# Portfolio Selection with Qualitative Input

Anant Chiarawongse[a], Seksan Kiatsupaibul[b,], Sunti Tirapat[a], Benjamin Van Roy[c]

[a]*Department of Banking and Finance, Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok 10330, Thailand*
[b]*Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok 10330, Thailand*
[c]*Department of Management Science and Engineering, Stanford University, CA 94305-4023, USA*

**Abstract**

We formulate a mean-variance portfolio selection problem that accommodates qualitative input about expected returns and provide an algorithm that solves the problem. This model and algorithm can be used, for example, when a portfolio manager determines that one industry will benefit more from a regulatory change than another but is unable to quantify the degree of difference. Qualitative views are expressed in terms of linear inequalities among expected returns. Our formulation builds on the Black-Litterman model for portfolio selection. The algorithm makes use of an adaptation of the hit-and-run method for Markov chain Monte Carlo simulation. We also present computational results that illustrate advantages of our approach over alternative heuristic methods for incorporating qualitative input.

*Keywords:* portfolio selection, Bayesian inference

*Email addresses:* `anant@acc.chula.ac.th` (Anant Chiarawongse), `seksan@acc.chula.ac.th` (Seksan Kiatsupaibul), `sunti@acc.chula.ac.th` (Sunti Tirapat), `bvr@stanford.edu` (Benjamin Van Roy)

## 1. Introduction

Portfolio managers acquire and process information from multiple sources in order to form expectations of future returns, also known as *alpha*. The Bayesian paradigm provides a coherent framework for synthesizing such information. In this spirit, Black and Litterman (1992) proposed an approach for combining multiple subjective views to generate alpha.

In the Black-Litterman model, each view is expressed in terms of a linear combination of alphas. For example, such a view may assert that the alpha of stock $A$ exceeds that of stock $B$ by $v$ percent. A view in the Black-Litterman model may also prescribe uncertainty to the estimate, for example, by stating that the alpha of stock $A$ exceeds that of stock $B$ by *approximately* $v$. In this case, the view must assign a level of variance $\sigma^2$ to quantify uncertainty around the estimate $v$.

Views of the kind we have described are *quantitative* in that they articulate numerical estimates. In this paper, we propose a methodology that also accommodates *qualitative* views that are expressed in terms of linear inequalities. As an example, consider a view that the alpha of stock $A$ exceeds that of stock $B$. This kind of view does not provide any quantitative estimate of degree but only ranks the alphas. Therefore, this model broadens the range of problems that can be solved by Bayesian methodology.

Linear inequalities can be used to express a wide variety of qualitative views. Let us consider a few examples involving different asset classes:

1. Equity Analysis: In equity analysis, the current condition and future prospects of multiple sectors of the economy are assessed and compared. Views that certain sectors are likely to outperform or to underperform

others can be encoded in terms of linear inequalities among stock alphas and used to guide sector rotation strategies.

2. Fixed Income Analysis: When the economy is expected to perform well, the credit spread between corporate bonds and government bonds should become narrower. As a result, corporate bonds should outperform government bonds. Such a view can be represented in terms of a linear inequality involving bond alphas.

3. Currency Analysis: Changes in relative economic conditions and regulations cause the value of some currencies to change relative to others. For example, rising inflation should lead to depreciation in foreign exchange rates. Such a view can be captured using linear inequalities among foreign exchange rates.

The work of Black and Litterman (1992) offers an approach to incorporating external information, i.e. investors' views of the future returns, into parameter estimation. An investor's view provides information in addition to the prior and observed historical returns. It can be acquired through analyses of external information and fused into a quantitative view, which is in the form of a system of linear equations of subjective forecast of returns and noises. However, a quantitative view demands details that are difficult to obtain. A qualitative view in the form of a system of linear inequalities, such as ranking or partial ranking is more natural to acquire, given that the information is not a direct observation of returns. In addition, the posterior distribution of the recent Bayesian estimations can be incorporated as the prior to the proposed model.

The ability to factor qualitative views into portfolio decisions also opens

the door to leveraging the growing body of work on robust comparative statics (e.g. Veinott, 1992; Vives, 1990; Topkis, 1978; Milgrom and Roberts, 1990b,a, 1994; Milgrom and Shannon, 1994). This line of economic research provides tools that can be used to draw qualitative conclusions about equilibria in games among multiple strategic agents even when parameters of the system are not known. Such tools are used, for example, to understand which companies or industries will benefit or be adversely affected by a regulatory change. Conclusions can be expressed in terms of linear inequalities and factored into portfolio decisions using our methodology.

In the next section we study a simple example involving a portfolio of two assets. The purpose is to illustrate our model and method in a simple context and to develop intuition for how qualitative input influences portfolio decisions. In this simple context, relevant calculations are amenable to closed form solution. Models involving many assets and multiple qualitative views, on the other hand, give rise to computational challenges.

The primary computational challenge in applying our approach involves computing the conditional expectation of returns, conditioned on qualitative views. This is generally a difficult inference problem, requiring integration of a normal distribution over a high-dimensional polytope. To deal with this problem, we propose an adaptation of the hit-and-run algorithm (e.g. Smith, 1984; Lovász, 1999; Lovász and Vempala, 2003, 2006b,a). One contribution of this paper is to demonstrate the efficacy of this algorithm in solving our portfolio optimization problem with qualitative input.

Our formulation generalizes the Black-Litterman model, which we review in Section 3. We then introduce in Section 4 our model, which incorporates

qualitative input. We present in Section 5 our solution method together with computational results that demonstrate its efficacy and merits relative to some alternative heuristic approaches that deal with qualitative input. Our methodology also allows for uncertainty in qualitative views. This captures, for example, a situation where the view is based on accurate information with some probability $\kappa$ and irrelevant information with some probability $1 - \kappa$. Here, $\kappa$ can be thought of as a level of confidence assigned to a qualitative view. We describe and study this model of uncertainty in Section 6.

## 2. A Simple Example

Suppose we wish to construct a portfolio $w \in \Re^2$ of two stocks. Each component is the dollar position in a stock. The return is given by $r^\mathsf{T} w$, where the vector $r$ of asset returns is generated according to

$$\mathbf{r} = \boldsymbol{\alpha} + \boldsymbol{\varepsilon},$$

with $\boldsymbol{\varepsilon} \sim N(0, I)$. Our prior beliefs about $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2]^\mathsf{T}$ take the form of a normal distribution $N(\bar{\alpha}, \tau I)$, for a vector $\bar{\alpha} = [-0.1, 0.1]^\mathsf{T}$ and a small positive scalar[1] $\tau \leq 0.1$. (As a convention, we use bold face to denote random quantities such as $\mathbf{x}$ and $\boldsymbol{\alpha}$ as opposed to their realizations $x$ and $\alpha$.)

It is natural to construct a portfolio by optimizing risk-adjusted expected return:

$$\max_{w \in \Re^2} \left( \bar{\alpha}^\mathsf{T} w - \frac{\rho}{2} w^\mathsf{T} w \right), \tag{1}$$

[1]From empirical Bayes' point of view, $\tau$ is the inverse of the number of observations in prior distribution construction. 10 observations imply $\tau = 0.1$. 100 observations imply $\tau = 0.01$.

where $\rho$ is a risk-aversion parameter. Note that we have ignored the risk introduced by our uncertainty about $\boldsymbol{\alpha}$. This is of negligible consequence because $\tau$ is small[2]. The solution to this portfolio optimization problem is given by

$$w = \frac{\bar{\alpha}}{\rho}.$$

Since $\bar{\alpha} = [-0.1, 0.1]^\mathsf{T}$, the optimal portfolio includes a short position in asset 1 and an equally sized long position in asset 2.

Now suppose we obtain a new piece of exogenous information, called a *view*, indicating that $\boldsymbol{\alpha}_1 \geq \boldsymbol{\alpha}_2$. The natural optimization problem then becomes

$$\max_{w \in \Re^2} \left( \hat{\alpha}^\mathsf{T} w - \frac{\rho}{2} w^\mathsf{T} w \right),$$

where $\hat{\alpha} = \mathbb{E}\left[\boldsymbol{\alpha} | \boldsymbol{\alpha}_1 \geq \boldsymbol{\alpha}_2\right]$. Computing the conditional expectation gives us $\hat{\alpha} = [a, \ -a]$, where

$$a = \frac{\bar{\alpha}_1 - \bar{\alpha}_2}{2} + \frac{1}{2}\sqrt{\frac{\tau}{\pi}} \frac{\exp(-(\bar{\alpha}_1 - \bar{\alpha}_2)^2/(4\tau))}{(1 - \Phi(-(\bar{\alpha}_1 - \bar{\alpha}_2)/\sqrt{2\tau}))},$$

and the optimal portfolio becomes

$$w = \frac{\hat{\alpha}}{\rho}.$$

Note that the conditional expected returns, and hence the portfolio composition, depend critically on $\tau$, the variance of the prior. Figure 1 illustrates how $\tau$ influences $a$, the magnitude of the conditional expected returns given the view.

[2]Note that the risk term technically should be $\frac{\rho(1+\tau)}{2} w^\mathsf{T} w$.

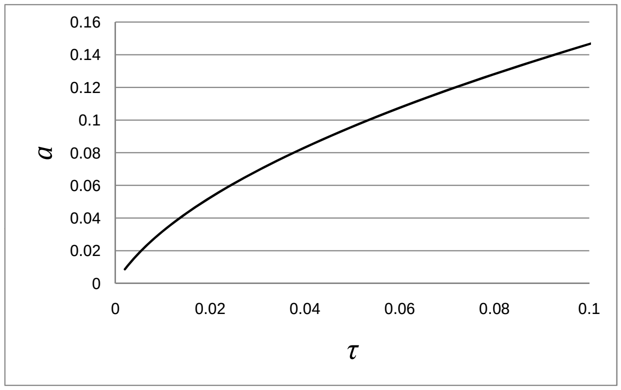


Figure 1: The conditional expected return given the view ($a$) versus the prior variance parameter ($\tau$)

As an example, consider the case where $\tau = 0.1$. Here, $\bar{\alpha} = [-0.1, 0.1]^\mathsf{T}$, $\hat{\alpha} = [0.147,\ -0.147]^\mathsf{T}$, and $w = [0.147/\rho,\ -0.147/\rho]^\mathsf{T}$. Note that this portfolio involves a long position in asset 1 and a short position in asset 2. This is opposite to the portfolio constructed in the absence of the qualitative view, which involved a short position in asset 1 and a long position in asset 2. The fact that these portfolios are diametrically opposed makes sense because while the prior suggests that asset 2 will outperform asset 1, the qualitative view rules this out.

Consider an alternative situation where the qualitative view takes the form of an inequality $\boldsymbol{\alpha}_1 \leq \boldsymbol{\alpha}_2$, reinforcing the ranking suggested by the prior. If $\tau = 0.1$, the conditional expectation of returns becomes $\hat{\alpha} = [-0.220,\ 0.220]^\mathsf{T}$.

Figure 2 illustrates the trade-offs between the expectation and variance of portfolio payoffs conditioned on three different information sets: 1) only the prior on $\boldsymbol{\alpha}$, 2) the opposing qualitative view $\boldsymbol{\alpha}_1 \leq \boldsymbol{\alpha}_2$, 3) the reinforcing

qualitative view $\boldsymbol{\alpha}_1 \geq \boldsymbol{\alpha}_2$. In this simple example, conditioning on either qualitative view improves the trade-off. As one would expect, a reinforcing view is preferable to an opposing view.

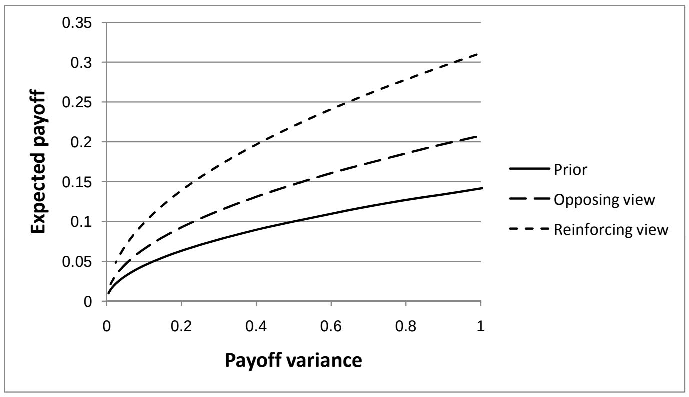


Figure 2: Trade-offs between the expectation and variance of portfolio payoffs with only the prior on $\boldsymbol{\alpha}$, with the opposing qualitative view $\boldsymbol{\alpha}_1 \leq \boldsymbol{\alpha}_2$, and with the reinforcing qualitative view $\boldsymbol{\alpha}_1 \geq \boldsymbol{\alpha}_2$

It is worth noting that in general qualitative views do not always improve the trade-off. Reinforcing views always improve the trade-off, but opposing views can either improve or hurt the trade-off. Figure 3 illustrates this. Two of the plots are identical to those of Figure 2; one based only on the prior on $\boldsymbol{\alpha}$ and the other involving conditioning on the opposing qualitative view $\boldsymbol{\alpha}_1 \leq \boldsymbol{\alpha}_2$. The other plots are also conditioned on this qualitative view, but are generated by reducing $\tau$. For smaller values of $\tau$, the trade-off deteriorates and becomes worse than the apparent trade-off in the absence of the qualitative view. Intuitively, this happens because a very small value of $\tau$ reflects high confidence in the prior expectation. In this case, the contradictory view results in a posterior expectation close to zero. When the stocks

generate no expected return, the trade-off becomes unattractive.

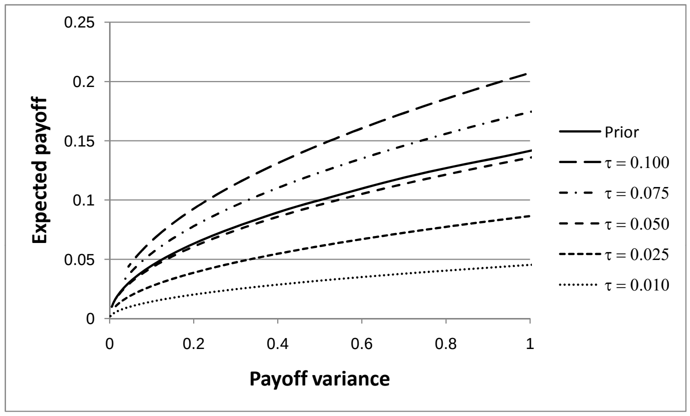


Figure 3: The trade-off between expected payoff and variance deteriorates as $\tau$ diminishes.

The above simple example illustrates that factoring a qualitative view into the estimation of alpha offers a potential benefit over the one without the view. One tends to expect a similar potential benefit in a more general setting involving multiple qualitative views and multiple assets. However, solving a portfolio selection problem with a more general information structure is not a trivial exercise. In fact, this can be accomplished by Markov chain Monte Carlo simulation, as we will discuss in Section 5. Before presenting our general model and the computational method in detail, we describe in the next section the Black-Litterman model, casting it in a Bayesian framework upon which our model is constructed.

## 3. The Black-Litterman Model

Black and Litterman (1991, 1992) provided a framework for synthesizing an investor's private views with a return distribution implied by market

equilibrium. Since its inception, the model has received significant attention among practitioners (Satchell and Scowcroft, 2000; Drobetz, 2001; Jones et al., 2007). We now review the Black-Litterman model, emphasizing its relation to a general Bayesian framework. This will facilitate comparison with our model for synthesizing qualitative views, which also fits in this framework.

The Black-Litterman model assigns a Gaussian prior to expected returns. This prior is either implied by a market equilibrium model or estimated from historical return data. Subjective views are represented as estimates of linear combinations of returns. Provided alongside each estimate is a variance parameter that reflects confidence in the estimate. A posterior distribution is computed and its mean is then used as input to a Markowitz-style mean-variance optimization problem.

Let $\boldsymbol{\alpha}$ denote a vector of expected returns for $N$ assets. Let the prior distribution of $\boldsymbol{\alpha}$ be a multivariate normal distribution with mean vector $\bar{\alpha}$ and covariance matrix $\tau\Sigma$, where $\tau$ is a scalar and $\Sigma$ is a positive definite matrix. That is,

$$\boldsymbol{\alpha} \sim N(\bar{\alpha},\ \tau\Sigma). \tag{2}$$

Given $\alpha$, the view $\mathbf{v}$ is a $K$-dimensional random vector with

$$\mathbf{v} \sim N(P\alpha,\ \Xi), \tag{3}$$

where $P$ is a $K \times N$ matrix. The posterior density of $\boldsymbol{\alpha}$, conditioned on the

view $v$, is then

$$\begin{aligned} f(\alpha \mid v) &\propto f(v \mid \alpha) f(\alpha) \\ &\propto \exp\left(-\frac{1}{2}(v - P\alpha)^\mathsf{T}\Xi^{-1}(v - P\alpha)\right)\exp\left(-\frac{1}{2}(\alpha - \bar{\alpha})^\mathsf{T}(\tau\Sigma)^{-1}(\alpha - \bar{\alpha})\right) \\ &\propto \exp\left(-\frac{1}{2}\left[(v - P\alpha)^\mathsf{T}\Xi^{-1}(v - P\alpha) + (\alpha - \bar{\alpha})^\mathsf{T}(\tau\Sigma)^{-1}(\alpha - \bar{\alpha})\right]\right). \end{aligned} \tag{4}$$

Observe that here $v$ plays the role of an observation in the Bayesian framework. Expression (4) implies that $\boldsymbol{\alpha}$ conditioned on the view $v$ is normally distributed. The conditional expectation of $\boldsymbol{\alpha}$ is given by a closed-form expression:

$$\mathbb{E}\left[\boldsymbol{\alpha} \mid v\right] = [P^\mathsf{T}\Xi^{-1}P + (\tau\Sigma)^{-1}]^{-1}[P^\mathsf{T}\Xi^{-1}v + (\tau\Sigma)^{-1}\bar{\alpha}] \tag{5}$$

This is the Black-Litterman expected return (Black and Litterman, 1992), which synthesizes subjective views with a prior.

## 4. Qualitative Input

Assume we have a Gaussian prior over expected returns as in Relation (2). We will consider qualitative views that can be expressed in terms of linear inequalities:

$$A\boldsymbol{\alpha} \leq b.$$

where $A$ is an $M \times N$ matrix and $b$ is an $M$-dimensional vector. A view that ranks assets provides one example that fits this framework. Such a view can be expressed as $N - 1$ linear inequalities of the form

$$\boldsymbol{\alpha}_{i_k} \geq \boldsymbol{\alpha}_{i_{k+1}},$$

for $k = 1, \ldots, N - 1$, where $(i_1, i_2, \ldots, i_N)$ is a permutation of $(1, \ldots, N)$.

We propose to use the conditional expectation

$$\hat{\alpha} = \mathbb{E}\left[\boldsymbol{\alpha} \mid A\boldsymbol{\alpha} \leq b\right]$$

in a Markowitz-style portfolio optimization problem. In particular, given this conditional expectation, the problem is to determine a vector of portfolio weights $w \in \Re^N$ that attains the maximum in

$$\max_{w \in \Re^N} \left(\hat{\alpha}^\mathsf{T} w - \frac{\rho}{2} w^\mathsf{T} \Sigma w\right).$$

This is entirely analogous to what is done in the Black-Litterman approach, except that we are conditioning on a different sort of information when computing $\hat{\alpha}$.

Let

$$\mathrm{R} = \{\alpha \in \Re^N : A\alpha \leq b\},$$

and let $I_\mathrm{R}$ be the indicator for the set R. The posterior density of returns, conditioned on qualitative views $A\boldsymbol{\alpha} \leq b$, is given by

$$f(\alpha \mid \boldsymbol{\alpha} \in \mathrm{R}) \propto I_\mathrm{R} \exp\left(-\frac{1}{2}\left[(\alpha - \bar{\alpha})^\mathsf{T}(\tau\Sigma)^{-1}(\alpha - \bar{\alpha})\right]\right). \tag{6}$$

The posterior expectation of returns is therefore

$$\begin{aligned} \hat{\alpha} &= \mathbb{E}\left[\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathrm{R}\right] \\ &= \frac{1}{K}\int_{x \in \mathrm{R}} x \exp\left(\frac{1}{2}(x - \bar{\alpha})^\mathsf{T}(\tau\Sigma)^{-1}(x - \bar{\alpha})\right) dx, \end{aligned} \tag{7}$$

where

$$K = \int_{x \in \mathrm{R}} \exp\left(\frac{1}{2}(x - \bar{\alpha})^\mathsf{T}(\tau\Sigma)^{-1}(x - \bar{\alpha})\right) dx.$$

The derivations of (6) and (7) can be found in the appendix.

The above integrals are over polytopes in a potentially high-dimensional space. As such, computing the estimate $\hat{\alpha}$ can be challenging. However, $\hat{\alpha}$ can be approximated by a Markov chain Monte Carlo method (MCMC) such as the hit-and-run algorithm (Smith, 1984; Bélisle et al., 1993), the Metropolis-Hasting algorithm, or the Gibbs sampler. In general, an MCMC generates a sequence of samples $\{\boldsymbol{\alpha}^i,\ i = 1, 2, \ldots\}$ that is an ergodic Markov chain with distribution converging in total variation to the posterior distribution $f(\cdot|\boldsymbol{\alpha} \in \mathrm{R})$. Furthermore,

$$\lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{T} \boldsymbol{\alpha}^i = \mathbb{E}\left[\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathrm{R}\right] \quad \text{a.s.}$$

Therefore, an average of MCMC samples can be employed as an estimate of the alpha. In Section 5, we devise an efficient MCMC based on the hit-and-run algorithm and demonstrate its efficacy.

## 5. Computational Study

In this section, we present results from a computational study aimed at assessing the merits of our approach to synthesizing qualitative input. This study compares our approach against simpler heuristics that offer alternative ways of dealing with qualitative input. Results from experiments involving randomly sampled problem instances indicate that our approach can yield substantial benefits over these heuristics.

### *5.1. Generative Model*

Our approach to estimating alpha requires the following problem parameters as input:

- number of assets $N$
- return variance $\Sigma$
- prior mean $\bar{\alpha}$
- prior variance scale parameter $\tau$
- constraint parameters $A$ and $b$

In this section, we present a generative model that will be used to sample the above problem parameters. Our generative model will additionally sample the realized alpha $\alpha$ for use in assessing the quality of portfolio decisions.

The generative model is itself parameterized by several inputs:

- number of assets $N$
- number of risk factors $M$
- variance of prior mean $\sigma_{\bar{\alpha}}^2$
- Wishart distribution variance parameter $\sigma_{\Sigma}^2$
- prior variance scale parameter $\tau$

Our generative model produces a covariance matrix $\Sigma = S + VDV^\mathsf{T}$, with a diagonal matrix $S \in \Re^{N \times N}$, $V \in \Re^{N \times M}$, and a diagonal matrix $D \in \Re^{M \times M}$, generated as follows. First, a matrix $\tilde{\Sigma}$ is sampled from a Wishart distribution $(\sigma_{\Sigma}^2 I, N)$. Then, a singular value decomposition $\tilde{\Sigma} = \tilde{V}\tilde{D}\tilde{V}^\mathsf{T}$ is computed, where $\tilde{D}$ contains the eigenvalues of $\tilde{\Sigma}$ sorted descendingly by magnitude along the diagonal. The matrix $V$ is taken to be the first $M$ columns of $\tilde{V}$, while the diagonal entries of $D$ are taken to be the first $M$ diagonal entries of $\tilde{D}$. Finally, for $n = 1, \ldots, N$, we let $S_{nn} = \tilde{\Sigma}_{nn} - \sum_{m=1}^{M} D_{mm} V_{nm}^2$. Note that

$\Sigma$ reflects risks associated with $M$ factors plus idiosyncratic asset-specific risks. The number of risk factors $M$ will typically be much smaller than $N$.

Our generative model produces the prior mean $\bar{\alpha}$ and realized alpha $\alpha$ as follows. First, $\bar{\alpha}$ is drawn from a $N(0, \sigma_{\bar{\alpha}}^2 I)$. Then, $\alpha$ is drawn from $N(\bar{\alpha}, \tau\Sigma)$.

For constraints, we will assume a particular kind of view that reflects a ranking of assets. In particular, let $i_1, \ldots, i_N$ be a permutation of $1, \ldots, N$ such that $\alpha_{i_1} \geq \alpha_{i_2} \geq \cdots \geq \alpha_{i_N}$. We consider $N-1$ constraints: $\alpha_{i_1} \geq \alpha_{i_2}, \alpha_{i_2} \geq \alpha_{i_3}, \ldots, \alpha_{i_{N-1}} \geq \alpha_{i_N}$. Such a complete ranking view may reflect practical contexts and also presents a challenging computational problem that serves to stress test our inference algorithm. In particular, as the number of assets increases, the number of constraints also increases, making the support of the posterior distribution increasingly complex. These constraints are encoded in terms of the matrix $A$ and vector $b$. Note that these constraints offer no information about alpha beyond restriction to a certain polytope. In particular, density of $\boldsymbol{\alpha}$ conditioned on observing such constraints is a multiple of the prior density for points that satisfy the constraints and zero for points that do not.

To the best of our knowledge, no prior work addresses how to incorporate such qualitative ranking information into a portfolio decision. With pre-existing methods, one can at best form a simple heuristic to adjust alpha estimates in response to ranking information. On the other hand, our proposed approach leads to coherent fusion of ranking information into the alpha estimation process. The estimated alpha is then used in a portfolio selection model to guide investment decisions. The estimation algorithm is introduced in the next section followed by computational results designed to

assess the performance of our approach relative to simple heuristics.

### *5.2. Solution Method*

Algorithm 1 computes a posterior expectation based on the model proposed in Section 4. Based on our computational experience, this algorithm efficiently processes asset rankings. It combines the hit-and-run algorithm with the Metropolis-Hasting algorithm and operates on polar coordinates. The algorithm takes as input parameters the number of assets $N$, the expectation $\bar{\alpha}$ and the variance $\tau\Sigma$ of the density $f$, parameters $A$ and $b$ that characterize the polytope R, an initial point $\alpha^0 \in$ R, and a number of iterations $n_{\text{iters}}$.

We will measure the efficiency of Algorithm 1, or PMHR, via Brooks and Gelman (1998)'s convergence criterion for an MCMC. For a Bayesian estimation problem, five sequences of samples are generated from the MCMC with five different starting points and are terminated when the estimates from the five sequences *converge*. The estimates from the five sequences are considered to converge when their *potential scale reduction factor* or *PSRF* is close to one. The number of iterations ($n_{\text{iters}}$) required until the five sequences converge determines the efficiency; the smaller the $n_{\text{iters}}$ the more efficient the algorithm.

To measure relative performance, we compare PMHR against the simple hit-and-run algorithm and the Gibbs sampler, two of the most commonly used MCMC methods in Bayesian inference literature. We try the three algorithms on problems generated by the generative model described above. For a number of assets $N$, 12 problems are generated. For each problem, we run the three algorithms to estimate the posterior expectation of the alpha

**Algorithm 1** Polar Metropolis Hit-and-Run (PMHR)

1: **for** $i = 0$ to $n_{\text{iters}}$ **do**

2: Let $\|\alpha^i\|$ be the length of $\alpha^i$ and define the current direction $\theta = \alpha^i/\|\alpha^i\|$. Let L be a line segment defined by

$$\mathrm{L} = \{l \ : \ l = \alpha^i + s\theta, s \in \Re\} \cap \mathrm{R},$$

and let S be the surface of the intersection between the hypersphere with radius $\|\alpha^i\|$ and the polytope R.

3: Randomly select to perform either an L move or an S move with equal probability.

L move: Sample $\alpha$ from the conditional density $f(\cdot \mid \boldsymbol{\alpha} \in \mathrm{L})$ and accept it as $\alpha^{(i+1)}$ with probability $\min\{1, (\|\alpha\|/\|\alpha^i\|)^{(N-1)}\}$. In case of rejection, let $\alpha^{(i+1)} = \alpha^i$.

S move: Sample $\alpha$ from the uniform distribution on $S$ and accept it as $\alpha^{(i+1)}$ with probability $\min\{1, f(\alpha)/f(\alpha^i)\}$. In case of rejection, let $\alpha^{(i+1)} = \alpha^i$.

4: **end for**

of the first asset. Each algorithm proceeds in batches of 1000 iterations and terminates when either the *PSRF* is less than 1.2 or 2500 batches have been reached. Upon termination, the terminal $n_{\text{iters}}$ in a unit of 1000 iterations is recorded. The average of the 12 $n_{\text{iters}}$ of each algorithm for each problem dimension $N$ is shown in Table 1. Observe that, at each $N$, PMHR offers an average $n_{\text{iters}}$ that is much smaller than those offered by the simple hit-and-run algorithm or the Gibbs sampler. On several problems, with $N$ equal to 50 and 100, both the simple hit-and-run and Gibbs sampler reach 2,500 batches before the convergence criteria are met. On the other hand, PMHR manages to converge on all problems.

Table 1: Relative efficiency between the simple hit-and-run, Gibbs sampler and PMHR as measured by the average number of iterations required for the algorithms to converge to the solution (unit in 1000 iterations) at different numbers of assets $N$.

| $N$ | Simple Hit-and-run | Gibbs | PMHR |
|---|---|---|---|
| 10 | 8.83 | 5.25 | 3.67 |
| 50 | 2411.83 | 243.25 | 3.83 |
| 100 | 2500.00 | 976.42 | 17.17 |

*5.3. Estimation of Alpha*

We will assess the advantages of our approach by comparing against simpler heuristics that factor in the side information of asset ranking. Each heuristic differs in the way it estimates alpha from the prior and side information. We now describe each method that we consider and how it estimates alpha based on data produced by the generative model:

1. **Bayesian estimation.** This is the approach we have proposed. To compute an estimate $\hat{\alpha} \approx \mathbb{E}[\boldsymbol{\alpha}|A\boldsymbol{\alpha} \leq b]$, we execute PMHR algorithm with $n_{\text{iters}}$ suggested by Table 1 and let $\hat{\alpha}$ be the sample average of iterates generated in the course of running the algorithm.

2. **Projection.** A simple alternative to Bayesian estimation is to project the prior mean $\bar{\alpha}$ onto the polytope defined by our constraints according to

$$\hat{\alpha} = \text{argmin}_{Aa \leq b}(a - \bar{\alpha})^{\mathsf{T}}\Sigma^{-1}(a - \bar{\alpha}).$$

   The resulting estimate $\hat{\alpha}$ is the mode of the posterior distribution used in our Bayesian estimation algorithm. The Bayesian estimation algorithm computes the mean, which poses a far greater computational challenge than computing the mode.

3. **Black-Litterman adaptation.** The Black-Litterman model does not accommodate qualitative views that rank assets. However, we will consider an adaptation that does. Observe that the asset ranking can be represented by $N-1$ inequalities of the form $\alpha_{i_n} - \alpha_{i_{n+1}} \geq 0$. Such an inequality can be rewritten as $\alpha_{i_n} - \alpha_{i_{n+1}} = v_n$ with $v_n \geq 0$. Consider now relaxing the inequality $v_n \geq 0$ and instead letting

$$v_n = \mathbb{E}[\boldsymbol{\alpha}_{i_n} - \boldsymbol{\alpha}_{i_{n+1}} \mid \boldsymbol{\alpha}_{i_n} - \boldsymbol{\alpha}_{i_{n+1}} \geq 0]. \tag{8}$$

   Further, let

$$\sigma_{v_n}^2 = \text{Var}[\boldsymbol{\alpha}_{i_n} - \boldsymbol{\alpha}_{i_{n+1}} \mid \boldsymbol{\alpha}_{i_n} - \boldsymbol{\alpha}_{i_{n+1}} \geq 0]. \tag{9}$$

   Now we can apply the Black-Litterman approach with $N-1$ views, each view $v_n$ representing an estimate of $\alpha_{i_n} - \alpha_{i_{n+1}}$ with variance $\sigma_{v_n}^2$,

leading to an alpha estimate $\hat{\alpha}$. In essence, this approach infers Black-Litterman's quantitative views from an asset ranking.

4. **Prior.** To provide a sanity check for other approaches, we consider the option of ignoring the views provided as side information and simply letting $\hat{\alpha}$ be equal to our prior mean $\bar{\alpha}$. We should expect this to perform worse than other methods that leverage the side information.
5. **Clairvoyance.** As an upper-bound on performance, we consider also the case where $\hat{\alpha}$ is taken to be the realized alpha $\alpha$. This is not a practical approach because in practice we would not know the value of $\alpha$. However, $\alpha$ is produced as a by-product from our generative model and we can use it in this context as a tool for our conceptual study.

### *5.4. Results*

We carried out a computational study with generative model inputs set to $M = 3$, $\sigma^2_{\bar{\alpha}} = 2.5 \times 10^{-7}$ , $\sigma^2_{\Sigma} = 10^{-3}$, and $\tau = 0.1$. We varied the number of assets $N$ across experiments. The mean-variance portfolio optimization was carried out with a risk aversion parameter $\rho = 4$. Algorithm 2 specifies the procedure used for evaluating any given method of estimating alpha. This procedure samples a hundred problem instances, and in each case computes the portfolio that would be selected and the certainty equivalent of the payoff given knowledge of $\alpha$. These values are averaged over the problem instances. We used common random numbers so that, for any fixed number of assets, the same problem instances were generated when evaluating each of the different methods for estimating $\alpha$.

This study is executed on a PC with an Intel Core i5 2.67GHz CPU and Windows 7 operating system. The procedure is written in and run on

**Algorithm 2** Evaluation procedure

**for** $i = 1$ to $100$ **do**

Sample $\Sigma$, $\bar{\alpha}$, $\alpha$, $A$, and $b$, based on the generative model

Compute alpha estimate $\hat{\alpha}$

Solve portfolio optimization problem

$$w = \underset{\sum_n w_n \leq 1}{\operatorname{argmax}} \left( \hat{\alpha}^{\mathsf{T}} w - \frac{\rho}{2} w^{\mathsf{T}} \Sigma w \right)$$

(Note that $w_0 = 1 - \sum_n w_n$ is assumed to be invested in a risk free asset with no expected return.)

Assess performance

$$u^i = \alpha^{\mathsf{T}} w - \frac{\rho}{2} w^{\mathsf{T}} \Sigma w$$

**end for**

Let $\hat{u} = (1/100) \sum_{i=1}^{100} u^i$

the R statistical software platform. Table 2 reports the compute time (in seconds) required for Bayesian estimation, the adapted Black-Litterman approach, and the projection method. Assuming clairvoyance or just using the prior eliminates the need for computational inference, an as such, we do not report compute times for these approaches. Bayesian estimation takes the longest time in average, while the adapted Black-Litterman and the projection method require very little time. Nevertheless, the time taken by the Bayesian approach is acceptable for practical use.

Table 2: Average computational time (in seconds) per problem required by the Bayesian estimation, the adapted Black-Litterman, and the projection method.

| $N$ | Bayesian | Black-Litterman | Projection |
|---|---|---|---|
| 10 | 1.65 | $<0.001$ | $<0.01$ |
| 50 | 3.62 | $<0.001$ | $<0.01$ |
| 100 | 52.50 | 0.001 | 0.02 |

Table 3 presents results from evaluating each method with 10, 50, and 100 assets. The columns are sorted so better performing methods appear to the left. As one would expect, clairvoyance leads to superior performance. After that, our Bayesian estimation approach works best, offering a certainty equivalent several times greater than that offered by the adapted Black-Litterman approach. The benefit of adopting the Bayesian approach over the adapted Black-Litterman approach, as measured by the ratio between the performance of the former and that of the latter, increases as the number of assets increases. The projection method looks far worse. As one would expect, ignoring side information and using only the prior leads to the worst

performance.

Table 3: Performance of alternative methods for estimating alpha measured by the certainty equivalent of the payoff.

| $N$ | Clairvoyance | Bayesian | Black-Litterman | Projection | Prior |
|---|---|---|---|---|---|
| 10 | 0.12850 | 0.09683 | 0.03369 | 0.00190 | 0.00004 |
| 50 | 0.64026 | 0.45056 | 0.05691 | 0.00197 | 0.00014 |
| 100 | 1.26021 | 0.76391 | 0.06435 | 0.00203 | 0.00013 |

Figure 4 plots the data from Table 3. These plots help to illustrate the large performance differences and how they grow with the number of assets in the portfolio. These results confirm the benefit of factoring in the qualitative view in terms of alpha ranking into the portfolio selection model. It is worth noting that this benefit results from that the ranking view is correct. In the next section, we further investigate the situation when there is uncertainty in the ranking view.

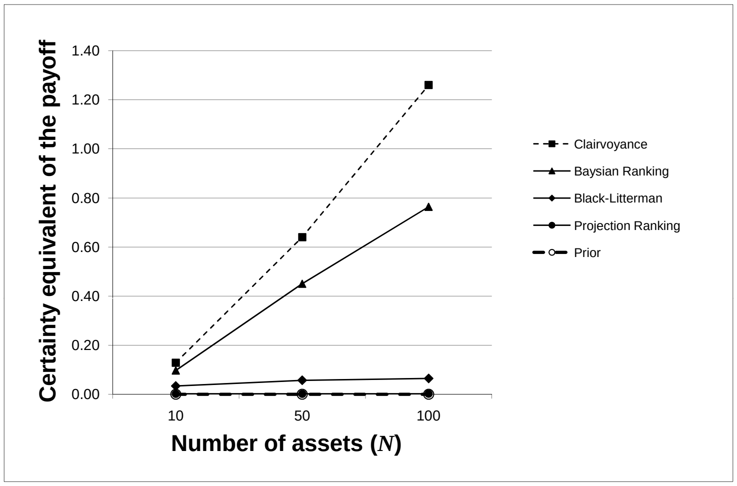


Figure 4: Certainty equivalent of payoff generated using (a) Clairvoyance, (b) Bayesian estimation, (c) the adapted Black-Litterman approach, (d) projection and (e) only prior information

## 6. Uncertain Qualitative Views

When an analyst offers a qualitative view but is uncertain about its validity, it is useful for the decision maker to be provided with a measure of confidence. This could take the form of a probability that the view is valid. In particular, an uncertain qualitative view may be represented in terms of a polytope

$$\mathrm{R} = \{\alpha \in \Re^N : A\alpha \leq b\}.$$

together with a probability $\kappa \in [0, 1]$.

To facilitate a coherent decision process we must have a precise way of interpreting the probability $\kappa$. For this purpose, we will treat the observed view $\mathbf{R}$ as a random variable. Letting R denote a realization, we define $\kappa$ as

follows

$$\kappa = \frac{\mathbb{P}\left(\boldsymbol{\alpha} \in X | \mathbf{R} = \mathrm{R}\right) - \mathbb{P}\left(\boldsymbol{\alpha} \in X\right)}{\mathbb{P}\left(\boldsymbol{\alpha} \in X | \boldsymbol{\alpha} \in \mathrm{R}\right) - \mathbb{P}\left(\boldsymbol{\alpha} \in X\right)}.$$

Our definition is best understood through an equivalent relation:

$$\mathbb{P}\left(\boldsymbol{\alpha} \in X | \mathbf{R} = \mathrm{R}\right) = \kappa \mathbb{P}\left(\boldsymbol{\alpha} \in X | \boldsymbol{\alpha} \in \mathrm{R}\right) + (1-\kappa)\mathbb{P}\left(\boldsymbol{\alpha} \in X\right). \tag{10}$$

To interpret this relation, first consider the case of $\kappa = 1$, where it reduces to $\mathbb{P}\left(\boldsymbol{\alpha} \in X | \mathbf{R} = \mathrm{R}\right) = \mathbb{P}\left(\boldsymbol{\alpha} \in X | \boldsymbol{\alpha} \in \mathrm{R}\right)$, which says that the probability distribution of $\boldsymbol{\alpha}$ conditioned on the observed view is equal to the distribution conditioned on the view being correct. This is equivalent to our earlier interpretation of the view in the absence of uncertainty. At the other extreme, when $\kappa = 0$, we have $\mathbb{P}\left(\boldsymbol{\alpha} \in X | \mathbf{R} = \mathrm{R}\right) = \mathbb{P}\left(\boldsymbol{\alpha} \in X\right)$. In this case, the view does not influence the conditional distribution of $\boldsymbol{\alpha}$, which remains equal to the prior distribution. When $\kappa \in (0,1)$, the conditional distribution lies between the extremes.

The following result provides a simple formula for computing the conditional expectation of $\boldsymbol{\alpha}$ given an uncertain qualitative view R and the associated degree of confidence $\kappa$. As in the previous sections, we let $\hat{\alpha} = \mathbb{E}\left[\boldsymbol{\alpha} | \boldsymbol{\alpha} \in \mathrm{R}\right]$. Further, we let $\tilde{\alpha} = \mathbb{E}\left[\boldsymbol{\alpha} \mid \mathbf{R} = \mathrm{R}\right]$.

**Theorem 1.** *For any view* R *with confidence* $\kappa$*,*

$$\tilde{\alpha} = \kappa\hat{\alpha} + (1-\kappa)\bar{\alpha}. \tag{11}$$

This result follows immediately from Equation (10).

Theorem 1 establishes that the Bayesian estimate with an uncertain qualitative view is a convex combination between the estimate that would be generated if the view were certain and the estimate that would be generated

in the absence of any view. This convex combination can be computed with virtually no effort beyond what is required in the case of a certain qualitative view.

To assess the influence of the confidence parameter $\kappa$ on the value of the view, we carried out experiments using data sampled in a way similar to that described in Section 5.1. In particular, to generate each realization, we simulate the generative model described in that section twice to obtain two independent data instances:

- base instance: $N$, $\Sigma$, $\bar{\alpha}$, $\alpha$, $A$, $b$
- auxiliary instance: $N$, $\Sigma$, $\bar{\alpha}$, $\alpha^\dagger$, $A^\dagger$, $b^\dagger$

From this data, we sample a realization. With probability $1 - \kappa$ this realization is the base instance. With probability $\kappa$ this realization is the base instance but with the view parameters $A$ and $b$ replaced by $A^\dagger$ and $b^\dagger$. It is easy to see that a realization generated in this way satisfies Equation (10).

To evaluate the effect of the confidence level $\kappa$ on the performance of the Bayesian inequalities model for portfolio optimization problem, we perform experiments on a special case when the side information we obtain is a complete ranking as in Section 5. We set up the simulation environment as that in Section 5 and modify Step 2, the estimation step, as follows.

Given the simulated instance, we obtain the estimate $\tilde{\alpha}$ by first computing $\hat{\alpha}$ using the hit-and-run algorithm as described earlier. The resulting estimate $\hat{\alpha}$ represents what the expectation would be if we were certain about the view. To obtain the expectation $\tilde{\alpha}$ given the uncertainty, we compute the convex combination $\tilde{\alpha} = \kappa\hat{\alpha} + (1 - \kappa)\bar{\alpha}$.

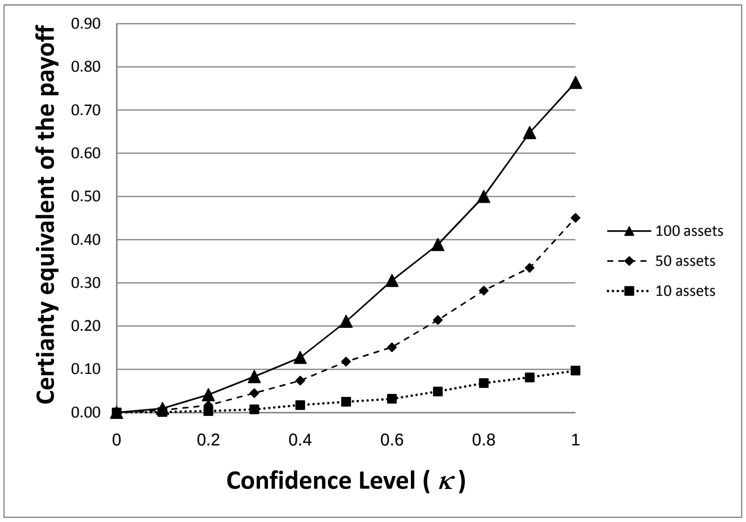


Figure 5: Certainty equivalent of the payoff as a function of the confidence level $\kappa$

We simulated and evaluated the certainty equivalents of the payoffs resulting from our estimation procedure for simulated instances using the same generative model parameter values as in Section 5.4. and with numbers of assets $N$ and values of $\kappa$. Results are plotted in Figure 5. When $\kappa = 1$, results are based on our expectation $\hat{\alpha}$ which assigns certainty to the view. When $\kappa = 0$, results are identical to those obtained by our prior expectation $\bar{\alpha}$. For intermediate values, performance gains from incorporating the view increases with $\kappa$. The slope also appears to increase, which indicates that performance is most sensitive to $\kappa$ when $\kappa$ is close to one.

## 7. Conclusion

We have introduced a new approach for incorporating qualitative views, taking the form of linear inequalities, into a mean-variance portfolio optimization framework. This approach involves computing the expectation of

alpha conditioned on qualitative views, which can be provided together with a degree of confidence. Computing the conditional expectation is a challenging problem, but we have devised an algorithm that addresses this problem effectively. We have compared the approach to other natural heuristics that may be used to adjust alpha based on qualitative views, and have found that our Bayesian approach offers substantial gains over the alternatives. In conclusion, our framework provides an optimal mechanism to utilize this form of information that is feasible to attain in practice.

## Appendix A. Derivations of Equations (6) and (7)

Let

$$\mathrm{R} = \{\alpha \in \Re^N : A\alpha \leq b\},$$

and let $I_{\mathrm{R}}$ be the indicator for the set R. The posterior density of returns, conditioned on qualitative views $A\boldsymbol{\alpha} \leq b$, is given by

$$\begin{aligned}
f(\alpha \mid \boldsymbol{\alpha} \in \mathrm{R}) &= \begin{cases} \frac{f(\alpha)}{\mathbb{P}(\boldsymbol{\alpha} \in \mathrm{R})}, & \text{if } \alpha \in \mathrm{R}, \\ 0, & \text{otherwise} \end{cases}, \\
&= \frac{I_{\mathrm{R}} f(\alpha)}{\mathbb{P}(\boldsymbol{\alpha} \in \mathrm{R})}, \\
&= \frac{I_{\mathrm{R}} (2\pi)^{-N/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}\left[(\alpha - \bar{\alpha})^{\mathsf{T}} (\tau\Sigma)^{-1} (\alpha - \bar{\alpha})\right]\right)}{\int_{x \in \mathrm{R}} (2\pi)^{-N/2} |\Sigma|^{-1/2} \exp\left(\frac{1}{2}(x - \bar{\alpha})^{\mathsf{T}} (\tau\Sigma)^{-1} (x - \bar{\alpha})\right) dx}, \\
&= \frac{I_{\mathrm{R}} \exp\left(-\frac{1}{2}\left[(\alpha - \bar{\alpha})^{\mathsf{T}} (\tau\Sigma)^{-1} (\alpha - \bar{\alpha})\right]\right)}{K},
\end{aligned}$$

where $K = \int_{x \in \mathrm{R}} \exp\left(\frac{1}{2}(x - \bar{\alpha})^{\mathsf{T}} (\tau\Sigma)^{-1} (x - \bar{\alpha})\right) dx$. Hence,

$$f(\alpha \mid \boldsymbol{\alpha} \in \mathrm{R}) \propto I_{\mathrm{R}} \exp\left(-\frac{1}{2}\left[(\alpha - \bar{\alpha})^{\mathsf{T}} (\tau\Sigma)^{-1} (\alpha - \bar{\alpha})\right]\right).$$

The posterior expectation of returns is therefore

$$\begin{aligned}\hat{\alpha} &= \mathbb{E}\left[\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathrm{R}\right] \\ &= \int x f(x \mid \boldsymbol{\alpha} \in \mathrm{R}) dx, \\ &= \frac{1}{K} \int_{x \in \mathrm{R}} x \exp\left(\frac{1}{2}(x - \bar{\alpha})^{\mathsf{T}} (\tau \Sigma)^{-1} (x - \bar{\alpha})\right) dx.\end{aligned}$$

**References**


Bélisle, C. J. P., Romeijn, H. E., Smith, R. L., 1993. Hit-and-Run algorithm for generating multivariate distributions. Mathematics of Operations Research 18, 255–266.

Black, F., Litterman, R., June 1991. Asset allocation: Combining investors views with market equilibrium. Journal of Fixed Income 1 (1), 7–18.

Black, F., Litterman, R., September-October 1992. Global portfolio optimization. Financial Analalysts Journal, 28–43.

Brooks, S. P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7 (4), 434–455.

Drobetz, W., 2001. How to avoid the pitfalls in portfolio optimization? putting the Black-Litterman approach at work. Journal of Financial Markets and Portfolio Management 15 (1), 59–75.

Jones, R., Lim, T., Zangari, P. J., Winter 2007. The Black-Litterman model for structured equity portfolios. Journal of Portfolio Management 33 (2), 24–33.

Lovász, L., 1999. Hit-and-run mixes fast. Mathematical Programming 86 (3), 443–461.

Lovász, L., Vempala, S. S., 2003. Hit-and-run is fast and fun. Tech. Rep. MSR-TR-2003-05, Microsoft Research.

Lovász, L., Vempala, S. S., 2006a. Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. In: Proceedings of the 47th IEEE Symposium on Foundations of Computer Science (FOCS '06). pp. 57–68.

Lovász, L., Vempala, S. S., 2006b. Hit-and-run from a corner. SIAM Journal of Computing 35 (4), 985–1005.

Milgrom, P., Roberts, J., 1990a. The economics of modern manufacturing: Technology, strategy, and organization. American Economic Review 80 (3), 511–528.

Milgrom, P., Roberts, J., 1990b. Rationalizability, learning, and equilibrium in games with strategic complementarities. Econometrica 58 (6), 1255–1277.

Milgrom, P., Roberts, J., 1994. Comparing equilibria. American Economic Review 84 (3), 441–459.

Milgrom, P., Shannon, C., 1994. Monotone comparative statics. Econometrica 62 (1), 157–180.

Satchell, S., Scowcroft, A., 2000. A demystification of the Black-Litterman

model: Managing quantitative and traditional portfolio construction. Journal of Asset Management 1 (2), 138–150.

Smith, R. L., 1984. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded convex regions. Operations Research 32 (6), 1296–1308.

Topkis, D., 1978. Minimizing a submodular function on a lattice. Operations Research 26 (2), 305–321.

Veinott, A. F., 1992. Lattice programming: Qualitative optimization and equilibria, course notes.

Vives, X., 1990. Nash equilibrium with strategic complementarities. Journal of Mathematical Economics 19, 305–321.