Measuring Portfolio Diversification Based on Optimized Uncorrelated Factors

Attilio Meucci
attilio.meucci@symmys.com

Alberto Santangelo
alberto.santangelo@hotmail.com

Romain Deguest
romain.deguest@edhec-risk.com

this revision: 09 August 2013
latest revision and code: symmys.com/node/599

Abstract[1]

We improve the "Effective Number of Bets" methodology in [Meucci, 2009a] (where the uncorrelated effective bets are the principal components of the market), using the Minimum-Torsion Bets, namely a set of uncorrelated factors, optimized to closely track the factors used to allocate the portfolio. This way we introduce a novel notion of "absolute risk contributions", which generalizes the "marginal contributions to risk" in traditional risk parity. We discuss the advantages of the Minimum-Torsion Bets over the traditional approach to diversification based on marginal contributions to risk. We present a case study in the S&P 500.

Fully documented code is available at symmys.com/node/599.

JEL Classification: C1, G11

Keywords: Effective Number of Bets, PCA, Diversification Distribution, marginal risk contributions, Procrustes problem

[1] The authors are grateful to Marcello Colasante, Bruno Dupire, Yashin Gopi, Lionel Martellini and Thierry Roncalli

1 Introduction

In recent years the practitioners and academic financial community has witnessed a surge in interest in the concept of risk parity, as well as the broader concept of diversification management, see [Roncalli, 2013] for a review and references.

In traditional risk parity, diversification is measured in terms of marginal risk contributions from each individual risk factor. Such contributions are spurious, because in reality they contain effects from all the factors at once. Furthermore, there exist no clear metric to quantify the diversification represented by the marginal risk contributions.

In this article we propose an alternative approach to risk parity based on the Effective Number of Bets in [Meucci, 2009a]: instead of the marginal contributions from correlated factors, we measure the true contributions from uncorrelated bets. Then the Effective Number of Bets precisely quantify the diversification level, summarizing in one number the fine structure of diversification contained in the set of uncorrelated bets in our portfolio.

In the original paper, the uncorrelated bets are the market's principal components. The Principal Components Bets have spurred interest and called for extensive empirical analysis, see [Frahm and Wiechers, 2011], [DFine, 2011], [Lohre et al., 2011], [Lohre et al., 2012], [Deguest et al., 2013]. However, the principal components are suboptimal, because they are purely statistical entities, not related to the investment process.

In this article we introduce a natural set of uncorrelated bets to manage diversification, the Minimum-Torsion Bets, which are the uncorrelated factors closest ("minimum-torsion") to the factors used by the portfolio manager. The contributions to risk from the Minimum-Torsion Bets constitutes a generalization of the marginal contributions to risk used in traditional risk parity.

The remainder of the paper is organized as follows: in Section 2 we revisit the general Effective Number of Bets framework in the context of factor-based risk parity and diversification management. In Section 3 we review the suboptimal implementation of the Effective Number of Bets, namely when the bets are represented by the principal components. In Section 4 we introduce the natural implementation of the Effective Number of Bets, namely when the bets are the Minimum-Torsion Bets. In Section 5 we highlight the key advantages of our approach to risk parity, based on Effective Number of Minimum-Torsion Bets, and the traditional approach to risk parity, based on marginal contributions to risk. In Section 6 we test our approach in a practical case study in the S&P 500. In Section 7 we conclude. In the Appendix A we detail all the technical proofs.

2 Effective Number of Bets

Here we review the Effective Number of Bets approach in [Meucci, 2009a], using a notation more suitable for the generalizations to follow. Refer to the original paper for all the details.

Consider an arbitrary portfolio which gives rise to a yet to be realized projected return R. In asset-based portfolio management, a portfolio is a combination of $\bar{n}$ correlated assets (stocks, options, bonds, futures, ...), and the portfolio return is a weighted average of the return of each asset $R = \sum_{n=1}^{\bar{n}} w_n R_n$, where w_n represent the weight of the n-th asset in the portfolio.

More in general, in factor-based portfolio management, and in factor-based risk parity, a portfolio is a combination of $\bar{n}$ *correlated* factors (momentum, value, etc.), and the portfolio return is a combination of the factor returns

$$R = \sum_{n=1}^{\bar{n}} b_n F_n, \tag{1}$$

where b_n represent the exposure of the portfolio to the n-th factor. Clearly, asset-based portfolio management represents a special case of factor-based management (1), where the factors are the asset returns $F_n = R_n$ and the exposures are the portfolio weights $b_n = w_n$.

Let us assume for now an ideal, apparently non-realistic scenario, where we can express the portfolio return as a combination of $\bar{n}$ *Bets*, or *uncorrelated* factors

$$R = \sum_{n=1}^{\bar{n}} \mathring{b}_n \mathring{F}_n, \tag{2}$$

where $\mathbb{C}r\{\mathring{F}_n, \mathring{F}_m\} = 0$ if $n \neq m$. Then, we can compute the *Diversification Distribution*, namely true relative contributions to total risk from each bet

$$p_n \equiv \frac{\mathbb{V}\{\mathring{b}_n \mathring{F}_n\}}{\mathbb{V}\{R\}}, \qquad n = 1, \ldots, \bar{n}, \tag{3}$$

where $\mathbb{V}$ denotes the variance. Notice that, as for any distribution, the masses p_n sum to one and are non-negative

$$\sum_{n=1}^{\bar{n}} p_n = 1, \qquad p_n \geq 0, \quad n = 1, \ldots, \bar{n}, \tag{4}$$

The Diversification Distribution (3) provides a detailed picture of the portfolio concentration structure. A portfolio is well diversified among the $\bar{n}$ factors, and thus achieves risk parity, if the masses $p_1, \ldots, p_{\bar{n}}$ are equal, or equivalently if the Diversification Distribution is uniform. To quantify diversification precisely, we use the exponential of the entropy, a tool from information theory, that measures the uniformity of a distribution. Accordingly, we define the *Effective Number of Bets* as follows

$$\mathbb{N} \equiv e^{-\sum_{n=1}^{\bar{n}} p_n \ln p_n}. \tag{5}$$

In the case of full concentration, i.e. when all the risk loads on one single factor, that factor's risk contribution (3) is 1, while all the other contributions are 0, and therefore $\mathbb{N} = 1$, the lowest possible value. At the opposite extreme, in the case of full diversification, the contributions to risk (3) from all the factors are equal, and $\mathbb{N} = \bar{n}$, the maximum possible value. For the intermediate cases $1 \leq \mathbb{N} \leq \bar{n}$.

3 Principal Components bets

To measure diversification via the Effective Number of Bets, we need to express our portfolio returns as a combination of uncorrelated terms, as in (2). To do so, one option is to use the principal components of the original factors $\boldsymbol{F}$ in (1), as suggested in [Meucci, 2009a].

Accordingly, we compute the covariance matrix of the factors $\boldsymbol{\Sigma}_F \equiv \mathbb{C}v\{\boldsymbol{F}\}$. Next, we perform the principal component decomposition $\boldsymbol{e}\boldsymbol{\lambda}^2\boldsymbol{e}' = \boldsymbol{\Sigma}_F$, where $\boldsymbol{\lambda}$ is the diagonal matrix of the singular values (square-root of eigenvalues) of $\boldsymbol{\Sigma}_F$ and $\boldsymbol{e}$ is the matrix whose columns are the eigenvectors of $\boldsymbol{\Sigma}_F$, which are orthogonal, and are normalized with length one $\boldsymbol{e}\boldsymbol{e}' = \boldsymbol{e}'\boldsymbol{e} = \boldsymbol{I}$, the identity matrix.

The eigenvectors generate $\bar{n}$ uncorrelated factors $\mathring{\boldsymbol{F}}_{PC}$ and $\bar{n}$ new portfolio exposures $\mathring{\boldsymbol{b}}_{PC}$, which allow us to express the portfolio return (1) in the uncorrelated format (2), as follows

$$\mathring{\boldsymbol{F}}_{PC} \equiv \boldsymbol{e}'\boldsymbol{F}, \qquad \mathring{\boldsymbol{b}}_{PC} \equiv \boldsymbol{e}'\boldsymbol{b}, \qquad R = \boldsymbol{b}'\boldsymbol{F} = \mathring{\boldsymbol{b}}'_{\boldsymbol{PC}}\mathring{\boldsymbol{F}}_{PC}, \tag{6}$$

where the last equality follows from $\boldsymbol{b}'\boldsymbol{F} = \boldsymbol{b}'\boldsymbol{e}\boldsymbol{e}'\boldsymbol{F}$.

Then given the exposures $\boldsymbol{b}$, we can compute the Effective Number of Bets (5)

$$\boldsymbol{p}_{PC}(\boldsymbol{b}) = \frac{(\boldsymbol{e}'\boldsymbol{b}) \circ (\boldsymbol{e}'\boldsymbol{\Sigma}_F\boldsymbol{b})}{\boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}} \quad \Rightarrow \quad \mathbb{N}_{PC}(\boldsymbol{b}) = e^{-\boldsymbol{p}_{PC}(\boldsymbol{b})' \ln \boldsymbol{p}_{PC}(\boldsymbol{b})}, \tag{7}$$

where $\circ$ denotes the term-by-term product, see Appendix A.1.

The principal components approach provides a formal set of uncorrelated factors from which to compute the Effective Number of Bets. However, it presents several problems.

First, the principal components bets tend to be statistically unstable, especially those relative to the lowest eigenvalues.

Second, the principal components bets are not invariant under simple scale transformations. To illustrate, suppose that we want to measure the diversification of the portfolio P&L, which is the portfolio return normalized by a constant. Clearly the diversification of the P&L and of the return must be the same. However, the Effective Number of Bets based on principal components of the P&L and of the return are different. To illustrate with a second example, suppose that we wish to measure some returns in basis points, instead of percentage points. The ensuing Effective Number of principal components bets will change dramatically, which is unacceptable. Such issue could be addressed in part by performing PCA on the correlation matrix.

Third, the principal components bets are not unique. Indeed, as discussed in [Deguest et al., 2013], if $\boldsymbol{e}_n$ is one of the $\bar{n}$ eigenvectors, so is its opposite $-\boldsymbol{e}_n$, and thus there are basically $2^{\bar{n}}$ possible combinations of principal components bets.

Fourth, the principal components bets are in general not easy to interpret, and hence disconnected from the decision process. In particular, in a dynamic setting, the meaning of the PCA factors changes from one date to another date, except possibly for the very first few factors.

Fifth, the principal components bets give rise to counter-intuitive results.

Example 1 *To illustrate the counter-intuitive results obtained with the principal component framework, consider the equal-load (equal-weight, in asset-based allocation) portfolio* $\boldsymbol{b}_{eq} \equiv (\frac{1}{\bar{n}}, \ldots, \frac{1}{\bar{n}})'$. *Also, consider an idealized, though non-realistic, market where all the factors have equal volatility and equal, positive pair-wise correlation*

$$[\boldsymbol{\Sigma}_F]_{m,n} \equiv \mathbb{C}v\{F_m, F_n\} = \begin{cases} \rho\sigma^2 & \text{for all } m \neq n \\ \sigma^2 & \text{for } m = n \end{cases} \tag{8}$$

In such a homogeneous market, if the correlation $\rho > 0$ *is very small, we would expect the equal-load portfolio to be highly diversified, giving rise to a number of uncorrelated bets close to the number of factors* $\mathbb{N}_{PC}(\boldsymbol{b}_{eq}) \approx \bar{n}$. *Instead, the equal-load portfolio always displays maximum concentration, i.e. only one (!) bet*

$$\mathbb{N}_{PC}(\boldsymbol{b}_{eq}) = 1. \tag{9}$$

The counter-intuitive full-concentration effect (9) follows because the equal-load portfolio is fully exposed to the first principal component and not exposed to any other principal component, see the proof in Appendix A.2.

4 Minimum-Torsion Bets

The Effective Number of Bets approach to risk parity builds on the *uncorrelated* decomposition (2). Hence, unlike the standard approach to risk parity based on *marginal* contributions to risk, the Effective Number of Bets approach highlights the contributions from truly separate sources of risk.

However, if the uncorrelated portfolio decomposition (2) is achieved via the principal components bets (6), we obtain suboptimal results for the several reasons highlighted in Section 3.

Fortunately, the principal components bets are not the only zero-correlation transformation of the original factors $\boldsymbol{F}$ that allows to express the portfolio as in the uncorrelated decomposition (2). There exist several alternative linear transformations $\mathring{\boldsymbol{F}} = \mathring{\boldsymbol{t}}\boldsymbol{F}$, or torsions, of the original factors $\boldsymbol{F}$, that are uncorrelated, and that are represented by a suitable $\bar{n} \times \bar{n}$ decorrelating torsion matrix $\mathring{\boldsymbol{t}}$.

For instance, we could use the lower-triangular Cholesky decomposition $\boldsymbol{\Sigma}_F \equiv \mathbb{C}v\{\boldsymbol{F}\} \equiv \boldsymbol{l}\boldsymbol{l}'$, where $\mathring{\boldsymbol{t}} \equiv \boldsymbol{l}^{-1}$. Indeed $\mathbb{C}v\{\boldsymbol{l}^{-1}\boldsymbol{F}\} = \boldsymbol{l}^{-1}\boldsymbol{l}\boldsymbol{l}'\boldsymbol{l}^{-1\prime} = \boldsymbol{I}$, and thus $\mathring{\boldsymbol{F}} = \boldsymbol{l}^{-1}\boldsymbol{F}$ are uncorrelated. However, such transformations display the same problems as the principal component approach. Most notably, the resulting uncorrelated factors $\mathring{\boldsymbol{F}}$ are not interpretable, as in general they bear no relationship with the original factors $\boldsymbol{F}$ that are used to manage the portfolio.

Here, we propose a natural, interpretable definition for the de-correlating transformation and the resulting uncorrelated factors: we choose the *minimum torsion* linear transformation that least disrupts the original factors $\boldsymbol{F}$. More precisely, among all the torsions $\mathring{\boldsymbol{t}}$ that ensure that the new factors are uncorrelated, we select the one that minimizes the tracking error with respect to the original factors

$$\mathring{\boldsymbol{t}}_{MT} \equiv \underset{\mathbb{C}r\{\boldsymbol{tF}\}=\boldsymbol{I}_{\bar{n}\times\bar{n}}}{\operatorname{argmin}} NTE\{\boldsymbol{tF}, \boldsymbol{F}\}, \tag{10}$$

where NTE denotes the multi-entry normalized tracking error

$$NTE\{\boldsymbol{Z}, \boldsymbol{F}\} \equiv \sqrt{\tfrac{1}{\bar{n}}\sum_n \mathbb{V}\{\frac{Z_n - F_n}{\mathbb{S}d\{F_n\}}\}}. \tag{11}$$

The normalization allows us not to worry about non-homogenous factors $\boldsymbol{F}$ measured in completely different units, such as interest rates and implied volatilities.

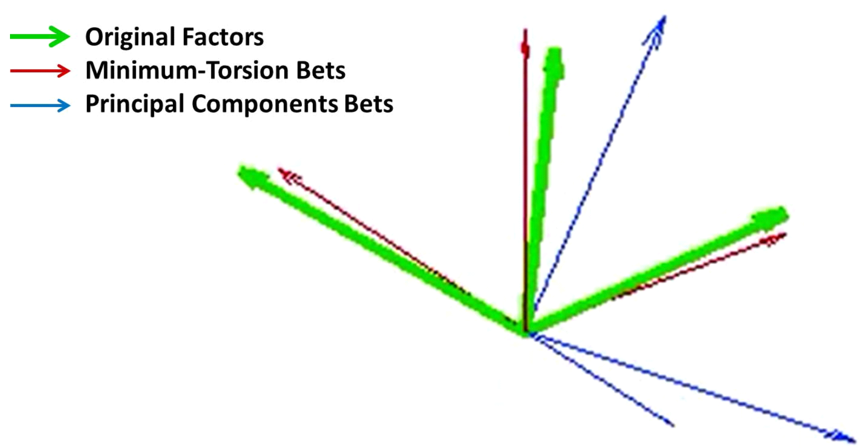

Figure 1: Minimum-Torsion Bets are the uncorrelated (orthogonal) factors closest to the original factors, Principal Component Bets are uncorrelated (orthogonal) factors with no clear connection to the original factors

Suppose that we can solve the minimum torsion optimization (10), which we do further below in this section. Then we introduce the *Minimum-Torsion Bets* $\mathring{\boldsymbol{F}}_{MT}$ and the respective *Minimum-Torsion Exposures* $\mathring{\boldsymbol{b}}_{MT}$, and use them to express the portfolio return (1) in the uncorrelated format (2), as follows

$$\mathring{\boldsymbol{F}}_{MT} \equiv \mathring{\boldsymbol{t}}_{MT}\boldsymbol{F}, \qquad \mathring{\boldsymbol{b}}_{MT} \equiv \mathring{\boldsymbol{t}}_{MT}^{\prime -1}\boldsymbol{b}, \qquad R = \mathring{\boldsymbol{b}}_{MT}^{\prime}\mathring{\boldsymbol{F}}_{MT}. \tag{12}$$

The interpretation of the Minimum-Torsion Bets $\mathring{\boldsymbol{F}}_{MT}$ is easily visualized geometrically in Figure 1, by equating factors to vectors, and no-correlation to orthogonality. Whereas the Minimum-Torsion Bets $\mathring{\boldsymbol{F}}_{MT}$ are the closest to the original factors $\boldsymbol{F}$ used to manage the portfolio, the PCA bets $\mathring{\boldsymbol{F}}_{PC}$ defined in (6) bear no close relationship with the management factors $\boldsymbol{F}$.

The Minimum-Torsion Bets and loadings (12) allow us to compute the Effective Number of Minimum-Torsion Bets

$$\boldsymbol{p}_{MT}(\boldsymbol{b}) = \frac{(\mathring{\boldsymbol{t}}_{MT}^{\prime -1}\boldsymbol{b}) \circ (\mathring{\boldsymbol{t}}_{MT}\boldsymbol{\Sigma}_F\boldsymbol{b})}{\boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}} \quad \Rightarrow \quad \mathbb{N}_{MT}(\boldsymbol{b}) = e^{-\boldsymbol{p}_{MT}(\boldsymbol{b})' \ln \boldsymbol{p}_{MT}(\boldsymbol{b})}, \tag{13}$$

see Appendix A.1.

The Minimum-Torsion Bets (12) address and solve all the problems of the principal components bets (6).

Example 2 *To provide a first taste of the intuitive nature of the Minimum-Torsion Bets, let us consider again the idealized homogeneous market with equal volatilities and arbitrary small homogeneous correlations (8). Furthermore, let us consider again the equal-load (equal-weight, in asset-based allocation) portfolio* $\boldsymbol{b}_{eq} \equiv (\frac{1}{\bar{n}}, \ldots, \frac{1}{\bar{n}})'$. *Unlike with Principal Component Bets (9), the Number of Minimum-Torsion Effective Bets (13) for the equal-load portfolio is, as intuition suggests, the largest possible*

$$\mathbb{N}_{MT}(\boldsymbol{b}_{eq}) = \bar{n}, \tag{14}$$

see Appendix A.2. This result is very intuitive: in an uncorrelated market each position of equal size represents a separate bet.

The solution of the minimum torsion optimization (10) is a special instance of a quadratically constrained quadratic program [W], related to the solution of the orthogonal Procrustes problem [W]. Adapting from [Everson, 1997], we obtain the solution by first computing analytically a starting guess, and then perturbing the starting guess via an efficient recursive algorithm, as follows.

Let us denote by $\boldsymbol{\sigma}_F$ the vector of the factors standard deviations, extracted with the correlation matrix $\boldsymbol{C}_F$ from the covariance matrix $\boldsymbol{\Sigma}_F \equiv \mathbb{C}v\{\boldsymbol{F}\}$, as follows

$$\boldsymbol{\Sigma}_F \equiv dg(\boldsymbol{\sigma}_F)\, \boldsymbol{C}_F \, dg(\boldsymbol{\sigma}_F). \tag{15}$$

Then, let us factor as in [Meucci, 2009b] the correlation matrix via its Riccati root $\boldsymbol{c}$, namely the symmetric matrix such that

$$\boldsymbol{C}_F = \boldsymbol{c}\boldsymbol{c}' = \boldsymbol{c}^2. \tag{16}$$

The Riccati root of the correlation matrix $\boldsymbol{C}_F$ is easily computed in terms of its PCA decomposition $\boldsymbol{C}_F = \boldsymbol{e}\boldsymbol{\lambda}^2\boldsymbol{e}'$ and reads $\boldsymbol{c} = \boldsymbol{e}\boldsymbol{\lambda}\boldsymbol{e}'$.

Next, we perform the principal component decomposition $= \boldsymbol{\Sigma}_F$, where $\boldsymbol{\lambda}$ is the diagonal matrix of the singular values (square-root of eigenvalues) of $\boldsymbol{\Sigma}_F$ and

$\boldsymbol{e}$ is the matrix whose columns are the eigenvectors of $\boldsymbol{\Sigma}_F$, which are orthogonal, and are normalized with length one $\boldsymbol{e}\boldsymbol{e}' = \boldsymbol{e}'\boldsymbol{e} = \boldsymbol{I}$, the identity matrix.

Next, let us compute a perturbation matrix $\boldsymbol{\pi}$ recursively with the algorithm below (refer to Appendix A.3.2 for the rationale)

$$\begin{array}{|ll|}\hline \multicolumn{2}{|c|}{\boldsymbol{\pi} = Perturb(\boldsymbol{c})} \\ \multicolumn{2}{|c|}{\text{Minimum-Torsion recursion}} \\ \hline \text{0. Initialize} & \boldsymbol{d} \leftarrow \boldsymbol{I} \\ \text{1. Riccati root} & \boldsymbol{u} \leftarrow (\boldsymbol{d}\boldsymbol{c}^2\boldsymbol{d})^{\frac{1}{2}} \\ \text{2. Rotation} & \boldsymbol{q} \leftarrow \boldsymbol{u}^{-1}\boldsymbol{d}\boldsymbol{c} \\ \text{3. Stretching} & \boldsymbol{d} \leftarrow dg(dg^{-1}(\boldsymbol{q}\boldsymbol{c})) \\ \text{4. Perturbation} & \boldsymbol{\pi} \leftarrow \boldsymbol{d}\boldsymbol{q} \\ \multicolumn{2}{|l|}{\text{5. If convergence, output } \boldsymbol{\pi}\text{; else go to 1}} \\ \hline \end{array} \tag{17}$$

where the operator $dg^{-1}(\boldsymbol{m})$ extracts the $\bar{n} \times 1$ vector on the principal diagonal of the $\bar{n} \times \bar{n}$ matrix $\boldsymbol{m}$; and the operator $dg(\boldsymbol{v})$ embeds the $\bar{n} \times 1$ vector $\boldsymbol{v}$ into the principal diagonal of a square matrix which is zero anywhere else.

Example 3 *For instance from a correlation*

$$\boldsymbol{C}_F = \begin{pmatrix} 1 & 0.5 & 0.3 \\ 0.5 & 1 & 0.5 \\ 0.3 & 0.1 & 1 \end{pmatrix}$$

we obtain

$$\boldsymbol{c} = \begin{pmatrix} 0.9544 & 0.2580 & 0.1503 \\ 0.2580 & 0.9656 & 0.0313 \\ 0.1503 & 0.0313 & 0.9881 \end{pmatrix}, \quad \boldsymbol{\pi} = \begin{pmatrix} 0.9535 & 0.0016 & 0.0026 \\ 0.0017 & 0.9661 & 0.0001 \\ 0.0027 & 0.0001 & 0.9886 \end{pmatrix}$$

The algorithm (17) converges extremely fast, within a dozen iterations, or fractions of a second, even when $\boldsymbol{\pi}$ has $\bar{n}^2 \approx 150{,}000$ entries (!), the dimensions required to handle the S&P case study in Section 6.

The solution of the minimum torsion problem (10) then reads

$$\mathring{\boldsymbol{t}}_{MT} = dg(\boldsymbol{\sigma}_F)\,\boldsymbol{\pi}\boldsymbol{c}^{-1}\,dg(\boldsymbol{\sigma}_F)^{-1}. \tag{18}$$

The minimum torsion transformation (18) defines the Minimum-Torsion Bets and the Minimum-Torsion Exposures, as in (12); and the Minimum-Torsion Diversification Distribution and the Effective Number of Minimum-Torsion Bets, as in (13).

If we switch off the numerical perturbation term, which amounts to setting $\boldsymbol{\pi} \equiv \boldsymbol{I}$ in the minimum torsion transformation (18), we obtain an approximate minimum-torsion

$$\mathring{\boldsymbol{t}} \equiv dg(\boldsymbol{\sigma}_F)\,\boldsymbol{c}^{-1}\,dg(\boldsymbol{\sigma}_F)^{-1} \approx \mathring{\boldsymbol{t}}_{MT}. \tag{19}$$

As we show in see Appendix A.3, the approximate minimum-torsion (19) solves the original minimum-torsion problem (10) under the additional constraint that the volatilities of the new factors be the same as the volatilities of the original factors

$$\mathbb{S}d\{\mathring{\boldsymbol{t}}\boldsymbol{F}\} = \boldsymbol{\sigma}_F. \tag{20}$$

The approximate minimum-torsion (19) is essentially the solution of the orthogonal Procrustes problem first derived by [Schoenemann, 1966]. Geometrically, the approximate minimum-torsion generates orthogonal bets $\mathring{\boldsymbol{t}}\boldsymbol{F}$ by rotating the original factors $\boldsymbol{F}$. However, we verified numerically that the discrepancy between the approximate minimum-torsion bets $\mathring{\boldsymbol{t}}\boldsymbol{F}$ and the true Minimum-Torsion Bets $\mathring{\boldsymbol{t}}_{MT}\boldsymbol{F}$ can become relevant in highly correlated markets.

5 Bets versus marginal contributions to risk

Here we reflect on the main conceptual differences between the traditional approach to risk parity, based on marginal contributions to risk, and the present approach, based on the Effective Number of Minimum-Torsion Bets, which we summarize in the table below

	Traditional	**Effective Number of Bets**
Risk contrib.	Marginal Contributions	Diversification Distributions
Expression	$\boldsymbol{m} \equiv \frac{\boldsymbol{b}\circ(\boldsymbol{\Sigma}_F\boldsymbol{b})}{\boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}}$	$\boldsymbol{p} \equiv \frac{(\mathring{\boldsymbol{t}}_{MT}'^{-1}\boldsymbol{b})\circ(\mathring{\boldsymbol{t}}_{MT}\boldsymbol{\Sigma}_F\boldsymbol{b})}{\boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}}$
Meaning	spurious contributions from original factors	proper contributions from Minimum-Torsion Bets
Properties	$\sum_n m_n = 1, \quad m_n \lesseqgtr 0$	$\sum_k p_n = 1, \quad p_n \geq 0$

(21)

The key to the traditional risk parity approach are the *Marginal Contributions to Risk*

$$\boldsymbol{m} \equiv \frac{\boldsymbol{b}\circ(\boldsymbol{\Sigma}_F\boldsymbol{b})}{\boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}}, \tag{22}$$

see e.g. [Roncalli, 2013]. In the traditional risk parity approach, a portfolio is diversified if the Marginal Contributions to Risk $\boldsymbol{m}$ are uniform. The key to our approach to risk parity based on Effective Number of Bets is the Diversification Distribution of the Minimum-Torsion bets $\boldsymbol{p}$, defined in (13): a portfolio is diversified if $\boldsymbol{p}$ is uniform.

The main difference, and at the same time the main weakness of the Marginal Contributions to Risk $\boldsymbol{m}$ with respect to the Diversification Distribution $\boldsymbol{p}$, is the fact that $\boldsymbol{m}$ represent differential contributions, i.e. they represent the sensitivity of risk to a small change in exposure to a given factor. Hence, unlike the Diversification Distribution $\boldsymbol{p}$, they do not represent the separate contributions to risk from a given factor.

Furthermore, the Marginal Contributions to Risk $\boldsymbol{m}$ sum to one, just like the Diversification Distribution $\boldsymbol{p}$. However, unlike the Diversification Distribution $\boldsymbol{p}$, the Marginal Contributions to Risk $\boldsymbol{m}$ are not necessarily positive, due to either negative correlations or the presence of negative exposures to factors.

As a result, when enforcing traditional risk parity, one focuses on the absolute values of the Marginal Contributions to Risk $|m_n|$, which no longer sum to one.

Finally, notice that if the factors $\boldsymbol{F}$ are uncorrelated, then the Minimum-Torsion transformation $\mathring{\boldsymbol{t}}_{MT}$ is the identity, and thus the Marginal Contributions to Risk and the Diversification Distribution are one and the same, $\boldsymbol{m} = \boldsymbol{p}$.

To summarize, both traditional risk parity and Effective Number of Minimum-Torsion Bets apply in full generality across markets with arbitrary factors. When the factors in the traditional risk parity approach are uncorrelated, the Marginal Contributions to Risk display all the palatable features of the Diversification Distribution, namely they sum to one, they are positive, and they are the true contributors to risk. In the more general case where the factors are correlated, the traditional risk parity approach incurs problems.

Hence, we can interpret the Effective Number of Minimum-Torsion Bets approach as a generalization of the traditional risk parity approach, which addresses the issues of the latter approach.

6 Case study

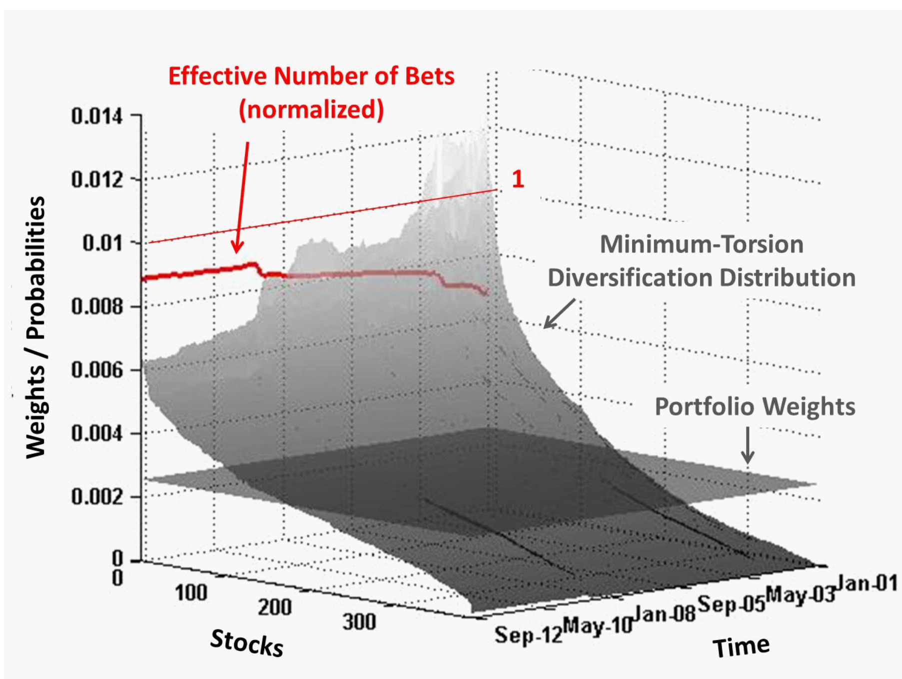

Figure 2: Diversification through time of an equal-weight portfolio of stocks

To illustrate, we consider an investment in $\bar{n} = 392$ stocks in the S&P 500 Index (for simplicity, we consider the stocks alive through the whole analysis period). In this case the factors $\boldsymbol{F}$ and respective exposures $\boldsymbol{b}$ in (1) are the $\bar{n}$ stock returns and the $\bar{n}$ portfolio weights respectively.

Let us consider a simple equal-load portfolio $\boldsymbol{b}_{eq} \equiv (\frac{1}{\bar{n}}, \ldots, \frac{1}{\bar{n}})'$. We estimate every month the covariance matrix of the of the S&P stock returns $\boldsymbol{\Sigma}_F$ using a one-year rolling window of daily observations, and filtering the smallest eigenvalues to ensure positive definiteness. Using the portfolio loadings and the returns covariance we compute the Minimum-Torsion Diversification Distribution $\boldsymbol{p}_{MT}(\boldsymbol{b}_{eq})$ and the Minimum-Torsion Effective Number of Bets $\mathbb{N}_{MT}(\boldsymbol{b}_{eq})$ of the equal-load portfolio, as in (13).

In Figure 2 we display the results. Each month, we sort the stocks in decreasing order of risk contribution, as measured by the Minimum-Torsion Diversification Distribution $\boldsymbol{p}_{MT}(\boldsymbol{b}_{eq})$. The fairly homogeneous structure of the S&P 500 stocks and the equal-load allocation provide a portfolio that is not as diversified as one would expect. Unlike the portfolio weights, the Minimum-Torsion Diversification Distribution is not flat: some stocks are riskier than others, a fact well known to portfolio managers.

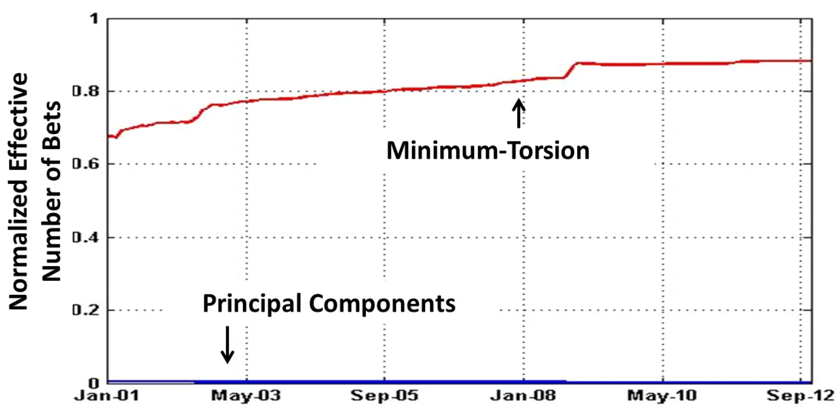

Figure 3: Intuitive diversification measure via Mimumum-Torsion versus counter-intuitive diversification measure via principal components

The fine structure of diversification represented by the Minimum-Torsion Diversification Distribution is summarized into the Minimum-Torsion Effective Number of Bets $\mathbb{N}_{PC}(\boldsymbol{b}_{eq})$. The Effective Number of Bets is strictly less than the number of stocks $\bar{n}$, because of the non-homogeneous contributions to risk from each stock. However, as intuition suggests, the Effective Number of Bets is of the order of a few hundreds, given that the Minimum-Torsion Diversification Distribution is not too steep:

$$0 \ll \frac{\mathbb{N}_{MT}(\boldsymbol{b}_{eq})}{\bar{n}} < 1. \tag{23}$$

The intuitive result (23) is the empirical counterpart of the similar theoretical full-diversification result (14).

For comparison, we also compute the principal component Diversification Distribution $\boldsymbol{p}_{PC}(\boldsymbol{b}_{eq})$ and the principal component Effective Number of Bets $\mathbb{N}_{PC}(\boldsymbol{b}_{eq})$ of the equal-load portfolio, as in (7). The Diversification Distribution is now too steep to display, loading basically in full on the first entry (which is no longer interpretable and not directly comparable with the portfolio weights). As a result, as we see in Figure 3, the normalized Effective Number of Bets is almost zero

$$0 \approx \frac{\mathbb{N}_{PC}(\boldsymbol{b}_{eq})}{\bar{n}} \ll 1. \tag{24}$$

The counter-intuitive result (24) is the empirical counterpart of the similar, extreme, theoretical full-concentration result (9).

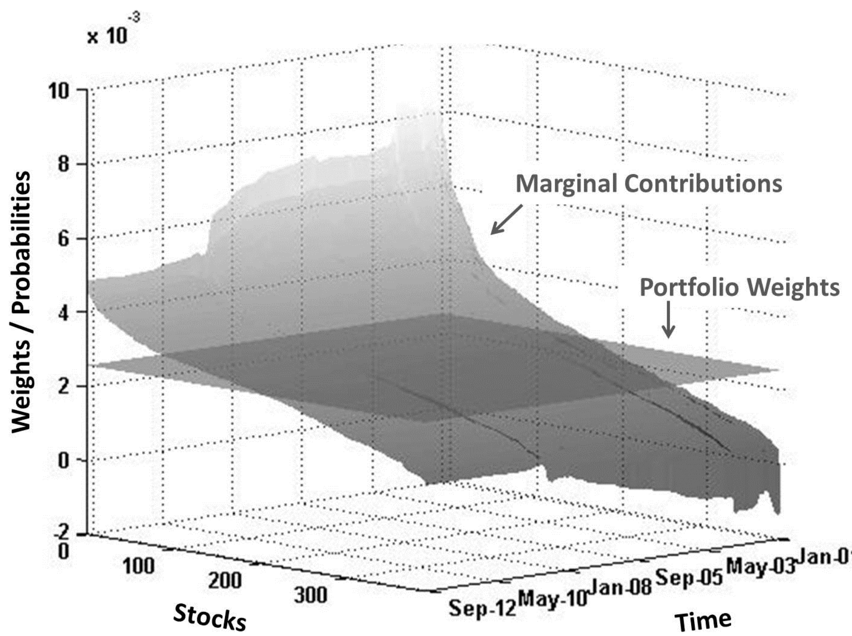

Figure 4: Marginal contributions to risk through time of an equal-weight portfolio of stocks

Finally, we compute the marginal risk contributions of the equal-load portfolio $\boldsymbol{m}(\boldsymbol{b}_{eq})$, as in (22) and display them in Figure 4. The overall risk profile is similar to the Diversification Distribution $\boldsymbol{p}_{MT}(\boldsymbol{b}_{eq})$ in Figure 2. However, even with such a homogeneous market and portfolio, we notice possibly spurious, sharp negative contributions at the beginning of the sample.

7 Conclusions

We have introduced the Minimum-Torsion Bets, the set of uncorrelated factors which closely tracks the factors used in the allocation process. With the Minimum-Torsion Bets we have given new life to the Effective Number of Bets approach to risk parity and, more in general, diversification management. Indeed, unlike the Principal Component Bets originally used to measure the Effective Number of Bets, the Minimum-Torsion Bets are easily interpretable and give rise to intuitive results.

We have highlighted the improvements of the Effective Number of Minimum-Torsion Bets over the standard approach to risk parity, which relies on marginal, rather than total, contributions to risk.

We have illustrated how to use the Effective Number of Minimum-Torsion Bets to measure the diversification of a portfolio of stocks in the S&P500.

References

[Deguest et al., 2013] Deguest, R., Martellini, L., and Meucci, A. (2013). Risk parity and beyond - from asset allocation to risk allocation decisions. *EDHEC-Risk Working Paper*.

[DFine, 2011] DFine, G. (2011). Asset allocation under a conditional diversification measure. https://www.maths.ox.ac.uk/system/files/attachments/300911.pdf.

[Everson, 1997] Everson, R. (1997). Orthogonal, but not orthonormal, procrustes problems.

[Frahm and Wiechers, 2011] Frahm, G. and Wiechers, C. (2011). On the diversification of portfolios of risky assets. http://hdl.handle.net/10419/45354.

[Lohre et al., 2011] Lohre, H., Neugebauer, U., and Zimmer, C. (2011). Diversifying risk parity. http://ssrn.com/abstract=1974446.

[Lohre et al., 2012] Lohre, H., Neugebauer, U., and Zimmer, C. (2012). Diversified risk parity strategies for equity portfolio selection.

[Meucci, 2009a] Meucci, A. (2009a). Managing diversification. http://symmys.com/node/199.

[Meucci, 2009b] Meucci, A. (2009b). Simulations with exact means and covariances. http://symmys.com/node/162.

[Roncalli, 2013] Roncalli, T. (2013). *Introduction to Risk Parity and Budgeting*. Chapman and Hall.

[Schoenemann, 1966] Schoenemann, P. H. (1966). A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31:1–10.

A Appendix

In this appendix we discuss technical results that can be skipped at first reading.

A.1 Explicit expressions for Effective Number of Bets

The denominator in the Diversification Distribution (3) follows from the expression of the portfolio return as a combination of factors (1) and reads

$$\mathbb{V}\{R\} = \mathbb{V}\{\boldsymbol{b}'\boldsymbol{F}\} = \boldsymbol{b}'\mathbb{V}\{\boldsymbol{F}\}\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}. \tag{25}$$

Assume that the Bets are an uncorrelated linear transformation of the factors

$$\mathring{\boldsymbol{F}} \equiv \boldsymbol{t}\boldsymbol{F}, \qquad \mathring{\boldsymbol{b}} \equiv \mathring{\boldsymbol{t}}'^{-1}\boldsymbol{b}. \tag{26}$$

Then numerator of the Diversification Distribution (3) reads

$$\begin{aligned}\mathbb{V}\{\mathring{\boldsymbol{b}} \circ \mathring{\boldsymbol{F}}\} &= \mathring{\boldsymbol{b}} \circ \mathbb{V}\{\mathring{\boldsymbol{F}}\} \circ \mathring{\boldsymbol{b}} = \mathring{\boldsymbol{b}} \circ (\mathbb{C}v\{\mathring{\boldsymbol{F}}\}\mathring{\boldsymbol{b}}) \\ &= (\mathring{\boldsymbol{t}}'^{-1}\boldsymbol{b}) \circ (\mathring{\boldsymbol{t}}\boldsymbol{\Sigma}_F\mathring{\boldsymbol{t}}'\mathring{\boldsymbol{t}}'^{-1}\boldsymbol{b}) = (\mathring{\boldsymbol{t}}'^{-1}\boldsymbol{b}) \circ (\mathring{\boldsymbol{t}}\boldsymbol{\Sigma}_F\boldsymbol{b})\end{aligned} \tag{27}$$

Dividing the numerator (27) by the denominator (25) we obtain the explicit formula Diversification Distribution (3), as follows

$$\boldsymbol{p} = \frac{(\mathring{\boldsymbol{t}}'^{-1}\boldsymbol{b}) \circ (\mathring{\boldsymbol{t}}\boldsymbol{\Sigma}_F\boldsymbol{b})}{\boldsymbol{b}'\boldsymbol{\Sigma}_F\boldsymbol{b}}. \tag{28}$$

The Principal Components expression (7) follows from setting $\mathring{\boldsymbol{t}} = \boldsymbol{e}'$ in the Diversification Distribution (28), as prescribed by the Principal Components transformation (6), which is a special case of (26).

The Minimum-Torsion expression (13) follows from setting $\mathring{\boldsymbol{t}} = \mathring{\boldsymbol{t}}_{MT}$ in the Diversification Distribution (28), as prescribed by the Minimum-Torsion transformation (12), which is a special case of (26).

A.2 Number of bets in homogeneous markets

Consider $\bar{n}$ factors $\boldsymbol{F} \equiv (F_1, \ldots, F_{\bar{n}})'$. Assume that the variances are all equal to σ^2 and all the pair-wise correlations are equal to $\rho > 0$, as in (8).

As proved in Appendix A.4 of [Meucci, 2009a], the equally-loading portfolio $\boldsymbol{b_{eq}}$ is the eigenvector of the covariance matrix $\boldsymbol{\Sigma}_F$ relative to the largest eigenvalue λ_1^2, or $\boldsymbol{\Sigma}_F\boldsymbol{b_{eq}} = \lambda_1^2\boldsymbol{b_{eq}}$ (the eigenvectors are defined modulo a scale factor, so $\boldsymbol{b_{eq}}$ need not satisfy $\boldsymbol{b}'_{\boldsymbol{eq}}\boldsymbol{b_{eq}} = 1$). As a result, the Diversification Distribution (7) reads $\boldsymbol{p}_{PC}(\boldsymbol{b_{eq}}) = (1, 0, \ldots, 0)'$ and thus $\mathbb{N}_{PC}(\boldsymbol{b_{eq}}) = 1$, as in (9).

On the other hand, the minimum torsion transformation acts equally on all the factors, which are indistinguishable. Hence, for symmetry reasons $(\mathring{\boldsymbol{t}}'^{-1}_{MT}\boldsymbol{b_{eq}}) \propto \boldsymbol{1}$ is a vector of equal entries, and so is $(\mathring{\boldsymbol{t}}_{MT}\boldsymbol{\Sigma}_F\boldsymbol{b_{eq}}) \propto \boldsymbol{1}$. Therefore the Diversification Distribution (13) reads $\boldsymbol{p}_{MT} = (\frac{1}{\bar{n}}, \ldots, \frac{1}{\bar{n}})'$ and thus $\mathbb{N}_{MT}(\boldsymbol{b_{eq}}) = \bar{n}$, as in (14).

A.3 Minimum-torsion optimization

Let us denote the vector of the standard deviations in the covariance matrix $\boldsymbol{\Sigma}_F$ by $\boldsymbol{\sigma} \equiv (dg^{-1}\boldsymbol{\Sigma}_F)^{1/2}$, where the operator $dg^{-1}\boldsymbol{x}$ extracts the diagonal from a matrix $\boldsymbol{x}$. Let us denote by $dg(\boldsymbol{v})$ a diagonal matrix with the vector $\boldsymbol{v}$ on the diagonal.

Solving (10) is equivalent to solving

$$\mathring{\boldsymbol{t}}_{MT} \equiv \underset{\mathbb{C}r\{\boldsymbol{tF}\}=\boldsymbol{I}}{\operatorname{argmin}}\ tr(\mathbb{C}v\{dg(\boldsymbol{\sigma})^{-1}(\boldsymbol{tF}) - dg(\boldsymbol{\sigma})^{-1}\boldsymbol{F}\}). \tag{29}$$

Let us define the normalized factors $\boldsymbol{Z} \equiv dg(\boldsymbol{\sigma})^{-1}\boldsymbol{F}$, whose covariance is the correlation matrix

$$\mathbb{C}v\{\boldsymbol{Z}\} = \boldsymbol{C}_F \equiv dg(\boldsymbol{\sigma})^{-1}\,\boldsymbol{\Sigma}_F\,dg(\boldsymbol{\sigma})^{-1}. \tag{30}$$

Noting that $\mathbb{C}r\{\boldsymbol{t}dg(\boldsymbol{\sigma})\boldsymbol{Z}\} = \boldsymbol{I} \Leftrightarrow \mathbb{C}r\{dg(\boldsymbol{\sigma})^{-1}\boldsymbol{t}dg(\boldsymbol{\sigma})\boldsymbol{Z}\} = \boldsymbol{I}$, we write (29) as follows

$$\mathring{\boldsymbol{t}}_{MT} \equiv \underset{\mathbb{C}r\{dg(\boldsymbol{\sigma})^{-1}\boldsymbol{t}dg(\boldsymbol{\sigma})\boldsymbol{Z}\}=\boldsymbol{I}}{\operatorname{argmin}}\ tr(\mathbb{C}v\{dg(\boldsymbol{\sigma})^{-1}\,\boldsymbol{t}\,dg(\boldsymbol{\sigma})\boldsymbol{Z} - \boldsymbol{Z}\}). \tag{31}$$

Equivalently, we can write

$$\mathring{\boldsymbol{t}}_{MT} = dg(\boldsymbol{\sigma})\,\mathring{\boldsymbol{x}}\,dg(\boldsymbol{\sigma})^{-1}, \tag{32}$$

where $\mathring{\boldsymbol{x}}$ solves

$$\begin{aligned}\mathring{\boldsymbol{x}} &\equiv \underset{\mathbb{C}r\{\boldsymbol{xZ}\}=\boldsymbol{I}}{\operatorname{argmin}}\ tr(\mathbb{C}v\{(\boldsymbol{x}-\boldsymbol{I})\boldsymbol{Z}\}) \\ &= \underset{\mathbb{C}r\{\boldsymbol{xZ}\}=\boldsymbol{I}}{\operatorname{argmin}}\ tr((\boldsymbol{x}-\boldsymbol{I})\boldsymbol{C}_F(\boldsymbol{x}'-\boldsymbol{I})) \\ &= \underset{\mathbb{C}r\{\boldsymbol{xZ}\}=\boldsymbol{I}}{\operatorname{argmin}}\ tr(\boldsymbol{x}\boldsymbol{C}_F\boldsymbol{x}' - \boldsymbol{x}\boldsymbol{C}_F - \boldsymbol{C}_F\boldsymbol{x}' + \boldsymbol{C}_F) \\ &= \underset{\mathbb{C}r\{\boldsymbol{xZ}\}=\boldsymbol{I}}{\operatorname{argmin}}\ tr(\boldsymbol{x}\boldsymbol{C}_F\boldsymbol{x}' - 2\boldsymbol{x}\boldsymbol{C}_F) + \bar{n}\end{aligned} \tag{33}$$

where the last equality follows from the symmetry of $\boldsymbol{C}_F$.

Let us denote by $\mathcal{D}$ the set of diagonal matrices with full rank, and let us introduce the Riccati root of the correlation matrix

$$\boldsymbol{c} = \boldsymbol{c}' \equiv (\boldsymbol{C}_F)^{\frac{1}{2}}. \tag{34}$$

Then $\mathring{\boldsymbol{x}}$ in (32) is the solution of

$$\mathring{\boldsymbol{x}} = \underset{\boldsymbol{xcc}'\boldsymbol{x}'\in\mathcal{D}}{\operatorname{argmin}}\ tr(\boldsymbol{xcc}'\boldsymbol{x}' - 2\boldsymbol{xcc}) \tag{35}$$

A.3.1 Constrained analytical solution

Let us impose the stronger constraint that volatilities are preserved (20), which amounts to

$$\mathbb{C}v\{\boldsymbol{x}\boldsymbol{Z}\} = \boldsymbol{x}\boldsymbol{C}_F\boldsymbol{x}' = \boldsymbol{I}. \tag{36}$$

We emphasize that the constraint (36) does restrict the solution, see (A.3.2) below. From the constraint (36) we obtain $tr(\boldsymbol{x}\boldsymbol{C}_F\boldsymbol{x}') = tr(\boldsymbol{I}) = \bar{n}$. As a result, the minimization (35) becomes

$$\mathring{\boldsymbol{x}} = \underset{\boldsymbol{x}\boldsymbol{c}\boldsymbol{c}'\boldsymbol{x}'=\boldsymbol{I}}{\text{argmax}}\; tr(\boldsymbol{x}\boldsymbol{c}\boldsymbol{c}), \tag{37}$$

Defining $\boldsymbol{y} \equiv \boldsymbol{x}\boldsymbol{c}$, we can write

$$\mathring{\boldsymbol{x}} \equiv [\underset{\boldsymbol{y}\boldsymbol{y}'=\boldsymbol{I}}{\text{argmax}}\, tr(\boldsymbol{y}\boldsymbol{c})]\boldsymbol{c}^{-1} = \boldsymbol{c}^{-1}, \tag{38}$$

where $\text{argmax}_{\boldsymbol{y}\boldsymbol{y}'=\boldsymbol{I}}\, tr(\boldsymbol{y}\boldsymbol{c}) = \boldsymbol{I}$ because $\boldsymbol{c}$ is symmetric with nonnegative eigenvalues, see (61) below.

A.3.2 Unconstrained numerical solution

To solve the general problem (35) without the additional constraint on volatilities (36), let us define $\boldsymbol{\pi} \equiv \boldsymbol{x}\boldsymbol{c}$. Then

$$\mathring{\boldsymbol{x}} = \mathring{\boldsymbol{\pi}}\boldsymbol{c}^{-1}, \tag{39}$$

where

$$\mathring{\boldsymbol{\pi}} \equiv \underset{\boldsymbol{\pi}\boldsymbol{\pi}'\in\mathcal{D}}{\text{argmin}}\, tr(\boldsymbol{\pi}\boldsymbol{\pi}' - 2\boldsymbol{\pi}\boldsymbol{c}). \tag{40}$$

Adapting from [Everson, 1997], we can address the optimization (40) with an iterative algorithm that solves two alternating steps. Let us write

$$\boldsymbol{\pi} \equiv \boldsymbol{d}\boldsymbol{q}, \tag{41}$$

where $\boldsymbol{d}$ is diagonal with full rank and $\boldsymbol{q}$ is orthonormal, in such a way that the constraint $\boldsymbol{\pi}\boldsymbol{\pi}' = \boldsymbol{d}\boldsymbol{q}\boldsymbol{q}'\boldsymbol{d}' = \boldsymbol{d}^2 \in \mathcal{D}$ is satisfied.

Step 1. Assume we know the diagonal matrix $\boldsymbol{d} \in \mathcal{D}$ such that $\boldsymbol{\pi}\boldsymbol{\pi}' = \boldsymbol{d}^2$. Then (40) becomes

$$\mathring{\boldsymbol{\pi}} \equiv \underset{\boldsymbol{\pi}\boldsymbol{\pi}'=\boldsymbol{d}^2}{\text{argmin}}\, tr(\boldsymbol{\pi}\boldsymbol{\pi}' - 2\boldsymbol{\pi}\boldsymbol{c}) = \underset{\boldsymbol{\pi}\boldsymbol{\pi}'=\boldsymbol{d}^2}{\text{argmax}}\, tr(\boldsymbol{c}\boldsymbol{\pi}), \tag{42}$$

where we used the symmetry of the Riccati root $\boldsymbol{c}$. The problem (42) is in the same format as (49) below. The solution then follows from (60) and reads

$$\mathring{\boldsymbol{\pi}} = \boldsymbol{d}((\boldsymbol{d}\boldsymbol{c}^2\boldsymbol{d})^{\frac{1}{2}})^{-1}\boldsymbol{d}\boldsymbol{c}. \tag{43}$$

Since $\boldsymbol{d}$ is invertible, from (41) we obtain

$$\mathring{\boldsymbol{q}} = \boldsymbol{d}^{-1}\mathring{\boldsymbol{\pi}} = ((\boldsymbol{d}\boldsymbol{c}^2\boldsymbol{d})^{\frac{1}{2}})^{-1}\boldsymbol{d}\boldsymbol{c}. \tag{44}$$

Step 2. Assume we know the orthogonal matrix $\boldsymbol{q}$ such that $\boldsymbol{q}\boldsymbol{q}' = \boldsymbol{I}$ and $\boldsymbol{\pi} = \boldsymbol{d}\boldsymbol{q}$. Then (40) becomes

$$\mathring{\boldsymbol{d}} \equiv \underset{\boldsymbol{d}\boldsymbol{q}\boldsymbol{q}'\boldsymbol{d} \in \mathcal{D}}{\operatorname{argmin}}\, tr(\boldsymbol{d}\boldsymbol{q}\boldsymbol{q}'\boldsymbol{d} - 2\boldsymbol{d}\boldsymbol{q}\boldsymbol{c}) = \underset{\boldsymbol{d}^2 \in \mathcal{D}}{\operatorname{argmin}}\, tr(\boldsymbol{d}^2 - 2\boldsymbol{d}\boldsymbol{q}\boldsymbol{c}). \tag{45}$$

In order to solve (45), we differentiate its objective function

$$f(\boldsymbol{d}) \equiv tr(\boldsymbol{d}^2) - 2tr(\boldsymbol{d}\boldsymbol{q}\boldsymbol{c}) = tr(\boldsymbol{d}^2) - 2tr(\boldsymbol{c}\boldsymbol{d}\boldsymbol{q}) \tag{46}$$

with respect to each entries on the diagonal of $\boldsymbol{d} = dg(d_1, \ldots, d_{\bar{n}})$

$$\begin{aligned} \frac{\partial f}{\partial d_n} &= \frac{\partial}{\partial d_n}(\textstyle\sum_m d_m^2 - 2\sum_{m,k} c_{m,k} d_k q_{k,m}) \\ &= 2(d_n - \textstyle\sum_m q_{n,m} c_{m,n}) \\ &= 2(d_n - [\boldsymbol{q}\boldsymbol{c}]_{n,n}), \end{aligned} \tag{47}$$

Setting to zero the derivatives we obtain $\mathring{d}_n = [\boldsymbol{q}\boldsymbol{c}]_{n,n}$ for all $n = 1, \ldots, \bar{n}$, or

$$\mathring{\boldsymbol{d}} = dg(dg^{-1}(\boldsymbol{q}\boldsymbol{c})). \tag{48}$$

Alternating (44) and (48) we arrive at the algorithm (17), which we initialized with $\boldsymbol{d} \equiv \boldsymbol{I}$.

A.4 The constrained Procrustes problem

Adapting from [Schoenemann, 1966], here we present the solution to the orthogonal Procrustes problem

$$\mathring{\boldsymbol{z}} \equiv \underset{\boldsymbol{z}\boldsymbol{z}' = \boldsymbol{d}^2}{\operatorname{argmax}}\, tr(\boldsymbol{k}\boldsymbol{z}), \tag{49}$$

where $\boldsymbol{k}$ is a real matrix and $\boldsymbol{d}^2$ is diagonal with full rank.

Consider the singular value decomposition of the product

$$\boldsymbol{k}\boldsymbol{d} \equiv \boldsymbol{p}\, dg(\boldsymbol{\theta})\boldsymbol{s}', \tag{50}$$

where $\boldsymbol{p}$ and $\boldsymbol{s}$ are orthonormal matrices and $\boldsymbol{\theta}$ is a vector with nonnegative entries. Using (50) and since $\boldsymbol{d}$ is invertible, we can define the change of variables

$$\boldsymbol{y} \equiv \boldsymbol{s}'\boldsymbol{d}^{-1}\boldsymbol{z}\boldsymbol{p}. \tag{51}$$

Then the optimization target in (49) reads

$$\begin{aligned} tr(\boldsymbol{k}\boldsymbol{z}) &= tr(\boldsymbol{p}\, dg(\boldsymbol{\theta})\boldsymbol{s}'\boldsymbol{d}^{-1}\boldsymbol{d}\boldsymbol{s}\boldsymbol{y}\boldsymbol{p}') = tr(\boldsymbol{p}\, dg(\boldsymbol{\theta})\boldsymbol{y}\boldsymbol{p}') \\ &= tr(dg(\boldsymbol{\theta})\boldsymbol{y}\boldsymbol{p}'\boldsymbol{p}) = tr(dg(\boldsymbol{\theta})\boldsymbol{y}) = \textstyle\sum_{n=1}^{\bar{n}} \theta_{n,n} y_{n,n}, \end{aligned} \tag{52}$$

and the constraint in (49) is equivalent to

$$\boldsymbol{y}\boldsymbol{y}' = \boldsymbol{s}'\boldsymbol{d}^{-1}\boldsymbol{z}\boldsymbol{p}\boldsymbol{p}'\boldsymbol{z}'\boldsymbol{d}^{-1}\boldsymbol{s} = \boldsymbol{s}'\boldsymbol{d}^{-1}\boldsymbol{z}\boldsymbol{z}'\boldsymbol{d}^{-1}\boldsymbol{s} = \boldsymbol{s}'\boldsymbol{d}^{-1}\boldsymbol{d}^2\boldsymbol{d}^{-1}\boldsymbol{s} = \boldsymbol{s}'\boldsymbol{s} = \boldsymbol{I}. \tag{53}$$

Hence the solution of (49) is

$$\mathring{\boldsymbol{z}} = \boldsymbol{d}\boldsymbol{s}\mathring{\boldsymbol{y}}\boldsymbol{p}', \tag{54}$$

where

$$\mathring{\boldsymbol{y}} \equiv \underset{\boldsymbol{y}\boldsymbol{y}'=\boldsymbol{I}}{\operatorname{argmax}} \textstyle\sum_{n=1}^{\bar{n}} \theta_{n,n} y_{n,n}. \tag{55}$$

Since $\theta_{n,n} \geq 0$ for all $n = 1, \ldots, \bar{n}$, the maximum is attained by $\mathring{y}_{n,n} = 1$ for all $n = 1, \ldots, \bar{n}$, which implies $\mathring{\boldsymbol{y}} = \boldsymbol{I}$. Substituting this in (54) we obtain the solution to (49)

$$\mathring{\boldsymbol{z}} = \boldsymbol{d}\boldsymbol{s}\boldsymbol{p}'. \tag{56}$$

To further simplify the solution (56), notice

$$\boldsymbol{s}\boldsymbol{p}' = \boldsymbol{s}dg(\boldsymbol{\theta})^{-1}\underbrace{\boldsymbol{s}'\boldsymbol{s}}_{\boldsymbol{I}}dg(\boldsymbol{\theta})\boldsymbol{p}' \overset{(50)}{=} \boldsymbol{s}dg(\boldsymbol{\theta})^{-1}\boldsymbol{s}'\boldsymbol{d}\boldsymbol{k}' \overset{(58)}{=} \boldsymbol{u}^{-1}\boldsymbol{d}\boldsymbol{k}', \tag{57}$$

where

$$\boldsymbol{u} \equiv \boldsymbol{s}dg(\boldsymbol{\theta})\boldsymbol{s}'. \tag{58}$$

It is easy to see that $\boldsymbol{u}$ is the Riccati root of

$$\begin{aligned} \boldsymbol{U} \equiv \boldsymbol{d}\boldsymbol{k}'\boldsymbol{k}\boldsymbol{d} &\overset{(50)}{=} \boldsymbol{s}dg(\boldsymbol{\theta})\boldsymbol{p}'\boldsymbol{p}dg(\boldsymbol{\theta})\boldsymbol{s}' = \boldsymbol{s}dg(\boldsymbol{\theta})dg(\boldsymbol{\theta})\boldsymbol{s}' = \boldsymbol{s}dg(\boldsymbol{\theta})\underbrace{\boldsymbol{s}'\boldsymbol{s}}_{\boldsymbol{I}}dg(\boldsymbol{\theta})\boldsymbol{s}' \\ &= \boldsymbol{u}\boldsymbol{u}' = \boldsymbol{u}^2. \end{aligned} \tag{59}$$

Hence we finally obtain the solution to (49) as follows

$$\mathring{\boldsymbol{z}} \overset{(56)\text{-}(57)}{=} \boldsymbol{d}\boldsymbol{u}^{-1}\boldsymbol{d}\boldsymbol{k}' \overset{(59)}{=} \boldsymbol{d}((\boldsymbol{d}\boldsymbol{k}'\boldsymbol{k}\boldsymbol{d})^{\frac{1}{2}})^{-1}\boldsymbol{d}\boldsymbol{k}'. \tag{60}$$

Notice that if $\boldsymbol{k}$ is symmetric with nonnegative eigenvalues, then $\boldsymbol{k}$ is the Riccati root of $\boldsymbol{k}'\boldsymbol{k}$. Hence

$$\left.\begin{array}{c} \boldsymbol{k} = \boldsymbol{k}' \\ \boldsymbol{d} = \boldsymbol{I} \end{array}\right\} \Rightarrow \mathring{\boldsymbol{z}} = ((\boldsymbol{k}'\boldsymbol{k})^{\frac{1}{2}})^{-1}\boldsymbol{k}' = (\boldsymbol{k})^{-1}\boldsymbol{k} = \boldsymbol{I}. \tag{61}$$