

# Portfolio optimization in incomplete markets and price constraints determined by maximum entropy in the mean

Argimiro Arratia · Henryk Gzyl

Cite this article: Arratia, A., Gzyl, H. Portfolio Optimization in Incomplete Markets and Price Constraints Determined by Maximum Entropy in the Mean. *Computational Economics*, 56, 929-952 (2020). <https://doi.org/10.1007/s10614-019-09954-3>

**Abstract** A solution to a portfolio optimization problem is always conditioned by constraints on the initial capital and the price of the available market assets. If a risk neutral measure is known, then the price of each asset is the discounted expected value of the asset's price under this measure. But if the market is incomplete, the risk neutral measure is not unique, and there is a range of possible prices for each asset, which can be identified with bid-ask ranges. We present in this paper an effective method to determine the current prices of a collection of assets in incomplete markets, and such that these prices comply with the cost constraints for a portfolio optimization problem. Our workhorse is the method of maximum entropy in the mean to adjust a distortion function from bid-ask market data. This distortion function plays the role of a risk neutral measure, which is used to price the assets, and the distorted probability that it determines reproduces bid-ask market values. We carry out numerical examples to study the effect on portfolio returns of the computation of prices of the assets conforming the portfolio with the proposed methodology.

**Keywords** portfolio optimization · maximum entropy in mean · distortion function · risk neutral measures · asset pricing

**Mathematics Subject Classification (2010)** 91B24 · 91B28 · 91B30 · 90C25

## 1 Introduction

There are many ways of optimizing portfolios. One can safely say that all of them, except for the naive portfolio diversification, consist of some form of trade off be-

---

A. Arratia  
Computer Science, Barcelona Tech (UPC), Barcelona, Spain  
E-mail: [argimiro@cs.upc.edu](mailto:argimiro@cs.upc.edu), ORCID: 0000-0003-1551-420X

H. Gzyl  
Center for Finance, IESA, Caracas, Venezuela  
E-mail: [henryk.gzyl@iesa.edu.ve](mailto:henryk.gzyl@iesa.edu.ve)

tween return and risk subject to an available capital constraint. One of the standard versions of the portfolio optimization problem, originally due to [Markowitz \(1952\)](#), has many variants, all of which use the volatility (or the standard deviation) of the portfolio as risk measure, and the trade off between risk and return is set up in a direct form, like for example: find the portfolio of minimum variance that yields a given expected return. This last constraint is clearly equivalent to a cost constraint upon the future values of the assets when a risk neutral probability is available. A few standard references for this matter are [Pliska \(1997\)](#); [Karatzas \(1998\)](#); [Föllmer and Schied \(2016\)](#).

Motivated by the impossibility of using non-arbitrage arguments when the market is incomplete, or when the class of risk neutral measures is not known, or more importantly, when the initial price for each asset can not be specified, an alternative theoretical methodology, which has been termed “conic portfolio theory”, was developed recently by [Madan \(2016\)](#) (see also [Madan and Schoutens \(2016\)](#)), and others. In this proposal, ideas coming from risk analysis in the insurance industry are used to provide the trade off between return and risk, by letting the role of the risk constraint to be played by either the bid or ask prices of the assets, which are estimated using a distortion function.

Here we shall use distortion functions in a different, almost opposite way. For this we think of the price of each asset as the price of a collection of risks specified up to a bid-ask range, and then we determine a risk pricing measure that yields prices within the bid-ask ranges. The prices of assets so determined will be the current prices which are to be used to compute the returns of the portfolio.

The numerical procedure we propose to obtain the distortion function is based on the method of maximum entropy in the mean. This is a convenient method to solve inverse problems subject to convex constraints upon the solution, which allows for the data to be specified in intervals. This procedure has the important advantage of being model free, which in particular does not call for the calibration of parameters.

The historical antecedents of the maximum entropy in the mean method can be traced back to [Jaynes \(1957\)](#), who proposed and used a standard maximum entropy methodology in statistical physics. This form of the maximum entropy method is later used as a valuable approximation scheme to solutions of some cases of the problem of moments in [Mead and Papanicolaou \(1984\)](#), and the mathematical foundation of the maximum entropy in the mean method is developed in [Dacunha-Castelle and Gamboa \(1990\)](#).

Here the maximum entropy in the mean method is used to assign a price to a market asset, which in a certain sense corresponds to the discounted expected value of the asset under some (unknown) risk-neutral measure, but what can only be said is that this price is within some market bid-ask range. We termed these prices, determined by the market, as *conservative market prices*, in concordance with similar terminology use in the literature ([Madan, 2016](#); [Madan and Schoutens, 2016](#)). Once these conservative market prices for the assets comprising our portfolio are at hand, we can do two things. First, we can use the prices to compute returns, and second, to assign value to the price constraint for the portfolio optimization problem.

Interrelationships between derivative valuation in incomplete markets and risk pricing in insurance have been brought up many times. When the market model is incomplete, it was noted that instead of a given price, one could assign a price

range to any given asset. One could also use methods from utility (or risk) theory to determine a market price of the asset. In the insurance industry a similar problem exists: How to assign a price to a potential loss, or a price to risk. For this purpose, a risk pricing axiomatics and a risk pricing methodology were developed, which in many ways parallels and overlaps that developed to value the risk of financial assets. For a few references about this see [Young \(2006\)](#) and [Laeven and Goovaerts \(2008\)](#).

### 1.1 Plan of our work

We begin with briefly describing the market models in finance and in insurance, at the same time that we introduce the necessary notation in [Section 2](#). There are two issues related to bid-ask ranges. On one hand, bid-ask ranges for assigning prices to derivatives are related to market incompleteness. When the market is incomplete but the class of risk neutral measures is known, a bid-ask range can be applied (but see [Cerny and Hodges \(2000\)](#) for a characterization in term of no good deal prices). On the other hand, the existence of bid-ask ranges for current prices does not allow the construction of the class of risk neutral probabilities, which impedes risk neutral pricing. But equally important, since to solve the problem of portfolio optimization we need to be able to calculate asset returns, we have to device a way of going around the absence of current asset price.

We briefly describe the way this is taken care of in [Madan and Schoutens \(2016\)](#), [Madan \(2016\)](#). Since their proposal involves the use of distorted probabilities and their connection to coherent risk measures, we devote some space to coherent risk measures, risk pricing in insurance, and eventually establish a connection between distorted risk measures and risk neutral probabilities, that can be summarised as follows (see [Section 2.6](#)). We take the bid-ask range as the only available market data, and determine a distortion function that yields asset prices within the bid-ask range. Since the distorted probability law so obtained is equivalent to a risk neutral probability yielding the computed price as actual price, we shall consider that price as the base price against which we compute the returns of the basic assets for the purpose of portfolio optimization.

In [Section 3](#) we describe the numerical procedure to obtain the distortion function, show how the discretization is carried out, and write down the solution of the problem by the method of maximum entropy in the mean. It will turn out that there will be two possible solutions, depending on whether we do, or do not, impose bounds upon the derivatives of the distortion function. We present both results there, but we work out the details of this procedure in the [Appendix 6.1](#).

We consider in [Section 4](#) a portfolio optimization problem. For that we shall consider an exponential utility and make use of the connection between that problem and entropy maximization, in a way similar to that developed by [Föllmer and Schied \(2016\)](#).

In [Section 5](#) we sum it all up, and in the [Appendices 6.1](#) and [6.2](#) we collect all material set aside for not to interrupt the main flow of ideas.

## 2 Bid-ask prices and distortion functions

Before entering into the subject of the section, it is convenient to establish the market models in finance and insurance in order to introduce the basic notations.

### 2.1 The market models

In this work we shall consider a two time market model consisting of a probability space  $(\Omega, \mathcal{F}, P)$  and a finite collection of  $M$  basic assets whose prices at time  $t = 0$  are denoted by  $S_1(0), \dots, S_M(0)$  and whose prices at time  $t = 1$  are positive, continuous random variables  $S_1(1), \dots, S_M(1)$ , each distributed according to a probability density function  $F_i$ ,  $i = 1, \dots, M$ .

In this model we might as well suppose that  $\Omega = [0, \infty)^M$ , and that  $\mathcal{F} = \mathcal{B}(\mathbf{S}(1))$ , the Borel  $\sigma$ -algebra generated by  $\mathbf{S}(1) = (S_1(1), \dots, S_M(1))$ . Here  $P$  denotes a probability measure on  $(\Omega, \mathcal{F})$ , the so called “physical law” which describes the statistical behavior of  $\mathbf{S}(1)$ . A standard financial asset in this model will be a finite valued,  $\mathcal{F}$ -measurable function  $X : \Omega \rightarrow \mathbb{R}$ , and the problem for the financial analyst is how to price  $X$ .

The simplest model used in the insurance industry goes as follows. We start with a probability space  $(\Omega, \mathcal{F}, P)$ , and a collection of positive random variables denoting the severity (or monetary value) of the losses occurring in a given period of time. We do not have the analogue of the initial prices, just the loss incurred in the period. And we suppose that there is a finite number  $L_1, \dots, L_M$  of them, denoted collectively by  $\mathbf{L}$ . As before, we suppose that  $\Omega = [0, \infty)^M$ , and that  $\mathcal{F} = \mathcal{B}(\mathbf{L})$ . In this case we are interested in pricing the losses or functions of the losses, which can be thought as the prices of the possible insurance products.

### 2.2 Bid-ask prices

A key tool for the pricing of financial assets is that of risk neutral probability. The class of such probabilities is defined (in a one period) market model by

$$\mathcal{M}_e = \{Q \in \mathcal{M}(P) \mid E_Q[e^{-r} S_i(1)] = S_i(0)\} \quad i = 1, \dots, M.$$

A market model is complete when this class reduces to one element. Otherwise, the market model is said to be incomplete, and financial assets do not have a unique price. The theoretical determination of the bid-ask price of an asset of future value  $X$  at time  $t = 1$  is given by (see [Pliska \(1997\)](#) or [Karatzas \(1998\)](#)):

$$\begin{aligned} b(X) &= \exp(-r) \inf\{E_Q[X] : Q \in \mathcal{M}_e\} \\ a(X) &= \exp(-r) \sup\{E_Q[X] : Q \in \mathcal{M}_e\} \end{aligned} \quad (1)$$

As mentioned, the class  $\mathcal{M}_e$  may not be known, then in this case (1) can not be used for obtaining a price range. One reason for which  $\mathcal{M}_e$  may not be known is that the basic assets do not have a specified price at  $t = 0$ , but are known only up to a bid-ask range. That is, we are in what is called a two price economy in [Madan \(2016\)](#) or [Madan and Schoutens \(2016\)](#). These authors propose a different market model in which there are no basic assets, and no notion of risk neutral measure.

To be specific, they propose a notion of bid-ask range that goes as follows. To begin with consider a convex cone  $\mathcal{A}$  of acceptable risks, and a class of probability laws  $\mathcal{P} = \{Q \sim P | E_Q[X] \geq 0; \forall X \in \mathcal{A}\}$  (see Chapter IV of Föllmer and Schied (2016) for these notions), and propose the following definition of bid-ask range for each  $X$  as follows:

$$\begin{aligned} b(X) &= \inf\{E_Q[X] : Q \in \mathcal{P}\} \\ a(X) &= \sup\{E_Q[X] : Q \in \mathcal{P}\} \end{aligned} \quad (2)$$

At this point we mention again that here we shall regard the bid-ask range of the basic assets as a datum obtained from the market, either as quoted quantities or to be determined from the daily price data.

### 2.3 Coherent and convex risk measures

The need to appropriately price risk resulted in an rather precise (axiomatic) characterization of what is a risk measure. Actually, there already exist several interrelated classes of risk measures, which are already a text book material, and what follows about coherent and convex risk measures is taken from Föllmer and Schied (2016), although originally proposed by Artzner et al. (1999). Let  $\Xi$  be the collection of random variables representing the changes of value of an asset. A *risk measure* is any function  $\rho : \Xi \rightarrow \mathbb{R}$ , and a *coherent risk measure* is defined as follows.

**Definition 1** A coherent risk measure is any risk measure satisfying the following conditions:

1. (Monotonicity) For  $X_1, X_2 \in \Xi$ , and  $X_1 \leq X_2$  then  $\rho(X_2) \leq \rho(X_1)$ .
2. (Sub-additivity) For  $X_1, X_2 \in \Xi$ , it holds that  $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$ .
3. (Homogeneity) For any  $a \geq 0$  and any  $X \in \Xi$ , we have  $\rho(aX) = a\rho(X)$ .
4. (Translation invariance) For any  $m \in \mathbb{R}$  and any  $X \in \Xi$ ,  $\rho(X + m) = \rho(X) - m$ .

When the sub-additivity and homogeneity conditions are merged into the convexity requirement:

$$\text{For } X_1, X_2 \in \Xi \text{ and } t \in [0, 1], \quad \rho(tX_1 + (1-t)X_2) \leq t\rho(X_1) + (1-t)\rho(X_2)$$

we say that  $\rho$  is a *convex risk measure*.

The following result, which can be derived from (Artzner et al., 1999, Prop. 4.1), suggests a similarity between bid-ask prices and coherent risk measures.

**Theorem 1** A lower semi-continuous risk measure  $\rho$  is coherent if and only if there exists a class  $\mathcal{Q} \subset \mathcal{M}(P)$  such that

$$\rho(X) = \inf\{E_Q[X] | Q \in \mathcal{Q}\}.$$

The probabilities in  $\mathcal{Q}$  are called “generalized scenarios”. Observe that, according to (1), when the class of generalized scenarios is the class of risk neutral measures, this theorem renders the bid price  $b(X)$  as a risk measure (if  $X$  could be thought of as a change of value of some asset). Similarly for  $-a(-X)$ . This analogy is behind the portfolio optimization methodology propounded in Madan (2016).

## 2.4 Risk pricing measures

Risk pricing in insurance exists from a long time ago, and many ways to compute premiums exist. Eventually various premium principles were established, and depending on the premium principles chosen, computational methods consistent with them were proposed. See [Wang et al. \(1997\)](#), [Young \(2006\)](#) or [Laeven and Goovaerts \(2008\)](#), for example. To stress the analogies with risk axiomatics, let us list the premium principles.

**Definition 2** Let  $\Xi$  be a collection of positive, integrable random variables modeling losses. A premium principle is a mapping  $\Pi : \Xi \rightarrow [0, \infty]$  satisfying:

1. (Law invariance)  $\Pi(X)$  depends only on  $F_X$ .
2. (Homogeneity) For any  $X \in \Xi$  and  $a > 0$ , we have  $\Pi(aX) = a\Pi(X)$ .
3. (Translation invariance) For any  $X \in \Xi$  and  $a > 0$ , we have  $\Pi(a + X) = a + \Pi(X)$ .  
Conditions 2 and 3 give the (*no unjustified risk loading*) principle that states that the price of a known loss is the amount of the loss;  $\Pi(a) = a$  for any  $a \geq 0$ .
4. (Risk loading) For any  $X \in \Xi$ ,  $\Pi(X) \geq E[X]$ .
5. (No rip-off) For any  $X \in \Xi$ ,  $\Pi(X) \geq \text{esssup}(X)$ .
6. (Sub-additivity) For any  $X_1, X_2 \in \Xi$  it holds that  $\Pi(X_1 + X_2) \leq \Pi(X_1) + \Pi(X_2)$ . Additivity and super-additivity can be similarly defined.
7. (Additivity for comonotonic losses). Two random losses  $X_1, X_2 \in \Xi$  are comonotonic if for every  $\omega_1, \omega_2 \in \Omega$  we have  $(X_1(\omega_2) - X_1(\omega_1))(X_2(\omega_2) - X_2(\omega_1)) \geq 0$ . For comonotonic variables,  $\Pi$  is comonotonic additive if  $\Pi(X_1 + X_2) = \Pi(X_1) + \Pi(X_2)$ .
8. (Monotonicity) For  $X_1, X_2 \in \Xi$  such that  $X_1 \leq X_2$ , then  $\Pi(X_1) \leq \Pi(X_2)$ .
9. (Stochastic ordering compatibility) A premium is said to be compatible with first stochastic dominance if  $F_{X_1} \leq F_{X_2}$ , then  $\Pi(X_1) \leq \Pi(X_2)$ . Similarly, it is said to be second order stochastic dominance compatible, if whenever  $E[(X_1 - d)^+] \leq E[(X_2 - d)^+]$ , for all  $d > 0$ , then  $\Pi(X_1) \leq \Pi(X_2)$ . (Here  $(X - a)^+$  denotes as usual  $\max(X - a, 0)$ .)

The following result from [Wang et al. \(1997\)](#) ties coherent risk theory with risk pricing in insurance and will be the basis of our proposal to obtain an alternative to risk neutral measures from bid-ask prices. (Recall that  $X \wedge a$  stands for  $\min(X, a)$ .)

**Theorem 2** Suppose that a risk pricing measure satisfies Properties (1), (2), (3), (7), (8) and the following continuity property:

$$\text{For } X \in \Xi, \quad \lim_{a \downarrow 0} \Pi((X - a)^+) = \Pi(X), \quad \lim_{a \downarrow 0} \Pi(X \wedge a) = \Pi(X).$$

Then there exists a non-decreasing (distortion) function  $g : (0, 1) \rightarrow (0, 1)$  such that

$$\Pi(X) = \int_0^\infty g(\widehat{F}_X(x)) dx. \quad (3)$$

Here  $\widehat{F}_X = 1 - F_X$ . Also, if  $\Xi$  contains all Bernoulli random variables,  $g$  is unique and  $g(p)$  is given by the price of the Bernoulli( $p$ ) random loss (remember law invariance).

Another proof of this representation result appears in [Wu and Wang \(2003\)](#).

## 2.5 Relation to bid and ask prices

If we think of the prices  $\{S_i(1), i = 1, \dots, M\}$  of the basic assets at time  $t = 1$  as positively valued risks, and think of the bid-ask price ranges  $(b_i, a_i)$  as being given by the market, we may ask ourselves: What is the distortion function producing these prices? That is, we want to solve the following problem:

Determine a continuously differentiable concave distortion function  $g(u)$  such that

$$e^{-r} \int_0^\infty g(\widehat{F}_{S_i}(x)) dx \in (b_i, a_i), \quad i = 1, \dots, M. \quad (4)$$

Hence forward to simplify the notation we use  $F_i$  instead of  $F_{S_i}$ . Let  $\psi(u) = 1 - g(1 - u)$ . Then

$$\begin{aligned} \int_0^\infty x d\psi(F_i(x)) &= \int_0^\infty \left( \int_0^\infty I_{[0,x]}(s) ds \right) d\psi(F_i(x)) \\ &= \int_0^\infty \left( \int_0^\infty I_{[s,\infty)}(x) d\psi(F_i(x)) \right) ds = \int_0^\infty g(\widehat{F}_i(s)) ds. \end{aligned}$$

But also in account of the assumptions on the  $S_i$  and  $g$ , making the change of variables  $u = F_i(x)$  or  $x = F_i^{-1}(u) = q_i(u)$ , we obtain

$$\begin{aligned} \int_0^\infty x d\psi(F_i(x)) &= \int_0^\infty x g'(\widehat{F}_i(x)) d\widehat{F}_i(x) \\ &= \int_0^1 q_i(u) g'(1 - u) du = \int_0^1 q_i(u) \phi(u) du. \end{aligned}$$

Here we set  $\phi(u) = g'(1 - u)$ . Since we want  $g$  to be concave, we have to require  $\phi(u)$  to be increasing. Thus with these changes of notation problem (4) becomes:

Determine a continuous, increasing  $\phi(u)$  on  $[0, 1]$  such that

$$e^{-r} \int_0^1 q_i(u) \phi(u) du \in (b_i, a_i), \quad i = 1, \dots, M. \quad (5)$$

## 2.6 Distortion functions and risk neutral densities

Here we establish a relation between the two valuation methods. Let us consider the case of a market model with one random asset and denote by  $F$  the cumulative distribution function of its future value  $S(1)$ . Let  $h(x)$  be either a positive or a bounded measurable function defined on  $[0, \infty)$ . We define the probability  $P^\psi$  on  $(\Omega, \sigma(S(1)))$  by

$$E^\psi[h(S(1))] = \int_0^\infty h(x) d\psi(F(x)) = \int_0^\infty h(x) \psi'(F(x)) dF(x).$$

Let us put  $\Lambda = \psi'(F(S(1)))$ . Observe that when  $\psi$  is concave and increasing  $\psi > 0$ , thus  $\Lambda > 0$ . Now define  $Q \sim P$  by  $dQ/dP = \Lambda$ . Then clearly

$$E_Q[h(S(1))] = E[\Lambda h(S(1))] = \int_0^\infty h(x) \psi'(F(x)) dF(x) = E^\psi[h(S(1))].$$

To obtain a distortion function from a risk neutral density, we do as follows. Suppose  $Q \sim P$  is a risk neutral probability. Let  $k$  be a positive Borel measurable function such that  $k(S(1)) = E[dQ/dP|S(1)]$ . Notice that  $k(x)$  can be obtained as the Radon-Nikodym derivative of  $Q(S(1) \leq x)$  with respect to  $P(S(1) \leq x)$  or as the regular conditional probability  $E[dQ/dP|S(1) = x]$ . Now define  $\psi_1 : [0, 1] \rightarrow [0, \infty)$  by  $\psi_1(u) = k(q(u))$  where  $q = F^{-1}$ , and set  $\psi(t) = \int_0^t \psi_1(u) du$ . Then  $\psi' = \psi_1$  for almost all  $u \in [0, 1]$ . Notice that  $\psi'(F(S(1))) = k(S(1))$  almost surely with respect to the Lebesgue measure on  $[0, \infty]$  due to the assumptions on  $S(1)$ . Furthermore, as above

$$E^\psi[h(S(1))] = E_Q[h(S(1))].$$

Observe that from the interpretation of  $k$  mentioned above, it is clear that  $\psi(1) = 1$ . Now set  $g(u) = 1 - \psi(1 - u)$ . With the notations as just introduced, we can gather these comments under the statement of the following theorem.

**Theorem 3** *Let  $g(u)$  be a concave, twice continuously differentiable distortion function satisfying*

$$e^{-r} \int_0^\infty g(\widehat{F}(x)) dx \in (b, a)$$

*with  $b$  and  $a$  bid and ask prices at time  $t = 0$ . Let  $\psi(u) = 1 - g(1 - u)$ , and define  $\Lambda = \psi'(F(S))$  so that  $dQ = \Lambda dP$ . Then  $Q \sim P$  and*

$$e^{-r} E_Q[S(1)] \in (b, a).$$

*Furthermore, for  $h$  as above,*

$$E^\psi[h(S(1))] = \int_0^\infty h(x) d\psi(F(x)) = \int_0^\infty h(x) \psi'(F(x)) dF(x) = E_Q[h(S(1))].$$

*Conversely, if  $Q \sim P$  satisfies  $e^{-r} E_Q[S(1)] \in (b, a)$ , define  $k(x) = dQ(S(1) \leq x)/dF(x)$ , and  $\psi(t) = \int_0^t k(q(u)) du$ . Since  $\psi'(F(S(1))) = k(S(1))$ , clearly*

$$E^\psi[h(S(1))] = E_Q[h(S(1))]$$

*and furthermore*

$$e^{-r} E^\psi[h(S(1))] \in (b, a).$$

The proof of this theorem is given in the Appendix 6.2.

### 3 Discretization of problem (5)

Observe that either of problems (4) or (5) is an integral equation with convex constraints upon the solution and with finite data given in an interval. In order to solve them numerically, the first step is to discretize the problem. Let us show how to discretize (5). First we partition  $[0, 1]$  into  $N$  intervals of equal length and put  $\phi_j = \frac{1}{2}(\phi((j-1)/N) + \phi(j/N))$ , for  $j = 1, \dots, N$ . Now put  $B(i, j) = \frac{e^{-r}}{N} q_i(j/N)$ , for  $i = 1, \dots, M$ , and  $j = 1, \dots, N$ . Then, problem (5) becomes:



$$\begin{aligned} \text{Solve the system} \quad \mathbf{B}\boldsymbol{\phi} \in \mathcal{K} &= \prod_{i=1}^M [b_i, a_i] \\ \text{Subject to} \quad 0 &< \phi_1 < \phi_2 < \dots < \phi_N. \end{aligned} \quad (6)$$

To change the nature of this constraint we proceed as follows. Let  $\mathbf{T}$  be the  $M \times N$  lower triangular matrix with all entries equal to 1 and let  $\boldsymbol{\xi} \in [0, \infty)^N$ , put  $\boldsymbol{\phi} = \mathbf{T}\boldsymbol{\xi}$  and  $\mathbf{A} = \mathbf{B}\mathbf{T}$ . Then (6) becomes

$$\begin{aligned} \text{Solve the system} \quad \mathbf{A}\boldsymbol{\xi} \in \mathcal{K} &= \prod_{i=1}^M [b_i, a_i] \\ \text{Subject to} \quad \boldsymbol{\xi} &\in [0, \infty)^N. \end{aligned} \quad (7)$$

This is an under-determined linear system with convex constraints on the solution and data in intervals. The method of Maximum Entropy in the Mean (MEM for short) takes care nicely of this type of problems.

### 3.1 Estimating the distortion functions by Maximum Entropy in the Mean

Actually, MEM allows for flexibility in the choice of the constraints. We could consider for example, that either  $\boldsymbol{\xi} \in [0, L]^N$ , or even in some more elaborate domain, or as above,  $\boldsymbol{\xi} \in [0, \infty)$ . The two choices lead to somewhat different representation of the solutions.

Here we shall only list the results and postpone the details to the appendices section. When we require the solution to problem (5) to be bounded (we shall refer to this case as the MEM bounded), the MEM will provide us with a representation like

$$\xi_j^* = L \frac{e^{-L(\mathbf{A}^t \boldsymbol{\lambda})_j}}{1 + e^{-L(\mathbf{A}^t \boldsymbol{\lambda})_j}}, \quad j = 1, \dots, N. \quad (8)$$

Here  $\mathbf{A}^t$  stands for the transpose of the matrix  $\mathbf{A}$ , and  $\boldsymbol{\lambda}^*$  is a particular value of a Lagrange multiplier that we shall specify how to determine in the appendices. Once  $\boldsymbol{\xi}^*$  is obtained by any method, we can determine the value  $\mathbf{S}(0)$  that enters in (12) according to

$$\mathbf{S}(0) = \mathbf{A}\boldsymbol{\xi}^* \in \prod_{i=1}^M [b_i, a_i]. \quad (9)$$

## 4 Portfolio optimization

The approach to the portfolio optimization problem that we are considering consists on constructing a portfolio which maximizes some utility function subject to some initial capital constraints. The utility function is supposed to incorporate the balance between return and risk of the investor.

Therefore, given a utility function  $U$ , the physical measure  $P$  and the risk neutral measure  $Q$ , a standard presentation of the portfolio optimization problem

consists in solving the following optimization problem (cf. Föllmer and Schied (2016)):

$$\begin{aligned} \text{Find : } & \sup_{\mathbf{h}} E^P[U(\langle \mathbf{h}, \mathbf{S}(1) \rangle)] \\ \text{Subject to : } & \langle \mathbf{h}, \mathbf{S}(0) \rangle = C_0; \quad h_i \geq 0, \quad i = 1, \dots, M. \end{aligned} \quad (10)$$

where we are considering  $M$  assets, whose prices at time  $t = 0$  are  $\mathbf{S}(0) = e^{-r} E_Q[\mathbf{S}(1)]$ , where as above,  $\mathbf{S}(1)$  denotes the random prices at  $t = 1$ , and where  $C_0$  is the available capital and, for the sake of making the analysis simpler (but without loss of generality), we consider only portfolios  $\mathbf{h} = (h_1, \dots, h_M)$  with long positions in the assets.

Alternatively, if one wants to work with unit-less weights, we may consider the gross return of the assets, defined by  $X_i(1) = S_i(1)/S_i(0)$ ,  $i = 1, \dots, M$ , and setting  $w_i = h_i S_i(0)/C_0$ , we can rewrite the portfolio optimization problem as

$$\begin{aligned} \text{Find : } & \sup_{\mathbf{w}} E^P[U(C_0 \langle \mathbf{w}, \mathbf{X}(1) \rangle)] \\ \text{Subject to : } & \langle \mathbf{w}, \mathbf{1} \rangle = 1; \quad 0 \leq w_i, \quad i = 1, \dots, M. \end{aligned} \quad (11)$$

Observe that this setup presupposes that we know the prices  $S_i(0)$  of the random assets at  $t = 0$ . Not only that, the sheer notions of gross return or rate of return, imply that we know an initial price of each asset.

To overcome these two issues we propose to use a distortion function  $\Psi$  to define *conservative market prices*  $S_i^d(0)$  by

$$S_i^d(0) = e^{-r} \int x d\Psi(F_{S_i}(x)) \in (b_i(x), a_i(x)), \quad i = 1, \dots, M \quad (12)$$

as specified in (9), and use them to restate the portfolio optimization problem as follows:

$$\begin{aligned} \text{Find : } & \sup_{\mathbf{h}} E^P[U(\langle \mathbf{h}, \mathbf{S}(1) \rangle)] \\ \text{Subject to : } & \langle \mathbf{h}, \mathbf{S}^d(0) \rangle = C_0; \quad h_i \geq 0, \quad i = 1, \dots, M. \end{aligned} \quad (13)$$

As explained in the previous section, we do not need to know (or impose) a particular distortion function, but instead we estimate this from bid and ask prices by the method of maximum entropy in the mean.

#### 4.1 An example of portfolio optimization

In this section we carry out the portfolio optimization procedure for a particular example of utility function. This function is chosen because some of the details of the process overlap those of the entropy maximization procedure for the determination of the distortion function. We shall consider the utility function  $U(c) = 1 - \exp(-c)$ , and then to find

$$\sup_{\mathbf{h}} E^P[U(\langle \mathbf{h}, \mathbf{S}(1) \rangle)]$$

which is equivalent to find

$$\inf_{\mathbf{h}} E^P[\exp(-\langle \mathbf{h}, \mathbf{S}(1) \rangle)]$$

subject to  $\langle \mathbf{h}, \mathbf{S}^d(0) \rangle = C_0$ , where (to insist)  $S_i^d(0) = e^{-r} \int x d\Psi(F_{S_i}(x))$ ,  $i = 1, \dots, M$ , and we know that  $S_i^d(0) \in (b_i, a_i)$ . The function to minimize is log-convex, so it is rather convenient to write  $V(\mathbf{h}) = E^P[\exp(-\langle \mathbf{h}, \mathbf{S}(1) \rangle)]$  and consider the following equivalent convex minimization problem:

$$\begin{aligned} \text{Find : } & \inf_{\mathbf{h}} \ln V(\mathbf{h}) \\ \text{Subject to : } & \langle \mathbf{h}, \mathbf{S}^d(0) \rangle = C_0; \quad h_i \geq 0, \quad i = 1, \dots, M. \end{aligned} \quad (14)$$

## 4.2 Numerical Experiments

In order to show the merits of our proposed portfolio optimization model (14) in practice, we build a portfolio with real data for seven companies trading in the NYSE market, chosen randomly and subject only to have a sufficiently long daily price history for making extensive simulations. The tickers of the companies considered are ALB, AAPL, APC, GE, WFC, HAL, and SHW, and for each we have daily OHLC prices history from 2000-01-01 to 2016-12-29.

For our experiments we fix the initial capital  $C_0 = 10^5$  units of the market currency. We sampled monthly returns and observed prices at the end of a month, for each asset, so that  $\mathbf{S}(1)$  represents the assets' prices after one month, and our goal is to compare the capital gain (or gross return) of the portfolio  $\mathbf{h} = (h_1, \dots, h_M)$  at  $t = 1$ , as given by  $\frac{\langle \mathbf{h}, \mathbf{S}(1) \rangle}{C_0}$ , when  $\mathbf{h}$  is obtained from the MEM (bounded or unbounded) estimation  $\mathbf{S}^d(0)$  of the initial prices (i.e. as solution of Equation 14), against the portfolio  $\mathbf{h}$  obtained from current initial price, denoted by  $\tilde{\mathbf{S}}(0)$ . We stress the fact that we are doing single period optimization at different epochs to assess our portfolio optimization method in different market scenarios. We are not doing multi-period rebalancing.

### 4.2.1 Estimating the bid and ask prices from data

Suppose we have, besides the historical prices, also the High and Low prices for each asset (as we do in our experimental data). Let  $\tau > 0$  be a number of periodic (monthly, in our experiments) observations of prices. Then we estimate the value  $Bid(t)$  of the bid at time  $t$ , as the mean of the daily Lows in the epoch  $[t - \tau, t]$ ; correspondingly the value  $Ask(t)$  of the ask at time  $t$  is taken as the mean of the daily Highs in the epoch  $[t - \tau, t]$ .

We denote by  $\mathbf{S}(0)_b$  and  $\mathbf{S}(0)_{unb}$  the initial conservative market prices, which should lie in between these bid-ask intervals (i.e. in  $[Bid(0), Ask(0)]$ ), using respectively the MEM Bounded (Eqs. 19 and 20) and the MEM Unbounded method (Eqs. 21 and 22).

Table 1 shows these different estimates, at a given date, for the bid-ask intervals and the initial conservative market price  $\mathbf{S}(0)$ , for each of the stocks considered in

our experiment. One can see that in practice both MEM methods (unbounded or bounded with  $L = 10^2$ ) give similar values for  $\mathbf{S}(0)$ , and in both cases these values verify to be in, or close to, the interior of the Bid-Ask interval. Therefore, for our portfolio optimizations we shall only make use of the MEM unbounded method for obtaining initial price  $\mathbf{S}^d(0)$ .

Asset	$\hat{\mathbf{S}}(0)$	$\mathbf{S}(0)_b$ ( $10^2$ )	$\mathbf{S}(0)_{unb}$	<i>Bid</i>	<i>Ask</i>
ALB	20.04	27.296	27.301	26.976	27.659
AAPL	13.53	11.765	11.765	11.405	11.765
APC	40.07	45.125	45.133	44.594	45.858
GE	16.37	33.628	33.711	33.558	34.074
WFC	28.37	31.389	31.389	31.389	32.087
HAL	17.38	29.563	29.569	29.168	29.968
SHW	59.91	50.979	51.013	50.863	51.997

**Table 1** Price  $\hat{\mathbf{S}}(0)$  in 2008-12-29, MEM-Bounded price  $\mathbf{S}(0)_b$  ( $L = 10^2$ ), MEM-Unbounded price  $\mathbf{S}(0)_{unb}$ , and bid-ask ranges.

#### 4.2.2 Estimating $\mathbf{S}(1)$

The estimation  $\hat{\mathbf{S}}(1)$  of  $\mathbf{S}(1)$ , necessary for solving the problem (14), can be done through a model of prices, for example a Geometric Brownian Motion (GBM) process, using the information on estimated initial prices  $\hat{\mathbf{S}}(0)$  and other parameters. We rather choose to estimate  $\mathbf{S}(1)$  by a data-driven method, consisting on using a sufficiently long past history of gross returns up to time  $t = 0$  for each asset. Let  $\mathbf{R} = (R_{ij})_{1 \leq i \leq N, 1 \leq j \leq M}$ , where  $R_{ij}$  is sample gross return of asset  $j$  at time  $i \leq t = 0$ , and set

$$\hat{\mathbf{S}}(1) = \mathbf{R}(\hat{\mathbf{S}}(0))^t$$

(where  $t$  in the formula means transpose).

#### 4.2.3 Mean Variance Portfolio selection.

For the sake of testing our portfolio methodology with other classical form of portfolio optimization, we consider optimizing portfolios with respect to the Markowitz's mean-variance utility function. Recall that the Mean-Variance (long only) portfolio optimization is given by the solution to the following problem:

$$\begin{aligned} \max_{\mathbf{w} \in \mathbb{R}^M} \quad & \mu^t \mathbf{w} - \frac{\gamma}{2} \mathbf{w}^t \Sigma \mathbf{w} \\ \text{subject to} \quad & \langle \mathbf{w}, \mathbf{1} \rangle = 1, \quad \mathbf{w} \geq 0. \end{aligned} \tag{15}$$

where  $\mu_i = E[X_i]$ ,  $i = 1, \dots, M$ ,  $\Sigma = \text{Var}(\mathbf{X})$  and  $\gamma \geq 0$  is the risk aversion parameter. We observe that in most general cases of real data,  $\mu$  is of the order of  $10^{-2}$ , and consequently  $\Sigma \sim 10^{-4}$ . Therefore, in order to build the Minimum Variance (MinVar) portfolio set  $\gamma > 10^2$ .

#### 4.2.4 General results

To simulate portfolio's monthly returns we use 5 years of monthly data (60 observations), taken from the dataset described above with dates ranging from 2000-01-01 to 2016-12-29. In order to obtain up to 13 different scenarios (monthly periods where to assess the portfolio performance) we apply a rolling window analysis, each window period of 5 years, and using increments of 1 year of data between successive windows. Figures 1 and 2 show the portfolios returns, for different monthly periods, with respect to the method chosen to optimize the portfolio (exponential utility or mean-variance maximization), and the way of computing the initial values of portfolio (the  $\mathbf{S}(0)$ ). We can see that for all but one case (the month ending in 2008-12-29, the date of the big market crash), optimizing the portfolios with initial prices computed by the MEM method improve returns substantially. In Table 2 below we report the averages quartiles of monthly portfolio returns, over all 13 different months, for portfolios optimized with respect to the exponential utility (EU) objective and the minimum variance (MV) objective, and taking  $\mathbf{S}(0)$  as current value ( $\mathbf{S}(0)_{current}$ ), or estimated by the MEM Unbounded method ( $\mathbf{S}(0)_{mem}$ ). The results show that, in general, estimating initial prices  $\mathbf{S}(0)$  with the MEM method translates in better portfolio performance.

EU	$Q_0$	$Q_1$	Median	$Q_3$	$Q_4$
$\mathbf{S}(0)_{current}$	0.923	0.985	1.015	1.045	1.118
$\mathbf{S}(0)_{mem}$	1.261	1.361	1.407	1.451	1.561

MV	$Q_0$	$Q_1$	Median	$Q_3$	$Q_4$
$\mathbf{S}(0)_{current}$	0.919	0.986	1.014	1.039	1.105
$\mathbf{S}(0)_{mem}$	1.251	1.352	1.392	1.433	1.526

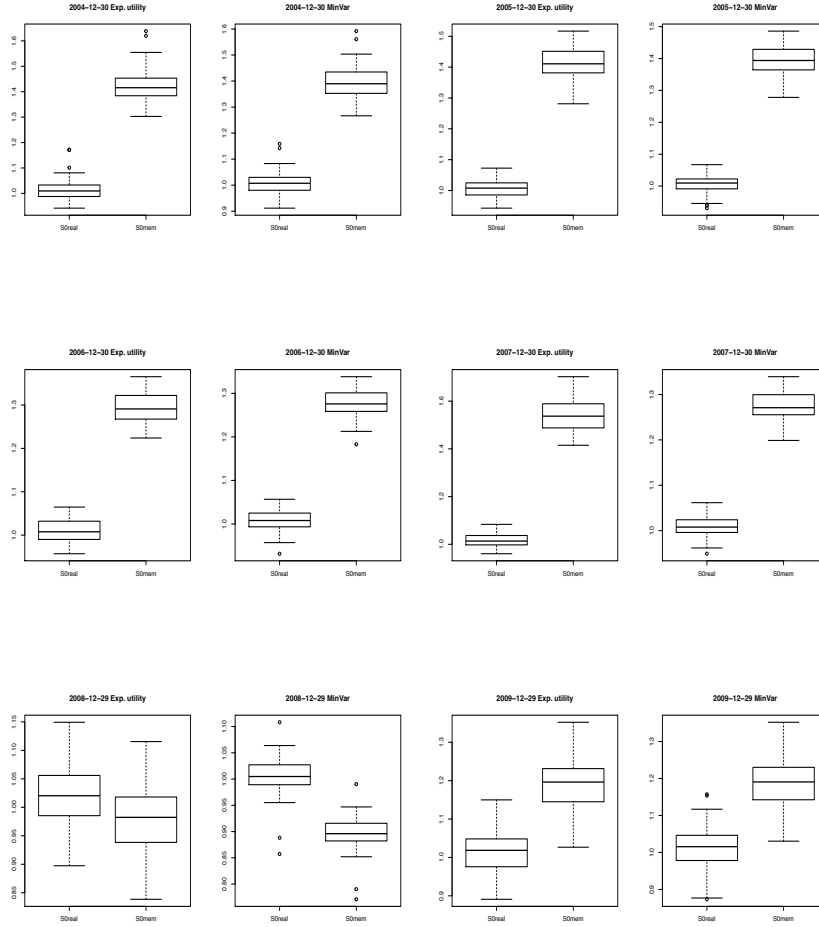
**Table 2** Average quartiles of monthly portfolio returns throughout all months considered

We use the numerical software CVX with solver ECOS (Domahidi, Chu and Boyd, 2013), and the programming language R (R Core Team, 2015) to solve our optimization problems.

## 5 Conclusions

As said at the beginning, the underlying idea behind this paper is to provide a method to fill in a methodological detail in the process of portfolio optimization. Namely, that in order to comply with the cost constraint, we need to be able to assign prices to the assets at  $t = 0$ , or in the notation used throughout, to be able to estimate  $\mathbf{S}(0)$ , of which we formally know that  $\mathbf{S}(0) = e^{-r} E_Q[\mathbf{S}(1)]$ , for  $\mathbf{S}(1)$  the random prices at  $t = 1$  and  $Q$  some risk-neutral measure. We proposed an idea coming from insurance risk pricing, which turns out to be equivalent to the notion of risk neutral probability.

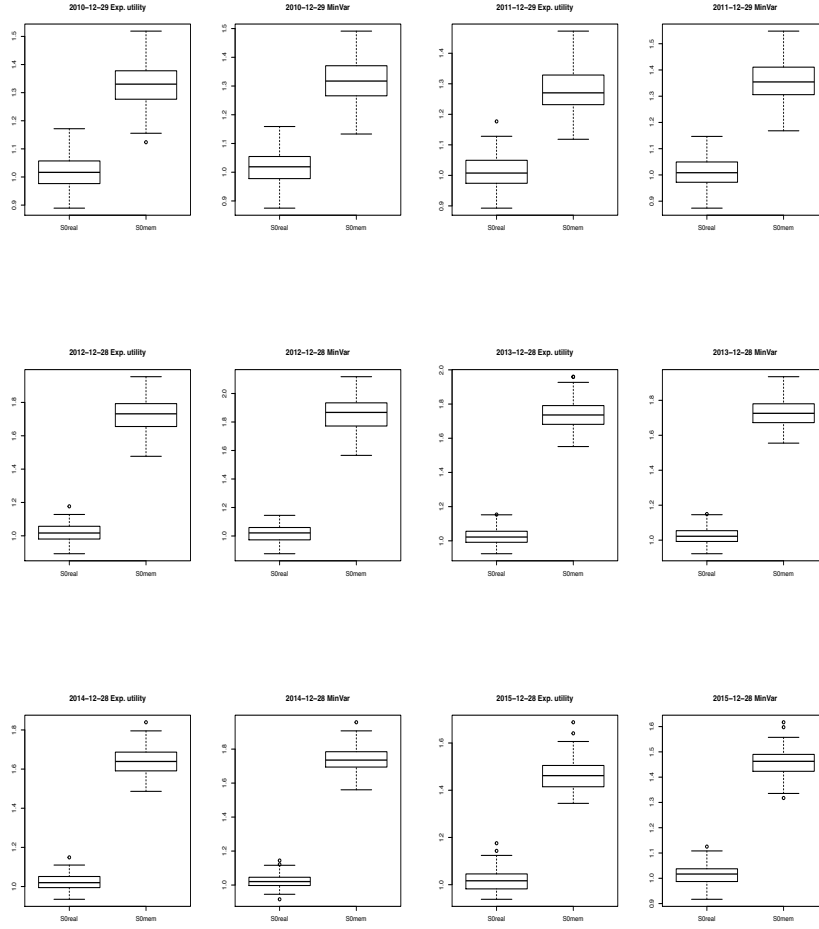
That part of the project produces nice results: that is, we are able to use a model free, non parametric approach to consistently estimate  $\mathbf{S}(0)$ . The second part of the project, consists of using that estimate, or the constraint based on it, as part of an optimization process. Here a standard issue appears: How to model



**Fig. 1** Boxplots of EU-type and MV-type portfolio's monthly returns (month end date indicated in header of each figure), for  $S(0)$  obtained by: current price ( $S0_{real}$ ), MEM method ( $S0_{mem}$ ).

the tradeoff between risk and return, which is the initial starting point of every portfolio optimization problem.

We chose to exemplify the optimal portfolio configuration using two different methods: On one hand, a utility maximization procedure, and on the other a simple return-risk tradeoff in the Markowitz (volatility, return)-plane. In both cases, it is clear the improvement in performance by the portfolios designed with the initial prices of its constituent assets, estimated by the method based on maxi-



**Fig. 2** Continuation of Figure 1 showing comparison of EU-type and MV-type portfolio's monthly returns according to the way of computing  $S(0)$ .

mum entropy in the mean that we have proposed in this paper. Finally, note that the MEM method provides us with a non-parametric model free procedure to go around the lack of knowledge about the risk neutral probabilities.

## 6 Appendices

### 6.1 Details about MEM

In this section we establish the basic formalism of the method of Maximum Entropy in the Mean (MEM). To begin, let us consider (7) in a context free notation. The problem that we want to solve is to find  $\xi \in \mathcal{C}$  such that

$$\mathbf{A}\xi \in \mathcal{K}$$

where  $\mathbf{A}$  is an  $M \times N$ -matrix. We suppose that the sets  $\mathcal{C}$  and  $\mathcal{K}$  are convex and have non-empty interiors. Eventually we shall consider  $\mathcal{C} = [0, L]^N$  or  $[0, \infty)^N$ , and  $\mathcal{K} = \prod_i [b_i, a_i]$ . The parameter  $L$  in our situation is related to a bound on the distortion function: it is given by the assumption that the values  $\phi_j$ , for  $j = 1, \dots, N$ , are bounded. If we do not want to preassign that value, we can set  $L = \infty$ . We shall give solutions for both cases of  $L$ , bounded or unbounded. However, in our applications to portfolio optimization we shall see that both cases of  $L$  give similar results; hence by the principle of parsimony one should prefer, in practice, to solve problem (7) over  $\mathcal{C} = [0, \infty)^N$ .

The idea behind MEM can be stated as follows: Consider an auxiliary random variable  $\mathbf{X}$ , taking values in  $\mathcal{C}$  such that  $\xi = E_{\mathbb{P}}[\mathbf{X}]$  with respect to an unknown probability measure  $\mathbb{P}$  that is supported by  $\mathcal{C}$ . If such probability measure can be found, the constraint upon  $\xi$  is automatically satisfied. To define  $\mathbf{X}$  we consider the probability space  $(\mathcal{C}, \mathcal{B}(\mathcal{C}), \mathbb{Q})$ , where  $\mathcal{B}(\mathcal{C})$  denotes the Borel subsets of  $\mathcal{C}$ , and  $\mathbb{Q}$  is any ( $\sigma$ -finite) measure such that the convex hull  $\text{con}(\text{supp}(\mathbb{Q}))$  generated by its support equals  $\mathcal{C}$ . What this last requirement achieves for us, is that any probability  $\mathbb{Q}$  having a strictly positive density  $\rho(\xi)$  with respect to  $\mathbb{Q}$ , will satisfy  $E_{\mathbb{P}}[\mathbf{X}] \in \mathcal{C}$ . The standard maximum entropy (SME) procedure enters as a stepping stone at this stage. On the class of densities (with respect to  $\mathbb{Q}$ ) we define the entropy function by

$$S(\rho) = - \int_{\mathcal{C}} \rho(\xi) \ln \rho(\xi) d\mathbb{Q}(\xi),$$

or  $-\infty$  if the integral of  $|\ln \rho(\xi)|$  is not convergent. The choice of sign is conventional, we want to maximize entropy. The entropy function is strictly concave due to the strict concavity of the logarithm. To find a  $\rho^*(\xi)$  such that  $d\mathbb{P}^* = \rho d\mathbb{Q}$  satisfies (5) we solve the problem

$$\text{Find } \rho^* \text{ at which } \sup\{S(\rho) | \mathbf{A}E_{\mathbb{Q}}[\mathbf{X}] \in \mathcal{K}\} \text{ is achieved} \quad (16)$$

Since the maximization can be split into a sequence of two steps:

$$\sup_{\mathbf{y} \in \mathcal{K}} \sup\{S(\rho_{\mathbf{y}}) | \mathbf{A}E_{\mathbb{P}}[\mathbf{X}] = \mathbf{y}\}$$

in which the inner maximization has a standard solution given by

$$\rho_{\mathbf{y}}(\xi) = \frac{e^{-\langle \lambda^*, \mathbf{A}_c \xi \rangle}}{Z(\lambda^*)} \quad (17)$$

where the normalization factor is clearly given by

$$Z(\lambda) = \int_{\mathcal{C}} e^{-\langle \lambda^*, \mathbf{A}_c \xi \rangle} d\mathbb{Q}(\xi).$$



The  $\lambda^*$  appearing in (17) is to be determined minimizing the convex function (dual entropy) given by

$$\Sigma(\lambda, \mathbf{y}) = \ln Z(\lambda) + \langle \lambda, \mathbf{y} \rangle$$

over  $\{\lambda \in \mathbb{R}^M \mid Z(\lambda) < \infty\}$ , which in all cases of interest for us will be  $\mathbb{R}^M$  itself. Furthermore, and very important,

$$S(\rho_{\mathbf{y}}^*) = \Sigma(\lambda^*, \mathbf{y}).$$

Therefore, the double maximization described above becomes

$$\sup_{\mathbf{y} \in \mathcal{K}} S(\rho_{\mathbf{y}}^*) = \sup_{\mathbf{y} \in \mathcal{K}} \inf_{\lambda \in \mathbb{R}^M} \{\Sigma(\lambda, \mathbf{y})\} = \inf_{\lambda \in \mathbb{R}^M} \{\ln Z(\lambda) + \sup_{\mathbf{y} \in \mathcal{K}} \{\langle \lambda, \mathbf{y} \rangle\}\}. \quad (18)$$

To make the pending computations explicit, note that

$$\sup_{\mathbf{y} \in \mathcal{K}} \{\langle \lambda, \mathbf{y} \rangle\} = \sum_{i=1}^M \left( \frac{a_i - b_i}{2} |\lambda_i| + \frac{a_i + b_i}{2} \lambda_i \right).$$

Now we shall complete the details in the two cases of interest to us, corresponding to  $\mathcal{C} = [0, L]^N$  or  $\mathcal{C} = [0, \infty)^N$ . In each case, we have to specify  $Q$  and compute  $Z(\lambda)$ .

#### 6.1.1 MEM: Bounded case

For  $L < \infty$ , consider the following measure  $\mathbb{Q}$  on  $\mathcal{C} = [0, L]^N$ :

$$\mathbb{Q}(d\boldsymbol{\xi}) = \prod_{j=1}^N (\epsilon_0(dx_j) + \epsilon_L(dx_j)).$$

We use the notation  $\epsilon_a(dx)$  to denote the measure that puts a unit mass at the point  $a$ . It takes a simple computation to verify that

$$Z(\lambda) = \prod_{j=1}^N \left( 1 + e^{-(L\mathbf{A}^t \lambda)_j} \right)$$

where the superscript “t” denotes transposition and  $(\mathbf{A}^t \lambda)_j = \sum_{i=1}^M \lambda_i A_{i,j}$ . With all this, the dual entropy to be minimized in (18) becomes

$$\Sigma(\lambda) = \ln Z(\lambda) + \sum_{i=1}^M \left( \frac{a_i - b_i}{2} |\lambda_i| + \frac{a_i + b_i}{2} \lambda_i \right) \quad (19)$$

Once the  $\lambda^*$  that minimizes the right hand side of (19) has been found, then

$$\xi_j^* = L \frac{e^{-(L\mathbf{A}^t \lambda^*)_j}}{1 + e^{-(L\mathbf{A}^t \lambda^*)_j}} \quad j = 1, \dots, N \quad (20)$$

At this point we add that when  $a_i = y_i = b_i$  for  $i = 1, \dots, M$ , that is when the data is pointwise, then the last term in (19) becomes  $\langle \lambda, \mathbf{y} \rangle$ . In this case the first order condition for  $\lambda^*$  to be a minimizer of (19) is equivalent to the fact that  $\mathbf{A}\xi^* = \mathbf{y}$  as can be easily verified.

To conclude the section we mention that the dual entropy  $\Sigma(\boldsymbol{\lambda}, \mathbf{y})$  may not have a minimum. For example,  $\Sigma(\boldsymbol{\lambda}, \mathbf{y})$  is bounded below in  $\boldsymbol{\lambda}$  for every  $\mathbf{y}$ , but the infimum is reached at infinity; or when the problem is not well specified and data does not fall in the range of  $\mathbf{A}$ , in which case  $\Sigma(\boldsymbol{\lambda}, \mathbf{y})$  is not bounded below. For further details on this, see [Borwein and Lewis \(2000\)](#) or [Gzyl and Velásquez \(2011\)](#).

### 6.1.2 MEM: Unbounded case

The problem that we want to solve now is to find  $\boldsymbol{\xi} \in \mathcal{C} = [0, \infty)^N$  such that

$$\mathbf{A}\boldsymbol{\xi} \in \mathcal{K}$$

where  $\mathbf{A}$  is an  $M \times N$ -matrix. As in the bounded case,  $\mathcal{C}$  and  $\mathcal{K}$  are convex and have non-empty interiors. Everything goes almost verbatim as above, except that the reference measure that we consider now is a product of un-normalized Poisson measures:

$$d\mathbb{Q} = \prod_{j=1}^N \left( \sum_{n=1}^{\infty} \frac{1}{n!} \epsilon_n d\xi_j \right).$$

This time the function  $Z(\boldsymbol{\lambda})$  is given by

$$Z(\boldsymbol{\lambda}) = \prod_{j=1}^N \exp \left( e^{-(\mathbf{A}^t \boldsymbol{\lambda})_j} \right)$$

To continue as above we form the function  $\Sigma(\boldsymbol{\lambda})$  and obtain

$$\Sigma(\boldsymbol{\lambda}) = \sum_{j=1}^N e^{-(\mathbf{A}^t \boldsymbol{\lambda})_j} + \sum_{i=1}^M \left( \frac{a_i - b_i}{2} |\lambda_i| + \frac{a_i + b_i}{2} \lambda_i \right) \quad (21)$$

Once again, when the  $\boldsymbol{\lambda}^*$  that minimizes the right hand side of (21) has been found, then the solution to  $\mathbf{A}\boldsymbol{\xi} \in \mathcal{K}$  is given by

$$\xi_j^* = e^{-(\mathbf{A}^t \boldsymbol{\lambda}^*)_j}, \quad j = 1, \dots, N \quad (22)$$

Notice as well that when  $a_i = y_i = b_i$  then

$$\Sigma(\boldsymbol{\lambda}) = \sum_{j=1}^N e^{-(\mathbf{A}^t \boldsymbol{\lambda})_j} + \sum_{i=1}^M \lambda_i y_i.$$

In this case, the first order condition on  $\boldsymbol{\lambda}^*$  to be a minimizer looks like

$$-\sum_{j=1}^N A_{i,j} e^{-(\mathbf{A}^t \boldsymbol{\lambda}^*)_j} + y_i = 0 \quad i = 1, \dots, M$$

from which it is clear that (22) satisfies the equation  $\mathbf{A}\boldsymbol{\xi} = \mathbf{y}$ .

### 6.1.3 Sub-differential of (19)

As a first step towards a solution, we have the following:

**Lemma 1** *With the notations introduced above, the function  $\Sigma(\boldsymbol{\lambda})$  defined in (19) is strictly convex*

Observe that  $\partial_i \Sigma(\boldsymbol{\lambda}) := \partial \Sigma(\boldsymbol{\lambda}) / \partial \lambda_i$  is defined except at  $\lambda_i = 0$ , where it is only sub-differentiable (Borwein and Lewis, 2000). More explicitly,

$$\partial_i \Sigma(\boldsymbol{\lambda}) = \begin{cases} \partial \ln Z(\boldsymbol{\lambda}) / \partial \lambda_i + a_i & \text{when } \lambda_i > 0, \\ \partial \ln Z(\boldsymbol{\lambda}) / \partial \lambda_i + b_i & \text{when } \lambda_i < 0, \\ \partial \ln Z(\boldsymbol{\lambda}) / \partial \lambda_i \in (b_i, a_i) & \text{when } \lambda_i = 0, \end{cases} \quad (23)$$

Moreover, in the bounded case,  $\ln Z(\boldsymbol{\lambda}) = \sum_{j=1}^N \ln(1 + e^{-(L\mathbf{A}^t \boldsymbol{\lambda})_j})$ , and using the expression  $(\mathbf{A}^t \boldsymbol{\lambda})_j = \sum_{k=1}^M \lambda_k A_{k,j}$ , we have that

$$\partial \ln Z(\boldsymbol{\lambda}) / \partial \lambda_i = - \sum_{j=1}^N L A_{i,j} \frac{e^{-(L\mathbf{A}^t \boldsymbol{\lambda})_j}}{1 + e^{-(L\mathbf{A}^t \boldsymbol{\lambda})_j}} \quad (24)$$

and in the unbounded case

$$\ln Z(\boldsymbol{\lambda}) = \sum_{j=1}^N e^{-(\mathbf{A}^t \boldsymbol{\lambda})_j}$$

and we have

$$\partial \ln Z(\boldsymbol{\lambda}) / \partial \lambda_i = - \sum_{j=1}^N A_{i,j} e^{-(\mathbf{A}^t \boldsymbol{\lambda})_j} \quad (25)$$

Finally, we have the general statement

**Theorem 4** *Suppose that the infimum  $\boldsymbol{\lambda}^*$  of  $\Sigma(\boldsymbol{\lambda})$  is reached at some  $\boldsymbol{\lambda}^*$  in the interior of  $\{\boldsymbol{\lambda} \in \mathbb{R}^M : Z(\boldsymbol{\lambda}) < \infty\}$ . Then*

$$\partial_i \Sigma(\boldsymbol{\lambda}) \in \{0\}, \quad i = 1, \dots, M.$$

## 6.2 Distortion functions and risk neutral densities: General case

Here we shall extend the result presented in Section 2.6 to the multidimensional case. Remember that we are considering an  $M$ -dimensional, positive random vector  $\mathbf{S}(1)$  describing the values of the assets at  $t = 1$ . Recall as well that we are supposing each  $S_i(1)$  to be continuously distributed having cumulative distribution function  $F_i(x)$  with strictly positive densities. Let us denote by  $F(\mathbf{x}) = P(S_1(1) \leq x_1, \dots, S_M(1) \leq x_M)$  the joint distribution function of  $\mathbf{S}(1)$ . Let  $g$  be the distortion function introduced in Section 2.6 and again  $\psi(u) = 1 - g(1 - u)$ . We begin by establishing some properties of the distorted cumulative distribution function  $G(\mathbf{x}) = \psi(F(\mathbf{x}))$ . Let us denote by  $G_i(x_i) = \psi(F_i(x_i))$  the one dimensional distorted cumulative density of  $S_i(1)$ .

**Lemma 2** *With the notations introduced above,  $G(\mathbf{x})$  has marginals  $G_i(x_i)$ .*

The proof for the general case is similar to that of the 2-dimensional case, but notationally simpler. Notice, for example, that, for fixed  $x_1 \in [0, \infty)$

$$\begin{aligned} \int_0^\infty d_2 G(x_1, x_2) &= \int_0^\infty d_2 \psi(F(x_1, x_2)) \\ &= \psi(F(x_1, \infty)) - \psi(F(x_1, 0)) = \psi(F_1(x_1)) = G_1(x_1). \end{aligned}$$

From this it is clear that, for any positive or bounded, measurable  $h$  defined on  $[0, \infty)$ , we have  $E[h(S_i)] = E^\psi[h(S_i)]$ , and in particular,

$$e^{-r} E^\psi[S_i] = e^{-r} \int_{\mathbb{R}_+^M} x_i d^M G(\mathbf{x}) = e^{-r} \int_0^\infty x_i d_i G_i(x_i), \quad (26)$$

where  $d^M := d_1 d_2 \cdots d_M$ . Therefore, we could have stated Problem (4) as consisting of finding a distortion function  $g$  such that

$$e^{-r} E^\psi[S_i] = e^{-r} \int_{\mathbb{R}_+^M} x_i d^M \psi(F(\mathbf{x})) \in (b_i, a_i), \quad \text{for } i = 1, \dots, M. \quad (27)$$

To continue, according to Sklar's theorem (Sklar, 1959), there exists a (unique under our assumptions)  $N$ -dimensional distribution function  $C_F(\mathbf{u})$  on  $[0, 1]^M$  such that  $F(\mathbf{x}) = C_F(F_1(x_1), \dots, F_M(x_M))$ . We shall now verify that the copula  $C_G$  of the distorted distribution function is a  $\psi$ -distorted copula of  $C_F$ . See Jaworski et al. (2010) for an introduction to the subject.

**Proposition 1** *With the notations introduced in the previous paragraph we have,*

$$C_G(\mathbf{u}) = \psi(C_F(\psi^{-1}(u_1), \dots, \psi^{-1}(u_M))).$$

The proof goes as follows: Recall that  $G_i(x_i) = \psi(F_i(x_i))$ , which implies that  $F_i(x_i) = \psi^{-1}(G_i(x_i))$ . Therefore

$$\begin{aligned} G(\mathbf{x}) &= \psi(C_F(F_1(x_1), \dots, F_M(x_M))) \\ &= \psi(C_F(\psi^{-1}(G_1(x_1)), \dots, \psi^{-1}(G_M(x_M)))) \end{aligned}$$

Since  $G(\mathbf{x}) = C_G(G_1(x_1), \dots, G_M(x_M))$ , the desired result follows by setting  $G_i(x_i) = u_i$ .

In order to relate  $G(\mathbf{x})$  to a risk neutral density, we have to establish the absolute continuity of  $G(\mathbf{x})$  with respect to  $\prod_{i=1}^M d_i F_i(x_i)$ . To motivate the result, consider the following computations.

For  $M = 1$ ,

$$d_1 G(x_1) = \psi^{(1)}(F_1) d_1 F_1(x_1).$$

For  $M = 2$ , drop the subscript  $F$  from  $C_F$  and  $C_i$  denotes  $\partial C / \partial u_i$ , etc. The chain rule applied to  $G(\mathbf{x}) = \psi(C(F(\mathbf{x})))$  yields

$$\begin{aligned} d_2 d_1 G(x_1, x_2) &= \left( \psi^{(2)}(C(F)) C_1(F) C_2(F) \right. \\ &\quad \left. + \psi^{(1)}(C(F)) C_{1,2}(F) \right) d_1 F_1(x_1) d_2 F_2(x_2). \end{aligned}$$

For  $M = 3$ ,

$$\begin{aligned} d_3 d_2 d_1 G(\mathbf{x}) = & \left( \psi^{(3)}(C(F)) C_1(F) C_2(F) C_3(F) + \psi^{(2)}(C_{1,2}(F) C_3(F) \right. \\ & \left. + C_{2,3}(F) C_1(F) + C_{1,3}(F) C_2(F)) + C_{1,2,3}(F) \right) d_1 F_1(x_1) d_2 F_2(x_2) d_3 F_3(x_3). \end{aligned}$$

Note that the symmetry with respect to the exchange of the subscripts is related to the fact that we can integrate with respect to any variable in any order. We shall now split Theorem 3 into two parts (to keep its length under control). But let us first introduce some more notations. Let  $[M] = \{1, \dots, M\}$  and put

$$\mathcal{P}(M, k) = \{\text{Partitions of } [M] \text{ into } k \text{ blocks}\}.$$

For each subset  $J = \{i_1, i_2, \dots, i_n\}$  of  $[M]$  we write

$$C_J = \frac{\partial^N C}{\partial u_{i_1} \dots \partial u_{i_n}}.$$

**Theorem 5** *With the notations introduced above, suppose that  $\psi(u)$  is  $M$ -times continuously differentiable with positive derivatives  $\psi^{(k)}$ . Suppose that  $C$  has continuous derivatives  $C_J$  for all  $J \subset [M]$ . (These are non-negative since  $C$  is a cumulative distribution function and we differentiate with respect to different variables). Then we have*

$$d^M G(\mathbf{x}) = \left[ \sum_{k=1}^M \psi^{(k)}(C(F)) \left( \sum_{J \in \mathcal{P}(M, k)} C_J(F) \right) \right] \prod_{i=1}^M d_i F_i(x_i). \quad (28)$$

That is,  $G(\mathbf{x})$  is absolutely continuous with respect to  $\prod_{i=1}^M d_i F_i(x_i)$ , with density given by the term in square brackets. Denote that by  $\rho(F(\mathbf{x}))$ . Furthermore (consider  $\psi(u) = u$  if you want),  $d^M F(\mathbf{x}) = C_{[M]}(F) \prod_{i=1}^M d_i F_i(x_i)$ .

**Corollary 1** *If both  $\rho(F) > 0$  and  $C_{[M]} > 0$ , then  $G \sim F$  with density  $\frac{\rho(F(\mathbf{x}))}{C_{[M]}(F(\mathbf{x}))}$ . Furthermore, setting  $\Lambda = \frac{\rho(F(\mathbf{S}))}{C_{[M]}(F(\mathbf{S}))}$  and  $dQ = \Lambda dP$ , then for any measurable, positive or bounded  $h(\mathbf{x})$ , we have  $E_Q[h(\mathbf{S})] = E^\psi[h(\mathbf{S})]$ , and in particular*

$$e^{-r} E_Q[S_i] \in (b_i, a_i), \quad \text{for } i = 1, \dots, M.$$

The statement of the converse will come at the end of the following observations. Note that if  $Q$  is a probability on  $(\Omega, \mathcal{F})$  which is equivalent to  $P$ , if all functions that we are going to integrate are  $\sigma(\mathbf{S})$ -measurable, we might suppose that there is a function  $\Lambda : \mathbb{R}_+^M \rightarrow (0, \infty)$  such that  $d^M G(\mathbf{x}) = \Lambda(\mathbf{x}) d^M F(\mathbf{x})$ . Here  $G(\mathbf{x}) = Q(\mathbf{S} \leq \mathbf{x})$ . We shall suppose that  $d^M F(\mathbf{x})$  is equivalent to  $\prod_{i=1}^M d_i F_i(x_i)$  with strictly positive density given by  $C_M(F(\mathbf{x}))$ . Write  $\Lambda(\mathbf{x}) = k(F_1(x_1), \dots, F_M(x_M))$ , with  $k(u_1, \dots, u_M) = \Lambda(F_1^{-1}(u_1), \dots, F_M^{-1}(u_M))$ , then

$$d^M G(\mathbf{x}) = k(F_1(x_1), \dots, F_M(x_M)) C_M(F_1(x_1), \dots, F_M(x_M)) \prod_{i=1}^M d_i F_i(x_i).$$

Note as well that the term in square brackets in (28) is obtained successively differentiating  $\psi(C(\mathbf{u}))$  with respect to  $u_1, \dots, u_M$ . Thus to obtain  $\psi$  from  $k(\mathbf{u}) C_M(\mathbf{u})$

all we have to do is to consider  $d_1 \cdots d_M \psi(C(\mathbf{u})) = k(\mathbf{u})C_M(\mathbf{u})$  as an equation to be solved for  $\psi$  by successive integrations and set

$$\psi(t) = \int_0^t \left( \int_0^1 \left( \int_0^1 \cdots \left( \int_0^1 k(\mathbf{u})C_M(\mathbf{u})du_M \right) du_{M-1} \right) \cdots \right) du_2 \Big) du_1.$$

When carrying out the integrations, one has to keep in mind that

$$C^{(M-k)}(u_1, \dots, u_{M-k}) = C(u_1, \dots, u_{M-k}, 1, 1, \dots, 1)$$

is an  $M - k$ -dimensional copula obtained by setting the last  $k$  components equal to 1, and that  $C^{(1)}(u_1) = u_1$ . We collect the previous remarks under the following theorem.

**Theorem 6** *With the notations just introduced, suppose that there exists a probability  $Q \sim P$  such that*

$$e^{-r} E_Q[S_i] \in (b_i, a_i), \quad \text{for } i = 1, \dots, M.$$

*Let  $G(\mathbf{x}) = Q(\mathbf{S} \leq \mathbf{x})$  and suppose that*

$$d^M G(\mathbf{x}) = k(F_1(x_1), \dots, F_M(x_M)) C_M(F_1(x_1), \dots, F_M(x_M)) \prod_{i=1}^M d_i F_i(x_i).$$

*Then there is an increasing function  $\psi(t)$  defined on  $[0, 1]$  by*

$$\psi(t) = \int_0^t \int_0^1 \int_0^1 \cdots \int_0^1 k(\mathbf{u})C_M(\mathbf{u})du_M du_{M-1} \cdots du_2 du_1,$$

*which necessarily satisfies  $\psi(0) = 0$  and  $\psi(1) = 1$  such that  $G(\mathbf{x}) = \psi(F(\mathbf{x}))$ , and therefore*

$$e^{-r} E^\psi[S_i] \in (b_i, a_i), \quad \text{for } i = 1, \dots, M.$$

**Acknowledgements** The authors gratefully acknowledge the fruitful comments of the anonymous reviewers that lead to an improvement of the original manuscript.

A. Arratia research is supported by grant TIN2017-89244-R from MINECO (Ministerio de Economía, Industria y Competitividad) and the recognition 2017SGR-856 (MACDA) from AGAUR (Generalitat de Catalunya).

## References

- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent Measures of Risk, *Mathematical Finance* **9**, 203–228.
- Borwein, J., Lewis, A. (2000). *Convex Analysis and Nonlinear Optimization*, CMS Springer.
- Cerny, A., Hodges, S. (2000). The Theory of Good Deal Pricing in Financial Markets. In: *H. Geman et al, Bachelier Congress, Mathematical Finance*.
- Dacunha-Castelle, D., Gamboa, F. (1990). Maximum d'entropie et problème des moments, *Annales de l'IHP Probabilités et Statistiques* **26**, 4, 567–596.
- Domahidi, A., Chu, E., Boyd, S. (2013). ECOS: An SOCP solver for embedded systems, *European Control Conference (ECC)*, 3071–3076.

- Föllmer, H., Schied, A. (2016). *Stochastic Finance: An introduction in discrete time (4th ed.)*, De Gruyter.
- Gzyl, H., Velásquez, Y. (2011). *Linear Inverse Problems: The Maximum Entropy Connection*, World Scient. Pubs, Singapore.
- Jaworski, P., Durante, F., Hardle, W. K., Rychlik, T. (2010). *Copula Theory and its Applications*, Springer.
- Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical Review* **106**, 4, 620.
- Karatzas, I., Shreve, S. (1998). *Methods of Mathematical Finance*, Springer.
- Laeven, R. J. A., Goovaerts, M. J. (2008). Premium calculation and insurance pricing, *Encyclopedia of Quantitative Risk Analysis and Assessment* **3**, 1302–1314.
- Madan, D. (2016). Conic portfolio theory. *Int. J. of Theor. and Applied Finance* **19**, 3.
- Madan, D., Schoutens, W. (2016). *Applied Conic Finance*, Cambridge Univ. Press.
- Markowitz, H. M. (1952). Portfolio Selection, *The Journal of Finance* **7**, 1, 77–91.
- Mead, L. R., Papanicolaou, N. (1984). Maximum entropy in the problem of moments, *Journal of Mathematical Physics* **25**, 8, 2404–2417.
- Pliska, S. R. (1997). *Introduction to Mathematical Finance*, Blackwell Pubs, Oxford.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.
- Wang, S., Young, V. R., Panjer, H. H. (1997). Axiomatic characterization of insurance prices, *Insurance: Mathematics and Economics* **21**, 2, 173–183.
- Wu, X., Wang, J. (2003). On characterization of the distortion premium principle, *Astin Bull.* **33**, 1–10.
- Young, V. (2006). *Premium Principles*, Encyclopedia of Actuarial Sciences, John Wiley.