A Note on Correlation and Rank Reduction *

Damiano Brigo

Product and Business Development Group
Banca IMI, San Paolo–IMI Group
Corso Matteotti 6
20121 Milano, Italy
Fax: + 39 02 76019324
E-mail: damiano.brigo@bancaimi.it Web site: http://www.damianobrigo.it

First Version: October 2001. This Version: May 10, 2002

Abstract

We review both full-rank and reduced-rank parameterizations for correlation matrices. In particular, the full-rank parameterizations of Rebonato (1999d) and Schoenmakers and Coffey (2000), the reduced-rank angle parameterization of Rebonato (1999d) and Rebonato and Jäckel (1999), and the well known zeroed-eigenvalues reduced-rank approximation are described in detail. Numerical examples are provided for the reduced-rank parameterizations and results are compared.

1 Introduction

This technical note is meant to present some known and established results in rank-reduction techniques for correlation matrices, with a few new developments. Besides being an interesting problem in its own right, rank reduction is an important technique in several applications, including image processing and, when applied to correlation matrices, mathematical finance (interest rate modeling, risk management). Here we will deal with the application to the LIBOR market model (as in Brigo and Mercurio (2001), Chapter 6), although the results reported here can be used also in different fields.

We start with a given $M \times M$ full-rank correlation matrix. In general the full instantaneous correlation matrix is characterized by $M\ (M-1)/2$ entries, given symmetry and the ones in the diagonal. This number of entries can be too high for practical purposes. Therefore, a parsimonious parametric form has to be found for ρ, based on a reduced number of parameters.

---

*Draft. To be checked again. This note is meant to be an expansion of Section 6.9 of the first edition of the book "Interest-Rate Models: Theory and Practice", Springer, 2001, by D. Brigo and F. Mercurio. This note is available at http://www.damianobrigo.it. Helpful suggestions and references provided by Raymond Lacey are gratefully acknowledged.

As a first possibility, we may decide to maintain a *full-rank* correlation matrix involving a number of parameters that is smaller than $M(M-1)/2$. Such an approach has been followed by Schoenmakers and Coffey (2000), in connection to swaptions calibration in the LIBOR market model. We summarize some of their findings in Section 2.

Alternatively, *reduced-rank* correlations can be obtained, either through Rebonato's angles parameterization, or through eigenvalues zeroing. We review both possibilities in Section 3.

Finally, we mention Zhang and Wu (2001) for their work on optimal low-rank approximations of correlation matrices based on Lagrange multipliers techniques.

2 Some convenient full rank parameterizations

Schoenmakers and Coffey (2000) propose the following full-rank parametric form for the correlation matrix ρ. They consider a finite sequence of positive real numbers

$$1 = c_1 < c_2 < \ldots < c_M, \quad \frac{c_1}{c_2} < \frac{c_2}{c_3} < \ldots < \frac{c_{M-1}}{c_M},$$

and they set

$$\rho^F(c)_{i,j} := c_i/c_j, \quad i \leq j, \quad i,j = 1,\ldots,M.$$

Notice that the correlation between changes in adjacent rates is $\rho^F_{i+1,i} = c_i/c_{i+1}$. The above requirements on c's translate into the requirement that the sub-diagonal of the resulting correlation matrix $\rho^F(c)$ be increasing when moving from NW ("North-West") to SE ("South-East"). This bears the interpretation that when we move along the yield curve, the larger the tenor, the more correlated changes in adjacent forward rates become. This corresponds to the experienced fact that the forward curve tends to flatten and to move in a more "correlated" way for large maturities than for small ones. Correlations in the sub-diagonal (i.e. referring to changes in adjacent rates) thus tend to grow as the maturity increases (as the index in the sub-diagonal increases). This holds also for lower levels below the diagonal: In general the map

$$i \mapsto \rho_{i+p,i}$$

is increasing for all p. All changes between equally spaced forward rates become more correlated as their expiries increase, not only changes in adjacent forward rates.

A second property that can be important for applications to the LIBOR market model is that correlations decrease when moving away from the "1" diagonal term along a row or a column. Indeed, we see that

$$\rho_{i,j} < \rho_{i,j+1}, \quad j < i.$$

This follows immediately from the fact that the c's are increasing, since

$$\rho_{i,j} = \rho_{j,i} = c_j/c_i < c_{j+1}/c_i = \rho_{i,j+1}, \quad j < i.$$

An analogous behaviour is found when moving off the diagonal in the opposite direction:

$$\rho_{i,j} < \rho_{i,j-1}, \quad j > i.$$

The number of parameters needed in this formulation is M, versus the $M(M-1)/2$ number of entries in the general correlation matrix. It is easy to prove that $\rho^F(c)$ is always a viable correlation matrix if defined as above (symmetric, positive semidefinite and with ones in the diagonal).

Schoenmakers and Coffey (2000) observe also that this parameterization can be always characterized in terms of a finite sequence of non-negative numbers $\Delta_2, \ldots, \Delta_M$:

$$c_i = \exp\left[\sum_{j=2}^{i} j\Delta_j + \sum_{j=i+1}^{M} (i-1)\Delta_j\right].$$

Some particular cases in this class of parameterizations that Schoenmakers and Coffey (2000) consider to be promising can be formulated through suitable changes of variables as follows. The first is the case where all Δ's are equal ($\Delta_2 = \ldots = \Delta_{M-1}$) except the last one: by a change of variable one has

Stable, full rank, two-parameters, "increasing along sub-diagonals" parameterization for instantaneous correlation:

$$\rho_{i,j} = \exp\left[-\frac{|i-j|}{M-1}\left(-\ln\rho_\infty + \eta\frac{M-1-i-j}{M-2}\right)\right].$$

Stability here is meant to point out that relatively small movements in the c-parameters connected to this form (and thus in the correlations themselves) cause relatively small changes in ρ_∞ and η. Notice that $\rho_\infty = \rho_{1,M}$ is the correlation between the farthest forward rates in the family considered, whereas η is related to the first non-zero Δ, i.e. $\eta = \Delta_{M-1}(M-1)(M-2)/2$.

A three-parameters form is obtained with the Δ_i's following a straight line (two parameters) for $i = 2, 3, \ldots, M-1$ and set to a third parameter for $i = M$. Again through a change of variable, this form can be expressed in a laborious expression generalizing the previous one. The calibration experiments of Schoenmakers and Coffey (2000) pointed out, however, that the parameter associated with the final point Δ_{M-1} of our straight line in the Δ's is practically always close to zero. Setting thus $\Delta_{M-1} = 0$ and maintaining the other characteristics of the last parameterization leads to the following

Improved, stable, full rank, two-parameters, "increasing along sub-diagonals" parameterization for instantaneous correlations:

$$\begin{aligned}\rho_{i,j} = \exp\Bigg[& - \frac{|i-j|}{M-1}\Bigg(- \ln\rho_\infty \\ & + \eta\frac{i^2+j^2+ij-3Mi-3Mj+3i+3j+2M^2-M-4}{(M-2)(M-3)}\Bigg)\Bigg].\end{aligned}$$

As before, $\rho_\infty = \rho_{1,M}$, whereas η is related to the steepness of the straight line in the Δ's.

Finally, consider the

Classical, two-parameters, exponentially decreasing parameterization

$$\rho_{i,j} = \rho_\infty + (1-\rho_\infty)\exp[-\beta|i-j|], \quad \beta \geq 0,$$

where now ρ_∞ is only asymptotically representing the correlation between the farthest rates in the family.

Schoenmakers and Coffey (2000) point out that Rebonato's (1999c,d) full-rank parameterization, consisting in the following perturbation of the classical structure:

Rebonato's three parameters full rank parameterization

$$\rho_{i,j} = \rho_\infty + (1-\rho_\infty)\exp[-|i-j|(\beta - \alpha\max(i,j))],$$

has still the desirable property of being increasing along sub-diagonals. However, this form does not fit the above general framework based on the c parameters, and moreover the domain of positivity for the resulting matrix is not specified "off-line" in terms of the parameters $\alpha, \beta, \rho_\infty$. One has to check at every step of a hypothetical calibration/optimization that the resulting matrix is positive semidefinite. On the contrary, the above formulations based on the c parameters are automatically positive semidefinite and thus do not require an iterative check in a calibration session. Since an unconstrained optimization is preferable to a constrained one, the above parameterizations of Schoenmakers and Coffey (2000) can be preferred from this point of view.

3 Reduced-rank formulations: Angles and eigenvalues zeroing

We know that, being ρ a positive definite symmetric matrix, it can be written as

$$\rho = PHP',$$

where P is a real orthogonal matrix, $P'P = PP' = I_M$, and H is a diagonal matrix of the positive eigenvalues of ρ. *The columns of P are the eigenvectors of ρ, associated to the eigenvalues located in the corresponding position in the diagonal matrix H.* Let Λ be the diagonal matrix whose entries are the square roots of the corresponding entries of H, so that if we set $A := P\Lambda$ we have both

$$AA' = \rho, \quad A'A = H \ .$$

We can try and mimic the decomposition $\rho = AA'$ by means of a suitable n-rank $M \times n$ matrix B such that BB' is an n-rank correlation matrix, with typically $n << M$.

The advantage of doing so is that we may take as new noise a standard n-dimensional Brownian motion W and replace the original M-dimensional random shocks $dZ(t)$ by $B\ dW(t)$. In other terms, we move from a noise correlation structure

$$dZ\ dZ' = \rho\ dt$$

to

$$B\ dW(B\ dW)' = B\ dW dW'\ B' = BB' dt.$$

Therefore, with noise given by $B\ dW$ our new instantaneous noise-correlation matrix is BB' whose rank is $n << M$, and the dimension of our random shocks has decreased to n. We set

$$\rho^B = BB'.$$

If we decide to adopt indeed a reduced-rank approach, we are left with the problem of choosing a suitable parametric form for the B matrix, such that BB' is a possible correlation matrix.

3.1 Rebonato's angles formulation

Rebonato (1999d) suggests the following general form for the i-th row of the above B matrix:

$$\begin{aligned} b_{i,1} &= \cos\theta_{i,1} \\ b_{i,k} &= \cos\theta_{i,k} \sin\theta_{i,1} \cdots \sin\theta_{i,k-1}, \quad 1 < k < n, \\ b_{i,n} &= \sin\theta_{i,1} \cdots \sin\theta_{i,n-1}, \end{aligned} \tag{1}$$

for $i = 1, 2, \ldots, M$. Notice that with this parameterization $\rho^B =: \rho(\theta)$ is clearly positive semidefinite and its diagonal terms are ones. It follows that $\rho(\theta)$ is a possible correlation matrix. The number of parameters in this case is $M \times (n-1)$.

A particular case of the above formulation (Morini (2002)) is obtained when considering $\theta_{i,j} = \theta_i$ for all j, resulting in angles that are homogeneous with respect to the column index. This formulation lowers the number of angles considerably. One may wonder whether there is a compact formula for the correlation resulting from this suggestion. We now briefly investigate this problem. The M parameters n-rank correlation matrix $\rho(\theta)$ reads, by straightforward calculations,

$$\rho_{i,j}(\theta) = \cos\theta_i \ \cos\theta_j + \sum_{k=1}^{n-2} \cos\theta_i \ \cos\theta_j \ (\sin\theta_i \ \sin\theta_j)^k + (\sin\theta_i \ \sin\theta_j)^{n-1}.$$

By recalling that $1 + x + \ldots + x^m = (1 - x^{m+1})/(1-x)$ and setting $\sin\theta_i = \alpha_i$, we obtain easily that the above expression is equivalent to the following n-rank parameterization:

n-rank, column-homogeneous angles parameterization:

$$(\rho_H)_{i,j} = \sqrt{(1-\alpha_i^2)(1-\alpha_j^2)} \ \frac{1-(\alpha_i\alpha_j)^{n-1}}{1-\alpha_i\alpha_j} + (\alpha_i\alpha_j)^{n-1}, \quad \alpha_k \in (-1,1),$$

with $\rho_H = B_H B_H'$, where

$$(b_H)_{i,j} = (1-\alpha_i^2)^{\frac{1\{j<n\}}{2}} \alpha_i^{(j-1)},$$

$1\{\cdot\}$ denoting the indicator function of the argument inside curly brackets.

Let us now consider a simple two-factor structure, $n = 2$, consisting of M parameters. This is obtained as

$$b_{i,1} = \cos\theta_{i,1}, \quad b_{i,2} = \sin\theta_{i,1} \ . \tag{2}$$

Dropping the second subscript for θ, we have

$$\rho^B_{i,j} = b_{i,1}b_{j,1} + b_{i,2}b_{j,2} = \cos(\theta_i - \theta_j). \tag{3}$$

This structure consists of M parameters $\theta_1, \ldots, \theta_M$.

Notice the following interesting feature of the general angle parameterization. Assume we use this parameterization while keeping full-rank $n = M$. The number of angle parameters is then $M(M-1)$, and is twice the $M(M-1)/2$ significant entries in the correlation matrix. The angle parameterization in this case increases the dimension, and this is usually considered a drawback. One has to remember, however, that the angle parameterization embeds the constraints on the one-entries diagonal and on symmetry and positive semidefiniteness. The excess in parameters is understood with a toy example in dimension two. Take $n = M = 2$. Then we have only one correlation parameter $\rho_{1,2} = \rho_{2,1} = c$, but two angles:

$$\begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} = \begin{bmatrix} 1 & \cos(\theta_2 - \theta_1) \\ \cos(\theta_2 - \theta_1) & 1 \end{bmatrix}.$$

As we see, what matters here is the difference between the angles (1 parameter) and not the particular angles themselves (2 parameters).

Incidentally, we see that the reduction of the number of parameters in the correlation matrix cannot be solved simply by slightly reducing the rank through the angle formulation. If we wish the number $M(n-1)$ of parameters of the angle parameterization to be much smaller than the significant number of $M(M-1)/2$ entries in the full rank matrix, we need ask not only $n < M$, as the rank reduction implies, but rather

$$n << (M+1)/2.$$

This is a possibly disturbing feature, making also the correspondence $\theta \mapsto \rho(\theta)$ not one to one, since in the full-rank case the domain has dimension $M(M-1)$, while the image space has dimension $M(M-1)/2$. However, Rapisarda (2002) describes a way to "standardize" the angles specification so as to restore the one-to-one nature of the correspondence, and we will hint at this result in Section 3.4.

Now let us assume we have selected $n = 2$ as above. If M is large (typically 20), we can still have troubles with a too large number of parameters θ. We can then select a subparameterization for the θ's of the type

$$\theta_k = \vartheta(k) \ ,$$

where $\vartheta(\cdot)$ is a function depending on a small (say four or five) number of parameters. Such a function could be for example a linear-exponential combination. However, in our experience keeping the θ's free from a subparameterization can be necessary if we are calibrating the LIBOR market model to a large number of swaptions, when using instantaneous correlations as swaptions-fitting parameters.

In general, when we calibrate the LIBOR market model to swaptions using instantaneous correlations ρ as fitting parameters, we are free to select a priori a parametric form for the correlation matrix. We may for example take the n rank matrix $\rho^B = BB'$ defined above in terms of angles θ. Here the reduced rank is built into the parameterization and we are sure that, once the model has been calibrated, an n-dimensional independent shocks structure $dW(t)$ will be sufficient to perform simulations with M dimensional correlated shocks $B\ dW(t)$.

However, in some parameterizations of the instantaneous covariance structure of the LIBOR market model, such as for example in Brigo and Mercurio (2001) Section 7.4, and Brigo and Mercurio (2002), the correlation matrix is given exogenously (say through historical estimation), instead of being calibrated to the swaption market. When the correlation matrix ρ is given as an exogenous input to the calibration, instead of being kept as a fitting parameter, it will have full rank M in general. The problem in this situation is that when decomposing ρ as BB', the resulting B will be a $M \times M$ matrix, so that the independent random shocks structure $dW(t)$ in the LIBOR market model will have full dimension M, exactly as the correlated noise structure. When in need of performing Monte Carlo simulation to value exotics, this full-dimensionality is an undesirable feature.

This problem can be alleviated if we can find a way to obtain a reduced rank correlation matrix $\rho^{(n)}$ that is, in some sense, the best approximation of the exogenously-given full-rank correlation matrix ρ. Then the swaption calibration proceeds by keeping $\rho^{(n)}$ as exogenous correlation matrix.

We now list some possible approaches to this problem,

$$M\text{-rank-}\rho \quad \longrightarrow n\text{-rank-}\rho^{(n)}.$$

Summing up, we are given in input a full rank $M \times M$ correlation matrix ρ. We aim at approximating this matrix through an at most n-rank correlation matrix $\rho^{(n)}$, $n < M$.

3.2 The approach of zeroing the smallest eigenvalues and rescaling

As before, we can try and mimic the decomposition $\rho = AA'$ by means of a suitable n-rank $M \times n$ matrix B such that BB' is an n-rank correlation matrix, with typically $n << M$. But this time, instead of taking B as an angle-parameterized matrix, we define B as follows.

Consider the diagonal matrix $\bar{\Lambda}^{(n)}$ defined as the matrix Λ with the $M-n$ smallest diagonal terms set to zero.

Define then $\bar{B}^{(n)} := P\bar{\Lambda}^{(n)}$, and the related candidate correlation matrix

$$\bar{\rho}^{(n)} := \bar{B}^{(n)}(\bar{B}^{(n)})'.$$

Notice that we can also equivalently define $\Lambda^{(n)}$ as the $n \times n$ (instead of $M \times M$) diagonal matrix obtained from Λ by taking away (instead of zeroing) the $M-n$ smallest diagonal elements and shrinking the matrix correspondingly. Analogously, we can define the $M \times n$ matrix $P^{(n)}$ as the matrix P from which we take away the columns corresponding to the diagonal elements we took away from Λ. More precisely, if we call

$$\Lambda_{i_1,i_1} \ , \ \Lambda_{i_2,i_2} \ , \ \ldots, \ \Lambda_{i_{M-n},i_{M-n}}$$

the $M-n$ smallest diagonal elements of Λ, then $P^{(n)}$ is obtained from P by taking away the columns $i_1, i_2, \ldots, i_{M-n}$. The result does not change, in that if we define the $M \times n$ matrix $B^{(n)} = P^{(n)}\Lambda^{(n)}$ we have

$$\bar{\rho}^{(n)} = \bar{B}^{(n)}(\bar{B}^{(n)})' = B^{(n)}(B^{(n)})'.$$

We keep the $B^{(n)}$ formulation. Now the problem is that, in general, while $\bar{\rho}^{(n)}$ is positive semidefinite, it does not feature ones in the diagonal. Throwing away some eigenvalues from Λ has altered the diagonal. The solution is to interpret $\bar{\rho}^{(n)}$ as a *covariance* matrix, and to derive the correlation matrix associated with it. We can do this immediately by defining

$$\rho^{(n)}_{i,j} := \frac{\bar{\rho}^{(n)}_{i,j}}{\sqrt{\bar{\rho}^{(n)}_{i,i} \ \bar{\rho}^{(n)}_{j,j}}}.$$

Now $\rho^{(n)}_{i,j}$ is an n-rank approximation of the original matrix ρ. But how good is the approximation, and are there more precise methods to approximate a full rank correlation matrix with a n-rank matrix?

3.3 The approach of optimizing on a low rank parametric form

Rebonato and Jäckel (1999) suggest the following alternative to the above procedure and then compare the two results.

We can start from an n-rank matrix $\rho(\theta) = B(\theta)B'(\theta)$ defined in terms of the angles as before: the i-th row of B is again

$$\begin{aligned} b_{i,1}(\theta) &= \cos\theta_{i,1} \\ b_{i,k}(\theta) &= \cos\theta_{i,k}\sin\theta_{i,1}\cdots\sin\theta_{i,k-1}, \quad 1 < k < n, \\ b_{i,n}(\theta) &= \sin\theta_{i,1}\cdots\sin\theta_{i,n-1}, \end{aligned}$$

where now the dependence on θ has been pointed out explicitly.

When a target full-rank correlation matrix $\hat{\rho}$ is given as input, we can try and find the parameters θ that minimize a norm of the difference between the target matrix ρ and our parameterized matrix $\rho(\theta)$. In other terms, we try and minimize with respect to θ the quantity

$$\sum_{i,j=1}^{M} (|\rho_{i,j} - \rho_{i,j}(\theta)|^2).$$

An important feature of this formulation is that the θ's are completely free: The resulting matrix $\rho(\theta)$ is *always* symmetric, positive semidefinite and with ones in the diagonal, so that we need to pose no constraints in the optimization.

Rebonato and Jäckel (1999) argue that the differences between this method, giving the optimal solution, and the "eigenvalues zeroing" method above is typically small, and show some examples. We are going to confirm these results with a few numerical experiments of our own in the following.

Assume we have a 10×10 full-rank correlation matrix coming from the full-rank classical parametric form:

$$\hat{\rho}_{i,j} = 0.5 + (1 - 0.5) \exp[-0.05|i - j|],$$

and that we first try to fit this matrix with a rank-2 correlation structure.

The input matrix $\hat{\rho}$ to be fitted is

1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704	0.8523	0.8352	0.8188
0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704	0.8523	0.8352
0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704	0.8523
0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704
0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894
0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094
0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304
0.8523	0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524
0.8352	0.8523	0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756
0.8188	0.8352	0.8523	0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1

In this case we show also the orthogonal (eigenvectors) matrix P and the diagonal (eigenvalues) matrix H such that $\rho = PHP'$. The matrix P is

-0.0714	0.1398	0.2072	0.2657	0.3224	0.3652	-0.4045	-0.4261	0.4293	0.3081
0.2039	-0.3624	-0.4416	-0.424	-0.3099	-0.1325	-0.0844	-0.2703	0.4003	0.3135
-0.3164	0.4468	0.3137	-0.0024	-0.3193	-0.4466	0.3067	-0.0081	0.3247	0.3175
0.3982	-0.3612	0.0715	0.4252	0.3131	-0.1406	0.4401	0.2577	0.2113	0.3202
-0.4412	0.1379	-0.3981	-0.2623	0.3162	0.3603	0.2038	0.423	0.0733	0.3216
0.4412	0.1379	0.3981	-0.2623	-0.3162	0.3603	-0.2038	0.423	-0.0733	0.3216
-0.3982	-0.3612	-0.0715	0.4252	-0.3131	-0.1406	-0.4401	0.2577	-0.2113	0.3202
0.3164	0.4468	-0.3137	-0.0024	0.3193	-0.4466	-0.3067	-0.0081	-0.3247	0.3175
-0.2039	-0.3624	0.4416	-0.424	0.3099	-0.1325	0.0844	-0.2703	-0.4003	0.3135
0.0714	0.1398	-0.2072	0.2657	-0.3224	0.3652	0.4045	-0.4261	-0.4293	0.3081

whereas H is the 10×10 diagonal matrix whose diagonal elements ($\hat{\rho}$'s eigenvalues) are, from NW to SE:

0.0128 0.0138 0.0157 0.0191 0.0249 0.0359 0.0594 0.1268 0.4207 9.2709

In this case the smallest eight eigenvalues are the first eight. We are thus left with the matrix $P^{(2)}$ given by the last two columns of P, and $\Lambda^{(2)}$ given by

$$\begin{bmatrix} \sqrt{0.4207} & 0 \\ 0 & \sqrt{9.2709} \end{bmatrix}$$

By defining the 10×2 matrix $B^{(2)} := P^{(2)}\Lambda^{(2)}$ and by following the above procedure we obtain, through scaling of $\bar{\rho}^{(2)} = B^{(2)}(B^{(2)})'$ the final rank-2 correlation matrix $\rho^{(2)}$:

1	0.9997	0.9973	0.9889	0.9713	0.9437	0.9097	0.8761	0.8503	0.838
0.9997	1	0.9987	0.9921	0.9765	0.9511	0.919	0.887	0.8622	0.8503
0.9973	0.9987	1	0.9972	0.9863	0.9656	0.938	0.9094	0.887	0.8761
0.9889	0.9921	0.9972	1	0.9959	0.9824	0.9613	0.938	0.919	0.9097
0.9713	0.9765	0.9863	0.9959	1	0.9953	0.9824	0.9656	0.9511	0.9437
0.9437	0.9511	0.9656	0.9824	0.9953	1	0.9959	0.9863	0.9765	0.9713
0.9097	0.919	0.938	0.9613	0.9824	0.9959	1	0.9972	0.9921	0.9889
0.8761	0.887	0.9094	0.938	0.9656	0.9863	0.9972	1	0.9987	0.9973
0.8503	0.8622	0.887	0.919	0.9511	0.9765	0.9921	0.9987	1	0.9997
0.838	0.8503	0.8761	0.9097	0.9437	0.9713	0.9889	0.9973	0.9997	1

This is the rank-2 matrix resulting from the "eigenvalues zeroing" quick procedure. Notice that we can express this rank-2 matrix by means of its angles parameters. By inverting the relevant transformation, we see that this matrix is associated with the angles $\theta^{(2)}$ given by

$$\theta^{(2)} = [1.2886 \;\; 1.3081 \;\; 1.3586 \;\; 1.4333 \;\; 1.5233 \;\; 1.6183 \;\; 1.7083 \;\; 1.7830 \;\; 1.8335 \;\; 1.8530]'.$$

As a more refined alternative, we now consider the above rank-2 parameterization in the θ's and minimize the difference

$$\sum_{i,j=1}^{M} (|\hat{\rho}_{i,j} - \rho_{i,j}(\theta)|^2)$$

through a numerical optimization. The numerical optimization takes a few seconds in an interpreted language such as MATLAB and yields

$$\theta^{*(2)} = [1.2367 \;\; 1.2812 \;\; 1.3319 \;\; 1.3961 \;\; 1.4947 \;\; 1.6469 \;\; 1.7455 \;\; 1.8097 \;\; 1.8604 \;\; 1.9049].$$

Compare these angles $\theta^{*(2)}$ to $\theta^{(2)}$ to see how close the quickly-computed zeroed-eigenvalues matrix $\rho^{(2)} = \rho(\theta^{(2)})$ is to the optimal rank-2 correlation matrix $\rho(\theta^{*(2)})$.

The resulting optimal rank-2 matrix $\rho(\theta^{*(2)})$ is given by

1	0.999	0.9955	0.9873	0.9669	0.917	0.8733	0.8403	0.8117	0.7849
0.999	1	0.9987	0.9934	0.9773	0.9339	0.8941	0.8636	0.8369	0.8117
0.9955	0.9987	1	0.9979	0.9868	0.9508	0.9157	0.888	0.8636	0.8403
0.9873	0.9934	0.9979	1	0.9951	0.9687	0.9396	0.9157	0.8941	0.8733
0.9669	0.9773	0.9868	0.9951	1	0.9885	0.9687	0.9508	0.9339	0.917
0.917	0.9339	0.9508	0.9687	0.9885	1	0.9951	0.9868	0.9773	0.9669
0.8733	0.8941	0.9157	0.9396	0.9687	0.9951	1	0.9979	0.9934	0.9873
0.8403	0.8636	0.888	0.9157	0.9508	0.9868	0.9979	1	0.9987	0.9955
0.8117	0.8369	0.8636	0.8941	0.9339	0.9773	0.9934	0.9987	1	0.999
0.7849	0.8117	0.8403	0.8733	0.917	0.9669	0.9873	0.9955	0.999	1

To compare the difference between the original matrix $\hat{\rho}$, the zeroed-eigenvalues rank-2 matrix $\rho^{(2)}$ and finally the optimal rank-2 matrix $\rho(\theta^{*(2)})$, we plot the second columns of the three matrices in the left hand side of figure 1.

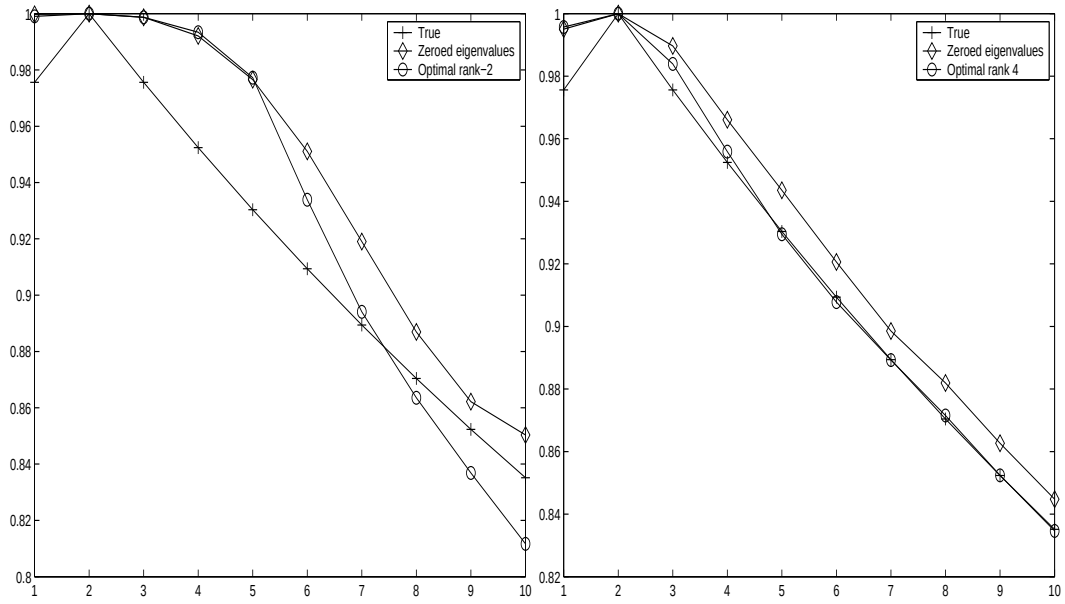

Figure 1: On the left: Second columns of $\hat{\rho}$, $\rho^{(2)}$, and $\rho(\theta^{*(2)})$. On the right: Second columns of $\hat{\rho}$, $\rho^{(4)}$, and $\rho(\theta^{*(4)})$

From the left figure we see that both rank-2 formulations are in troubles because of the sigmoid-like shape they are forced to assume. On the contrary, the original correlations move away from the diagonal in an almost straight pattern. The optimal rank-2 matrix is the best we can do with a rank-2 approximation, and we see that "on average" it recovers the true correlation columns, in that it stays partly above and partly below the true correlation (almost straight) plot. Indeed, being forced to assume sigmoid-like columns, the rank-2 optimal approximation oscillates from above to below said column. The rougher zeroed-eigenvalues rank-2 reduction does not even do that, remaining always above the true correlation column. Differences are not huge but are however not negligible. Moreover, if decorrelation is more pronounced, the approximation can worsen.

We now check the analogous approximations if we resort to rank 4 approximation. The resulting correlation matrices are as follows. The zeroed-eigenvalues matrix $\rho^{(4)}$ is given by

1	0.9951	0.9708	0.9379	0.9148	0.8979	0.8821	0.867	0.8448	0.8237
0.9951	1	0.9897	0.966	0.9435	0.9206	0.8985	0.8819	0.8627	0.8448
0.9708	0.9897	1	0.9922	0.9742	0.9463	0.9174	0.8986	0.8819	0.867
0.9379	0.966	0.9922	1	0.9925	0.9685	0.9388	0.9174	0.8985	0.8821
0.9148	0.9435	0.9742	0.9925	1	0.9904	0.9685	0.9463	0.9206	0.8979
0.8979	0.9206	0.9463	0.9685	0.9904	1	0.9925	0.9742	0.9435	0.9148
0.8821	0.8985	0.9174	0.9388	0.9685	0.9925	1	0.9922	0.966	0.9379
0.867	0.8819	0.8986	0.9174	0.9463	0.9742	0.9922	1	0.9897	0.9708
0.8448	0.8627	0.8819	0.8985	0.9206	0.9435	0.966	0.9897	1	0.9951
0.8237	0.8448	0.867	0.8821	0.8979	0.9148	0.9379	0.9708	0.9951	1

whereas the optimal four-rank matrix $\rho(\theta^{*(4)})$ based on the optimal angles is this time

1	0.9957	0.9633	0.9267	0.8999	0.8867	0.8752	0.8598	0.8346	0.8136
0.9957	1	0.9839	0.9559	0.9295	0.9078	0.8893	0.8715	0.8524	0.8346
0.9633	0.9839	1	0.9904	0.9673	0.9317	0.9012	0.8796	0.8715	0.8598
0.9267	0.9559	0.9904	1	0.9911	0.9603	0.9276	0.9012	0.8893	0.8752
0.8999	0.9295	0.9673	0.9911	1	0.9865	0.9603	0.9317	0.9078	0.8867
0.8867	0.9078	0.9317	0.9603	0.9865	1	0.9911	0.9673	0.9295	0.8999
0.8752	0.8893	0.9012	0.9276	0.9603	0.9911	1	0.9904	0.9559	0.9267
0.8598	0.8715	0.8796	0.9012	0.9317	0.9673	0.9904	1	0.9839	0.9633
0.8346	0.8524	0.8715	0.8893	0.9078	0.9295	0.9559	0.9839	1	0.9957
0.8136	0.8346	0.8598	0.8752	0.8867	0.8999	0.9267	0.9633	0.9957	1

The two matrices are associated respectively with the angles

$\theta^{(4)}$			$\theta^{*(4)}$		
1.6695	1.7239	1.2837	1.6844	1.7328	1.2775
1.5914	1.6672	1.3068	1.6088	1.6828	1.2965
1.496	1.5737	1.358	1.4688	1.581	1.3444
1.4634	1.4784	1.4319	1.4435	1.4708	1.4267
1.5211	1.4194	1.5226	1.5051	1.3957	1.5203
1.6205	1.4194	1.6189	1.6365	1.3957	1.6213
1.6782	1.4784	1.7097	1.6981	1.4708	1.7149
1.6456	1.5737	1.7836	1.6728	1.581	1.7972
1.5502	1.6672	1.8348	1.5328	1.6828	1.8451
1.4721	1.7239	1.8579	1.4571	1.7328	1.864

We can plot again the second columns of the three matrices in the right-hand side of figure 1.

As we see both from the figures and the tables, the approximation has improved.

When moving to rank 7 matrices, the approximation becomes very good and few differences are noticeable in the three matrices.

We now examine a more extreme case where the rates decorrelate very quickly and steeply from 1 to 0. Consider the 10×10 full-rank correlation matrix $\hat{\rho}$ coming from the full-rank classical parametric form:

$$\hat{\rho}_{i,j} = \exp[-|i-j|].$$

The matrix $\hat{\rho}$ is this time

1	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009	0.0003	0.0001
0.3679	1	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009	0.0003
0.1353	0.3679	1	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009
0.0498	0.1353	0.3679	1	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025
0.0183	0.0498	0.1353	0.3679	1	0.3679	0.1353	0.0498	0.0183	0.0067
0.0067	0.0183	0.0498	0.1353	0.3679	1	0.3679	0.1353	0.0498	0.0183
0.0025	0.0067	0.0183	0.0498	0.1353	0.3679	1	0.3679	0.1353	0.0498
0.0009	0.0025	0.0067	0.0183	0.0498	0.1353	0.3679	1	0.3679	0.1353
0.0003	0.0009	0.0025	0.0067	0.0183	0.0498	0.1353	0.3679	1	0.3679
0.0001	0.0003	0.0009	0.0025	0.0067	0.0183	0.0498	0.1353	0.3679	1

If we focus on a rank-4 approximation, the zeroed-eigenvalues procedure yields a matrix $\rho^{(4)}$ given by

1	0.9474	0.5343	-0.0116	-0.1967	-0.0427	0.1425	0.1378	-0.042	-0.1511
0.9474	1	0.775	0.2884	0.0164	-0.03	0.0316	0.0538	0	-0.042
0.5343	0.775	1	0.8137	0.4993	0.0979	-0.1229	-0.1035	0.0538	0.1378
-0.0116	0.2884	0.8137	1	0.8583	0.3725	-0.0336	-0.1229	0.0316	0.1425
-0.1967	0.0164	0.4993	0.8583	1	0.7658	0.3725	0.0979	-0.03	-0.0427
-0.0427	-0.03	0.0979	0.3725	0.7658	1	0.8583	0.4993	0.0164	-0.1967
0.1425	0.0316	-0.1229	-0.0336	0.3725	0.8583	1	0.8137	0.2884	-0.0116
0.1378	0.0538	-0.1035	-0.1229	0.0979	0.4993	0.8137	1	0.775	0.5343
-0.042	0	0.0538	0.0316	-0.03	0.0164	0.2884	0.775	1	0.9474
-0.1511	-0.042	0.1378	0.1425	-0.0427	-0.1967	-0.0116	0.5343	0.9474	1

The optimization on an angle-parameterized rank-4 matrix yields the following output matrix $\rho(\theta^{*(4)})$:

1	0.9399	0.4826	-0.0863	-0.2715	-0.0437	0.1861	0.1808	-0.077	-0.2189
0.9399	1	0.7515	0.234	-0.0587	-0.0572	0.0496	0.0843	-0.0135	-0.077
0.4826	0.7515	1	0.7935	0.4329	0.015	-0.1745	-0.1195	0.0843	0.1808
-0.0863	0.234	0.7935	1	0.8432	0.3222	-0.0872	-0.1745	0.0496	0.1861
-0.2715	-0.0587	0.4329	0.8432	1	0.7421	0.3222	0.015	-0.0572	-0.0437
-0.0437	-0.0572	0.015	0.3222	0.7421	1	0.8432	0.4329	-0.0587	-0.2715
0.1861	0.0496	-0.1745	-0.0872	0.3222	0.8432	1	0.7935	0.234	-0.0863
0.1808	0.0843	-0.1195	-0.1745	0.015	0.4329	0.7935	1	0.7515	0.4826
-0.077	-0.0135	0.0843	0.0496	-0.0572	-0.0587	0.234	0.7515	1	0.9399
-0.2189	-0.077	0.1808	0.1861	-0.0437	-0.2715	-0.0863	0.4826	0.9399	1

If we resort to a rank-7 approximation, the zeroed-eigenvalues approach yields the following matrix $\rho^{(7)}$:

1	0.5481	0.0465	0.0944	0.0507	-0.0493	0.034	0.0169	-0.0441	0.0284
0.5481	1	0.6737	0.0647	0.0312	0.112	-0.0477	-0.0162	0.0691	-0.0441
0.0465	0.6737	1	0.579	0.1227	0.0353	0.0562	0.0012	-0.0162	0.0169
0.0944	0.0647	0.579	1	0.5822	0.0674	0.0806	0.0562	-0.0477	0.034
0.0507	0.0312	0.1227	0.5822	1	0.6472	0.0674	0.0353	0.112	-0.0493
-0.0493	0.112	0.0353	0.0674	0.6472	1	0.5822	0.1227	0.0312	0.0507
0.034	-0.0477	0.0562	0.0806	0.0674	0.5822	1	0.579	0.0647	0.0944
0.0169	-0.0162	0.0012	0.0562	0.0353	0.1227	0.579	1	0.6737	0.0465
-0.0441	0.0691	-0.0162	-0.0477	0.112	0.0312	0.0647	0.6737	1	0.5481
0.0284	-0.0441	0.0169	0.034	-0.0493	0.0507	0.0944	0.0465	0.5481	1

The optimization on an angle-parameterized rank-7 matrix yields the following output matrix $\rho(\theta^{*(7)})$:

1	0.5592	-0.0177	0.1085	0.0602	-0.0795	0.0589	0.018	-0.0734	0.0667
0.5592	1	0.5992	0.0202	0.0277	0.1123	-0.0652	-0.008	0.0797	-0.0734
-0.0177	0.5992	1	0.5464	0.0618	0.0401	0.0561	-0.012	-0.008	0.018
0.1085	0.0202	0.5464	1	0.5556	0.018	0.0834	0.0561	-0.0652	0.0589
0.0602	0.0277	0.0618	0.5556	1	0.5819	0.018	0.0401	0.1123	-0.0795
-0.0795	0.1123	0.0401	0.018	0.5819	1	0.5556	0.0618	0.0277	0.0602
0.0589	-0.0652	0.0561	0.0834	0.018	0.5556	1	0.5464	0.0202	0.1085
0.018	-0.008	-0.012	0.0561	0.0401	0.0618	0.5464	1	0.5992	-0.0177
-0.0734	0.0797	-0.008	-0.0652	0.1123	0.0277	0.0202	0.5992	1	0.5592
0.0667	-0.0734	0.018	0.0589	-0.0795	0.0602	0.1085	-0.0177	0.5592	1

We plot the fifth columns of the original matrix and of the rank-4 approximations in the left part of figure 2. It is clear that, due to the steep decorrelation pattern when moving off the diagonal, both 4-rank approximations are worse than with our earlier example. To obtain a better approximation we have to resort to 7-rank matrices, and the related plots are on the right hand side of figure 2. Keep in mind, however, that this is an extreme case and that in general rank-4 approximations are satisfactory in most situations.

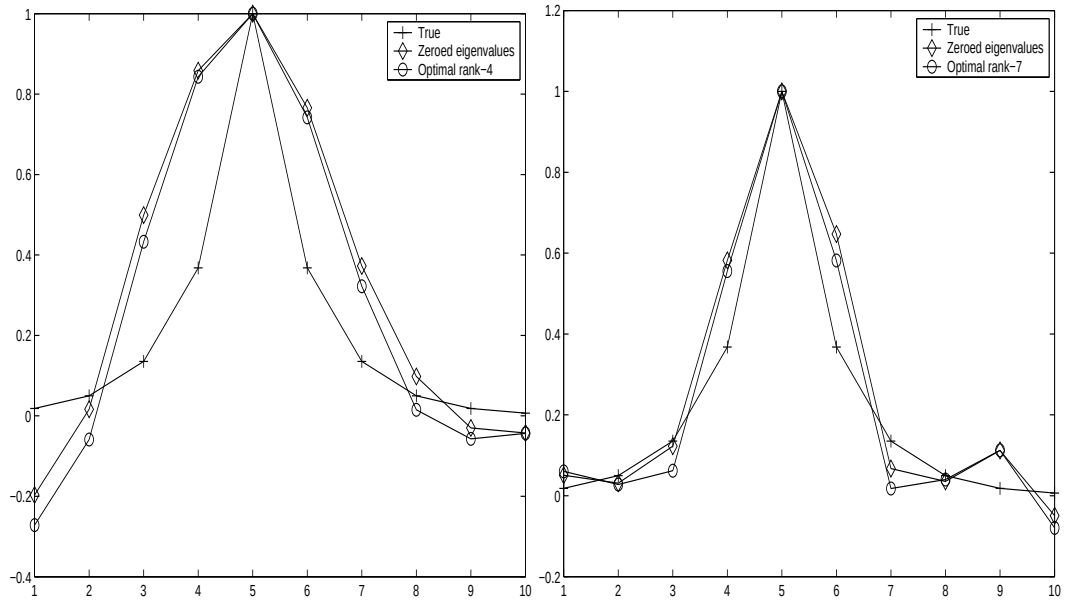

Figure 2: On the left: Fifth columns of $\hat{\rho}$, $\rho^{(4)}$, and $\rho(\theta^{*(4)})$. On the right: Fifth columns of $\hat{\rho}$, $\rho^{(7)}$, and $\rho(\theta^{*(7)})$

3.4 Reducing the angles

There are some problems left with the general angles parameterization. We pointed out in Section 3.1 that the θ parameters are redundant. Let us consider the full rank case $n = M$ for a start. In this case we have $M(M-1)$ angle parameters θ, whereas we just need half these parameters to characterize a correlation matrix. Recently, Rapisarda (2002) found an interesting solution to this problem, based on a geometric interpretation of the angles parameterization. Rapisarda (2002)

bases his interpretation on the idea of subsequent Jacobi rotations. Roughly speaking, it turns out that the correlation matrix can be obtained by mutual projections of versors obtained by subsequent rotations of a starting versor. The redundancy in the angles parameterization comes essentially from not fixing this initial versor. If we establish this initial versor to be for example $[1, 0, \ldots, 0]'$, then redundancy vanishes, and if further we take $\theta_{i,j} = 0$ for each $j \geq i$ when $i < n$, we obtain a "canonical" version of the angle parameterization that is as rich as the initial one in terms of correlation matrices. Incidentally, notice that now the number of nonzero angles is $M(M-1)/2$ in the full-rank case ($n = M$), exactly the number of different entries in a general full-rank correlation matrix. For example, in case $M = n = 5$, we have the angles matrix

0	0	0	0
$\theta_{2,1}$	0	0	0
$\theta_{3,1}$	$\theta_{3,2}$	0	0
$\theta_{4,1}$	$\theta_{4,2}$	$\theta_{4,3}$	0
$\theta_{5,1}$	$\theta_{5,2}$	$\theta_{5,3}$	$\theta_{5,4}$

to which corresponds a triangular matrix $B(\theta)$:

1	0	0	0	0
$b_{2,1}$	$b_{2,2}$	0	0	0
$b_{3,1}$	$b_{3,2}$	$b_{3,3}$	0	0
$b_{4,1}$	$b_{4,2}$	$b_{4,3}$	$b_{4,4}$	0
$b_{5,1}$	$b_{5,2}$	$b_{5,3}$	$b_{5,4}$	$b_{5,5}$

which is the Cholesky decomposition of the correlation matrix $\rho(\theta)$.

More generally, when the rank is smaller than the matrix size, $n < M$, we have a similar triangular matrix on top, followed below by a rectangular matrix. For example, in case $n = 5$ and $M = 10$, we have

0	0	0	0
$\theta_{2,1}$	0	0	0
$\theta_{3,1}$	$\theta_{3,2}$	0	0
$\theta_{4,1}$	$\theta_{4,2}$	$\theta_{4,3}$	0
$\theta_{5,1}$	$\theta_{5,2}$	$\theta_{5,3}$	$\theta_{5,4}$
$\theta_{6,1}$	$\theta_{6,2}$	$\theta_{6,3}$	$\theta_{6,4}$
$\theta_{7,1}$	$\theta_{7,2}$	$\theta_{7,3}$	$\theta_{7,4}$
$\theta_{8,1}$	$\theta_{8,2}$	$\theta_{8,3}$	$\theta_{8,4}$
$\theta_{9,1}$	$\theta_{9,2}$	$\theta_{9,3}$	$\theta_{9,4}$
$\theta_{10,1}$	$\theta_{10,2}$	$\theta_{10,3}$	$\theta_{10,4}$

and a similar structure holds for the $B(\theta)$ matrix.

The above standardization of the θ's (and the induced standardization on the B's) helps by reducing the number of parameters in the optimization and also in other respects. It is further possible to constrain the angles in suitable intervals in order to avoid further redundancy. These problems are discussed in Rapisarda (2002).

As an example, we consider again the matrix

$$\hat{\rho}_{i,j} = \exp[-|i-j|]$$

given earlier, with $M = 10$. If we take $n = 4$, the resulting canonical "triangularized" angles, replacing the partly redundant $\theta^{*(4)}$ given earlier, are

$\theta^{**(4)}$		
0	0	0
0.3387	0	0
1.0656	0.0221	0
1.6565	0.3039	0.6206
1.8519	0.8419	0.7275
1.6264	1.5406	0.9085
1.3974	1.8835	1.2993
1.3922	1.8061	1.9873
1.6417	1.4174	2.5489
1.7839	1.212	2.8069

The related matrices $\rho(\theta^{*(4)})$ and $\rho(\theta^{**(4)})$, although being very close, are not exactly equal, since the final result depends on approximations and details concerning the optimization procedure. Since by taking away the first (zeroed) row of this canonical θ's matrix one is left with a lower triangular matrix on top, we refer to this canonical θ parameterization as to the "triangular angles parameterization".

For more details and a geometric insight on this procedure see Rapisarda (2002) and Brigo, Mercurio and Rapisarda (2002).

References

[1] Brigo, D., Mercurio, F. (2001). *Interest-Rate Models: Theory and Practice*, Springer - Verlag, Berlin.

[2] Brigo, D., Mercurio, F. (2002). *Calibrating LIBOR*, Risk Magazine, 2002 January issue.

[3] Brigo, D., Mercurio, F., and Rapisarda, F. (2002). In progress.

[4] Morini, M. (2002). Private communication.

[5] Rapisarda, F. (2002). Parameterizing correlations: A geometric intepretation. Preprint.

[6] Rebonato, R. (1999c) On the Simultaneous Calibration of Multifactor Lognormal Interest Rate Models to Black Volatilities and to the Correlation Matrix. *The Journal of Computational Finance* 2, 5-27.

[7] Rebonato, R. (1999d) Volatility and Correlation. Wiley, Chichester.

[8] Rebonato, R., and Jäckel, P. (1999). The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. QUARC preprint, available at `http://www.rebonato.com`

[9] Schoenmakers, J., and Coffey, C. (2000). Stable implied calibration of a multi-factor LIBOR model via a semi-parametric correlation structure. Weierstrass Institute Preprint n. 611.

[10] Zhang, Z.-T., and Wu, L. (2001). Optimal Low-rank approximation of Correlation Matrices, working paper available at `http://www.math.ust.hk/∼malwu/`