Multivariate localization

Benjamin Ménétrier

Last update: November 4, 2024

Contents

1 Framework

1.1 State vector definition

The state vector $\mathbf{x} \in \mathbb{R}^n$ is split into Q concatenated sub-vectors $\mathbf{x}_q$ (called "groups"), each one gathering P_q sub-vectors $\mathbf{x}_{p,q}$ of size n_q (called "variables"). Each sub-vector $\mathbf{x}_{p,q}$ corresponds to the data for a given physical variable (e.g. zonal wind, temperature, humidity, etc.). In a given group, all variables have the same size (n_q). Thus:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \hline \vdots \\ \hline \mathbf{x}_Q \end{pmatrix} \quad \text{and} \quad \mathbf{x}_q = \begin{pmatrix} \mathbf{x}_{1,q} \\ \hline \vdots \\ \hline \mathbf{x}_{P_q,q} \end{pmatrix} \tag{1}$$

As a consequence:

- The total number of variables is $P = \sum_{q=1}^{Q} P_q$.

- Each group $\mathbf{x}_q$ has a size $P_q n_q$, so the full state vector size is $n = \sum_{q=1}^{Q} P_q n_q$.

1.2 Illustration setup

A simple example will be used as an illustration throughout this note. We define a vector $\mathbf{x}$ with $Q = 3$ groups, where $P_1 = 3$, $P_2 = 1$ and $P_3 = 2$:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{1,1} \\ \mathbf{x}_{2,1} \\ \mathbf{x}_{3,1} \\ \hline \mathbf{x}_{1,2} \\ \hline \mathbf{x}_{1,3} \\ \mathbf{x}_{2,3} \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} \mathbf{x} \\ \mathbf{x} \\ \mathbf{x} \end{matrix}}\right\} \mathbf{x}_1 \ (1^{\text{st}} \text{ group of variables}) \\ \left.\vphantom{\begin{matrix} \mathbf{x} \end{matrix}}\right\} \mathbf{x}_2 \ (2^{\text{nd}} \text{ group of variables}) \\ \left.\vphantom{\begin{matrix} \mathbf{x} \\ \mathbf{x} \end{matrix}}\right\} \mathbf{x}_3 \ (3^{\text{rd}} \text{ group of variables}) \end{array} \tag{2}$$

1.3 Localization square-root

The localization matrix $\mathbf{L} \in \mathbb{R}^{n \times n}$ is applied to $\mathbf{x}$. The main issue is to design a positive semi-definite $\mathbf{L}$ matrix, and the simplest solution is to build it as:

$$\mathbf{L} = \mathbf{U}\mathbf{U}^{\mathrm{T}} \tag{3}$$

where $\mathbf{U} \in \mathbb{R}^{n \times m}$ is called a "square-root" of the localization matrix $\mathbf{L}$. The number of colums m of the square-root $\mathbf{U}$ can be smaller or larger than the number of lines n. The control vector $\mathbf{v} \in \mathbb{R}^m$ is such that:

$$\mathbf{x} = \mathbf{U}\mathbf{v} \tag{4}$$

We can define a component $\mathbf{U}_q \in \mathbb{R}^{n_q \times m_q}$ for each group. The relationship between the number of columns m_q of each component $\mathbf{U}_q$ and the total number of columns m of $\mathbf{U}$ depends on the multivariate strategy.

The main idea behind this distribution of variables into groups is to apply a localization based on the same component to all the variables pairs within a given group. Thus, the number of distinct component can be reduced if the number of groups is smaller than the number of variables.

2 Zero cross-localization between groups

For the first three multivariate strategies, cross-correlations between groups are cut off (i.e. cross-localization is zero). Localization is non-zero only for auto-correlations and cross-correlations between variables of the same group.

2.1 Univariate localization

For the case where even cross-correlations between variables of the same group are cut off, the square-root $\mathbf{U}$ can be defined as:

$$
\mathbf{U} = \left(
\begin{array}{ccc|c|cc}
\mathbf{U}_1 & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{U}_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{U}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & \mathbf{U}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \mathbf{U}_3 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{U}_3
\end{array}
\right) \tag{5}
$$

Thus:

$$
\mathbf{L} = \left(
\begin{array}{ccc|c|cc}
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & \mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}}
\end{array}
\right) \tag{6}
$$

In this case, the total control vector size is $m = \sum\limits_{q=1}^{Q} P_q m_q$

2.2 Duplicated localization

It is possible to define the cross-localization between the variables of a same group as duplicates of the diagonal localization, by defining $\mathbf{U}$ as:

$$
\mathbf{U} = \left(\begin{array}{c|c|c}
\mathbf{U}_1 & 0 & 0 \\
\mathbf{U}_1 & 0 & 0 \\
\mathbf{U}_1 & 0 & 0 \\
\hline
0 & \mathbf{U}_2 & 0 \\
\hline
0 & 0 & \mathbf{U}_3 \\
0 & 0 & \mathbf{U}_3
\end{array}\right)
$$

$$
= \left(\begin{array}{c|c|c}
\mathbf{I}_1 & 0 & 0 \\
\mathbf{I}_1 & 0 & 0 \\
\mathbf{I}_1 & 0 & 0 \\
\hline
0 & \mathbf{I}_2 & 0 \\
\hline
0 & 0 & \mathbf{I}_3 \\
0 & 0 & \mathbf{I}_3
\end{array}\right)
\left(\begin{array}{c|c|c}
\mathbf{U}_1 & 0 & 0 \\
\hline
0 & \mathbf{U}_2 & 0 \\
\hline
0 & 0 & \mathbf{U}_3
\end{array}\right)
\tag{7}
$$

where $\mathbf{I}_q \in \mathbf{R}^{n_q, n_q}$ is the identity matrix of size n_q. Thus:

$$
\mathbf{L} = \left(\begin{array}{ccc|c|cc}
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & \mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \\
0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}}
\end{array}\right)
\tag{8}
$$

* In this case, the total control vector size is $m = \displaystyle\sum_{q=1}^{Q} m_q$

2.3 Duplicated and weighted localization

A variant of the previous case uses group-specific lower triangular matrices $\mathbf{S}^q \in \mathbb{R}^{P_q \times P_q}$ to define $\mathbf{U}$ as:

$$
\mathbf{U} = \left(\begin{array}{ccc|c|cc}
S^1_{1,1}\mathbf{U}_1 & 0 & 0 & 0 & 0 & 0 \\
S^1_{2,1}\mathbf{U}_1 & S^1_{2,2}\mathbf{U}_1 & 0 & 0 & 0 & 0 \\
S^1_{3,1}\mathbf{U}_1 & S^1_{3,2}\mathbf{U}_1 & S^1_{3,3}\mathbf{U}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & S^2_{1,1}\mathbf{U}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & S^3_{1,1}\mathbf{U}_3 & 0 \\
0 & 0 & 0 & 0 & S^3_{2,1}\mathbf{U}_3 & S^3_{2,2}\mathbf{U}_3
\end{array} \right)
$$

$$
= \left(\begin{array}{ccc|c|cc}
S^1_{1,1}\mathbf{I}_1 & 0 & 0 & 0 & 0 & 0 \\
S^1_{2,1}\mathbf{I}_1 & S^1_{2,2}\mathbf{I}_1 & 0 & 0 & 0 & 0 \\
S^1_{3,1}\mathbf{I}_1 & S^1_{3,2}\mathbf{I}_1 & S^1_{3,3}\mathbf{I}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & S^2_{1,1}\mathbf{I}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & S^3_{1,1}\mathbf{I}_3 & 0 \\
0 & 0 & 0 & 0 & S^3_{2,1}\mathbf{I}_3 & S^3_{2,2}\mathbf{I}_3
\end{array} \right)
\left(\begin{array}{ccc|c|cc}
\mathbf{U}_1 & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{U}_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{U}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & \mathbf{U}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \mathbf{U}_3 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{U}_3
\end{array} \right)
\tag{9}
$$

Thus:

$$
\mathbf{L} = \left(\begin{array}{ccc|c|cc}
W^1_{1,1}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{1,2}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{1,3}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
W^1_{2,1}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{2,2}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{2,3}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
W^1_{3,1}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{3,2}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{3,3}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & W^2_{1,1}\mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & W^3_{1,1}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & W^3_{1,2}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \\
0 & 0 & 0 & 0 & W^3_{2,1}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & W^3_{2,2}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}}
\end{array} \right)
\tag{10}
$$

where $\mathbf{W}^q = \mathbf{S}^q\mathbf{S}^{q\mathrm{T}}$ provides the weights applied to each covariance between variables within the group q. In practice, it is easier to set $\mathbf{W}^q$ and to compute $\mathbf{S}^q$ using a Cholesky decomposition. It also seems reasonable to set the diagonal elements of $\mathbf{W}^q$ to one.

In this case, the total control vector size is $m = \sum_{q=1}^{Q} P_q m_q$

3 Implicit cross-localization between groups

For the last multivariate strategy, cross-correlations between groups are defined implicitly as the product of components of both groups.

3.1 Crossed localization

In this case, $\mathbf{U}$ is reduced to a column:

$$
\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_1 \\ \mathbf{U}_1 \\ \hline \mathbf{U}_2 \\ \hline \mathbf{U}_3 \\ \mathbf{U}_3 \end{pmatrix}
$$

$$
= \left(\begin{array}{c|c|c} \mathbf{I}_1 & 0 & 0 \\ \mathbf{I}_1 & 0 & 0 \\ \mathbf{I}_1 & 0 & 0 \\ \hline 0 & \mathbf{I}_2 & 0 \\ \hline 0 & 0 & \mathbf{I}_3 \\ 0 & 0 & \mathbf{I}_3 \end{array} \right) \begin{pmatrix} \mathbf{U}_1 \\ \hline \mathbf{U}_2 \\ \hline \mathbf{U}_3 \end{pmatrix} \tag{11}
$$

An additional assumption is required here: the components $\mathbf{U}_q$ must all have the same number of columns: $m_1 = ... = m_Q = \overline{m}$. Thus:

$$
\mathbf{L} = \left(\begin{array}{ccc|c|cc} \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} \\ \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} \\ \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} \\ \hline \mathbf{U}_2\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_3^{\mathrm{T}} \\ \hline \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \\ \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \end{array} \right) \tag{12}
$$

In this case, the total control vector size is $m = \overline{m}$.

3.2 Cross-localization amplitude

The amplitude of cross-localization is not one in this case. For $1 \leq q, q' \leq Q$, we denote $\mathbf{u}$ the $\mathrm{k^{th}}$ column of $\mathbf{U}_q$ and $\mathbf{u}'$ the $\mathrm{k^{th}}$ column of $\mathbf{U}_{q'}$. The $\mathrm{k^{th}}$ diagonal coefficient of $\mathbf{U}_q\mathbf{U}_{q'}^{\mathrm{T}}$ is given by $\langle \mathbf{u}, \mathbf{u}' \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the canonical inner product. The Cauchy-Schwartz inequality requires that:

$$
|\langle \mathbf{u}, \mathbf{u}' \rangle| \leq \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{u}', \mathbf{u}' \rangle} \tag{13}
$$

Thus, the cross-localization amplitude between groups q and q' is necessarily smaller than the geometric mean of the auto-localization amplitudes for groups q and q'. Besides, the more the vectors $\mathbf{u}$ and $\mathbf{u}'$ have different shapes, the smaller their inner product and the smaller the cross-localization amplitude.
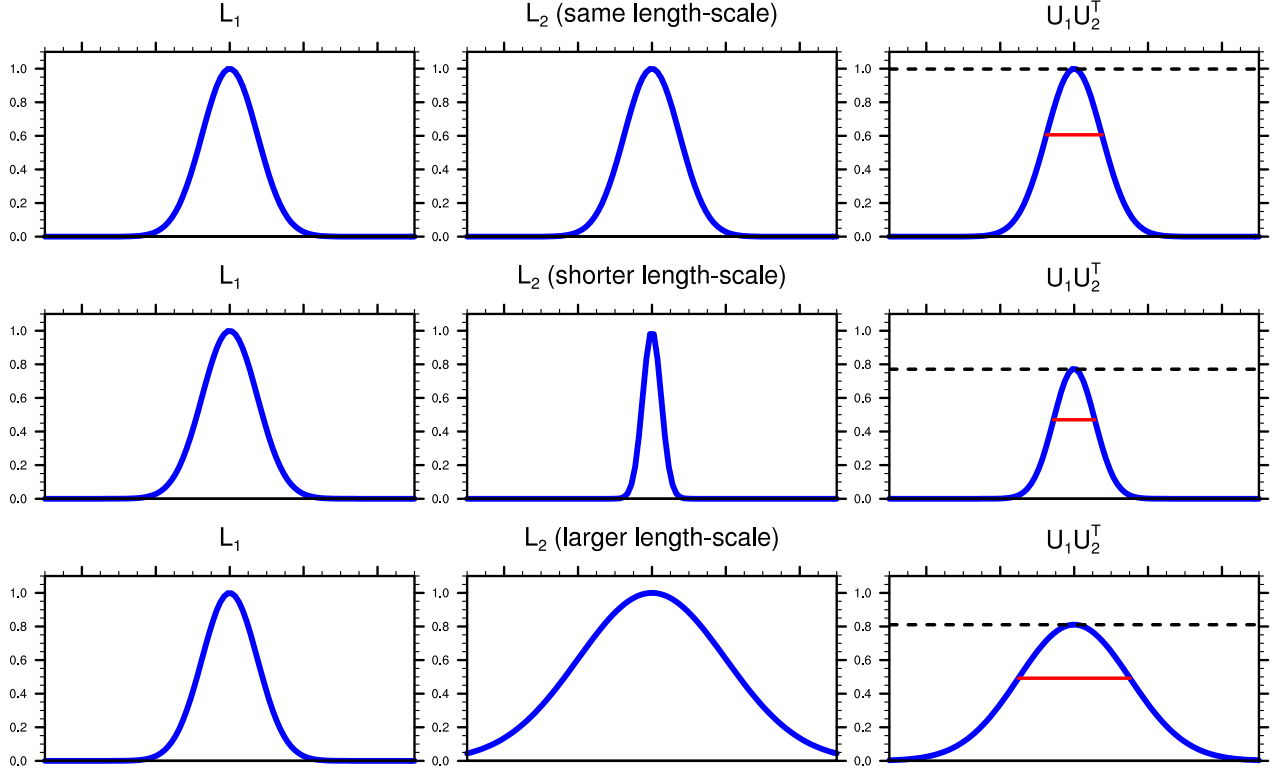
Figure 1: 1D illustration of the product of two localization square-root, with Gaussian localization functions. The dashed black line shows the function amplitude, while the solid red line shows the function length-scale.

3.3 Mixing 3D and 2D variables

In many applications, the state vector contains both 3D and 2D variables. To use the crossed method in this case, it is necessary to consider the 2D variables as 3D variables with only one level containing meaningful data (the first or the last level in general) and the other levels set to zero.

For sake of simplicity, we assume here that group 1 contains one 3D variable and group 2 contains one 2D variable. We also assume that both control variable $\mathbf{x}$ and control vector $\mathbf{v}$ can be split into K components, each one handling one vertical level:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{1,k=1} \\ \vdots \\ \mathbf{x}_{1,k=K} \\ \hline \mathbf{x}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} \mathbf{v}_{k=1} \\ \vdots \\ \mathbf{v}_{k=K} \end{pmatrix} \tag{14}$$

We can define an extended control variable:

$$
\mathbf{x}' = \begin{pmatrix} \mathbf{x}_{1,k=1} \\ \vdots \\ \mathbf{x}_{1,k=K} \\ \hline \mathbf{x}'_{2,k=1} \\ \vdots \\ \mathbf{x}'_{2,k=K} \end{pmatrix} = \mathbf{E}\mathbf{x} \tag{15}
$$

where $\mathbf{E}$ is the extension operator padding with zeros, so that both sub-vectors have the same size. In general:

- either $\mathbf{x}'_{2,k=1} = \mathbf{x}_2$ and $\mathbf{x}'_{2,k=2}$ to $\mathbf{x}'_{2,k=K}$ are set to zero,

- or $\mathbf{x}'_{2,k=K} = \mathbf{x}_2$ and $\mathbf{x}'_{2,k=1}$ to $\mathbf{x}'_{2,k=K-1}$ are set to zero.

The adjoint of $\mathbf{E}$, denoted $\mathbf{E}^{\mathrm{T}}$, is thus a reduction operator removing the zeroed levels from the extended control variable $\mathbf{x}'$ to yield the control variable $\mathbf{x}$:

$$
\mathbf{x} = \mathbf{E}^{\mathrm{T}}\mathbf{E}\mathbf{x} \tag{16}
$$

We define an extended square-root $\mathbf{U}'_2$:

$$
\mathbf{U}'_2 = \begin{pmatrix} \mathbf{U}_2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{or} \quad \mathbf{U}'_2 = \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \mathbf{U}_2 \end{pmatrix} \tag{17}
$$

depending on the location of the 2D level in the 3D grid, at level 1 or K, respectively. The extended square-root $\mathbf{U}$ is then given by:

$$
\mathbf{U}' = \begin{pmatrix} \mathbf{U}_1 \\ \hline \mathbf{U}'_2 \end{pmatrix} \tag{18}
$$

Finally, we get

$$
\mathbf{x}' = \mathbf{U}'\mathbf{v} \tag{19}
$$

so we can conclude that:

$$
\mathbf{U} = \mathbf{E}^{\mathrm{T}}\mathbf{U}' \tag{20}
$$

In practice, the operator $\mathbf{E}^{\mathrm{T}}\mathbf{U}'$ can be optimized to avoid the storage in memory of the extended variable $\mathbf{x}'$, keeping the relevant level only for the 2D variables.