as the set of solutions to a family of LQPs
- Asset and portfolio returns share the same closed-form distribution (in multivariate and univariate guises, respectively)

These similarities should make it popular with practitioners already well-acquainted with standard mean-variance technology. However, the GM approach has the following important advantages over the mean-variance approach:

- More flexible because of its ability to handle non-elliptic asset return distributions
- 'Scalable' solution: where sufficient data exist for calibration, any distribution can be modelled to arbitrary accuracy simply by increasing the number of regimes
- Appropriate objective functions are intuitive and favoured by practitioners, e.g. expected target shortfall
- Gives significantly reduced probability of shortfall relative to naïvemean-variance approach
- GM regime mixing parameter, w, gives rise to a new risk measures *objective distress sensitivity* and *weight distress sensitivity*, which measure the rate of change of the optimal objective and asset weight, respectively, with respect to the probability of the distressed regime occurring.

The primary disadvantage is the computational overhead of finding the global solution to an NLP with multiple local extrema. However, with two regimes the objective function does not possess more than two extrema in weight space, so our numerical examples have been robust and quick to solve.

We have compared the GM and mean-variance approaches. Natural questions to ask are whether the two approaches give different optimal weights from one another, and if so, whether holding the GM weights gives improved performance using measures preferred by practitioners. The response to both questions is in the affirmative. The optimal weights are indeed significantly different between the approaches.

Of course, it is tautological to say that probability of shortfall optimal portfolios have a lower probability of shortfall than mean-variance optimal portfolios. This is inevitable. However, what is important is that the margin of outperformance is significant: of the order of 4% for the parameter values that we used. This result will be of interest to anyone performing a feasibility study of the GM technology and seeking to justify the work required to implement the new approach.

The different 'GM-aware' objectives favour similar portfolios. The differences in performance between these are much smaller than the difference between any given GM-aware objective and the mean-variance objective. We conclude that it is not so important which GM-aware objective is used, just so long as one of them is.

Another rationale for performing this study is to estimate the model risk and performance losses that result from making do with the standard mean-variance approach in place of the GM approach (as a shortage of data for calibration purposes often dictates) in a setting in which non-Gaussian asset return behaviour, in particular correlation breakdown, is suspected.

The GM approach, using almost identical techniques to the mean-variance approach and sharing its tractability, offers an improvement in model flexibility for a modest increase in model complexity. The approach offers a good compromise between having the flexibility to capture important non-Gaussian features of the asset returns and yet having sufficient simplicity that calibration to realistic data sets is feasible. In the theoretical multiple Gaussian regime world investigated here, the portfolios given by the GM approach not unsurprisingly outperformed those from the mean-variance approach. If reality is better modelled by the GM distribution than the Gaussian distribution, which empirical studies suggest, then we expect that the GM approach will outperform the mean-variance approach for managing portfolios in the real world, too.

Appendix A. Gaussian mixture distribution

A.1. Definitions.

Definition A.1.1. A (scalar) random variable Z has the *univariate GM distribution* if its probability density function $f_Z(z)$ is of the form

$$f_Z(z) = \sum_{i=1}^{n} w_i f_{X_i}(z) = \sum_{i=1}^{n} w_i \phi(\frac{z - \mu_i}{\sigma_i}) \tag{A.1.1}$$

where the random variables X_i are normally distributed with normal probability density functions $\phi_{X_i}(x) = \phi(\frac{z-\mu_i}{\sigma_i})$, $\phi(z)$ is the standard normal probability density function and the weights w_i sum to one. The random variables X_i have means μ_i and variances σ_i^2. The finite sum is over the desired number of normal components to combine, n.

Remark A.1.2. The cumulative distribution function is trivially:

$$F_Z(z) = \sum_{i=1}^{n} w_i \Phi(\frac{z - \mu_i}{\sigma_i}) \tag{A.1.2}$$

where $\Phi(z)$ is the standard normal cumulative density function. We make use of this observation in Section 3.3.4 to define the probability of outperformance objective.

Similarly,

Definition A.1.3. A vector random variable $\boldsymbol{Z}$ has the *multivariate GM distribution* if its probability density function $f_{\boldsymbol{Z}}(\boldsymbol{z})$ is of the form

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \sum_{i=1}^{n} w_i f_{\boldsymbol{X}^{(i)}}(\boldsymbol{z}) = \sum_{i=1}^{n} w_i \phi_{\boldsymbol{\mu}^{(i)},\boldsymbol{V}^{(i)}}(\boldsymbol{z}) \tag{A.1.3}$$

where the (vector) random variables $\boldsymbol{X}^{(i)}$ are multivariate normally distributed with probability density functions $\phi_{\boldsymbol{\mu}^{(i)},\boldsymbol{V}^{(i)}}(\boldsymbol{z})$, and the weights w_i sum to one. The vector random variable $\boldsymbol{X}^{(i)}$ has mean $\boldsymbol{\mu}^{(i)}$ and variance-covariance matrix $\boldsymbol{V}^{(i)}$. E.g. if we take the ath and bth components of $\boldsymbol{X}^{(i)}$, their covariance is the element (a, b) of $\boldsymbol{V}^{(i)}$; i.e. $\mathrm{Cov}(X_a^{(i)}, X_b^{(i)}) = V_{ab}^{(i)}$. Also $\mathbb{E}[X_a^{(i)}] = \mu_a^{(i)}$.

Remark A.1.4. Note that by definition $\mathrm{Cov}(X_a^{(i)}, X_b^{(j)}) = 0$ for $i \neq j$.

Remark A.1.5. In the numerical experiments described later, the mixture distribution contains two normal components, describing asset returns under *tranquil* and *distressed* conditions. We shall refer to the weights w_i as *regime weights.*

A.2. Moments.

The mean of a random variable with the mixture distribution is simply expressed as a linear combination of the means of the component normal distributions.

Proposition A.2.1. *The expectation of a function f of a random variable with the GM distribution can be expressed in terms of the expectations of functions of the component normally distributed variables:*

$$\mathbb{E}[f(\boldsymbol{Z})] = \sum_{i=1}^{n} w_i \mathbb{E}[f(\boldsymbol{X}^{(i)})] \tag{A.2.1}$$

Remark A.2.2. In particular,

$$\mathbb{E}[\boldsymbol{Z}] = \sum_{i=1}^{n} w_i \boldsymbol{\mu}^{(i)} \tag{A.2.2}$$

N.B. This is a potential source of confusion given that it is *not* true in general that $\boldsymbol{Z} = \sum_{i=1}^{n} w_i \boldsymbol{X}^{(i)}$.

The variance depends not only on the variances of the components, but also on the differences between the means of the components.

Proposition A.2.3. *The variance of a random variable with the (univariate) GM distribution can be expressed in terms of the expectations and variances of the component normally distributed variables:*

$$\begin{aligned} \mathrm{Var}[Z_a] &= \sum_{i=1}^{n} w_i \mathrm{Var}[X_a^{(i)}] + \sum_{i,j<i}^{n,n} w_i w_j (\mathbb{E}[X_a^{(i)}] - \mathbb{E}[X_a^{(j)}])^2 \\ &= \sum_{i=1}^{n} w_i (\sigma_a^{(i)})^2 + \sum_{i,j<i}^{n,n} w_i w_j (\mu_a^{(i)} - \mu_a^{(j)})^2 \end{aligned} \tag{A.2.3}$$

Remark A.2.4. If we permit the variance and expectation operators to thread over the components of the vector arguments $f(\boldsymbol{X})_a := f(X_a)$, this can be written in alternative vector form as

$$\begin{aligned} \mathrm{Var}[\boldsymbol{Z}] &= \sum_{i=1}^{n} w_i \mathrm{Var}[\boldsymbol{X}^{(i)}] + \sum_{i,j<i}^{n,n} w_i w_j (\mathbb{E}[\boldsymbol{X}^{(i)}] - \mathbb{E}[\boldsymbol{X}^{(j)}])^2 \\ &= \sum_{i=1}^{n} w_i {\boldsymbol{\sigma}^{(i)}}^2 + \sum_{i,j<i}^{n,n} w_i w_j (\boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)})^2 \end{aligned} \tag{A.2.4}$$

The variance result, above is a special case of the following result for the covariance:

Proposition A.2.5. *The covariance between two elements of a vector random variable with the (multivariate) GM distribution can be expressed in terms of the expectations of functions of the component normally distributed variables:*

$$\begin{aligned} \mathrm{Cov}[Z_a, Z_b] = &\sum_{i=1}^{n} w_i \mathrm{Cov}[X_a^{(i)}, X_b^{(i)}] + \\ &\sum_{i,j<i}^{n,n} w_i w_j (\mathbb{E}[X_a^{(i)}] - \mathbb{E}[X_a^{(j)}])(\mathbb{E}[X_b^{(i)}] - \mathbb{E}[X_b^{(j)}]) \\ = &\sum_{i=1}^{n} w_i V_{ab}^{(i)} + \sum_{i,j<i}^{n,n} w_i w_j (\mu_a^{(i)} - \mu_a^{(j)})(\mu_b^{(i)} - \mu_b^{(j)}) \end{aligned} \tag{A.2.5}$$

Remark A.2.6. In matrix notation, where $W_{ab} := \mathrm{Cov}[Z_a, Z_b]$,

$$\boldsymbol{W} = \sum_{i=1}^{n} w_i \, \boldsymbol{V}^{(i)} + \sum_{i,j<i}^{n,n} w_i w_j (\boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)}).(\boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)})' \tag{A.2.6}$$

where $'$ indicates matrix transpose.

Remark A.2.7. Such expressions are useful if one wishes to find a Gaussian distribution with the same first and second moments as a GM distribution. The covariance matrix gives only partial information about the dependence structure of a collection of GM distributed random variables.

Moments of GM distributions are the sums of the component distribution moments. However, for central moments, such as the covariance, the expressions have additional terms.

A.3. **Linear combinations of random variables with the GM distribution.**

Proposition A.3.1. *Linear combinations of random variables with the (multivariate) GM distribution will themselves have a (univariate) mixture of normals distribution. In particular, the (scalar) random variable* $Y = \sum_{a=1}^{m} \theta_a Z_a$ *where the* m*-vector random variable* $\boldsymbol{Z}$ *has the multivariate GM distribution and* $\boldsymbol{\theta}$ *is an* m*-vector of real coefficients, has probability density function*

$$f_Y(y) = \sum_{i=1}^{n} w_i \phi_{\bar{\mu}_i, \bar{\sigma}_i^2}(y) \tag{A.3.1}$$

where $\bar{\mu}_i = \boldsymbol{\mu}_i.\boldsymbol{\theta}$ *and* $\bar{\sigma}_i^2 = \boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}$.

Similar identities may be found in [33].

Appendix B. Lower partial moment identities

We present some useful identities for first and second order lower partial moment objectives.

$$\int_{-\infty}^{a} (a-x)\ \phi(x)\,dx = a\,\Phi(a) + \phi(a) \tag{B.0.2}$$

$$= \int_{-\infty}^{a} \Phi(x)\,dx \tag{B.0.3}$$

$$\int_{-\infty}^{a} (a-x)^2\,\phi(x)\,dx = \left(1+a^2\right)\,\Phi(a) + a\,\phi(a) \tag{B.0.4}$$

$$\int_{-\infty}^{a} \frac{a-x}{\sigma}\,\phi_{\mu,\sigma}(x)\,dx = \frac{a-\mu}{\sigma}\,\Phi_{\mu,\sigma}(a) + \sigma\,\phi_{\mu,\sigma}(a) \tag{B.0.5}$$

$$\frac{1}{\sigma^2}\int_{-\infty}^{a} (a-x)^2\,\phi_{\mu,\sigma}(x)\,dx = (a-\mu)\ \phi_{\mu,\sigma}(a) + \left(1+\frac{(a-\mu)^2}{\sigma^2}\right)\,\Phi_{\mu,\sigma}(a) \tag{B.0.6}$$

Appendix C. Figures

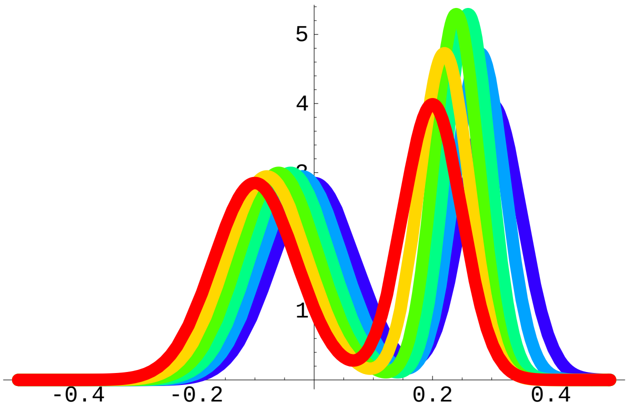

Figure 1. Portfolio return pdf for the GM distribution to show diversification within regimes. Each curve is the return pdf for a portfolio containing two assets in different ratios. The differences between the regime means have been exaggerated. The tranquil regime (right) displays more pronounced diversification effects than the distressed regime (left).

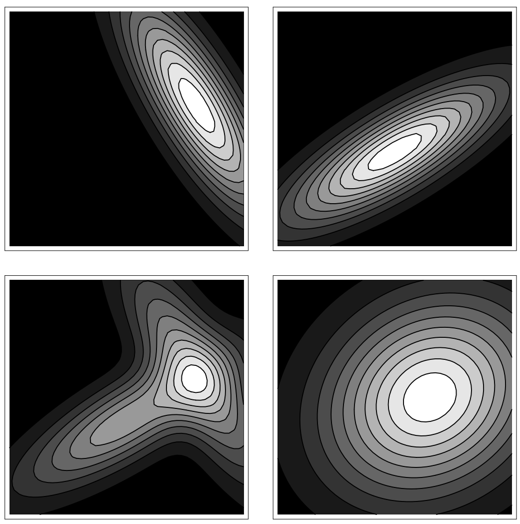

Figure 2. Contour plots of probability density functions. The top row contains two bivariate Gaussian distributions - potentially for the tranquil (left) and distressed (right) regimes. The bottom row illustrates the composite Gaussian mixture distribution obtained by mixing the two distributions from the top row (left) and a bivariate normal distribution with the same means and variance-covariance matrix as the composite (right).

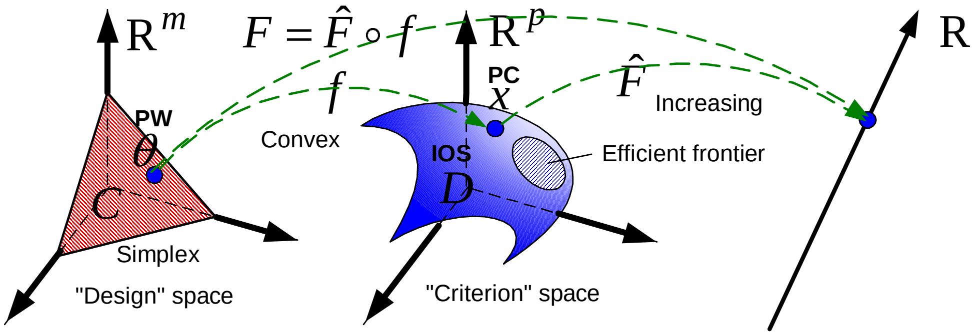

FIGURE 3. Schematic diagram to show the portfolio weight space $\mathbb{R}^m$ (left), portfolio coordinate space $\mathbb{R}^p$ (middle) and range of the objective function, $\mathbb{R}$ (right); and convex and increasing functions f and $\hat{F}$, respectively. The objective function with the portfolio weight space as its domain is the composition of these: $F(\boldsymbol{\theta}) = \hat{F}(f(\boldsymbol{\theta}))$.

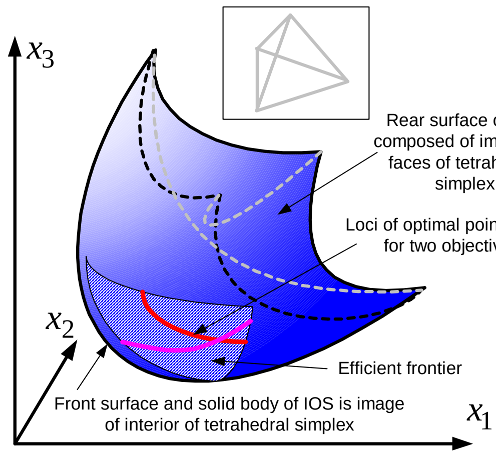

FIGURE 4. Schematic diagram to show the investment opportunity set in three-dimensional portfolio coordinate space for four assets. With four assets, the simplex has the topology of a tetrahedron (see inset box). The front face and solid body of the IOS are the images of parts of the interior of the simplex, under the convex portfolio coordinate map. The rear boundary of the investment opportunity set is composed of the images of the faces and edges of the tetrahedral simplex. With $p = 3$, the efficient frontier is two-dimensional. In this case, return target k-dependent objectives favour portfolios along a curve. The loci of optimal portfolios as k is varied for two different objectives are indicated by curves in the efficient frontier.

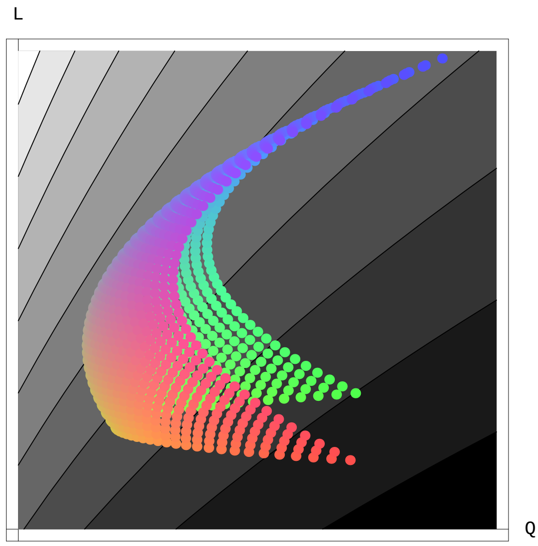

FIGURE 5. Investment opportunity set and contour plot for Sharpe ratio objective, in the standard mean-variance setting. The axes are the variance (Q) and the mean (L) of the portfolio return. This is a three asset example.

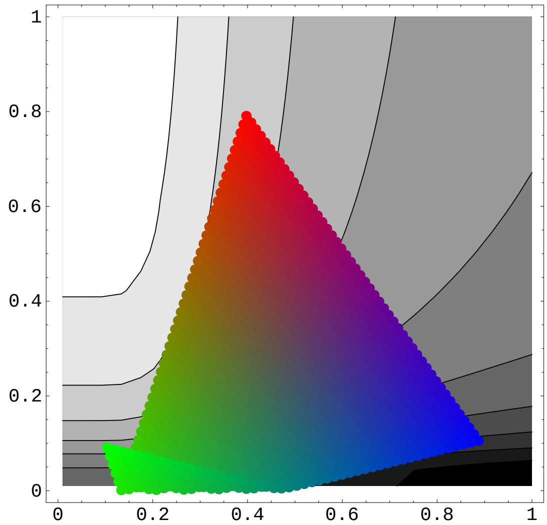

FIGURE 6. A counterexample to the false assertion that all optimal portfolios are solutions to linearised problems. The investment opportunity set for three assets with, in the tranquil regime (x axis) all assets perfectly correlated and in the distressed regime (y axis) two assets perfectly correlated and the third perfectly anti-correlated with this pair, in $(S_{\mathrm{T}}, S_{\mathrm{D}})$ space, where $S_i = \sqrt{Q_i}$. Contours are for the objective $\Phi(\frac{L_{\mathrm{T}}-k}{S_{\mathrm{T}}}) + \Phi(\frac{L_{\mathrm{D}}-k}{S_{\mathrm{D}}})$ for fixed $L_{\mathrm{T}}, L_{\mathrm{D}}$, and $L_{\mathrm{D}} < k < L_{\mathrm{T}}$, with $L_{\mathrm{T}} - 0.4 = k = L_{\mathrm{D}} + 0.1$. The optimal portfolio is to the top left of the investment opportunity set, and being behind a line joining the left and upper assets is not a (global) portfolio coordinate linear solution.

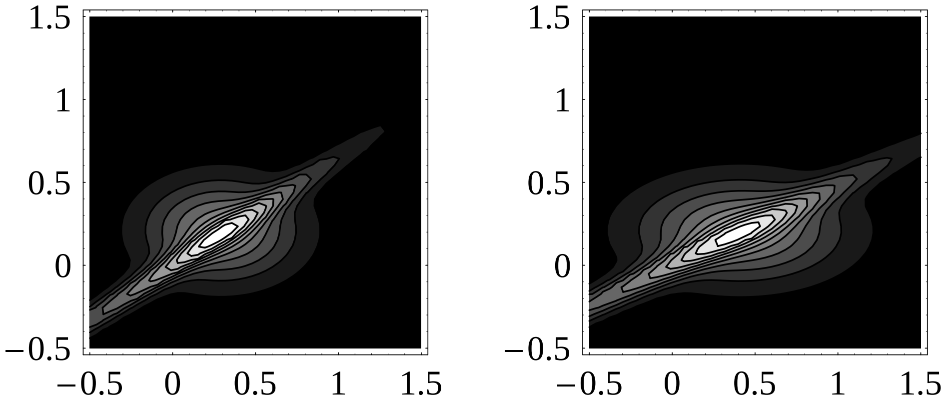

FIGURE 7. Contour plots of two-dimensional sections of asset return probability density function. Three asset case.

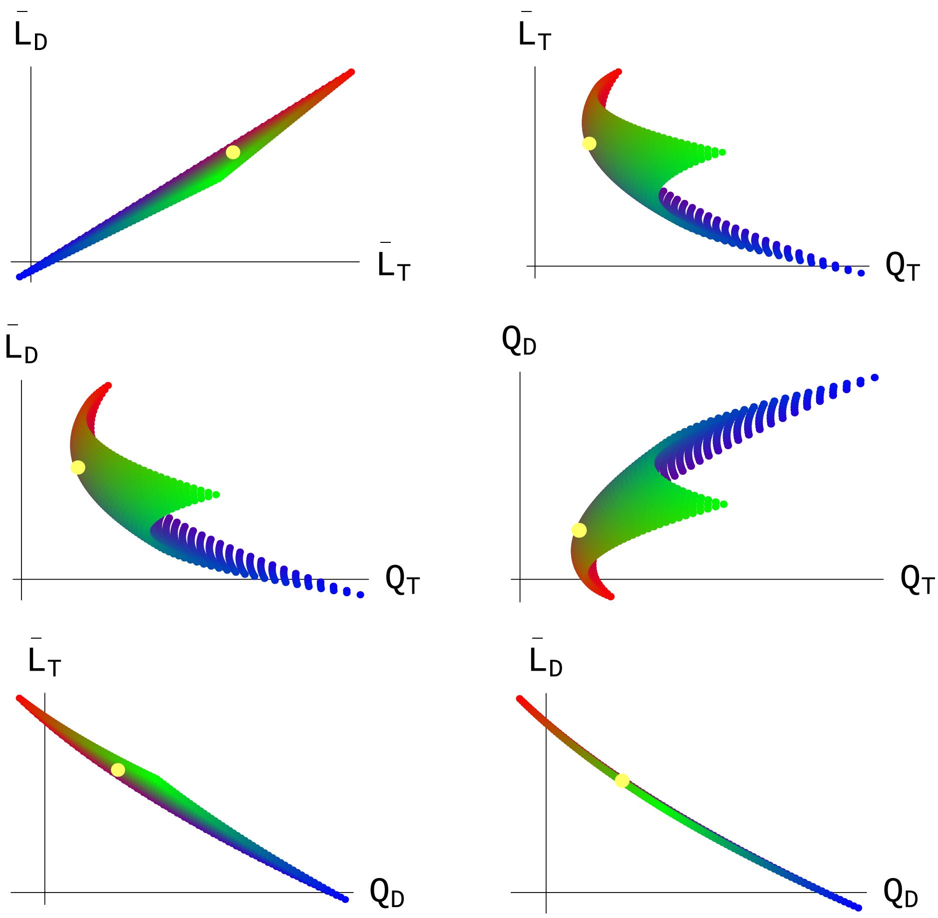

FIGURE 8. Investment opportunity set and PO objective optimal portfolio in four-dimensional space $\{\bar{L}_i, Q_i\}_{i \in \{\mathrm{T},\mathrm{D}\}}$, displayed as projections into two-dimensional planes obtained by taking the axes in pairs. Three asset case. The objective parameters are $k = 0$ and $w = 0.5$. Efficient portfolios tend to be found to the bottom-left of the image of each projected investment opportunity set.

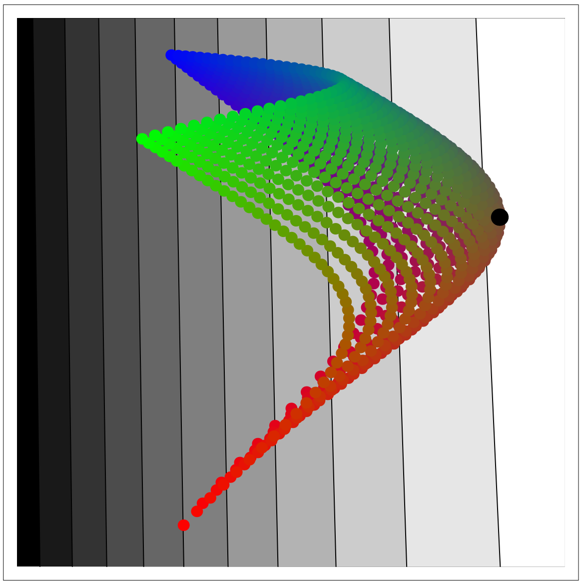

FIGURE 9. Investment opportunity set and probability of outperformance objective optimal portfolio in two-dimensional regime Sharpe ratio space $(\alpha_{\mathrm{T}}(k), \alpha_{\mathrm{D}}(k))$. Three asset case with return target $k = 0$ and mixing parameter $w = 0.5$.

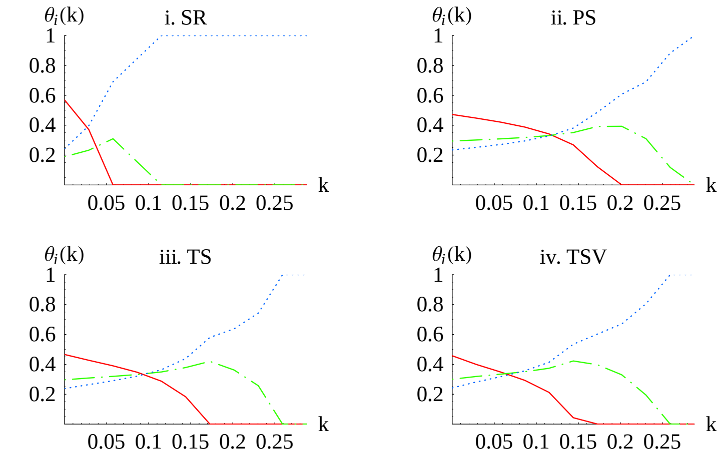

FIGURE 10. Optimal weight $\boldsymbol{\theta}_i(k)$ for asset i against return target k, for $w = 40\%$

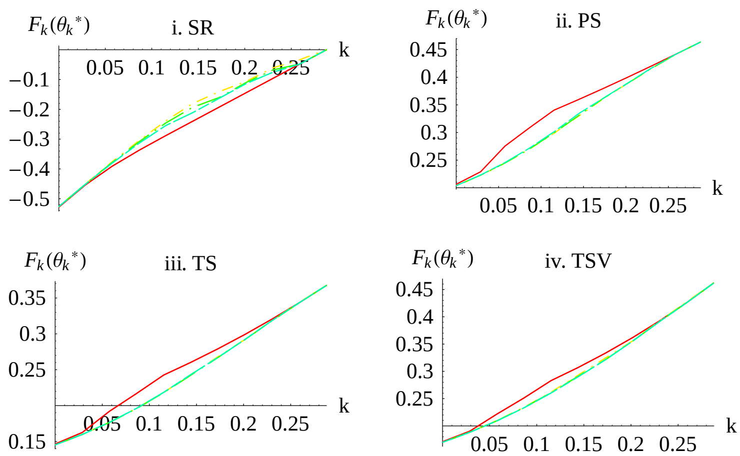

FIGURE 11. $F_k^a(\boldsymbol{\theta}_k^b)$: Objective a evaluated using the optimal weights for objective b against target k, for $w = 40\%$. Each plot is for an objective a; each curve is for an objective b, where $a, b \in \{\text{SR}, \text{PS}, \text{TS}, \text{TSV}\}$.

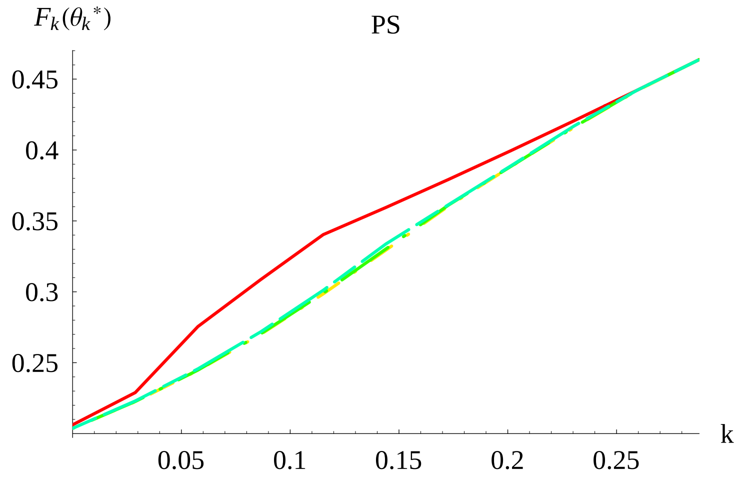

FIGURE 12. $F_k^{\mathrm{PS}}(\boldsymbol{\theta}_k^b)$: Objective PS against target, k, evaluated at the optimal weights for objective $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$.

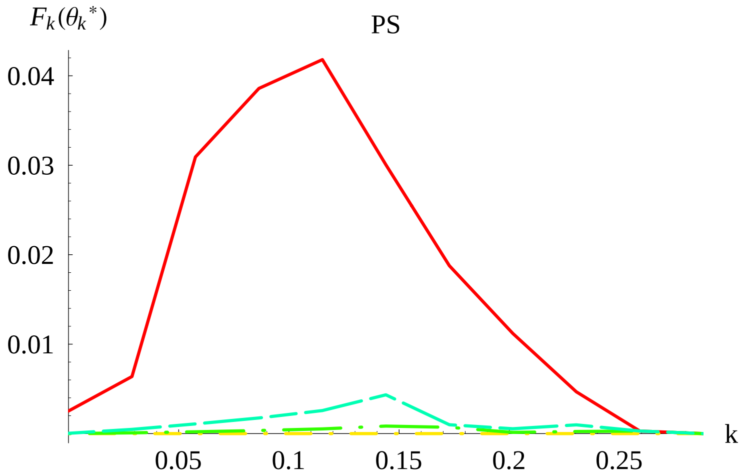

FIGURE 13. $F_k^{\mathrm{PS}}(\boldsymbol{\theta}_k^b) - F_k^{\mathrm{PS}}(\boldsymbol{\theta}_k^{\mathrm{PS}})$: Objective PS penalty against target k for using non-PS objective b to obtain optimal weights $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$. The penalties in terms of increased probability of failure to meet the target, for using the "wrong" objectives are in the order $\mathrm{SR} > \mathrm{TSV} > \mathrm{TS} > \mathrm{PS} = 0$

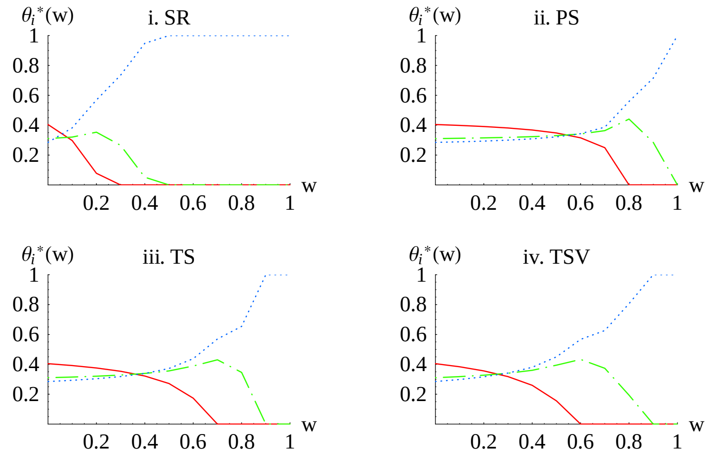

FIGURE 14. Optimal weight $\boldsymbol{\theta}_i(w)$ for asset i against mixing parameter w, for $k = 10\%$

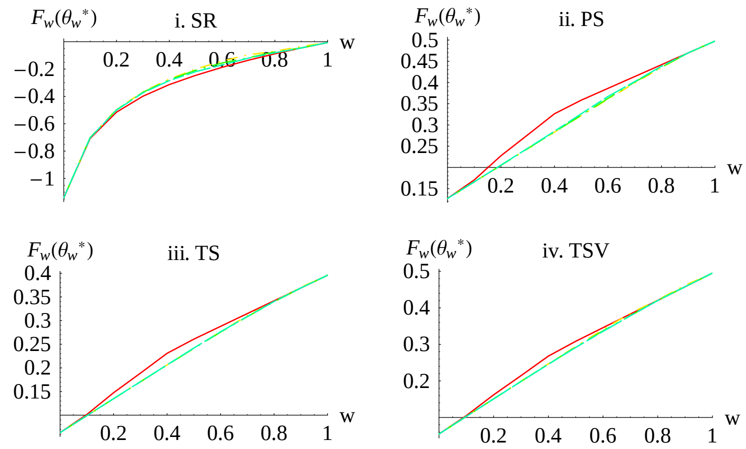

FIGURE 15. $F_w^a(\boldsymbol{\theta}_w^b)$: Objective a evaluated using the optimal weights for objective b against mixing parameter w, for $k = 10\%$. Each plot is for an objective a; each curve is for an objective b, where $a, b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$.

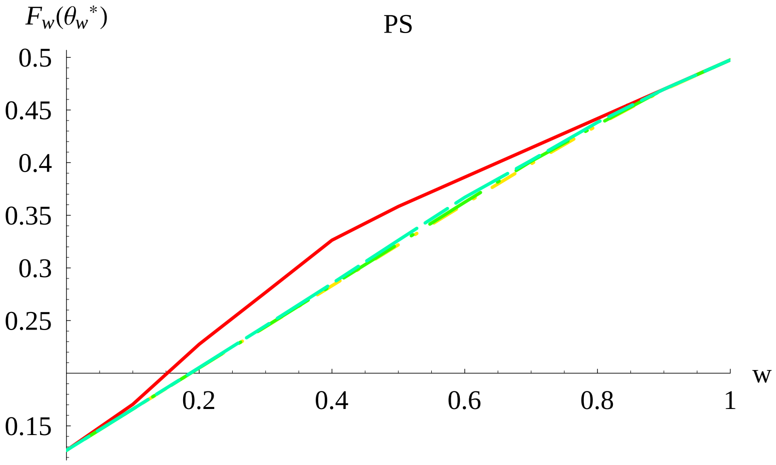

FIGURE 16. $F_w^{\mathrm{PS}}(\boldsymbol{\theta}_w^b)$: Objective PS against mixing parameter, w, evaluated at the optimal weights for objective $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$.

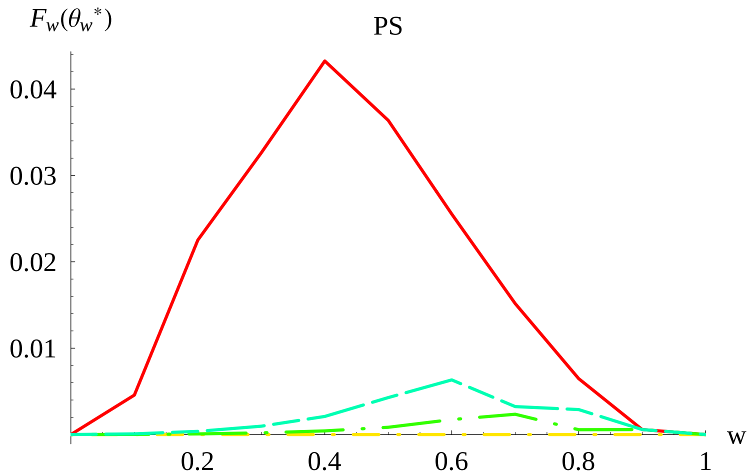

FIGURE 17. $F_w^{\mathrm{PS}}(\boldsymbol{\theta}_w^b) - F_w^{\mathrm{PS}}(\boldsymbol{\theta}_w^{\mathrm{PS}})$: PS penalty against mixing parameter w for using non-PS objective b to obtain optimal weights $b \in \{\mathrm{SR}, \mathrm{PS}, \mathrm{TS}, \mathrm{TSV}\}$. The penalties in terms of increased probability of failure to meet the target, for using the "wrong" objectives are in the order $\mathrm{SR} > \mathrm{TSV} > \mathrm{TS} > \mathrm{PS} = 0$

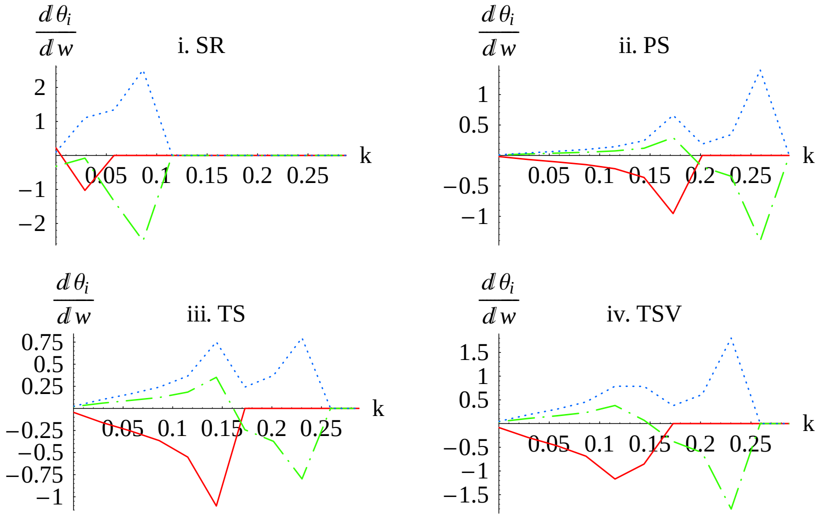

FIGURE 18. Distress sensitivities $\frac{d\theta_i}{dw}$ vs. k, for $w = 40\%$

References

1. A. Ang and G. Bekaert, *International asset allocation with regime shifts*, The Review of Financial Studies. **15** (2002), no. 4, 1137–1187.
2. ______, *How do regimes affect asset allocation?*, NBER Working Paper No. w10080, November 2003.
3. F. Arditti and H. Levy, *Portfolio efficiency analysis in three moments: The multi-period case*, Journal of Finance **30** (1975), 797–809.
4. P. Artzner, F. Delbaen, J. Eber, and D. Heath, *Thinking coherently*, RISK **10** (1997), no. 11, 68–71.
5. D.P. Bertsekas, *Non-linear programming*, Athena Scientific, 1995.
6. N.H. Bingham and R. Kiesel, *Semi-parametric modelling in finance*, Quantitative Finance **2** (2002), 241–250.
7. I.R.C. Buckley, G. Comezaña, B. Djerroud, and L. Seco, *Portfolio optimization for alternative investments*, Seminario de matematica financiera Instituto MEFF-RiskLab, Madrid. ISBN 84-688-2450-X (Carrillo-Sanchez Calle, ed.), vol. 3, 2003, pp. 1–15.
8. R. Campbell, K. Koedijk, and P. Kofman, *Increased correlation in bear markets*, Financial Analyst Journal (2002), 87–94.
9. L. Capiello and T.A. Fearnley, *International capm with regime switching garch parameters*, On-line at http://www.fame.ch:8080/research/papers/CapielloFearnley.pdf, July 2000.
10. D. Dowe, *David dowe's clustering, mixture modelling and unsupervised learning page*, On-line at http://www.csse.monash.edu.au/∼dld/mixture.modelling.page.html, 1999.
11. J. Driffill, T. Kenc, and M. Sola, *Merton-style option pricing under regime switching*, Computing in Economics and Finance (2002), no. 304.
12. E. Eberlein, *Lévy processes: Theory and applications*, ch. Applications of generalized hyperbolic Levy motions to finance, Boston, MA: Birkhauser, 2001.
13. C.B. Erb, C.R. Harvey, and T.E. Viskanta, *Forecasting international equity correlations*, Financial Analysts Journal (1994), 32–45.
14. R. Frey, A.J. McNeil, and M. Nyfeler, *Copulas and credit models*, RISK (2001), 111–114.
15. W. Fung and D.A. Hsieh, *Is mean-variance analysis applicable to hedge funds?*, Economic letters **62** (1998), 53–58.
16. A. Graflund and B. Nilsson, *Dynamic portfolio selection: The relevance of switching regimes and investment horizon*, European Financial Management (2003), 179–200.
17. J. Hamilton, *A quasi-bayesian approach to estimating parameters for mixtures of normal distributions*, Journal of Business and Economic Statistics **9** (1991), no. 1, 2739.
18. C.R. Harvey, J.C. Liechty, M.W. Liechty, and P. Müller, *Portfolio selection with higher moments*, Working paper W70, Fuqua School of Business, Duke University, July 2003.
19. S. Hodges, *A generalization of the sharpe ratio and its application to valuation bounds and risk measures*, University of Warwick Financial Options Research Centre, FORC Pre-Print 98/88, April 1998.
20. J. Hull and A. White, *Value at risk when daily changes in market variables are not normally distributed*, Journal of Derivatives **5** (1998), no. 3, 9–19.
21. A. Kraus and R. H. Litzenberger, *Skewness preference and the valuation of risk assets*, Journal of Finance **31** (1983), 1085–1100.
22. C. Labidi and T. An, *Revisiting the finite mixture of gaussian distributions with applications to futures markets*, Computing in Economics and Finance (2000), no. 67.
23. F. Longin and B. Solnik, *Correlation structure of international equity markets during extremely volatile periods*, Journal of Finance **56** (2001), no. 2, 649–676.
24. D.B. Madan and G.S. McPhail, *Investing in skews*, Journal of Risk Finance **2** (2000), no. 1, 10–18.
25. H. Markowitz, *Portfolio selection*, Journal of Finance **7** (1952), 77–71.
26. D. Pelletier, *Regime switching for dynamic correlations*, Working paper, Department of Economics, North Carolina State University, July 2003.
27. M. Pivato and L. Seco, *Estimating the spectral measure of a multivariate stable distribution via spherical harmonic analysis*, Journal of Multivariate Analysis **87** (2003), no. 2, 219–240.
28. W.F. Sharpe, *The sharpe ratio*, The Journal of Portfolio Management **Fall** (1994).
29. D. Sornette, P. Simonetti, and J.V. Andersen, *"nonlinear" covariance matrix and portfolio theory for non-gaussian multivariate distributions*, Physics Report **335** (2000), no. 2, 19–92.

30. M.C. Steinbach, *Markowitz revisited: Mean-variance models in financial portfolio analysis*, SIAM Review **43** (2001), no. 1, 31–85.
31. A. Suárez and S. Carrillo, *Computational tools for the analysis of market risk*, Computing in Economics and Finance (2000), no. 144.
32. S. Venkataraman, *Value at risk for a mixture of normal distributions: The use of quasi-bayesian estimation techniques*, Economic Perspectives (1997), 2–13.
33. J. Wang, *Modeling and generating daily changes in market variables using a multivariate mixture of normal distributions*, Working paper. Department of Math and Computer Science, Valdosta State University, January 2000.
34. P. Zangari, *An improved methodology for measuring var*, RiskMetrics Monitor **2Q** (1996).

E-mail address: i.buckley@imperial.ac.uk

Centre for Quantitative Finance, Tanaka Business School, Imperial College, London, SW7 2BX, UK

E-mail address: saunders@euler.math.pitt.edu

Dept. of Mathematics, THACK 617 Pittsburgh, University of Pittsburgh, Pittsburgh PA, 15260, USA

E-mail address: seco@math.toronto.edu

RiskLab, University of Toronto, Toronto, Canada
Current address: Sigmanalysis, Toronto, Canada