ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Lord, Roger; Pelsser, Antoon

Working Paper

Level-Slope-Curvature - Fact or Artefact?

Tinbergen Institute Discussion Paper, No. 05-083/2

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Lord, Roger; Pelsser, Antoon (2005) : Level-Slope-Curvature - Fact or Artefact?, Tinbergen Institute Discussion Paper, No. 05-083/2, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/86522

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2005-083/2
Tinbergen Institute Discussion Paper

Level-Slope-Curvature - Fact or Artefact?

Roger Lord[1,3]
Antoon Pelsser[1,2]

[1] *Erasmus University Rotterdam, and Tinbergen Institute,*
[2] *ING Group, Amsterdam, and Econometric Institute, Rotterdam.*
[3] *Rabobank International, Utrecht.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Level-slope-curvature – fact or artefact?

Roger Lord[1]
Antoon Pelsser[2]

First version: September 29th, 2004
This version: September 6th, 2005

ABSTRACT

The first three factors resulting from a principal components analysis of term structure data are in the literature typically interpreted as driving the level, slope and curvature of the term structure. Using slight generalisations of theorems from total positivity, we present sufficient conditions under which level, slope and curvature are present. These conditions have the nice interpretation of restricting the level, slope and curvature of the correlation surface. It is proven that the Schoenmakers-Coffey correlation matrix also brings along such factors. Finally, we formulate and corroborate our conjecture that the order present in correlation matrices causes slope.

Keywords: Principal components analysis, correlation matrix, total positivity, oscillation matrix, Schoenmakers-Coffey matrix.

The authors would like to express their gratitude to Abdel Lantere for some valuable insights. Thanks also go out to Allan Pinkus, Peter van de Ven, and finally Jan Brinkhuis, Maarten Janssen and Thierry Post for the organisation of the Unfinished Manuscripts seminar at the Erasmus University of Rotterdam, where the idea for this paper was initially presented by the second author.

[1] Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: lord@few.eur.nl, Tel. +31-(0)10-4088935) and Modelling and Research (UC-R-355), Rabobank International, P.O. Box 17100, 3500 HG Utrecht, The Netherlands (e-mail: roger.lord@rabobank.com, Tel. +31-(0)30-2166566).

[2] ING Group, Corporate Insurance Risk Management, P.O. Box 810, 1000 AV Amsterdam, The Netherlands and Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: pelsser@few.eur.nl, Tel. +31-(0)10-4081258).

1. Introduction

In an attempt to parsimoniously model the behaviour of the interest rate term structure, many studies find that using the first three principal components of the covariance or correlation matrix already accounts for 95-99% of the variability, a result first noted for interest rate term structures by Steeley [1990] and Litterman and Scheinkman [1991]. These results were also found to hold for the term structure of copper futures prices by Cortazar and Schwartz [1994], and also for the multiple-curve case, as shown by Hindanov and Tolmasky [2002].

This paper does not deal with the question of how many factors one should use to model the interest rate term structure, or any term structure for that matter, but addresses the shape of the first three factors. The shape hereof is such that many authors, starting from Litterman and Scheinkman, have attached an interpretation to each of these three factors. The first factor, or indeed eigenvector of the covariance or correlation matrix, is usually relatively flat. As such it is said to determine the *level* or *tilt* of the term structure. The second, which has opposite signs at both ends of the term structure, can be interpreted as determining the *slope* or *trend*. The third factor finally, having equal signs at both ends of the maturity spectrum, but an opposite sign in the middle, is said to determine the *curvature*, *twist* or *butterfly* of the term structure.

A question that comes to mind is whether the observed pattern is caused by some fundamental structure within term structures, or whether it is merely an artefact of principal components analysis (PCA). Alexander [2003] in fact claims that "… the interpretation of eigenvectors as trend, tilt and curvature components is one of the stylised facts of all term structures, particularly when they are highly correlated". In this paper we investigate sufficient conditions under which the level-slope-curvature effect occurs. To the best of our knowledge only one article has so far tried to mathematically explain this level-slope-curvature effect in the context of a PCA of term structures, namely that of Forzani and Tolmasky [2003]. They demonstrate that when the correlation between two contracts maturing at times t and s is of the form $\rho^{|t-s|}$, where ρ is a fixed positive correlation, the observed factors are perturbations of cosine waves with a period which is decreasing in the number of the factor under consideration. This correlation function is widely used as a parametric correlation function in e.g. the LIBOR market model, see Rebonato [1998].

We formulate the level-slope-curvature effect differently than Forzani and Tolmasky. As noted, the first factor is quite flat, the second has opposite signs at both ends of the maturity spectrum, and the third finally has the same sign at both ends, but has an opposite sign in the middle. This observation leads us to consider the number of sign changes of each factor or eigenvector. If the first three factors have respectively zero, one and two sign changes, we say that we observe level, slope and curvature.

Using a concept named total positivity, Gantmacher and Kreĭn considered the spectral properties of totally positive matrices in the first half of the twentieth century. One of the properties of a sub-class of these matrices, so-called oscillation matrices, is indeed that the nth eigenvector of such a matrix has exactly n-1 sign changes. These results can be found e.g. in their book [1960, 2002]. With a minor generalisation of their theorems, we find sufficient conditions under which a term structure indeed displays the level-slope-curvature effect. The conditions have the nice interpretation of placing restrictions on the level, slope and curvature of the correlation curves.

Subsequently we turn to a correlation parameterisation which was recently proposed by Schoenmakers and Coffey [2003]. In matrix theory the resulting correlation matrix is known as a *Green's* matrix. The exponentially decreasing correlation function considered by Forzani and Tolmasky is contained as a special case of the Schoenmakers-Coffey parameterisation. The resulting correlation matrix has the nice properties that correlations decrease when moving away from the diagonal term along a row or a column. Furthermore, the correlation between equally

spaced rates rises as their expiries increase. These properties are observed empirically in correlation matrices of term structures. Gantmacher and Kreĭn derived necessary and sufficient conditions for a Green's matrix to be an oscillation matrix, and hence to display level, slope and curvature. The Schoenmakers-Coffey parameterisation satisfies these restrictions, and hence also displays this effect. This actually confirms and proves a statement by Lekkos [2000], who numerically showed that when continuously compounded forward rates are independent, the resulting correlation matrix of zero yields displays level, slope and curvature.

Unfortunately total positivity and related concepts only provide a partial explanation of the level, slope and curvature phenomenon. We therefore end the paper with a conjecture that an ordered correlation matrix with positive elements will always display level and slope. This conjecture is not proven, but is corroborated by results from a simulation study.

The paper is organised as follows. In section 2 we first briefly introduce the terminology used in principal components analysis, and perform an empirical analysis of Bundesbank data[3], which contains interest rate data for the Euro market from 1972 onwards. Observing the same empirical pattern as in other studies, we mathematically formulate our criteria for the level-slope-curvature effect. In section 3 we present and slightly modify some theorems from theory on total positivity, which will lead to sufficient conditions for level, slope and curvature. We also provide an interpretation of these conditions. In the fourth section we turn to the Green's or Schoenmakers-Coffey correlation matrices, and show that they satisfy the conditions formulated in the third section. In the fifth section we consider sign regularity, a concept extending total positivity and end with our conjecture that positive and ordered correlation matrices will always display level and slope. Section 6 concludes.

2. Problem formulation

As stated before, we will in this paper investigate conditions under which we observe the level-slope-curvature effect. Before mathematically formulating the problem, we will, for the purpose of clarity, briefly review some concepts of principal components analysis in the first paragraph of this section. For a good introduction to PCA we refer the reader to Jackson [2003]. In the second paragraph we will review some empirical studies and conduct a PCA on historical data obtained from the Bundesbank database to illustrate the level-slope-curvature effect we will be analysing. Finally, the third and final paragraph will formulate our problem mathematically.

2.1. Principal components analysis

Suppose we are considering a model with N random variables, in our case prices of contracts within the term structure. These random variables will be contained in a column vector[4] $\mathbf{X}$. For notational purposes we will assume that these random variables are centered. The goal of PCA is to describe the data we have with $K < N$ orthogonal random variables, so-called principal components, which will be linear combinations of the original stochastic variables. We denote the kth principal component as:

$$Y_k = \mathbf{X}^T \mathbf{w}_k \text{ for } k = 1, \ldots, N \tag{2.1}$$

Having determined all weight vectors $\mathbf{w}_i$ for $i = 1, \ldots, k$, the weight vector w_k follows from the following maximisation problem:

[3] This data can be obtained by selecting the daily term structure of interest rates from the time series database, subsection capital market, at http://www.bundesbank.de/statistik/statistik_zeitreihen.en.php.

[4] As a matter of notation, vectors and matrices will be typeset in bold.

$$\max_{\mathbf{w}_k \in \mathbb{R}^N} \mathrm{Var}(\mathbf{X}^T \mathbf{w}_k) \quad \text{s.t.} \quad \forall_{i \le k}\ \mathbf{w}_i^T \mathbf{w}_k = 1_{[i=k]} \tag{2.2}$$

We maximise the variance of each principal component, so that each component describes as large a part of the total variability as possible. The restriction that each weight vector must have length 1 only serves to remove the indeterminacy; hence the vectors w_k form an orthonormal system. It is in fact easy to prove that (2.2) is solved by setting $\mathbf{w}_k$ to be the kth eigenvector, i.e. associated with the kth largest eigenvalue λ_k, of the covariance matrix $\mathbf{\Sigma}$. The variance of the kth principal component is therefore equal to $\mathrm{Var}(\mathbf{X}^T \mathbf{w}_k) = \mathbf{w}_k^T \mathbf{\Sigma} \mathbf{w}_k = \lambda_k$. In PCA the proportion of variance explained by the kth factor is calculated as:

$$\frac{\mathrm{Var}(Y_k)}{\mathrm{Var}(Y_1 + \ldots + Y_N)} = \frac{\lambda_k}{\lambda_1 + \ldots + \lambda_N} \tag{2.3}$$

Note that all eigenvalues of a covariance (or correlation) matrix obtained from data will be positive, since any proper covariance matrix will be positive definite[5].

Let us now denote the spectral decomposition of the covariance matrix as $\mathbf{\Sigma}=\mathbf{W\Lambda W}^T$. Then it is obvious that we can write $\mathbf{Y} = \mathbf{W}^T\mathbf{X}$, and this relationship can be inverted to find $\mathbf{X} = \mathbf{WY}$. In general we will however use K < N principal components, so that we will have:

$$\mathbf{Y}_{(K)} = \mathbf{W}_{(K)}^T \mathbf{X} \tag{2.4}$$

where the subscript (K) indicates that we are only using the first K principal components. We cannot invert this relation directly, but if we regress $\mathbf{Y}_{(K)}$ on $\mathbf{X}$, we can show that the ordinary least squares estimator for $\mathbf{A}$ in $\mathbf{X} = \mathbf{AY}_{(K)} + \boldsymbol{\varepsilon}$ is $\mathbf{W}_{(K)}$, so that we finally have:

$$\mathbf{X} = \mathbf{W}_{(K)} \mathbf{Y}_{(K)} + \boldsymbol{\varepsilon} \tag{2.5}$$

As a final note, we know by definition from (2.1) that the jth entry of a weight vector $\mathbf{w}_k$ contains the weight with which X_j is embedded within the kth principal component. Within PCA, the scaled eigenvectors $\sqrt{\lambda_k}\,\mathbf{w}_k$ are called *factors*, and its entries are referred to as *factor loadings*.

2.2. Empirical results

As mentioned in the introduction, many studies have dealt with a PCA of term structures, in particular term structures of interest rates. Although in this paper we will mainly focus on the level-slope-curvature effect for an arbitrary covariance matrix, and the work will be more mathematical than empirical, it is nevertheless interesting to review a number of results from recent empirical studies that could be important for this paper. After this brief review we will investigate whether we find the level-slope-curvature pattern in the Bundesbank dataset.

We first mention a recent study by Lardic, Priaulet and Priaulet [2003]. Noticing that many studies use quite different methodologies, they pose a number of questions in their paper. The

[5] There are situations where one can obtain an estimate for a covariance matrix that is not positive definite, e.g. when one has missing data for one or more of the observed variables. However, any proper covariance matrix must be positive definite, since otherwise we can construct a linear combination of our random variables that has a negative variance. This clearly cannot be the case.

first question is whether one should use interest rate changes or levels as an input to a PCA. Naturally, interest rate levels are much more correlated than interest rate changes. They find that interest rate changes are stationary and conclude that therefore a PCA should be implemented with interest rate changes. Secondly, they investigate whether one should use centered changes, or standardised changes, i.e. whether a PCA should be conducted on a covariance or a correlation. Since the volatility term structure is typically not flat, but either hump-shaped or hockey stick shaped, there certainly is a difference between both methods. They conclude that a PCA with a covariance matrix will overweight the influence of the more volatile short term rates, and hence that one should use a PCA only with correlation matrices. Later on we will show that, under certain restrictions, our definition of level, slope and curvature will be such that it is irrelevant whether we use a covariance or a correlation matrix. Their final questions address whether the results of a PCA are dependent on the rates that are included in the analysis, and on the data frequency. Both aspects certainly affect the results one obtains, but we feel these questions are less important, as they depend on the application under consideration.

The second study we mention is that of Lekkos [2000]. He criticises the conclusion of many authors, starting from Steeley [1990] and Litterman and Scheinkman [1991], that three factors, representing the level, slope and curvature of the term structure, are sufficient to describe the evolution of interest rates. He claims that the results are mainly caused by the fact that most studies focus on zero yields, as opposed to (continuously compounded) forward rates. We will explain this now. In mathematical models the price of a zero-coupon bond is often written as:

$$\begin{aligned} P(t,T) &= \exp\big(-R(t,T)\cdot(T-t)\big) \\ &= \exp\left(-\int_0^T f(t,u)\,du\right) \\ &= \exp\big(-\alpha\big(f(t,t,t+\alpha)+\ldots+f(t,T-\alpha,T)\big)\big) \\ &= \big(1+\alpha F(t,t,t+\alpha)\big)^{-1}\cdot\ldots\cdot\big(1+\alpha F(t,T-\alpha,T)\big)^{-1} \end{aligned} \tag{2.6}$$

where P(t,T) is the time t price of a zero-coupon bond paying 1 unit of currency at time T. The first formulation uses the zero yield R(t,T) of the zero-coupon bond. The second through fourth formulations are in terms of forward rates. The second uses instantaneous forward rates, typically only used in mathematical models such as the Heath, Jarrow and Morton framework. The third is in terms of continuously compounded forward rates, where f(t,T,S) indicates the time t forward rate over [T,S]. Finally, the fourth formulation uses discretely compounded forward rates, which is the way interest rates are typically quoted in the market. Lekkos works with the third formulation. Relating the zero yields to these forward rates, where we use a fixed tenor equal to α, we find that the zero yields are averages of these continuously compounded forward rates:

$$R(t,T) = \tfrac{\alpha}{T-t}\big(f(t,t,t+\alpha)+\ldots+f(t,T-\alpha,T)\big) \tag{2.7}$$

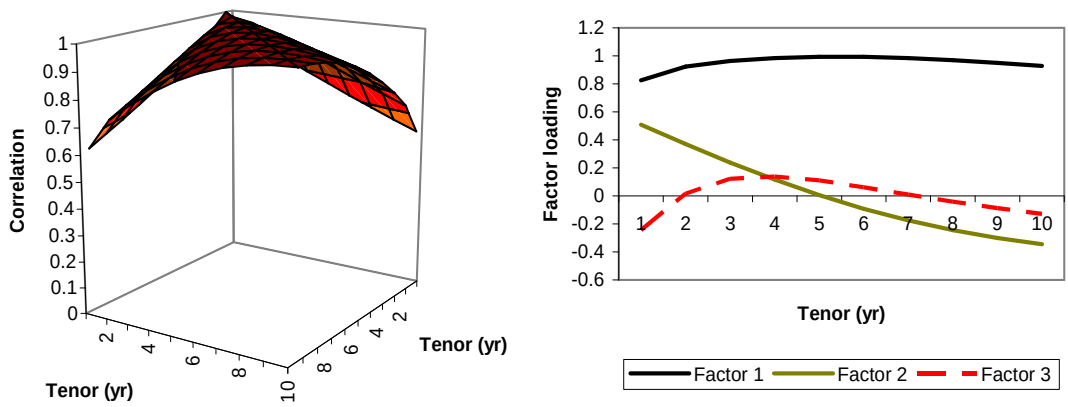
Lekkos claims that the high correlation found for interest rate changes is mainly caused by this averaging effect in (2.7), and that we should therefore analyse the spectral structure of α-forward rates instead. In a numerical example he shows that when these α-forward rates are independent, the correlation matrix of the zero yields still displays the level-slope-curvature effect. We will in fact prove this result later on, in section 4. Although forward rates are not found to be independent in his empirical analysis, the spectral structure for α-forward rates he finds is quite different than that of the zero yields. The second and third factors cannot be interpreted as driving the slope and curvature of the term structure, and furthermore up to five factors are required to account for 95% of the total variation.

The final study we consider is that of Alexander and Lvov [2003]. One of the things considered in their paper are the statistical properties of a time series of discretely compounded forward rates. The time series are obtained from quoted rates via three different yield curve fitting techniques, namely two spline methods, and the Svensson[6] [1994] method. The functional form of an instantaneous forward rate with time to maturity T in the Svensson model is given by:

$$f(T) = \beta_0 + \beta_1 \exp(-\tfrac{T}{\tau_1}) + \beta_2 \tfrac{T}{\tau_1} \exp(-\tfrac{T}{\tau_1}) + \beta_3 \tfrac{T}{\tau_2} \exp(-\tfrac{T}{\tau_2}) \tag{2.8}$$

where the six parameters β_0 through β_3 and τ_1 and τ_2 have to be estimated from the data. The equation (2.8) is an addition of an asymptotic value and several negative exponentials, which are able to create humps or U-shapes. This model is able to capture several facts found empirically in the term structure of forward rates. Alexander and Lvov conclude that the choice of the yield curve fitting technique affects the correlation matrix much more than the choice of sample size. In their study they find that the Svensson curve gives the best overall sample fit, and through its parametric form it also yields the smoothest correlation matrices. As an interesting note, the first three factors from their PCA can all be interpreted as driving the level, slope and curvature of the term structure, contrary to the study of Lekkos. Although Alexander and Lvov use discretely compounded forward rates, whereas Lekkos uses continuously compounded rates, we would not expect this to affect the results so markedly. Therefore, we suspect that the differences between Alexander and Lvov's results and those of Lekkos can mainly be attributed to the difference in yield-curve fitting technique. Lekkos uses the bootstrap method, which can give rise to less smooth yield curves, and hence also to a less smooth correlation surface.

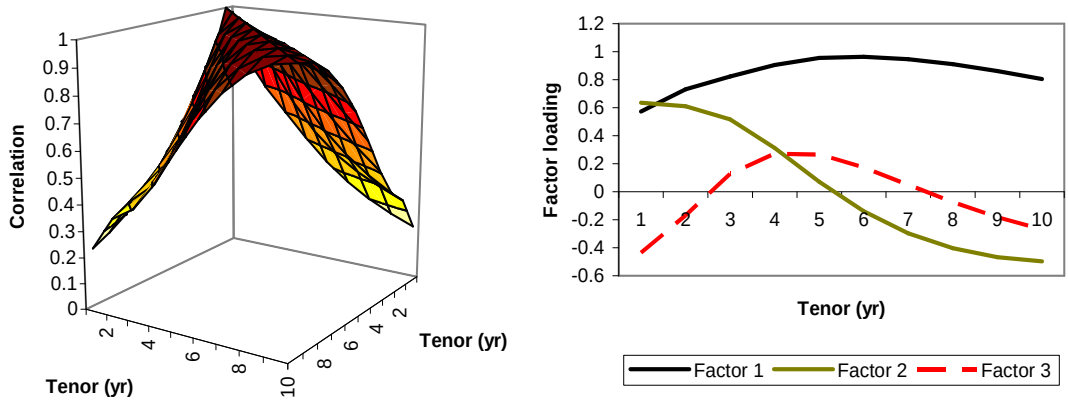
Using these insights, we will now ourselves conduct a PCA of Bundesbank data, which contains estimated Svensson curves for the Euro market from 1972 onwards. Until 1997 the curves have been estimated on a monthly basis. From August 1997 onwards, the curves are available on a daily basis. As we are only interested in reproducing the level-slope-curvature effect here, we ignore both the sample size and frequency issues, and use all end-of-month data from January 1980[7] up to and including June 2004. We calculated the correlations between



Graph 1: Estimated correlations between and first three factors of monthly log-returns on 1-10 year zero yields

[6] The Svensson model is also often referred to as the extended Nelson and Siegel model, as it is an extension of the original model by Nelson and Siegel [1987].

[7] Data from the seventies was not included as it changed the correlation estimates severely.



Graph 2: Estimated correlations between and first three factors of monthly log-returns on discretely compounded annual forward rates, with maturities ranging from 1-10 years

the correlations between logarithmic returns on both zero yields, with tenors from 1 to 10 years, as well as on discretely compounded annual forward rates, with maturities also ranging from 1 to 10 years. The estimated correlation surfaces, as well as the first three factors following from a PCA, can be found in graphs 1 and 2 on this and the previous page.

We indeed notice that that the resulting correlation surfaces are quite different for the zero yields than for the forward rates. The relation between zero yields and forward rates in (2.6) indicates that zero yields are averages of the forward rates. This relation by itself causes the correlations between the zero yields (or log-returns hereof) to be higher than those between the forward rates. For the full sample period, we find that in the zero yield case, the first three factors explain up to 99% of the total variability, whereas this number is reduced to 91% in the case of forward rates. However, if we consider a sample period similar to that of Lekkos, e.g. 1987-1995, we find that the first three factors explain more than 97% of the total variability in forward rates, which is a much higher number than found in this study. Furthermore, for any sample period we still find the level, slope and curvature pattern, contrary to Lekkos' study. The difference could, as mentioned before, be caused by the difference in yield-curve fitting technique.

We have seen that the observed pattern does not always occur, i.e. in the case of the estimated forward rate correlation matrices by Lekkos. A natural question one could ask however, is whether the noticed pattern always occurs in the case of highly correlated and ordered stochastic systems. To this end consider the following artificially constructed correlation matrix:

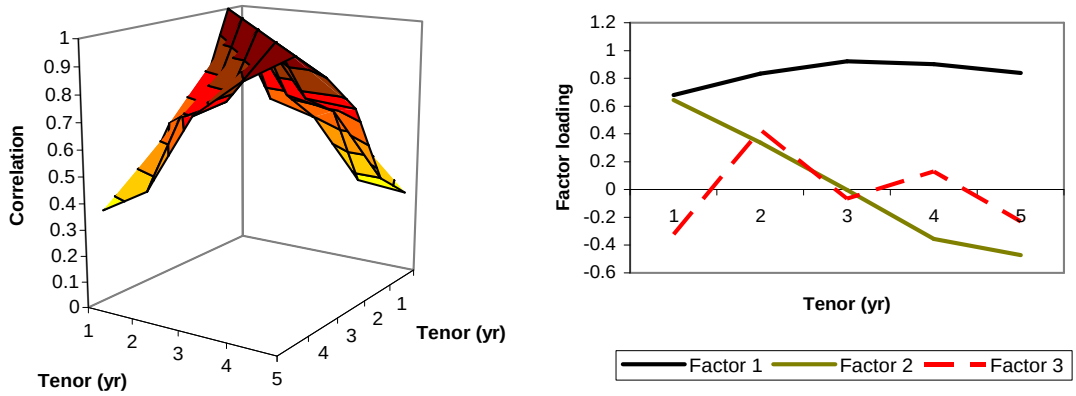
$$\mathbf{R} = \begin{pmatrix} 1 & 0.649 & 0.598 & 0.368 & 0.349 \\ 0.649 & 1 & 0.722 & 0.684 & 0.453 \\ 0.598 & 0.722 & 1 & 0.768 & 0.754 \\ 0.368 & 0.684 & 0.768 & 1 & 0.896 \\ 0.349 & 0.453 & 0.754 & 0.896 & 1 \end{pmatrix} \tag{2.9}$$

The matrix is a proper correlation matrix, and furthermore it satisfies certain properties which are typically found in empirical interest rate correlation matrices:

i) $\rho_{i,j+1} \leq \rho_{ij}$ for $j \geq i$, i.e. correlations decrease when we move away from the diagonal;

ii) $\rho_{i,j-1} \leq \rho_{ij}$ for $j \leq i$, same as i);

iii) $\rho_{i,i+j} \leq \rho_{i+1,i+j+1}$, i.e. the correlations increase when we move from northeast to southwest.



Graph 3: Artificially constructed correlation matrix which does not display curvature

In words property (iii) means that the correlation between two adjacent contracts or rates increases as the tenor of both contracts increases. For instance, the 4 and 5 year rate are more correlated than the 1 and 2 year rate. Hence, the matrix in (2.9) clearly is a correlation matrix of an ordered and highly correlated system, and could certainly be the correlation matrix of a term structure. The above graph, in which its correlation surface and first three factors are depicted, demonstrates however that conditions (i)-(iii) (i.e. (i), (ii) and (iii)) are insufficient for a matrix to display level, slope and curvature. Although the first two eigenvectors can certainly be interpreted as level and slope, the third eigenvector displays a different pattern than we usually find.

Concluding, although the correlation structure between either consecutive zero yields or forward rates is quite different, we find the level-slope-curvature effect in both cases, provided we use a smooth enough yield-curve fitting technique. Finally, the fact that we have a highly correlated system, in combination with certain properties that empirical interest rate correlation matrices satisfy, is not enough for the correlation matrix to display the observed pattern. Additional or different conditions are required, something we will investigate in the next section. Using these empirical findings we will first mathematically formulate level, slope and curvature in the next paragraph.

2.3. Mathematical formulation of level, slope and curvature

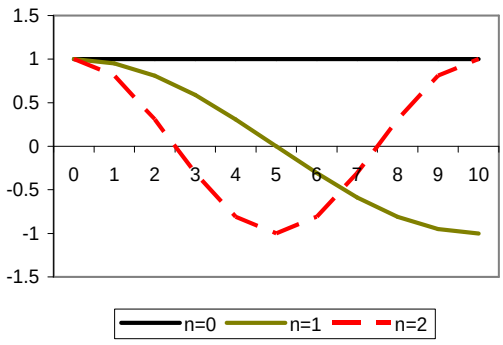
Regardless of whether we consider correlations between (returns of) zero yields or forward rates, we have seen the presence of level, slope and curvature. Before analysing this effect, we have to find a proper mathematical description. Forzani and Tolmasky [2003] analysed the effect in case the correlation structure between contracts maturing at times t and s is equal to $\rho^{|t-s|}$. Working with a continuum of tenors on [0,T], they analyse the eigensystem of:

$$\int_0^T \rho^{|y-x|} f(y)\,dy = \lambda f(x) \tag{2.10}$$

This problem is analogous to determining the eigenvectors of the correlation matrix, when we consider a discrete set of tenors. When ρ approaches 1, they find that the nth eigenfunction (associated with the nth largest eigenvalue), approaches the following function:

$$f(x) = \begin{cases} \cos(\frac{n\pi x}{T}) - \frac{2T\ln\rho}{n^2\pi^2 + T^2(\ln\rho)^2} & \text{n even} \\ \cos(\frac{n\pi x}{T}) & \text{n odd} \end{cases} \tag{2.11}$$

We notice that the first factor, corresponding to n = 0, approaches a constant, and hence will be relatively flat when the contracts in the term structure are highly correlated. Similarly, we notice that the nth eigenfunction has a period equal to 2T/n. Hence, the second factor (n = 1) will have half a period on [0,T], and the third factor (n = 2) will have a full period on [0,T]. In the following graph we display the functions 1, $\cos\left(\frac{\pi x}{T}\right)$ and $\cos\left(\frac{2\pi x}{T}\right)$ on [0,T] where T = 10:



Graph 4: Limits of eigenfunctions for $\rho^{|t-s|}$ when $\rho \to 1$

Indeed, these limiting functions do resemble our notion of level, slope and curvature. The true eigenfunctions are perturbations of these cosine waves.

For the exponentially decaying correlation function the analysis is much facilitated, as the eigenfunctions can be calculated explicitly. We are not able to do this in general. Therefore we use another definition of level, slope and curvature, which will not require the knowledge of the explicit form of the eigenvectors or eigenfunctions. We notice in graphs 1 and 2 that the first factor is quite flat, and in fact has equal sign for all tenors. The second factor has opposite signs at both ends of the maturity range. Finally, the third has equal signs at both extremes, but has an opposite sign in the middle. If we therefore look at the number of times each factor or eigenvector changes sign, we notice that the first factor has zero sign changes, the second has one, and the third has two. This does not give a full description of what we perceive as level, slope and curvature. For instance, if in graph 3 the third factor would be shifted slightly upwards, it would only have two sign changes, although it would still be dissimilar from the usual pattern. In all empirical studies we have seen however, our definition correctly signals the presence of level, slope and curvature, so that we expect it to be sufficient.

For a continuous eigenfunction, the number of sign changes is easily defined as the number of zeroes of this function. We will however mainly be working with a discrete set of tenors, which calls for a slightly different definition. For an $N \times 1$ vector $\mathbf{x}$ we mathematically define the number of sign changes as follows:

- $S^{-}(\mathbf{x})$ - the number of sign changes in $x_1, \ldots, x_N$ with zero terms discarded;
- $S^{+}(\mathbf{x})$ - the maximum number of sign changes in $x_1, \ldots, x_N$, with zero terms arbitrarily assigned either +1 or –1.

Both functions will only give a different number when the eigenvector contains zeroes and the non-zero elements at either side of a sequence of zeroes have the same sign. In the next chapter the distinction between both definitions will ultimately not be that important, as the sufficient conditions under which we will find the level-slope-curvature effect will imply that both definitions will give the same result when applied to the eigenvectors at hand. Ignoring zero terms

within an eigenvector, we therefore define level, slope and curvature as the following sign-change pattern within the first three eigenvectors:

- Level: $S^-(\mathbf{x}^1) = 0$
- Slope: $S^-(\mathbf{x}^2) = 1$
- Curvature: $S^-(\mathbf{x}^3) = 2$

where $\mathbf{x}^i$ is the ith eigenvector. In the next section we will consider total positivity theory, which will provides us with sufficient conditions under which we find level-slope-curvature.

3. Sufficient conditions for level, slope and curvature

In this section we turn to theory on total positivity, which, for our formulation of the level-slope-curvature effect, will yield the right tools to clarify its occurrence. In the first paragraph we introduce some notation and concepts that we will require in the remainder of this section. The second paragraph reviews some results from total positivity theory. Minor generalisations hereof will yield sufficient conditions under which level, slope and curvature occur. In the paragraph hereafter we rewrite these conditions, and show we can interpret them as being conditions on the level, slope and curvature of the correlation surface. We will mainly work with a discrete set of tenors, although we also touch upon the case where we have a continuum of tenors. The continuous case will greatly facilitate the interpretation of the conditions we find.

3.1. Notation and concepts

Before turning to some theorems from total positivity theory, we need to introduce some notation and concepts. First of all we will be dealing with covariance or correlation matrices. A covariance matrix $\mathbf{\Sigma}$ of size $N \times N$ satisfies the following properties:

1. $\mathbf{\Sigma}$ is symmetric, that is $\mathbf{\Sigma} = \mathbf{\Sigma}^T$;
2. $\mathbf{\Sigma}$ is positive definite, i.e. for any non-zero vector $\mathbf{x} \in \mathbb{R}^N$ we have $\mathbf{x}^T\mathbf{\Sigma}\mathbf{x} > 0$.

Any matrix satisfying these properties is invertible and can be diagonalized as $\mathbf{\Sigma} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^T$, where the eigenvectors of the matrix are contained in $\mathbf{X}$, and the eigenvalues in $\mathbf{\Lambda}$. All eigenvalues are furthermore strictly positive. The correlation matrix $\mathbf{R}$ associated with $\mathbf{\Sigma}$ is obtained as:

$$\mathbf{R} = \mathrm{diag}(\mathbf{\Sigma})^{-1/2}\,\mathbf{\Sigma}\,\mathrm{diag}(\mathbf{\Sigma})^{-1/2} \tag{3.1}$$

where diag($\mathbf{\Sigma}$) is a matrix of the same dimensions as $\mathbf{\Sigma}$, containing its diagonal and zeroes everywhere else. Naturally $\mathbf{R}$ is also a covariance matrix.

The theorems in the next section will require the following concepts. For a given positive integer N we define:

$$I_{p,N} = \left\{\mathbf{i} = (i_1,\ldots,i_p) \mid 1 \le i_1 < \ldots < i_p \le N\right\} \tag{3.2}$$

where of course $1 \le p \le N$. When $\mathbf{\Sigma}$ is an $N \times N$ matrix, we define for $\mathbf{i}, \mathbf{j} \in I_{p,N}$:

$$\mathbf{\Sigma}_{[p]}(\mathbf{i},\mathbf{j}) = \mathbf{\Sigma}\begin{pmatrix} i_1,\ldots,i_p \\ j_1,\ldots,j_p \end{pmatrix} = \det(a_{i_k j_\ell})_{k,\ell=1}^{p} \tag{3.3}$$

In terms of covariance matrices, definition (3.3) means we are taking the determinant of the covariance matrix between the interest rates indexed by vector **i**, and those indexed by vector **j**. The pth compound matrix $\mathbf{\Sigma}_{[p]}$ is defined as the $\binom{N}{p} \times \binom{N}{p}$ matrix with entries equal to $\left(\mathbf{\Sigma}_{[p]}(\mathbf{i},\mathbf{j})\right)_{\mathbf{i},\mathbf{j}\in I_{p,N}}$, where the $\mathbf{i} \in I_{p,N}$ are arranged in lexicographical order, i.e. $\mathbf{i} \geq \mathbf{j}$ ($\mathbf{i} \neq \mathbf{j}$) if the first non-zero term in the sequence $i_1 - j_1, \ldots, i_p - j_p$ is positive.

3.2. Sufficient conditions via total positivity

Before turning to the theory of total positivity, we will solve the level problem. Perron's theorem, which can be found in most matrix algebra textbooks, deals with the sign pattern of the first eigenvector.

Theorem 1 – Perron's theorem
Let A be an $N \times N$ matrix, all of whose elements are strictly positive. Then A has a positive eigenvalue of algebraic multiplicity equal to 1, which is strictly greater in modulus than all other eigenvalues of A. Furthermore, the unique (up to multiplication by a non-zero constant) associated eigenvector may be chosen so that all its components are strictly positive. □

The result of the theorem only applies to matrices with strictly positive elements. Since the term structures we are investigating are highly correlated, this is certainly not a restriction for our purposes. The result is valid for any square matrix, not only for symmetric positive definite matrices. As long as all correlations between the interest rates are positive, this means that the first eigenvector will have no sign changes.

This has solved the level problem. For the sign-change pattern of other eigenvectors we have to turn to the theory of total positivity. The results in this paragraph mainly stem from a paper by Gantmacher and Kreĭn [1937], which, in an expanded form, can be found in Gantmacher and Kreĭn [1960, 2002]. Most results can also be found in the monograph on total positivity by Karlin [1968]. For a good and concise overview of the theory of total positivity we refer the reader to Ando [1987] and Pinkus [1995]. The latter paper gives a good picture of the historical developments in this field, and the differences between the matrix and the kernel case.

A square matrix **A** is said to be totally positive (sometimes totally non-negative, TP), when for all $\mathbf{i}, \mathbf{j} \in I_{p,N}$ and $p \leq N$, we have:

$$\mathbf{A}_{[p]}(\mathbf{i},\mathbf{j}) \geq 0 \tag{3.4}$$

In the case of covariance matrices, this means that we require the covariance matrix between **i** and **j** to have a non-negative determinant. When $\mathbf{i} = \mathbf{j}$ this will clearly be the case, as the resulting matrix is itself a covariance matrix, and will be positive definite. In the other cases the meaning of this condition is less clear. In the next paragraph we will spend some time on interpreting these conditions. If strict inequality holds then we say that the matrix is strictly totally positive (STP). Furthermore, we say that a matrix is TP_k if (3.4) holds for $p = 1, \ldots, k$, and we define STP_k in a similar fashion. Hence, an $N \times N$ matrix is TP when it is TP_N, and STP when it is STP_N. Gantmacher and Kreĭn proved the following theorem for general STP matrices. A full version of their theorem also considers the so-called variation-diminishing property of such matrices, but we

will here only deal with the sign-change pattern of such matrices. We reformulate their theorem for covariance matrices that are not necessarily STP, but only STP_k. Reading their proof shows that it can be altered straightforwardly to cover this case. For completeness we have included the proof in the appendix.

Theorem 2 – Sign-change pattern in STP_k matrices
Assume $\mathbf{\Sigma}$ is an $N \times N$ positive definite symmetric matrix (i.e. a valid covariance matrix) that is STP_k. Then we have $\lambda_1 > \lambda_2 > \ldots > \lambda_k > \lambda_{k+1} \geq \ldots \lambda_N > 0$, i.e. at least the first k eigenvalues are simple. Denoting the jth eigenvector by $\mathbf{x}^j$, we have $S^-(\mathbf{x}^j) = S^+(\mathbf{x}^j) = j-1$, for $j = 1, ..., k$.

Proof: See the appendix. □

A consequence of theorem 2 is that a sufficient condition for a correlation matrix to display level, slope and curvature, is for it to be STP_3. Naturally all principal minors of a covariance matrix are determinants of a covariance matrix, and hence will be strictly positive. It is however not immediately clear what the remaining conditions mean – we will find an interpretation hereof in the following paragraph. The conditions in theorem 2 can be relaxed somewhat further via the concept of an oscillation or oscillatory matrix, again due to Gantmacher and Kreĭn. The name oscillation matrix arises from the study of small oscillations of a linear elastic continuum, e.g. a string or a rod. An $N \times N$ matrix $\mathbf{A}$ is an oscillation matrix if it is TP and some power of it is STP. As in theorem 2, we slightly alter the original theorem by using the concept of an oscillation matrix of order k.

Theorem 3 – Oscillation matrix of order k
Akin to the concept of an oscillation matrix, we define an oscillation matrix of order k. An $N \times N$ matrix $\mathbf{A}$ is oscillatory of the order k if:

1. $\mathbf{A}$ is TP_k;
2. $\mathbf{A}$ is non-singular;
3. For all $i = 1, ..., N-1$ we have $a_{i,i+1} > 0$ and $a_{i+1,i} > 0$.

For oscillatory matrices of the order k, we have that $\mathbf{A}^{N-1}$ is STP_k.

Proof: See the appendix. □

Gantmacher and Kreĭn proved theorem 3 and its converse for the STP case. As we are only interested in sufficient conditions for level, slope and curvature, we do not prove the converse. The proof of theorem 3 is included in the appendix for completeness, although the original proof carries over almost immediately.

Corollary 1
In theorem 2 we can replace the condition that the matrix is STP_k with the requirement that some finite power of it is oscillatory of order k.

Proof:
Suppose $\mathbf{\Sigma}$ is a positive definite symmetric N x N matrix, for which $\mathbf{\Sigma}^i$ is oscillatory of order k. As the matrix is invertible, we can write $\mathbf{\Sigma} = \mathbf{X\Lambda X}^T$, and hence:

$$\mathbf{\Sigma}^{i(N-1)} = \mathbf{X\Lambda}^{i(N-1)}\mathbf{X}^T \tag{3.5}$$

so that $\mathbf{\Sigma}^{i(N-1)}$ has the same eigenvectors as $\mathbf{A}$. Since $\mathbf{\Sigma}^{i(N-1)}$ is STP_k, we can apply theorem 2 to

first find that $S^-(\mathbf{x}^j) = S^+(\mathbf{x}^j) = j-1$, for j = 1, ..., k. In other words, we have the same sign-change pattern for matrices of which a finite power is oscillatory of order k. Finally, the eigenvalues can be ordered as $\lambda_1^{i(N-1)} > \ldots > \lambda_k^{i(N-1)} \geq \ldots \lambda_N^{i(N-1)} > 0$. This directly implies that the first k eigenvalues are simple. □

With this corollary the sufficient conditions from theorem 2 have been relaxed somewhat. Instead of requiring that the covariance or correlation is STP_3, we now only need some finite power of it to be TP_3, invertible, and to have a strictly positive super- and subdiagonal. The following corollary states that multiplying an oscillatory matrix by a totally positive and invertible matrix (both of the same order), yields a matrix which is again oscillatory.

Corollary 2
Let **A** and **B** be a square $N \times N$ matrices, where **A** is oscillatory of order k, and **B** is invertible and TP_k. Then **AB** and **BA** are oscillatory of order k.

Proof:
For a matrix to be oscillatory of order k, we have to check the three defining properties in our proposition. Obviously the first and second properties are satisfied for both matrices. We only have to check the third criterium, concerning the positivity of the super- and subdiagonal elements. For the superdiagonal we basically have:

$$(\mathbf{AB})_{i,i+1} = \sum_{j=1}^{N} a_{ij} b_{j,i+1} \tag{3.6}$$

which is certainly non-negative, due to the fact that both matrices are TP_k. One element contained in (3.6) is $a_{i,i+1}b_{i+1,i+1}$. For **A** we know that all superdiagonal elements are positive. Furthermore, since **B** is invertible, all its diagonal elements must be strictly positive, so that (3.6) is clearly strictly positive. The proof is identical for the subdiagonal. □

This corollary directly implies the following one, which implies that when analysing the sign change pattern of oscillatory matrices, it does not matter whether we analyse covariance or correlation matrices.

Corollary 3
A valid covariance matrix is oscillatory if and only if its correlation matrix is oscillatory.

Proof:
Suppose we have a valid covariance matrix which can be written as $\mathbf{\Sigma} = \mathbf{SRS}$, where **S** is a diagonal matrix containing the (strictly positive) standard deviations on its diagonal, and **R** is the correlation matrix. The "if" part now follows. Since $\mathbf{\Sigma}$ is invertible, so is **S**. An invertible diagonal matrix with strictly positive diagonal elements is clearly totally positive. Hence, if **R** is oscillatory, so will **SRS** by virtue of corollary 2. The "only if" part follows similarly. □

Corollary 3 states that the sign change pattern in the eigenvectors will be the same in covariance and correlation matrices. A graph of the eigenvectors will however look quite different in both matrices, due to the fact that the the term structure of volatilities is typically not flat. As argued in paragraph 2.3, the actual shape of the eigenvectors, e.g. that the first eigenvector is relatively flat, is caused by the fact that the term structure is highly correlated.

Having derived sufficient conditions under which a matrix displays level, slope and curvature, we try to interpret these conditions in the next paragraph.

3.3. Interpretation of the conditions

As we saw in the previous section, a sufficient condition for a covariance or correlation matrix to display level, slope and curvature, is for it to be oscillatory of order 3. We will here try to interpret these conditions. Remember that corollary 3 showed that our definition is invariant to whether we use a covariance or a correlation matrix, so that we opt to use correlation matrices for ease of exposure. For an $N \times N$ correlation matrix **R** to be oscillatory of order 3, we require that:

1. **R** is TP_3;
2. **R** is non-singular;
3. For all i = 1, ..., N-1 we have $\rho_{i,i+1} > 0$ and $\rho_{i+1,i} > 0$.

As any proper covariance or correlation matrix will be invertible, condition ii) is irrelevant. In the term structures we will be analysing, it seems natural to expect that all correlations ρ_{ij} are strictly positive. Condition iii) is immediately fulfilled, as is the case for the order 1 determinants from i). Under this mild condition we can already interpret the first eigenvector as driving the level of the term structure. Hence, the level of the correlations determines whether or not we have level.

Now we turn to the second order determinants. As the usual interpretation of a second order determinant as the signed area of a parallelogram is not very useful here, we need to find another one. Given that **R** is TP_1, it is also TP_2 if for $i \le j$ and $k \le \ell$:

$$\left|\begin{pmatrix} \rho_{ik} & \rho_{i\ell} \\ \rho_{jk} & \rho_{j\ell} \end{pmatrix}\right| = \rho_{ik}\rho_{j\ell} - \rho_{i\ell}\rho_{jk} \ge 0 \Leftrightarrow \rho_{ik}\rho_{j\ell} \ge \rho_{i\ell}\rho_{jk} \tag{3.7}$$

It is not immediately clear how this condition should be interpreted. However, since all correlations were assumed to be positive, we can rearrange (3.7) to find the following condition:

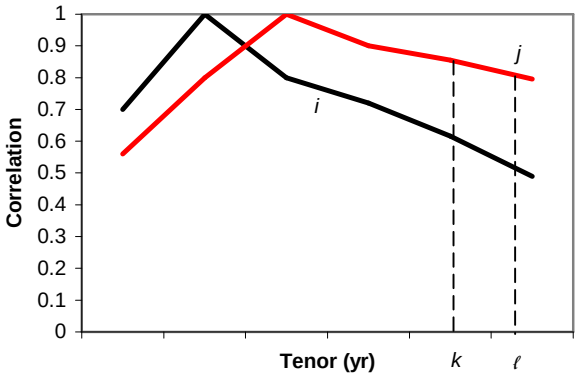
$$\frac{\rho_{ik}}{\rho_{i\ell}} \ge \frac{\rho_{jk}}{\rho_{j\ell}} \Leftrightarrow \frac{\rho_{j\ell} - \rho_{jk}}{\rho_{j\ell}} \ge \frac{\rho_{i\ell} - \rho_{ik}}{\rho_{i\ell}} \tag{3.8}$$

In words, condition (3.8) states that the relative change from moving from k to ℓ ($k \le \ell$), relative to the correlation with ℓ, should be larger on the correlation curve of j than on the curve of i, where $i \le j$. This says that on the right-hand side of the diagonal the relative change on correlation curves for larger tenors should be flatter than for shorter tenors, as is depicted in the graph on the following page. On the left-hand side of the diagonal this is reversed – the relative change there should be larger for shorter than for larger tenors. The derived condition clearly puts a condition on the slopes of the correlation curves.

In practice we usually have a continuous function from which we generate our correlation matrix. With a continuum of tenors we do not analyse the eigensystem of a covariance matrix, but of a symmetric and positive definite kernel $K \in C([0,T] \times [0,T])$. The eigenfunctions and eigenvalues satisfy the following integral equation:

$$\int_0^T K(x,y)\,\phi(y)\,dy = \lambda\,\phi(x) \tag{3.9}$$

This setting is also analysed in Forzani and Tolmasky [2003] for a specific choice of K.



Graph 5: Two correlation curves from a TP_2 matrix

Analysing a continuous problem sometimes makes life easier, but surprisingly the analysis here remains essentially the same. The kernel case was historically studied prior to the matrix case, by O.D. Kellogg. Kellogg [1918] noticed that sets of orthogonal functions often have the property "that each changes sign in the interior of the interval on which they are orthogonal once more than its predecessor". He noted that this property does not only depend on the fact that the functions are orthogonal. As in the discrete case, total positivity of order n is equivalent to:

$$K\begin{pmatrix} x_1,\ldots,x_n \\ y_1,\ldots,y_n \end{pmatrix} = \det\left(K(x_i, y_j)\right)_{i,j=1}^{n} \geq 0 \tag{3.10}$$

for all $x,y \in [0,T]$. When $n = 2$ we regain condition (3.7): $K(x_1,y_1)K(x_2,y_2) \geq K(x_1,y_2)K(x_2,y_1)$. If we in addition assume that K is twice differentiable, one can show that an equivalent condition is:

$$K(x,y)\frac{\partial^2 K(x,y)}{\partial x \partial y} - \frac{\partial K(x,y)}{\partial x}\frac{\partial K(x,y)}{\partial y} = K(x,y)^2 \frac{\partial^2 \ln K(x,y)}{\partial x \partial y} \geq 0 \tag{3.11}$$

Note that if we have a kernel that only depends on the difference of the two arguments, in other words if $K(x,y) = f(x-y)$, (3.11) states that f should be log-concave. A slightly stronger condition than (3.11) is obtained by considering the empirical properties of correlation matrices of term structures we mentioned in paragraph 2.2. Typically correlations are positive, i.e. $K(x,y) > 0$. Secondly, correlations decrease if we move away from the diagonal along a row or a column, implying that $\frac{\partial K(x,y)}{\partial x}\frac{\partial K(x,y)}{\partial y} < 0$. From (3.11) we then see that K is TP_2 if $\frac{\partial^2 K(x,y)}{\partial x \partial y} \geq 0$. Again, if K only depends on the difference of its two arguments, this property requires f to be concave.

Although the condition for slope allows for a clear interpretation, the condition for curvature is much more cumbersome. We just present the final result as the intermediate steps again just follow from rewriting the determinant inequality in (3.4) for $p = 3$. We first define the relative change from moving from k to ℓ ($k \leq \ell$), along correlation curve i as:

$$\Delta_i(k,\ell) = \frac{\rho_{i\ell} - \rho_{ik}}{\rho_{i\ell}} \tag{3.12}$$

Using this definition, the matrix is obviously TP_2 if and only if $\Delta_j(k,\ell) \geq \Delta_i(k,\ell)$ for all i < j and k < ℓ. The additional condition we must impose for the matrix to be TP_3 is then:

$$\frac{\left(\Delta_j(\ell,n)-\Delta_j(m,n)\right)-\left(\Delta_i(\ell,n)-\Delta_i(m,n)\right)}{\Delta_j(m,n)-\Delta_i(m,n)} \geq \frac{\left(\Delta_k(\ell,n)-\Delta_k(m,n)\right)-\left(\Delta_j(\ell,n)-\Delta_j(m,n)\right)}{\Delta_k(m,n)-\Delta_j(m,n)} \tag{3.13}$$

The terms $\Delta_j(\ell,n)-\Delta_j(m,n)$ are changes in relative slopes, and hence are a measure of curvature of correlation curve j. Although it is harder to visualise (3.13) than (3.8), the condition states that this (weighted) "curvature" is allowed to change more from i to j than from j to k.

Summarising we find that the derived sufficient conditions for level, slope and curvature are in fact conditions on the level, slope and curvature of the correlation surface. It seems that, provided the term structure is properly ordered, the conditions do not state much more than that the correlation curves should be flatter and less curved for larger tenors, and steeper and more curved for shorter tenors.

4. Parametric correlation surfaces

Many articles have proposed various parametric correlation matrices, either to facilitate the empirical estimation of correlation matrices or the calibration to market data. One example of this we have seen already is the exponentially decaying correlation function which features in many articles as a simple, but somewhat realistic correlation function. Other examples are the correlation parameterisations by Rebonato [1998], De Jong, Driessen and Pelsser [2004] and Alexander [2003]. The first two parameterisations are both formulated from an economically plausible perspective, but unfortunately are not always guaranteed to be positive definite. The latter is a rank three correlation matrix, defined by restricting the first three "eigenvectors" to be flat, linear and quadratic. We say "eigenvectors" because the constructed vectors are not chosen to be orthogonal, so that these vectors will not be the true eigenvectors. Since the resulting matrix is not of full rank, we will not consider it here.

The correlation matrices we consider in this section will be based on Green's matrices, which in the finance literature are probably better known as Schoenmakers-Coffey correlation matrices. In a continuous setting they already feature in Santa-Clara and Sornette [2001]. Schoenmakers and Coffey [2003] analysed the properties of its discrete analog and proposed various subparameterisations which they claim allow for a stable calibration to market swaption and caplet volatilities. A motivation for their matrix follows directly from the following construction. We will here take a slightly more general route than Schoenmakers and Coffey, leading to a more general correlation matrix. Let b_i, i = 1, …, N be an arbitrary sequence which is increasing in absolute value. We set $b_0 = b_1 = 1$ and $a_1 = 1$, $a_i = \sqrt{b_i^2 - b_{i-1}^2}$. Finally, let Z_i, i = 1, …, N be uncorrelated random variables, with unit variance. We now define:

$$Y_i = \mathrm{sgn}(b_i) \cdot \sum_{k=1}^{i} a_k Z_k \tag{4.1}$$

The covariance between Y_i and Y_j for for i ≤ j is equal to:

$$\mathrm{Cov}(Y_i, Y_j) = \mathrm{sgn}(b_i b_j) \sum_{k=1}^{i} a_j^2 = \mathrm{sgn}(b_i b_j) b_i^2 \tag{4.2}$$

implying that their correlation is equal to:

$$\mathrm{Corr}(Y_i, Y_j) = \frac{b_i}{b_j} = \mathrm{sgn}(b_i b_j) \cdot \frac{\min(|b_i|, |b_j|)}{\max(|b_i|, |b_j|)} \tag{4.3}$$

It is easy to see that we obtain the same correlation structure if the Z_i's do not have unit variance, and also when each Y_i is premultiplied with a non-zero constant c_i. The difference with the approach of Schoenmakers and Coffey is that we here allow the sequence b_i to take negative values, whereas they only considered non-negative correlations. Furthermore, they restricted the sequence b_i/b_{i+1} to be strictly increasing, which has a nice consequence as we will see shortly. Even without these restrictions, the above construction always yields a valid correlation matrix.

We note that an $N \times N$ correlation matrix of the above form, say $\mathbf{R} = (\rho_{ij})_{i,j=1}^{N}$, can also be written in the following form:

$$\rho_{ij} = \prod_{k=i}^{j-1} \rho_{k,k+1} \tag{4.4}$$

i.e. we can view it as a parameterisation in terms of super- or alternatively subdiagonal elements. Schoenmakers and Coffey showed that the above parameterisation of the correlation matrix (with positive b_i's and with the restriction that $\rho_{i,i+1} = b_i/b_{i+1}$ is increasing) satisfies properties (i) – (iii) from paragraph 2.2, properties that are commonly found in empirical correlation matrices of term structures. Sometimes it may be necessary to have a more flexible correlation structure at our disposal, in which case we can relax the restriction that b_i/b_{i+1} is to be increasing. This sacrifices property (iii), the property that the correlation between two adjacent contracts or rates increases as the tenor increases. Properties (i) – (ii) will however still hold.

Returning to the level-slope-curvature pattern, Gantmacher and Kreĭn [1960] prove total positivity for certain special matrices. One of these matrices is a Green's matrix, in which category the above correlation matrix falls.

Theorem 4 – Total positivity of a Green's matrix
An $N \times N$ Green's matrix A with elements:

$$a_{ij} = \begin{cases} u_i v_j & i \geq j \\ u_j v_i & i \leq j \end{cases} \tag{4.5}$$

where all u_i and v_j are different from zero, is totally nonnegative if and only if all u_i and v_j have the same sign and:

$$\frac{v_1}{u_1} \leq \ldots \leq \frac{v_N}{u_N} \tag{4.6}$$

The rank of **A** is equal to the number of times where the inequality in (4.6) is strict, plus one. □

We note that in correlation form the concept of a Green's matrix is not more general than the extended Schoenmakers-Coffey matrix in (4.3) or (4.4). The corresponding correlation matrix **R**

of the Green's matrix **A** from theorem 4 has elements equal to:

$$r_{ij} = \begin{cases} \dfrac{u_i v_j}{\sqrt{u_i v_i u_j v_j}} = \dfrac{\sqrt{u_i v_j}}{\sqrt{u_j v_i}} & i \geq j \\ \dfrac{u_j v_i}{\sqrt{u_i v_i u_j v_j}} = \dfrac{\sqrt{u_j v_i}}{\sqrt{u_i v_j}} & i \leq j \end{cases} \tag{4.7}$$

Indeed, setting $b_i = v_i/u_i$ shows that a Green correlation matrix and the extended Schoenmakers-Coffey correlation matrix are equivalent. This observation combined with theorem 4 leads to the following corollary.

Corollary 4 – Oscillatoriness of the Schoenmakers-Coffey matrix
The Schoenmakers-Coffey correlation matrix, and its more general formulation in (4.3) or (4.4), is oscillatory provided that all correlations on the superdiagonal are positive and smaller than 1. Hence, the matrix displays level, slope and curvature.

Proof:
The requirement that all correlations on the superdiagonal are positive amounts to requiring the sequence b_i to be strictly positive. The requirement that all entries on the superdiagonal are smaller than 1 implies the sequence b_i should be strictly increasing. Setting $b_i = v_i/u_i$ as mentioned, and substituting it into (4.6) yields:

$$b_1^2 \leq \ldots \leq b_N^2 \tag{4.8}$$

which is true due to the fact that the sequence b_i is strictly increasing. Furthermore, since the inequalities are strict, the correlation matrix is of full rank. The latter result still remains true if we allow the b_i's to take negative numbers, but still require that the sequence is strictly increasing in absolute value. Since all entries on the super- and subdiagonal are strictly positive, the matrices are oscillatory. By virtue of corollary 1 this implies that we have level, slope and curvature. □

Hence, if all correlations on the superdiagonal are positive and smaller than 1, the correlation matrix in (4.3) or (4.4) will display level, slope and curvature. We note that property (iii) clearly does not imply or affect level, slope or curvature for these matrices – the extended Schoenmakers-Coffey matrix displays level, slope and curvature regardless of whether property (iii) holds or not. A nice property of a Green's matrix is that its inverse is tridiagonal. Inversion of tridiagonal matrices requires only O(7N) arithmetic operations, and is therefore much more efficient than the $O(N^3/3)$ operations required for arbitrary matrices.

As a final point of interest we return to the claim of Lekkos [2000]. We remind the reader of equation (2.7), where zero yields were expressed as averages of continuously compounded α-forward rates:

$$R(t,T) = \tfrac{\alpha}{T-t}\left(f(t,t,t+\alpha) + \ldots + f(t,T-\alpha,T)\right) \tag{4.9}$$

In a numerical example Lekkos shows that if these forward rates are statistically independent, the correlation matrix of the zero yields displays level, slope and curvature. The way in which the Schoenmakers-Coffey matrix was constructed in equations (4.1) – (4.3) shows that if all forward rates in (4.9) are independent, the correlation matrix of R(t,t+α), …, R(t,t+Nα) will be a

Schoenmakers-Coffey correlation matrix, and as such will display level, slope and curvature. Lekkos' claim is therefore true. In fact, using the Schoenmakers-Coffey matrix for consecutive zero yields directly implies that all forward rates must be independent. Similarly, using the Schoenmakers-Coffey correlation matrix for changes in consecutive zero yields implies that the changes in consecutive forward rates are independent. As we have seen in paragraph 2.2 forward rates and forward rate changes are far from independent, so that one should be aware of these implications. Schoenmakers and Coffey suggest using their correlation matrix and parameterised versions thereof as an instantaneous correlation matrix within the LIBOR market model, where the above considerations do not apply directly.

5. Level, slope and curvature beyond total positivity

In the previous two sections we have turned to total positivity theory to provide us with sufficient conditions for level, slope and curvature. Obviously, this is only a partial answer to the question of what drives this phenomenon. In fact, if we look at the empirical correlation matrices from graphs 1 and 2, the theory that we treated up till now is only able to explain level and slope for both graphs, as both matrices contain only positive correlations, and the second power of both correlation matrices is oscillatory of order 2. The presence of curvature however still remains unexplained. Clearly there must be a more general theory that allows us to explain the presence of level, slope and curvature. Here we first take a brief look at the concept of sign regularity, which extends the notion of total positivity. However, we demonstrate that the only correlation matrices that were not already captured by the class of totally positive matrices are degenerate in some sense. Finally, we formulate a conjecture which we cannot prove, but which we suspect is true, based on an extensive simulation study. This conjecture directly relates the order present in correlation matrices to level and slope.

5.1. Sign regularity

In the literature the concept of total positivity has been extended to the notion of sign regularity. For a square $N \times N$ matrix $\mathbf{A}$ to be sign regular of order k, or SR_k, we require the existence of a sequence ε_1 through ε_k, all $\in \{1,-1\}$, such that for all $p \leq k$ and $\mathbf{i}, \mathbf{j} \in I_{p,N}$, such that:

$$\varepsilon_p \cdot \mathbf{A}_{[p]}(\mathbf{i}, \mathbf{j}) \geq 0 \tag{5.1}$$

Analogous to strict total positivity, strict sign regularity can be defined. Sign regularity hence requires all determinants of a certain order to have the same sign, whereas total positivity required them to be positive. The concept of an oscillatory matrix can easily be extended using sign regularity. We can consider a square invertible matrix $\mathbf{A}$ with non-zero diagonal, super- and subdiagonal elements, that is SR. In this case $\mathbf{A}^2$ is oscillatory, and $\mathbf{A}^{2(N-1)}$ will be strictly totally positive, so that we can again apply theorem 2 to this matrix. This extension is however not useful for our application, as we will see in the following theorem.

Theorem 6 – The set of SR_3 correlation matrices is degenerate
The volume of all square $N \times N$ (for any $N \geq 3$) invertible correlation matrices, that are not TP_3, but SR_3, is equal to zero.

Proof:
Let us first consider the case N = 3, where the correlation matrix can be written as:

$$\mathbf{R} = \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \tag{5.2}$$

where of course $-1 \le x, y, z \le 1$. Via a Cholesky decomposition we can easily check that this matrix is positive definite if and only if its determinant is strictly positive, i.e. when $1 - x^2 - y^2 - z^2 + 2xyz > 0$. For **A** to be TP_2 we would furthermore require:

- First order: $x \ge 0$, $y \ge 0$ and $z \ge 0$
- Second order: $x \ge yz$, $y \le xz$ and $z \ge xy$

Since the matrices we want to consider in this theorem should be positive definite and SR_3 (but not TP_3), we should consider the following three cases:

a) $x \le 0$, $y \le 0$ and $z \le 0$, and $x \ge yz$, $y \le xz$ and $z \ge xy$, and $1 - x^2 - y^2 - z^2 + 2xyz > 0$
b) $x \ge 0$, $y \ge 0$ and $z \ge 0$, and $x \le yz$, $y \ge xz$ and $z \le xy$, and $1 - x^2 - y^2 - z^2 + 2xyz > 0$
c) $x \le 0$, $y \le 0$ and $z \le 0$, and $x \le yz$, $y \ge xz$ and $z \le xy$, and $1 - x^2 - y^2 - z^2 + 2xyz > 0$

Ad a)
From $x,y,z \le 0$, $x \ge yz$ and $z \ge xy$ we have $x,z \ge 0$, which can only be satisfied if $x = z = 0$. This implies that $\iiint_{x=z=0, y^2<1} dx\, dy\, dz = 0$, i.e. the volume of all such matrices is equal to zero.

Ad b and c)
There are three 2×2 submatrices of **R** which are covariance matrices. The corresponding determinants are $1-x^2$, $1-y^2$ and $1-z^2$. If these are required to be negative or zero, as in b) and c), we must have $x, y, z \in \{-1,1\}$. We can check that the determinant of **R** can then only be 0 or -4, which means that the matrix can never be positive definite in these cases.

We have proven the theorem for $N = 3$. Since the assertion holds for any 3×3 invertible correlation matrix, it will also hold for any 3×3 principal submatrix of an $N \times N$ invertible correlation matrix. This will therefore place the same restrictions on these 3×3 submatrices, so that the volume of all invertible correlation matrices, that are not TP_3, but SR_3, is zero. □

This last theorem shows that the class of SR_3, but not TP_3, invertible correlation matrices is degenerate. As far as we know, no other classes of matrices are known which have the same sign change pattern as oscillatory matrices.

5.2. The relation between order, level and slope

As we already mentioned earlier, Alexander [2003] claims that "… the interpretation of eigenvectors as trend, tilt and curvature components is one of the stylised facts of all term structures, particularly when they are highly correlated". Based on an extensive simulation study we come up with a slightly different conjecture, which will follow shortly. The example in section 2.2 demonstrated that curvature is not always present, even though we have an ordered and highly correlated system. Similarly we can show that there are matrices for which no finite power is oscillatory of order 3, so that the theory from section 3 cannot be used to prove the presence of slope and curvature for these correlation matrices. One such example follows.

Example – Total positivity is not enough
Consider the following correlation matrix:

$$\mathbf{R} = \begin{pmatrix} 1 & 0.8396 & 0.8297 & 0.8204 \\ 0.8396 & 1 & 0.9695 & 0.901 \\ 0.8297 & 0.9695 & 1 & 0.9785 \\ 0.8204 & 0.901 & 0.9785 & 1 \end{pmatrix} \tag{5.3}$$

This correlation matrix itself is clearly not TP_2, consider for example its second order compound matrix $\mathbf{R}(\{1,2\},\{3,4\})$, i.e. the 2×2 matrix in the right-upper corner of $\mathbf{R}$. Its determinant is negative. From matrix theory we know that:

$$\lim_{k\to\infty} \frac{\mathbf{R}^k}{\lambda_1^k} = \mathbf{x}^1(\mathbf{x}^1)^T \tag{5.4}$$

where λ_1 is the largest eigenvalues and $\mathbf{x}^1$ is the corresponding eigenvector. Since higher powers of $\mathbf{R}$ are also not TP_2, and $\mathbf{R}^5$ is almost indistinguishable from the limiting matrix in (5.4), we can be sure that no finite power of $\mathbf{R}$ will be oscillatory of order 2.

Since the matrix in (5.3) satisfies properties (i)-(iii), as most empirical correlation matrices do, we came up with the following conjecture.

Conjecture – Sufficiency of properties (i)-(iii) for level, slope and curvature
A quasi-correlation matrix $\mathbf{R}$ with strictly positive entries, that satisfies:

i) $\rho_{i,j+1} \le \rho_{ij}$ for $j \ge i$, i.e. correlations decrease when we move away from the diagonal;

ii) $\rho_{i,j-1} \le \rho_{ij}$ for $j \le i$, same as i);

iii) $\rho_{i,i+j} \le \rho_{i+1,i+j+1}$, i.e. the correlations increase when we move from northeast to southwest.

displays level and slope.

By a quasi-correlation matrix we mean a matrix that resembles a correlation matrix, i.e. has ones on the diagonal and off-diagonal elements that are smaller than or equal to 1 in absolute value, but is not necessarily positive definite. We claim that the empirically observed properties (i)-(iii) are sufficient, although still not necessary, for a quasi or proper correlation matrix to display level and slope. The fact that these properties are not necessary is clear from the Green's matrix – certain Green's matrices are still totally positive even though property (iii) is not satisfied, as we saw in the previous section.

We extensively tested this conjecture by simulating random correlation matrices, satisfying properties (i)-(iii). Although several methods exist to simulate random correlation matrices, we are not aware of one that allows the aforementioned properties to be satisfied. Firstly, we present the algorithm we used to simulate a random quasi-correlation matrix with positive entries, that in addition satisfies (i)-(iii). Note that a smoothing factor α is included in the algorithm that essentially ensures that two consecutive elements on a row are at most $100\alpha\%$ apart. Finally, note that each correlation is drawn from a uniform distribution – this is obviously an arbitrary choice.

1. $\rho_{ii} = 1$ for $1 \le i \le N$.
2. For $1 < j \le N$ set LB_{1j} equal to $(\rho_{1,j-1}-\alpha)^+$ and UB_{1j} equal to $\rho_{1,j-1}$. Draw $\rho_{1j} \sim U[LB_{1j},UB_{1j}]$.
3. For $2 \le i \le N-1$ and $i < j$ set $LB_{ij} = \max\left((\rho_{i,j-1}-\alpha)^+, \rho_{i-1,j-1}\right)$ and $UB_{ij} = \rho_{i,j-1}$. If $LB_{ij} > UB_{ij}$, a valid matrix cannot be constructed, so that we have to restart our algorithm at step 1. Otherwise we draw $\rho_{ij} \sim U[LB_{ij},UB_{ij}]$.
4. Set $\rho_{ij} = \rho_{ji}$ for $i > j$.

Algorithm 1: Simulation of a quasi-correlation matrix with strictly positive entries, satisfying (i)-(iii)

Algorithm 1 can easily be adapted to generate a matrix that only satisfies (i)-(ii), by replacing $\rho_{i-1,j-1}$ in step 3 by $\rho_{i-1,j}$. Adapting this algorithm to yield a positive definite matrix can be achieved if we use the angles parameterisation of Rebonato and Jäckel [1999]. They show[8] that any correlation matrix $\mathbf{R} \in \mathbb{R}^{N\times N}$ can be written as $\mathbf{R}=\mathbf{BB}^T$, where $\mathbf{B} \in \mathbb{R}^{N\times N}$ is lower triangular and has entries equal to $b_{11} = 1$, and:

$$b_{ij} = \cos\theta_{ij}\prod_{k=1}^{j-1}\sin\theta_{ik} \qquad b_{ii} = \prod_{k=1}^{i-1}\sin\theta_{ik} \tag{5.5}$$

for $i > j$ and $i > 1$. Using this parameterisation it can be shown that the first row of the correlation matrix follows directly from the first column of the matrix with angles, i.e. $\rho_{1j} = \cos\theta_{j1}$ for $j > 1$. Hence, adapting step 2 is easy: we only have to solve for θ_{j1} in step 2. Adapting step 3 is slightly more involved. For $i < j$ we have:

$$\begin{aligned}\rho_{ij} &= \sum_{\ell=1}^{i} b_{i\ell}b_{j\ell} \\ &= \sum_{\ell=1}^{i-1}\cos\theta_{i\ell}\cos\theta_{j\ell}\prod_{k=1}^{\ell-1}\sin\theta_{ik}\sin\theta_{jk} + \cos\theta_{ji}\prod_{k=1}^{i-1}\sin\theta_{ik}\sin\theta_{jk}\end{aligned} \tag{5.6}$$

At entry (i,j) of the correlation matrix, we have already solved for the angles in columns 1 up to and including i-1, as well as angles θ_{jk} for $k < i$. The only new angle in (5.6) is thus θ_{ji}. Since we necessarily have $-1 \le \cos\theta_{ji} \le 1$, (5.6) places a lower and upper bound on ρ_{ij}. All we have to do is incorporate these additional restrictions into step 3 – this ensures that the new algorithm terminates with a positive definite correlation matrix. The algorithm hence becomes:

1. $\rho_{ii} = 1$ for $1 \le i \le N$.
2. For $1 < j \le N$ set LB_{1j} equal to $(\rho_{1,j-1}-\alpha)^+$ and UB_{1j} equal to $\rho_{1,j-1}$. Draw $\rho_{1j} \sim U[LB_{1j},UB_{1j}]$. Solve θ_{j1} from $\rho_{1j} = \cos\theta_{j1}$.
3. For $2 \le i \le N-1$ and $i < j$ set $LB_{ij} = \max\left((\rho_{i,j-1}-\alpha)^+, \rho_{i-1,j-1}\right)$ and $UB_{ij} = \rho_{i,j-1}$. Incorporate lower and upper bound from (5.6) into LB_{ij} and UB_{ij}. If we then have $LB_{ij} > UB_{ij}$, a valid matrix cannot be constructed, so we have to restart our algorithm at step 1. Otherwise we draw ρ_{ij} from $U[LB_{ij},UB_{ij}]$ and solve for θ_{ji} from (5.6).
4. Set $\rho_{ij} = \rho_{ji}$ for $i > j$.

Algorithm 2: Simulation of a valid correlation matrix with strictly positive entries, satisfying (i)-(iii)

Using algorithms 1 and 2 we performed a large amount of simulations, for various sizes of

[8] Their article uses N(N-1) angles, but it can be shown that it suffices to use as many angles as correlations, i.e. ½N(N-1).

matrices and values of α. In each simulation we kept track of the percentage of matrices without slope and/or curvature. The pattern was the same in each simulation, so that we here only display results for sizes equal to 3, 4 and 5 and α equal to 20%. The results can be found below:

	Properties (i)-(ii)		**Properties (i)-(iii)**	
Size	*No slope*	*No curvature*	*No slope*	*No curvature*
3	0%	0%	0%	0%
4	0.04%	19.1%	0%	23.05%
5	0.01%	27.98%	0%	43.81%

Table 1: Percentage of random quasi-correlation matrices w/o slope and/or curvature
Results based on 10,000 random matrices from algorithm 1, using α = 20%

	Properties (i)-(ii)		**Properties (i)-(iii)**	
Size	*No slope*	*No curvature*	*No slope*	*No curvature*
3	0%	0%	0%	0%
4	0.13%	14.91%	0%	18.31%
5	0.02%	23.1%	0%	35.38%

Table 2: Percentage of random proper correlation matrices w/o slope and/or curvature
Results based on 10,000 random matrices from algorithm 2, using α = 20%

As is clear from the tables, for both types of matrices properties (i)-(iii) seem to imply the presence of slope. Leaving out property (iii) causes some violations of the slope property, albeit in a very small number of cases. The results seem to indicate that our conjecture has some validity, although this is of course far from a formal proof.

6. Conclusions

In this article we analysed the so-called level, slope and curvature pattern which one frequently observes when conducting a principal components analysis of term structure data. The first factor which one usually finds is relatively flat, the second has opposite signs at both ends of the maturity spectrum, and the third finally has the same sign at both ends, but an opposite sign in the middle of the maturity range. As such, many authors have interpreted these factors as driving respectively the level, slope and curvature of the term structure. Forzani and Tolmasky [2003] analysed this pattern for an exponentially decaying correlation function and showed that, for high correlations, we indeed have the observed level, slope and curvature pattern.

We mathematically define the observed pattern by stating that if the first three factors or eigenvectors have respectively zero, one and two sign changes, the corresponding matrix displays level, slope and curvature. This characterisation leads us to the theory of total positivity. If all correlations are positive, the first factor will always have no sign changes, by virtue of Perron's theorem. For the remaining two factors, we turn to several theorems by Gantmacher and Kreĭn [1937, 1960, 2002]. Slight alterations hereof provide sufficient conditions for level, slope and curvature. The conditions can be interpreted as conditions on the level, slope and curvature of the correlation surface. In essence, the conditions roughly state that the correlation curves should be flatter and less curved for larger tenors, and steeper and more curved for shorter tenors.

Using the concepts of total positivity, we analysed the correlation matrix of Schoenmakers and Coffey [2003], and a generalisation thereof. Their parameterisation is in matrix theory better known as a Green's matrix. Provided all correlations on the superdiagonal are positive and smaller than 1, the matrix displays level, slope and curvature. In Lekkos [2000] it is claimed that since zero yields are averages of (continuously compounded) forward rates, the correlation matrix

of zero yields will always display level, slope and curvature if the forward rates are independent. Since the resulting correlation matrix is then a Schoenmakers-Coffey or Green's matrix, this claim has now been proven mathematically. Vice versa, when this correlation matrix is used as a correlation matrix of consecutive zero yields or changes hereof, the underlying assumption is that (continuously compounded) consecutive forward rates or changes thereof are independent.

Finally, we demonstrated that there are matrices for which we cannot prove the existence of slope and curvature using total positivity theory or extensions hereof. We ended the paper with an unproved conjecture that the order present in correlation matrices drives the slope pattern, and corroborated this with results from a simulation study.

Bibliography

ALEXANDER, C. (2003). "Common correlation and calibrating LIBOR", *Wilmott Magazine*, March 2003, pp. 68-78.

ALEXANDER, C. AND D. LVOV (2003). "Statistical properties of forward LIBOR rates", working paper, University of Reading.

ANDO, T. (1987). "Totally positive matrices", *Linear Algebra and Its Applications*, vol. 90, pp. 165-219.

CORTAZAR, G. AND E. SCHWARTZ (1994). "The valuation of commodity-contingent claims", *Journal of Derivatives*, vol. 1, no. 4, pp. 27-39.

FORZANI, L. AND C. TOLMASKY (2003). "A family of models explaining the level-slope-curvature effect", *International Journal of Theoretical and Applied Finance*, vol. 6, no. 3, pp. 239-255.

GANTMACHER, F.R. AND M.G. KREĬN (1937). "Sur les matrices complètement non negatives et oscillatoires", *Composito Math*, vol. 4, pp. 445-476.

GANTMACHER, F.R. AND M.G. KREĬN (1960). "Oszillationsmatrizen, Oszillationskerne und Kleine Schwingungen Mechanischer Systeme,", Akademie-Verlag, Berlin.

GANTMACHER, F.R. AND M.G. KREĬN (2002). "Oscillation matrices and kernels and small vibrations of mechanical systems", AMS Chelsea Publishing.

HINDANOV, D. AND C. TOLMASKY (2002). "Principal components analysis for correlated curves and seasonal commodities: The case of the petroleum market", *Journal of Futures Markets*, vol. 22, no. 11, pp. 1019-1035.

JACKSON, J.E. (2003). "A user's guide to principal components", *Wiley*.

JONG, F. DE, DRIESSEN, J. AND A. PELSSER (2004). "On the information in the interest rate term structure and option prices", *Review of Derivatives Research*, vol. 7., no. 2, pp. 99-127.

KARLIN, S. (1968). "Total positivity, Volume I", *Stanford University Press*.

KELLOGG, O.D. (1918). "Orthogonal function sets arising from integral equations", *American Journal of Mathematics*, vol. 40, pp. 145-154.

LARDIC, S., PRIAULET, P. AND S. PRIAULET (2003). "PCA of the yield curve dynamics: questions of methodologies", *Journal of Bond Trading and Management*, vol. 1, no. 4, pp. 327-349.

LEKKOS, I. (2000). “A critique of factor analysis of interest rates”, vol. 8, no. 1, pp. 72-83.

LITTERMAN, R. AND J. SCHEINKMAN (1991). “Common factors affecting bond returns”, *Journal of Fixed Income*, vol. 1, no. 1, pp. 54-61.

NELSON, C.R. AND A.F. SIEGEL (1987). “Parsimonious modeling of yield curves”, *Journal of Business*, vol. 60, no. 3, pp. 473-489.

PINKUS, A. (1995). “Spectral properties of totally positive kernels and matrices”, in: *Total positivity and its applications* (eds.: M. Gasca and C.A. Micchelli), Kluwer Academic Publishers, pp. 477-511.

REBONATO, R. (1998). *Interest-rate option models* (Second edition), John Wiley & Sons.

REBONATO, R. AND P. JÄCKEL (1999/2000). “The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes”, *Journal of Risk*, vol. 2, no. 2, pp. 1–16.

SANTA-CLARA, P. AND D. SORNETTE (2001). “The dynamics of the forward interest rate curve with stochastic string shocks”, *Review of Financial Stu*dies, vol. 14, no. 1, pp. 149-185.

SCHOENMAKERS, J.G.M. AND B. COFFEY (2003). “Systematic Generation of Correlation Structures for the Libor Market Model”, *International Journal of Theoretical and Applied Finance*, vol. 6, no. 4, pp. 1-13.

STEELEY, J.M. (1990). “Modelling the dynamics of the term structure of interest rates”, *Economic and Social Review*, vol. 21, no. 4, pp. 337-361.

SVENSSON, L. (1994). “Estimating and interpreting forward interest rates: Sweden 1992-94”, NBER working paper no. w4871.

Appendix - Proofs of various theorems

In this appendix we have included the proofs of theorem 2 and our proposition are included for the sake of completeness. Before this however we will require the following theorems from matrix algebra.

Theorem A.1 – Cauchy-Binet formula
Assume A = BC where A, B and C are N × N matrices. The Cauchy-Binet formula states that:

$$\mathbf{A}_{[p]}(\mathbf{i},\mathbf{j}) = \sum_{\mathbf{k}\in\mathbf{I}_{p,N}} \mathbf{B}_{[p]}(\mathbf{i},\mathbf{k})\mathbf{C}_{[p]}(\mathbf{k},\mathbf{j}) \tag{A.1}$$

In other words, $\mathbf{A}_{[p]} = \mathbf{B}_{[p]}\mathbf{C}_{[p]}$, the operations of matrix multiplication and compound are interchangeable. □

The next theorem is useful when studying the eigensystem of compound matrices.

Theorem A.2 – Part of Kronecker's theorem
Let $\mathbf{\Sigma}$ be an invertible N × N matrix with eigenvalues λ_1 , ... , λ_N listed to their algebraic multiplicity. The matrix can be decomposed as $\mathbf{\Sigma} = \mathbf{X\Lambda X}^T$, where $\mathbf{\Lambda}$ contains the eigenvalues on its diagonal, and $\mathbf{X}$ contains the eigenvectors. In this case, $\mathbf{\Lambda}_{[p]}$ and $\mathbf{X}_{[p]}$ contain respectively the eigenvalues and eigenvectors of $\mathbf{\Sigma}_{[p]}$. The $\binom{N}{p}$ eigenvalues of $\mathbf{\Sigma}_{[p]}$, listed to their algebraic multiplicity, are $\lambda_{i_1} \cdot \ldots \cdot \lambda_{i_p}$, for $\mathbf{i} \in I_{p,N}$.

Proof:
When $\mathbf{\Sigma}$ is a general square matrix, the theorem concerning the eigenvalues is known as Kronecker's theorem. Its proof is easy and can be found in e.g. Karlin [1968], Ando [1987] or Pinkus [1995]. In case $\mathbf{\Sigma}$ is invertible everything is simplified even further. By virtue of the Cauchy-Binet formula we have that $\mathbf{\Sigma X} = \mathbf{\Lambda}$ and $\mathbf{XX}^T = \mathbf{I}$ implies that $\mathbf{\Sigma}_{[p]}\mathbf{X}_{[p]}^T = \mathbf{\Lambda}_{[p]}$ and $\mathbf{X}_{[p]}\mathbf{X}_{[p]}^T = \mathbf{I}_{[p]} = \mathbf{I}$. Indeed, this means that $\mathbf{\Lambda}_{[p]}$ and $\mathbf{X}_{[p]}$ respectively contain the eigenvalues and eigenvectors of $\mathbf{\Sigma}_{[p]}$. □

We are now ready to prove theorem 2.

Theorem 2 – Sign-change pattern in STP_k matrices
Assume $\mathbf{\Sigma}$ is an N × N positive definite symmetric matrix (i.e. a valid covariance matrix) that is STP_k. Then we have $\lambda_1 > \lambda_2 > \ldots > \lambda_k > \lambda_{k+1} \geq \ldots \lambda_N > 0$, i.e. at least the first k eigenvalues are simple. Denoting the jth eigenvector by $\mathbf{x}^j$, we have $S^-(\mathbf{x}^j) = S^+(\mathbf{x}^j) = j-1$, for j = 1, ..., k.

Proof:
We will first prove that at least the largest k eigenvalues are distinct. We know that the eigenvalues of $\mathbf{\Sigma}$ are strictly positive, so we can write $\lambda_1 \geq \ldots \geq \lambda_N > 0$. Since $\mathbf{\Sigma}$ is STP_k, we can apply Perron's theorem to find that $\lambda_1 > \lambda_2 \geq \ldots \geq \lambda_N$. Assume we have proven the statement for the largest j-1 eigenvalues, j-1 < k, this means we know that $\lambda_1 > \lambda_2 > \ldots > \lambda_{j-1} > \lambda_j \geq \ldots \lambda_N > 0$. Since $\mathbf{\Sigma}$ is STP_j, each element of $\mathbf{\Sigma}_{[j]}$ is strictly positive. Perron's theorem (theorem 1) now states

that $\mathbf{\Sigma}_{[j]}$ has a unique largest eigenvalue. From Kronecker's theorem (theorem A.2) we deduce that this eigenvalue is $\lambda_1 \cdot \ldots \cdot \lambda_j$. Hence:

$$\lambda_1 \cdot \ldots \cdot \lambda_j > \lambda_1 \cdot \ldots \lambda_{j-1} \lambda_{j+1} \Leftrightarrow \lambda_j > \lambda_{j+1} \tag{A.2}$$

i.e., since we know that all eigenvalues are positive, this implies the property also holds for j. For j equal to 1 the property has already been proven, so by induction it follows that the largest k eigenvalues are distinct.

Now we turn to the number of sign changes of the eigenvectors associated with the largest k eigenvalues. Since $\mathbf{\Sigma}$ is positive definite, orthogonal eigenvectors $\mathbf{x}^1$ through $\mathbf{x}^N$ exist. Eigenvector $\mathbf{x}^i$ is associated with eigenvalue λ_i. We write $\mathbf{\Sigma} = \mathbf{X\Lambda\Lambda}^T$, where $\mathbf{X}$ is the matrix containing the eigenvectors as its columns, and $\mathbf{\Lambda}$ contains the eigenvalues on its diagonal. Let us first assume that $S^+(\mathbf{x}^j) \geq j$, for $j \leq k$. Then there are $1 \leq i_0 < \ldots < i_j \leq N$ and an $\in = \pm 1$ such that:

$$\varepsilon(-1)^\ell \mathbf{x}^j_{i_\ell} \geq 0 \text{ for } \ell = 0, \ldots, j \tag{A.3}$$

Now set $\mathbf{x}^0 = \mathbf{x}^j$, and we extend $\mathbf{X}$ as $\overline{\mathbf{X}}$ to include $\mathbf{x}_0$ as its first column. Obviously we must have that $\overline{\mathbf{X}}\begin{pmatrix} i_0, \ldots, i_j \\ 0, \ldots, j \end{pmatrix} = 0$. On the other hand, we can expand this determinant on the first column:

$$\overline{\mathbf{X}}\begin{pmatrix} i_0, \ldots, i_j \\ 0, \ldots, j \end{pmatrix} = \sum_{\ell=0}^{j} (-1)^\ell x^0_{i_\ell} \mathbf{X}\begin{pmatrix} i_0, \ldots, i_{\ell-1}, i_{\ell+1}, \ldots, i_j \\ 1, \ldots, j \end{pmatrix} \tag{A.4}$$

From (A.3) we have that the first part of the sum is of one sign. The determinant on the right-hand side is an element of the first column of $\mathbf{X}_{[j]}$. Via theorem 2 we know that this column contains the eigenvectors of $\mathbf{\Sigma}_{[j]}$, which contains strictly positive elements by the assumption that $\mathbf{\Sigma}$ is STP_j for $j \leq k$. Perron's theorem then implies that this first eigenvector is either strictly positive or negative, i.e. every element on the right-hand side is of one sign. If the left-hand side is therefore going to be zero, we must have that $\mathbf{x}^0_{i_\ell} = \mathbf{x}^j_{i_\ell}$ is zero for $\ell = 0, \ldots, j$. Note that the determinant on the right-hand side is the determinant of the submatrix formed by only using the rows i_0 through i_j, excluding i_ℓ, and columns 1 through j. If $\mathbf{x}^0_{i_\ell} = \mathbf{x}^j_{i_\ell}$ is zero for $\ell = 0, \ldots, j$, we would be taking the determinant of a matrix which contains a column filled with zeroes. Necessarily this determinant would be equal to zero, which is a contradiction as we just saw. We can therefore conclude that $S^+(\mathbf{x}^j) \leq j-1$, for $j \leq k$.

The second part of the proof is very similar. From the definitions of S^+ and S^- it is clear that we must have $S^-(\mathbf{x}^j) \leq S^+(\mathbf{x}^j) \leq j-1$. Let us assume that $S^-(\mathbf{x}^j) = p \leq j-2$ for $j \leq k$. This implies that there exist $1 \leq i_0 < \ldots < i_p \leq N$ and an $\in = \pm 1$ such that:

$$\varepsilon(-1)^\ell \mathbf{x}^j_{i_\ell} > 0 \text{ for } \ell = 0, \ldots, p \tag{A.5}$$

Again, we set $\mathbf{x}^0 = \mathbf{x}^j$. Then obviously the determinant $\overline{\mathbf{X}}\begin{pmatrix} i_0, \ldots, i_p \\ 0, \ldots, p \end{pmatrix} = 0$. As before, we find:

$$\overline{\mathbf{X}}\begin{pmatrix} i_0, \ldots, i_p \\ 0, \ldots, p \end{pmatrix} = \sum_{\ell=0}^{p} (-1)^\ell x^0_{i_\ell} \mathbf{X}\begin{pmatrix} i_0, \ldots, i_{\ell-1}, i_{\ell+1}, \ldots, i_p \\ 1, \ldots, p \end{pmatrix} \tag{A.6}$$

From before, we know that the determinants on the right-hand side can be chosen to be strictly positive. Together with (A.5) this implies that the right-hand side is positive, which is a contradiction. Therefore $S^-(\mathbf{x}^j)$ must be larger than j-2, and we have proven $S^-(\mathbf{x}^j) = S^+(\mathbf{x}^j) = j-1$ for $j \leq k$. □

To prove our proposition, we will also require the Hadamard inequality for positive semi-definite matrices. In Gantmacher and Kreĭn [1937] this was originally proven for TP matrices. In Karlin [1968] we can find the following formulation.

Theorem A.3 – Hadamard inequality
For $\mathbf{\Sigma}$ an $N \times N$ positive semi-definite matrix we have:

$$\mathbf{\Sigma}\begin{pmatrix}1,\ldots,N\\1,\ldots,N\end{pmatrix} \leq \mathbf{\Sigma}\begin{pmatrix}1,\ldots,k\\1,\ldots,k\end{pmatrix}\cdot\mathbf{\Sigma}\begin{pmatrix}k+1,\ldots,N\\k+1,\ldots,N\end{pmatrix} \tag{A.7}$$

for k = 1, …, N-1. □

Now we are ready to prove our theorem about oscillation matrices of order k.

Theorem 3 – Oscillation matrix of order k
Akin to the concept of an oscillation matrix, we define an oscillation matrix of order k. An $N \times N$ matrix $\mathbf{A}$ is oscillatory of the order k if:
1. $\mathbf{A}$ is TP_k;
2. $\mathbf{A}$ is non-singular;
3. For all i = 1, ..., N-1 we have $a_{i,i+1} > 0$ and $a_{i,i+1} > 0$.

For oscillatory matrices of the order k, we have that $\mathbf{A}^{N-1}$ is STP_k.

Proof:
First we prove that for all matrices satisfying i), ii) and iii), we have $\mathbf{A}_{[p]}(\mathbf{i},\mathbf{j}) > 0$, for $p \leq k$ and for all $\mathbf{i}$ and $\mathbf{j} \in I_{p,n}$ satisfying:

$$\left|i_\ell - j_\ell\right| \leq 1 \text{ and } \max(i_\ell, j_\ell) < \min(i_{\ell+1}, j_{\ell+1}) \qquad \ell = 1, \ldots, p \tag{A.8}$$

where $i_{p+1} = j_{p+1} = \infty$. Gantmacher and Kreĭn dubbed these minors quasi-principal minors. We will prove this by induction on p. For p = 1 all quasi-principal minors are all diagonal and super- and subdiagonal elements. The latter are positive by assumption. Furthermore, from the assumption of non-singularity and the Hadamard inequality (theorem A.3) for totally positive matrices, we have:

$$0 < \det \mathbf{A} \leq \prod_{i=1}^{N} a_{ii} \tag{A.9}$$

i.e. all diagonal elements are non-zero, and from the assumption of total positivity are hence positive. Now assume that the assertion holds for p-1, but that it does not hold for $p \leq k$. Hence, all quasi-principal minors of order smaller than p are positive, but there are $\mathbf{i}$ and $\mathbf{j} \in I_{p,N}$ satisfying (A.8) such that:

$$\mathbf{A}\begin{pmatrix} i_1,\ldots,i_p \\ j_1,\ldots,j_p \end{pmatrix} = 0 \tag{A.10}$$

From the induction assumption we have that:

$$\mathbf{A}\begin{pmatrix} i_1,\ldots,i_{p-1} \\ j_1,\ldots,j_{p-1} \end{pmatrix} \cdot \mathbf{A}\begin{pmatrix} i_2,\ldots,i_p \\ j_2,\ldots,j_p \end{pmatrix} > 0 \tag{A.11}$$

Consider the matrix (a_{ij}) with elements $i = i_1, i_1+1$ through i_p, and $j = j_1, j_1+1$ through j_p. These last two results and corollary 9.2 from Karlin [1968] imply that the rank of this matrix is p-1. Now set $h = \max(i_1,k_1)$. Then it follows from (A.8) that $h+p-1 \le \min(i_p,j_p)$. This implies that:

$$\mathbf{A}\begin{pmatrix} h, h+1,\ldots,h+p-1 \\ h, h+1,\ldots,h+p-1 \end{pmatrix} \tag{A.12}$$

is a principal minor of order p of **A**, and hence is equal to zero. Since the matrix **A** is positive definite, so is the square submatrix (a_{ij}) with i,j = h, …, h+p-1. Therefore (A.10) cannot hold, and since **A** is TP_k, we must have $A(\mathbf{i},\mathbf{j}) > 0$.
Now we will prove that $B=A^{N-1}$ is STP_k. Indeed, for any $\mathbf{i}, \mathbf{j} \in I_{p,N}$, where $p \le k$ we have, from the Cauchy-Binet formula:

$$\mathbf{B}(\mathbf{i},\mathbf{j}) = \sum_{\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(N-2)}} \prod_{\ell=1}^{N-1} \mathbf{A}\left(\boldsymbol{\alpha}^{(\ell-1)}, \boldsymbol{\alpha}^{(\ell)}\right) \tag{A.13}$$

where each $\boldsymbol{\alpha}^{(\ell)} \in I_{p,N}$ and we set $\boldsymbol{\alpha}^{(0)} = \mathbf{i}$ and $\boldsymbol{\alpha}^{(N-1)} = \mathbf{j}$. Since **A** is TP_k, $\mathbf{B}(\mathbf{i},\mathbf{j})$ is a sum of nonnegative determinants, and hence is itself nonnegative. Following Gantmacher and Kreĭn we can now construct a series of $\boldsymbol{\alpha}^{(\ell)}$ such that each determinant in (A.13) is a quasi-principal minor of order smaller than or equal to k, and hence by the previous result is strictly positive. The construction works as follows.

1. Set $\boldsymbol{\alpha}^{(0)} = \mathbf{i}$ and s = 1.
2. Compare **j** with $\boldsymbol{\alpha}^{(s-1)}$. Writing down $\boldsymbol{\alpha}^{(s-1)}$ in lexicographical order, we see it can be divided into consecutive parts, each of which contains elements $\alpha_\ell^{(s-1)}$ that are either smaller than, larger than, or equal to j_ℓ. We will refer to these parts as the positive, negative and zero parts.
3. We now construct $\boldsymbol{\alpha}^{(s)}$ from $\boldsymbol{\alpha}^{(s-1)}$ by adding 1 to the last s elements in each positive part and subtracting 1 from the first s indices in each negative part. Each zero part is left unchanged. If any part has less than s elements, we alter all elements.
4. Repeat 2 and 3 for s = 2, …, N-1.

As an example, consider the vectors (2, 3, 5, 8, 9) and (2, 4, 6, 7, 9). Let the first vector play the role of $\boldsymbol{\alpha}^{(0)}$ and the second the role of **j**. We group $\boldsymbol{\alpha}^{(0)}$ as ((2), (3,5), (8), (9)), and see that it consists of a zero, a positive, a negative, and a zero part, in that order. We are now ready to construct $\boldsymbol{\alpha}^{(1)}$, and find that it is equal to (2, 3, 6, 7, 9).
We will now prove that each pair $\boldsymbol{\alpha}^{(s-1)}$, $\boldsymbol{\alpha}^{(s)}$ for s = 1, …, N-1 satisfies (A.8). The first part of (A.8) is obviously satisfied, as each element of $\boldsymbol{\alpha}^{(s)}$ differs by at most 1 from the corresponding entry in $\boldsymbol{\alpha}^{(s-1)}$. That the constructed $\boldsymbol{\alpha}^{(s)} \in I_{p,N}$, i.e. that $1 \le \alpha_1^{(s)} < \ldots < \alpha_p^{(s)} \le N$ is easy to check; we

will omit this here. We will now prove that $\alpha_{\ell}^{(s)} < \alpha_{\ell+1}^{(s-1)}$ for each ℓ. If both entries come from the same part of $\boldsymbol{\alpha}^{(s-1)}$, this obviously holds. Hence, we need to check that this condition holds at the boundary of two parts. In fact we only have to check those cases where $\alpha_{\ell+1}^{(s-1)}$ belongs to a negative part. This means:

$$\alpha_{\ell+1}^{(s)} = \alpha_{\ell+1}^{(s-1)} - 1 \geq j_{\ell+1} \tag{A.14}$$

Since $\boldsymbol{\alpha}^{(s)} \in I_{p,N}$, we immediately have: $\alpha_{\ell}^{(s)} < \alpha_{\ell+1}^{(s)} < \alpha_{\ell}^{(s-1)}$. Hence, all $A(\boldsymbol{\alpha}^{(s-1)}, \boldsymbol{\alpha}^{(s)})$ we have constructed for s = 1, …, N-1, are quasi-principal minors. We will now prove that $\boldsymbol{\alpha}^{(N-1)} = \mathbf{j}$. Note that $\ell \leq i_{\ell}$ and $j_{\ell} \leq N-p+\ell$ are true. These inequalities imply:

$$\left| i_{\ell} - j_{\ell} \right| \leq N - p \tag{A.15}$$

From the construction we followed, it is clear that for any $i_{\ell} \neq j_{\ell}$, the ℓth element of $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\alpha}^{(1)}$, … will have the value i_{ℓ} up to a certain point, and will then approach j_{ℓ} with increments of 1. The convergence towards j_{ℓ} will start when s = p, at the latest. Due to (A.15) we will have certainly achieved $\boldsymbol{\alpha}^{(N-1)} = \mathbf{j}$, as we have then performed N-1 steps of the algorithm. This implies that we can construct an element of the summation in (A.13) where each element is a quasi-principal minor of order $p \leq k$. By the previous result, each of these minors is strictly positive, so that indeed the matrix $B=A^{N-1}$ is STP_k. □