Completing correlation matrices

Olaf Dreyer[*]   Horst Köhler [†]   Thomas Streuer [‡]

October 25, 2021

Abstract

We describe a way to complete a correlation matrix that is not fully specified. Such matrices arise in financial applications when the number of stochastic variables becomes large or when several smaller models are combined in a larger model. We argue that the proper completion to consider is the matrix that maximizes the entropy of the distribution described by the matrix. We then give a way to construct this matrix starting from the graph associated with the incomplete matrix. If this graph is chordal our construction will result in a proper correlation matrix. We demonstrate the construction using a cross-currency model with six stochastic variables.

Keywords: stochastic models, hybrid models, stochastic volatility, correlation matrices, graph theory, chordal graphs

JEL: C02, C63, G12

1 Introduction

Modern financial models are driven by a large number of stochastic variables. These stochastic variables are usually correlated which creates the need to determine the coefficients of the underlying correlation matrix. Since the number of correlation coefficients increases quadratically with the number of stochastic variables, this number can become very large. While it would be desirable to calibrate all these coefficients this is in practice not possible because such a calibration procedure would be too slow, too unstable, or would require calibration instruments for which quotes can not be obtained. Hybrid models are another source of undetermined correlation coefficients. Two calibrated models might be parts of a larger model. While the coefficients for the individual models have been determined in separate calibration, the coefficients that mix variables from the two different models are

[*]OD Consulting
[†]Commerzbank AG
[‡]RIVACON GmbH

undetermined. In a large model, with many smaller constituent models, the number of such coefficients can constitute the bulk of the matrix.

Faced with this situation, it is tempting to set the missing coefficients to zero. The idea here is that any other choice seems to add arbitrary information. This approach has several shortcomings though. The most important of which is that one might end up with a matrix that is not positive semidefinite and thus not a valid correlation matrix. To obtain a valid matrix further steps are required. One possibility is to use the fact that the set of positive definite correlation matrices is an open set that contains the unit matrix. It is thus always possible to scale the off-diagonal elements of the matrix with a common scale factor $\lambda < 1$ and obtain a positive definite matrix. Another possibility is to choose the nearest correlation matrix in a suitable topology (see e.g [Higham, 2002]). These additional steps might change all the coefficients of the matrix, including those coefficients, that have been obtained in previous calibrations. They also rely on somewhat arbitrary further assumptions like a particular notion of distance on the space of matrices.

The approach presented here does not suffer from these shortcomings. We start with a particular criterium that singles out one particular completion. The criterium is distinguished by the fact that it puts the least amount of additional constraints on the distribution described by the completed matrix. As it turns out, this is not the matrix that contains zeroes in the places of unknown coefficients. Rather, it is the inverse of the matrix that contains zeroes in those places that were undetermined before. This criterium removes the arbitrary nature of other constructions.

We then show how to construct this completion. The construction starts by looking at the graph associated with the incomplete matrix. The vertices of the graph are the rows (or columns) of the matrix. Since we are looking at a correlation matrix, the vertices correspond to the stochastic variables of the model. Two vertices are connected by an edge in the graph if the corresponding correlation coefficient in the matrix is know. It is here that we encounter the one condition that needs to be satisfied for the construction to work. The graph that we have just constructed needs to be chordal. A graph is chordal if every loop in the graph that has more than three edges contains a chord, i.e. an edge that connects two non-adjacent vertices in the loop. If this condition is satisfied we can follow a construction first introduced by [Grone et al., 1984] and refined by [Smith, 2008] to complete the matrix. Every step of the construction adds new vertices to the edge which in turn correspond to entries in the correlation matrix. The matrix obtained in this way will automatically be a positive-semidefinite so that no further steps need to be taken to obtain a valid correlation matrix.

The paper now consists of three parts. In the next section we discuss the completion criterium. The matrix can be characterized in three different ways and we present them here. We then discuss the construction of the

completion in section 3. We do this by following the concrete example of a cross-currency model with two currencies and one exchange rate. The two models for the interest rate and the model for the exchange rate are assumed to be stochastic volatility models so that we end up with a total of six stochastic variables. We describe all the required steps using this model but the steps can easily be generalized to more complicated models. One such model is the N-currency model with $(N-1)$ foreign exchange rates that is described in section 4.

2   The completion criterium

Given a correlation matrix with unknown coefficients what is the right way to complete it? To motivate the answer we need to introduce some notation. Let

$$X = (x_1, \ldots, x_k)^T \in \mathbb{R}^k \tag{1}$$

be a k dimensional random variable with a density function given by a k dimensional normal distribution

$$f(X) = N_k[\mu, H_X](X) \tag{2}$$

$$= (2\pi)^{-k/2} \det(H_X)^{-1/2} \exp\left(-\frac{1}{2}(X-\mu)^T H_X^{-1}(X-\mu)\right), \tag{3}$$

with mean μ and covariance matrix H_X. A crucial property of the normal distribution is that if we condition on some of the x_i, $i = 1, \ldots, k$, we again obtain a normal distribution. Let us assume that we want to condition on the last l components of X:

$$Z = (x_{k-l+1}, \ldots, x_k)^T \in \mathbb{R}^l. \tag{4}$$

We partition the matrix H in a way that reflects this partition of X:

$$H_X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}, \tag{5}$$

with $A \in M_{k-l}$, $C \in M_l$, and $B \in M_{k-l,l}$. Let $\bar{X}$ be the first $k-l$ components of X:

$$\bar{X} = (x_1, \ldots, x_{k-l})^T \in \mathbb{R}^{k-l} \tag{6}$$

If we assume that X has zero mean (i.e. $\mu = 0$), the density function for X conditioned on Z is then given by

$$f(X|Z) = N_{k-l}[BC^{-1}Z, H_X/C](\bar{X}), \tag{7}$$

(see [Rao, 2002, chapter 8] and [Cottle, 1974]) where H_X/C is the Schur complement of C in H_X:

$$H_X/C = A - BC^{-1}B^T \tag{8}$$

(Emilie Haynsworth introduced the name and highlighted its usefulness in [Haynsworth, 1968]. Schur originally made use of it in [Schur, 1917]. For an overview of the properties of the Schur complement see [Horn, 2005].) We see that we again obtain a normal distribution. We will meet the Schur complement again soon.

Now let us look at a second set of random variables

$$Y = (x_{k-l+1}, \ldots, x_k, x_{k+1}, \ldots, x_n)^T \in \mathbb{R}^{n-k}, \tag{9}$$

that has an overlap Z with the random variables X. Let the density function for Y be given by

$$f(Y) = N_{n-k}[0, H_Y](Y), \tag{10}$$

with

$$H_Y = \begin{pmatrix} C & D \\ D^T & E \end{pmatrix}. \tag{11}$$

Note that the submatrix C is shared with H_X. In our context the matrix C might have been obtained by a previous calibration of a model that is part of both X and Y. If we were to condition on Z we would find

$$f(Y|Z) = N_{n-k-l}[D^T C^{-1} Z, H_Y/C](\bar{Y}), \tag{12}$$

with

$$H_Y/C = E - D^T C^{-1} D \tag{13}$$

and $\bar{Y}$ is the part of Y that is not Z. If we now want to describe X and Y together we are looking at the matrix

$$H = \begin{pmatrix} A & B & W \\ B^T & C & D \\ W^T & D^T & E \end{pmatrix}, \tag{14}$$

with a yet to be determined matrix $W \in M_{k-l,n-l}$. Because X and Y share the random variables in Z we can not make them independent. The next best thing that we can do is to demand that if we fix Z the remaining parts of X and Y are independent. The combined density for X and Y when we condition on Z is given by

$$f(X, Y|Z) = N_{n-l}[(B^T, D)^T C^{-1} Z, H/C](\bar{X}, \bar{Y}), \tag{15}$$

with

$$H/C = \begin{pmatrix} A - BC^{-1}B^T & W - BC^{-1}D \\ W^T - D^T C^{-1} B^T & E - D^T C^{-1} D \end{pmatrix} \tag{16}$$

$$= \begin{pmatrix} H_X/C & W - BC^{-1}D \\ W^T - D^T C^{-1} B^T & H_Y/C \end{pmatrix}. \tag{17}$$

4

We see that if we set

$$W = BC^{-1}D \tag{18}$$

the above matrix becomes block diagonal and the conditional densities obey

$$f(X, Y|Z) = f(X|Z)f(Y|Z) \tag{19}$$

because the blocks in the diagonal are H_X/C and H_Y/C from equations (8) and (13). Given that Z is part of both X and Y this is as much independence as we can ask for. The way to choose the matrix W and complete H is to demand that X and Y, when conditioned on their shared part Z, are independent.

This choice of W can be characterized in two more ways (see [Smith, 2008] and [Dreyer et al., 2021]). Banachiewicz showed that the inverse of a matrix can be expressed using the Schur complement. Since the Schur complement H/C is block diagonal the inverse of H is given by

$$H^{-1} = \begin{pmatrix} (H_X/C)^{-1} & -(H_X/C)^{-1}BC^{-1} & 0 \\ -C^{-1}B^T(H_X/C)^{-1} & \Xi & -C^{-1}D(H_Y/C)^{-1} \\ 0 & -(H_Y/C)^{-1}D^TC^{-1} & (H_Y/C)^{-1} \end{pmatrix}, \tag{20}$$

with

$$\Xi = C^{-1} + C^{-1}B^T(H_X/C)^{-1}BC^{-1} + C^{-1}D(H_Y/C)^{-1}D^TC^{-1}. \tag{21}$$

We see that H^{-1} has zeroes in the places where W sits in H. The W from equation (18) is the unique choice with that property.

When looking at the determinant of H we find another characterization of W. Using the Schur complement again, we find

$$\det H = \det C \ \det H/C \tag{22}$$

$$\leq \det C \ \det H_X/C \ \det H_Y/C, \tag{23}$$

where we have equality in equation (23) if and only if the off-diagonal blocks vanish, i.e. if and only if W is given by equation (18). This choice of W is the unique choice that maximizes the determinant of H.

The entropy S of the normal distribution $N_n[\mu, H]$ is given by

$$S = \frac{1}{2}\log \det H + \text{ const.} \tag{24}$$

Because we are maximizing the determinant of H we are also maximizing the entropy S of the distribution described by H. The H chosen by the W from equation (18) is the one that puts the fewest constraints on the distribution of X and Y.

3 The completion procedure

In the last section we have completed the matrix H in the very simple case where there are two known matrices H_X and H_Y with just one shared part C. The general case is usually more complicated. Building on the result in [Grone et al., 1984] Smith [Smith, 2008] (see also [Dreyer et al., 2021] for a more general treatment of the semidefinite case) showed how to solve the general problem by reducing it to a repeated application of the result from the previous section. We will discuss the required steps in this section by looking at a particular example: the cross currency model.

Our model will consist of two interest rate models; one for the base currency E and one the foreign currency A. These two currencies are linked by an exchange rate X. We will model the two interest rates and the exchange rate with stochastic volatility models so that we have a total of six stochastic variables:

$$E, \nu_E, A, \nu_A, X, \nu_X \tag{25}$$

We will denote the corresponding correlation coefficients by

$$(a, b), \tag{26}$$

where a and b are one of the stochastic variables from equation (25). In total there are 15 different correlation coefficients in the correlation matrix H of our model. We assume that the coefficients

$$(E, \nu_E) \text{ and } (A, \nu_A) \tag{27}$$

have already been determined in separate calibrations of the interest rate models. That leaves us with 13 free coefficients in H. This number of coefficients is too large for a stable calibration. Instead we will only calibrate the coefficients

$$(E, A), (E, X), (A, X), (X, \nu_X), \tag{28}$$

and determine the remaining nine coefficients using the completion procedure described here (see figure 1).

The first step to complete the matrix H is to determine the graph γ corresponding to H. The vertices of the graph are given by the row (or column) indices of H. In our case these are identified with the stochastic variables in equation (25). We then obtain the graph for H by connecting the vertices for which we have coefficients. In our example we obtain the graph in figure 2.

Now that we have the graph we need to determine if the graph is chordal. A graph is chordal if every loop of length four or more has a chord (i.e. an edge connecting two non-consecutive vertices in the loop; see e.g. [Golumbic, 2004]). In our example this is easy to check. The largest loop is the central loop consisting of the vertices E, A, and X. This loop

$$
\begin{pmatrix}
1 & \boxed{(E, \nu_E)}_{\text{dashed}} & \boxed{(E,A)} & (E, \nu_A) & \boxed{(E,X)} & (E, \nu_X) \\
 & 1 & (\nu_E, A) & (\nu_E, \nu_A) & (\nu_E, X) & (\nu_E, \nu_X) \\
 & & 1 & \boxed{(A, \nu_A)}_{\text{dashed}} & \boxed{(A,X)} & (A, \nu_X) \\
 & & & 1 & (\nu_A, X) & (\nu_A, \nu_X) \\
 & & & & 1 & \boxed{(X, \nu_X)} \\
 & & & & & 1
\end{pmatrix}
$$

Figure 1: The correlation matrix for the cross currency model. The coefficients (E, ν_E) and (A, ν_A) (the boxes with the dashed lines) are determined in separate interest rate calibrations. The coefficients (E, X), (E, A), (A, X), and (X, ν_X) (the boxes with the solid lines) will be determined in the cross currency calibration itself. The remaining coefficients are the result of the completion procedure.

has just three elements and thus does not require a chord. The procedure that we are describing in this section requires a chordal graph.

The next step is to identify the cliques in our graph. A clique is a maximal set of vertices that are all connected to each other (the subgraph induced by the set of vertices is complete; again, see [Golumbic, 2004]). In our example there are four such cliques given by the central loop and the three arms of the graph:

$$\alpha_C = \{E, A, X\} \tag{29}$$
$$\alpha_E = \{E, \nu_E\} \tag{30}$$
$$\alpha_A = \{A, \nu_A\} \tag{31}$$
$$\alpha_X = \{X, \nu_X\} \tag{32}$$

Note that the cliques have non-trivial intersections. We can use these intersections to construct a new graph Γ. The vertices of this new graph are given by the cliques $\alpha_C, \ldots, \alpha_X$ and we connect two cliques if their intersection is non-empty (see figure 3).

Because the graph γ for the matrix H is chordal, there exists a clique tree, i.e. a spanning tree for the graph Γ that possesses the intersection property. A spanning tree for Γ is a tree that has the same vertices as the graph itself. It possesses the intersection property if for any two vertices
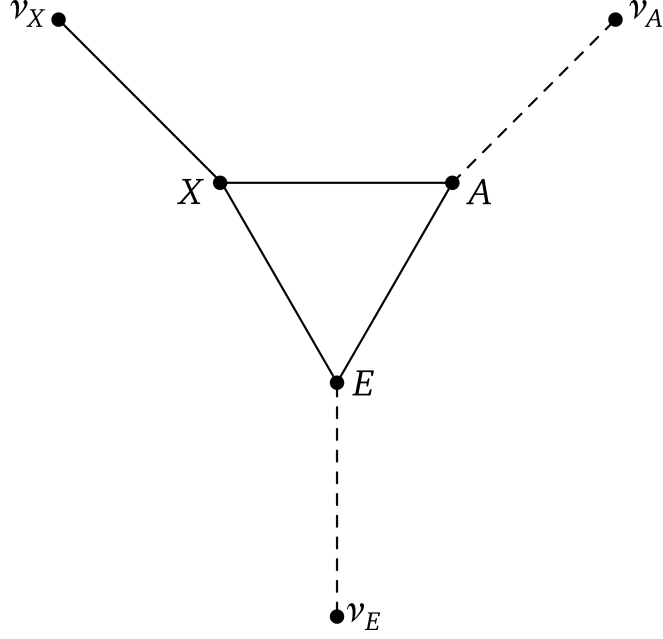
Figure 2: The initial graph γ for the cross currency calibration. The coefficients (E, ν_E) and (A, ν_A) (dashed edges) are determined in separate interest rate calibrations. The other edges are determined in the cross currency calibration itself.

α_1 and α_2 the intersection $\alpha_1 \cap \alpha_2$ is contained in all the vertices on the unique path in the tree that connects α_1 to α_2 (see [Blair et al., 1993] for the different characterizations of a chordal graph). In our example the graph Γ is itself a tree. It is also easy to check that it possesses the intersection property. In the next section we will see an example where the intersection property is not trivially fulfilled. We will choose $\alpha_E = \{E, \nu_E\}$ as the root of the tree.

We now start completing the matrix H by repeatedly applying equation (18) from the previous section. We start with the cliques on the top of the tree and work our way down. We start with the clique $\alpha_X = \{X, \nu_X\}$. It intersects the clique α_C in the element X. This leads us to make the following identifications:

$$H_X = \begin{pmatrix} 1 & (X, \nu_X) \\ & 1 \end{pmatrix} \tag{33}$$

$$H_Y = \begin{pmatrix} 1 & (E, X) & (A, X) \\ & 1 & (E, A) \\ & & 1 \end{pmatrix} \tag{34}$$

8

Note that the common matrix C is just given by 1. Equation (18) now gives:

$$W = ((E, \nu_X), (A, \nu_X)) \tag{35}$$
$$= BC^{-1}D \tag{36}$$
$$= (X, \nu_X)((E, X), (A, X))) \tag{37}$$
$$= ((X, \nu_X)(E, X), (X, \nu_X)(A, X))), \tag{38}$$

or

$$(E, \nu_X) = (X, \nu_X)(E, X) \tag{39}$$
$$(A, \nu_X) = (X, \nu_X)(A, X). \tag{40}$$

Figure 5 shows the first completion step in the graph γ.

We have one more clique with height two in the clique tree for γ. In the next step we take the clique α_A and combine it with the newly created clique from the last step. This time we make the identifications

$$H_X = \begin{pmatrix} 1 & (A, \nu_A) \\ & 1 \end{pmatrix} \tag{41}$$

$$H_Y = \begin{pmatrix} 1 & (A, E) & (A, X) & (A, \nu_X) \\ & 1 & (E, X) & (E, \nu_X) \\ & & 1 & (X, \nu_X) \\ & & & 1 \end{pmatrix} \tag{42}$$

We apply equation (18) again to obtain:

$$W = ((E, \nu_A), (X, \nu_A), (\nu_X, \nu_A)) \tag{43}$$
$$= BC^{-1}D \tag{44}$$
$$= (A, \nu_A)((E, A), (A, X), (A, \nu_X)), \tag{45}$$

or

$$(E, \nu_A) = (A, \nu_A)(E, A) \tag{46}$$
$$(X, \nu_A) = (A, \nu_A)(A, X) \tag{47}$$
$$(\nu_X, \nu_A) = (A, \nu_A)(A, \nu_X) \tag{48}$$
$$= (A, \nu_A)(X, \nu_X)(A, X) \tag{49}$$

Figure 6 shows the result of this step in the graph γ.

There are no more cliques of height two left. Looking at the graph in figure 6 we see that only the edges connecting ν_E to the large clique that is the result of the last step are missing. These missing edges will be created

in the last step. We make the following identifications:

$$H_X = \begin{pmatrix} 1 & (E, \nu_E) \\ & 1 \end{pmatrix} \tag{50}$$

$$H_Y = \begin{pmatrix} 1 & (E,A) & (E,\nu_A) & (E,X) & (E,\nu_X) \\ & 1 & (A,\nu_A) & (A,X) & (A,\nu_X) \\ & & 1 & (\nu_A,X) & (\nu_A,\nu_X) \\ & & & 1 & (X,\nu_X) \\ & & & & 1 \end{pmatrix} \tag{51}$$

Applying equation (18) one more time gives:

$$W = ((\nu_E, A), (\nu_E, \nu_A), (\nu_E, X), (\nu_E, \nu_X)) \tag{52}$$
$$= BC^{-1}D \tag{53}$$
$$= (E, \nu_E)((E, A), (E, \nu_A), (E, X), (E, \nu_X)), \tag{54}$$

or

$$(\nu_E, A) = (E, \nu_E)(E, A) \tag{55}$$
$$(\nu_E, \nu_A) = (E, \nu_E)(E, \nu_A) \tag{56}$$
$$= (E, \nu_E)(A, \nu_A)(E, A) \tag{57}$$
$$(\nu_E, X) = (E, \nu_E)(E, X) \tag{58}$$
$$(\nu_E, \nu_X) = (E, \nu_E)(E, \nu_X) \tag{59}$$
$$= (E, \nu_E)(X, \nu_X)(E, X) \tag{60}$$

With this step we have added the last missing edges in the graph γ and have obtained a positive definite completion of the matrix H.

4 Several currencies

In the previous section we have just looked at one foreign currency A with one exchange rate X. The procedure can easily be extended to an arbitrary number of foreign currencies. We will denote the foreign currencies by a capital letter:

$$A, B, C, \ldots \tag{61}$$

The exchange rate between the domestic currency E and the foreign currency A will now be denoted by

$$X_E^A. \tag{62}$$

If m_A is an amount in currency A then $X_E^A m_A$ is the corresponding amount in currency E. If we are given the same correlation coefficients as in the previous section we obtain the graph in figure 7. The corresponding graph

of cliques is shown in figure 8. The corresponding clique tree is obtained by erasing the lines that connect the cliques

$$\alpha_{CA}, \alpha_{CB}, \ldots, \alpha_{CG} \tag{63}$$

on the middle layer. The intersection of any two of these cliques is given by

$$\{E\} \subset \alpha_E. \tag{64}$$

Since the path connecting the two cliques passes through α_E this tree satisfies the intersection property.

5 Conclusion

In this paper we have described a method to complete an insufficiently specified correlation matrix. Because financial models are getting more complex such matrices are a common occurrence.

Acknowledgement

References

[Higham, 2002] , Higham, N. J. (2002). Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis* 22: 329 - 343.

[Grone et al., 1984] Grone R., Johnson C.R., Sa E., Wolkowicz H. (1984). Positive definite completion of partial Hermitian matrices. *Linear Algebra Appl.* 58: 109–124.

[Smith, 2008] Smith, R. L. (2008). The positive definite completion problem revisited. *Linear Algebra and its Applications* 429:1442 – 1452.

[Rao, 2002] C. R. Rao (2002). *Linear Statistical Inference and its Applications, 2nd edition.* John Wiley & Sons, Inc.

[Cottle, 1974] Cottle, R. W. (1974). Manifestations of the Schur complement. *Linear Algebra and its Applications* 8:189–211.

[Haynsworth, 1968] Haynsworth, E.V. (1968), Determination of the inertia of a partitioned Hermitian matrix, *Linear Algebra and its Applications* 1:73–81.

[Schur, 1917] Schur, J. (1917). Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind. *Journal für die reine und angewandte Mathematik* 147: 205–232.

[Horn, 2005] Horn, R. A. (2005). Basic Properties of the Schur Complement, 17–46. In Fuzhen Zhang (Ed.) *The Schur complement and its application.* Springer.

[Dreyer et al., 2021] our other paper ...

[Golumbic, 2004] Golumbic, M. C. (2004). *Algorithmic Graph Theory and Perfect Graphs.* Elsevier.

[Blair et al., 1993] Blair, J. R. S., Peyton, B. (1993). An introduction to chordal graphs and clique trees, 1 – 29. In George, A., Gilbert, J. R., Liu, J. W. H. (Eds.). *Graph Theory and Sparse Matrix Computation*, The IMA Volumes in Mathematics and its Applications 56. Springer.
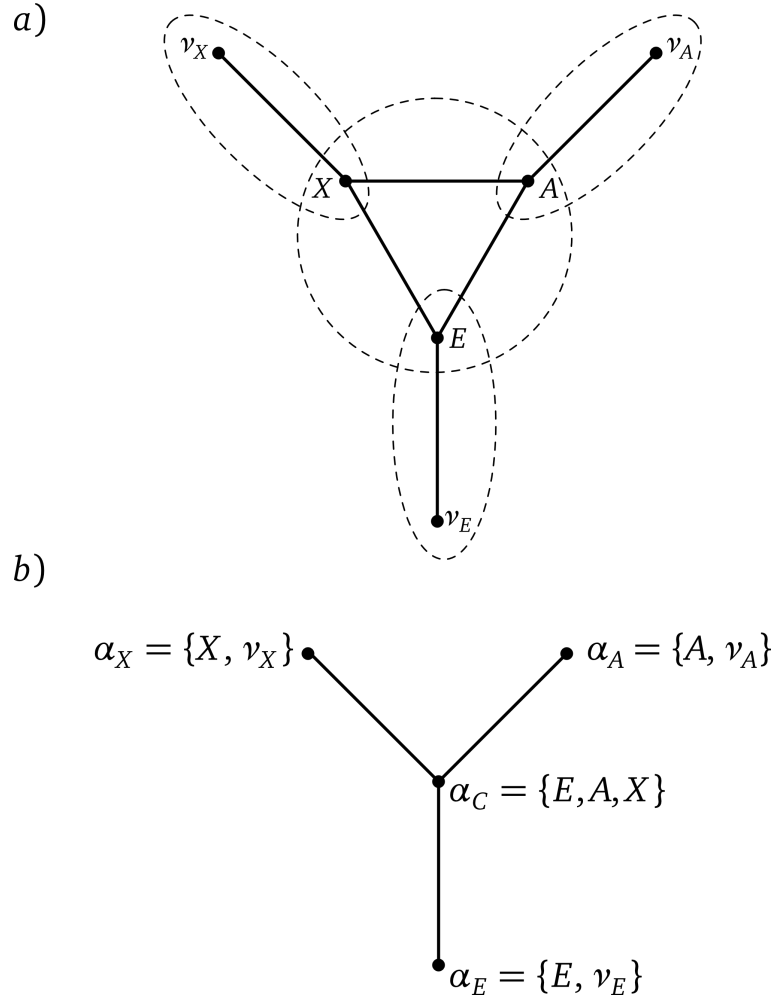
a)

v_X •   • v_A

X A

E

v_E •

b)

$\alpha_X = \{X, v_X\}$ •   • $\alpha_A = \{A, v_A\}$

$\alpha_C = \{E, A, X\}$

$\alpha_E = \{E, v_E\}$

Figure 3: a) The graph for the cross currency has four cliques. b) These cliques are the vertices of a new graph Γ. An edge in this new graph connects two cliques that have a non-empty intersection.
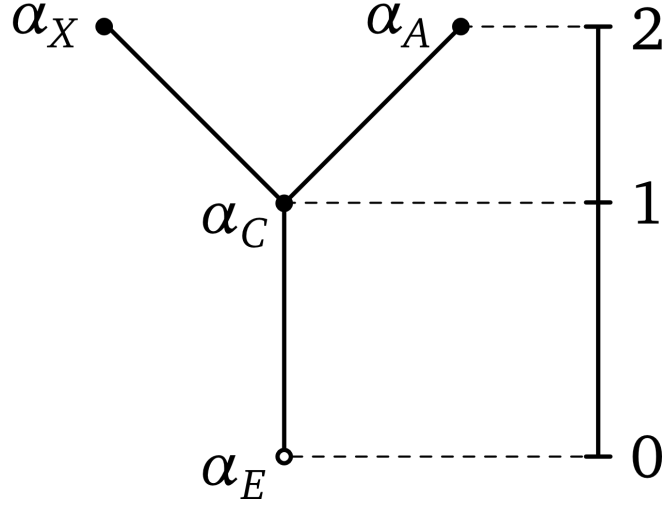
Figure 4: The clique tree for the graph γ. The numbers on the right denote the height in the tree if we choose α_E to be the root of the tree. The completion procedure proceeds from the top of the tree to the bottom. We complete the matrix by repeatedly using equation (18) from the previous section.
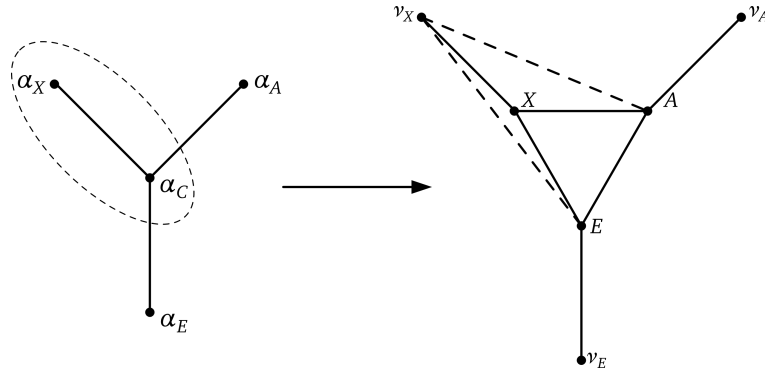

Figure 5: We start the completion of the matrix H with the cliques of height two. Applying equation (18) to the matrices given by the cliques α_X and α_C adds the dashed lines to the graph γ.
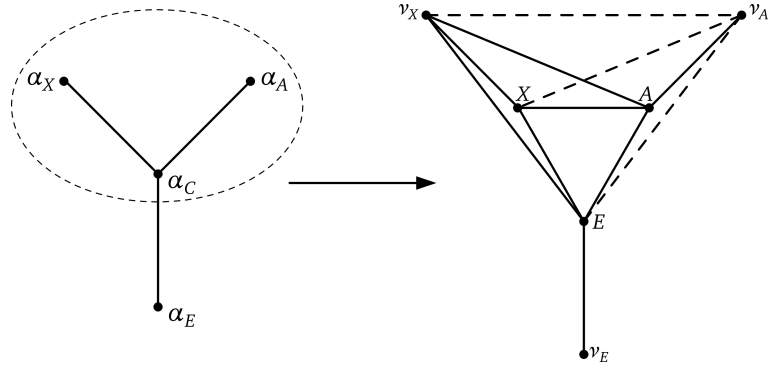
Figure 6: In the second step we combine the clique from the first step with the remaining clique of height two in the tree. This step adds the three edges (ν_X, ν_A), (X, ν_A), and (E, ν_A).
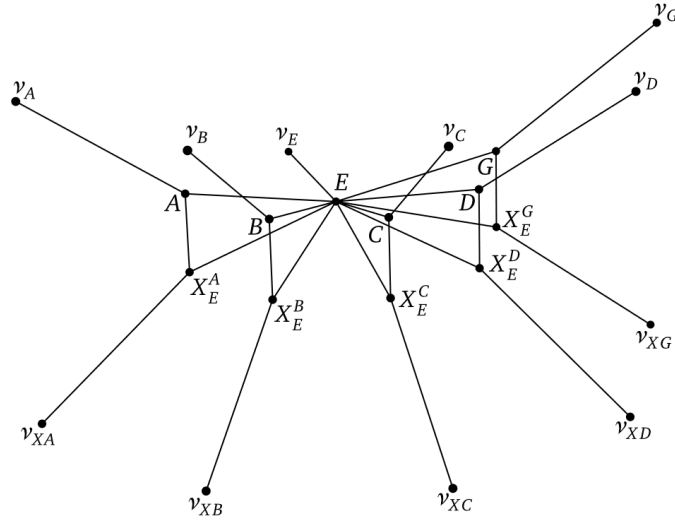

Figure 7: The graph corresponding to the incomplete correlation matrix for five foreign currencies.
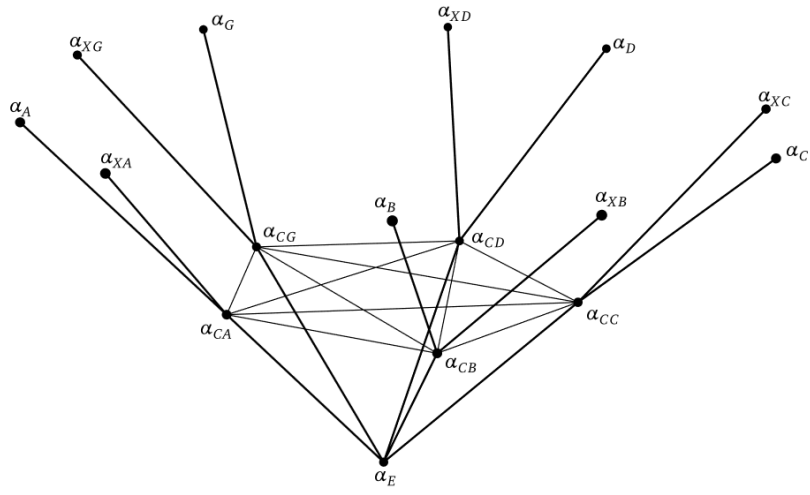
Figure 8: The graph of cliques for the case of several foreign currencies. A clique tree is obtained by removing the lighter edges in the middle layer. It is again a tree of height two.