# Supplement to "A persistence-based Wold-type decomposition for stationary time series"

(*Quantitative Economics*, Vol. 11, No. 1, January 2020, 203–230)

Fulvio Ortu
Department of Finance, Università Bocconi and IGIER

Federico Severino
Département de Finance, Assurance et Immobilier, FSA, Université Laval and Department of Economics, USI Lugano

Andrea Tamoni
Department of Finance and Economics, Rutgers Business School

Claudio Tebaldi
Department of Finance, Università Bocconi, IGIER, and Baffi-CAREFIN

## A. Online supplement

### A.1 *Proofs*

The notation employed here is taken from Section 2.1. Lemma A.1 is preparatory for the proof of Theorem 1.

Lemma A.1. *Let $\boldsymbol{\varepsilon}$ be a unit variance white noise. The Hilbert space $\mathcal{H}_t(\boldsymbol{\varepsilon})$ decomposes into the orthogonal sum $\mathcal{H}_t(\boldsymbol{\varepsilon}) = \bigoplus_{j=1}^{\infty} \mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}}$, where*

$$\mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}} = \left\{ \sum_{k=0}^{+\infty} b_k^{(j)} \varepsilon_{t-k2^j}^{(j)} \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : b_k^{(j)} \in \mathbb{R} \right\}$$

*and, for any $j \in \mathbb{N}$ and $t \in \mathbb{Z}$, $\varepsilon_t^{(j)}$ is given by equation* (6).

Proof. $\mathcal{H}_t(\boldsymbol{\varepsilon})$ is a Hilbert subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$, equipped with the inner product $\langle A, B\rangle = \mathbb{E}[AB]$ for all $A, B \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. We begin with showing that the scaling operator $\mathbf{R}$ is well-defined, linear, and isometric on $\mathcal{H}_t(\boldsymbol{\varepsilon})$.

Consider any $X = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}$ in $\mathcal{H}_t(\boldsymbol{\varepsilon})$, that is, $\|X\|^2 = \sum_{p=0}^{\infty} a_p^2 < +\infty$. Then

$$\|\mathbf{R}X\|^2 = \frac{1}{2}\sum_{k=0}^{+\infty} a^2_{\lfloor \frac{k}{2} \rfloor} = \frac{1}{2}\sum_{p=0}^{+\infty} a^2_{\lfloor \frac{2p}{2} \rfloor} + \frac{1}{2}\sum_{p=0}^{+\infty} a^2_{\lfloor \frac{2p+1}{2} \rfloor} = \sum_{p=0}^{+\infty} a_p^2 = \|X\|^2$$

Fulvio Ortu: fulvio.ortu@unibocconi.it
Federico Severino: federico.severino@fsa.ulaval.ca
Andrea Tamoni: andrea.tamoni.research@gmail.com
Claudio Tebaldi: claudio.tebaldi@unibocconi.it


Available at http://qeconomics.org. https://doi.org/10.3982/QE994

and this quantity is finite. As a result, $\mathbf{R}$ is a well-defined (and bounded) operator. The linearity of $\mathbf{R}$ is immediate. To prove that $\mathbf{R}$ is isometric, take any $X=\sum_{k=0}^{\infty} a_k \varepsilon_{t-k}$, $Y=\sum_{h=0}^{\infty} b_h \varepsilon_{t-h}$ in $\mathcal{H}_t(\boldsymbol{\varepsilon})$. By the white-noise properties of $\boldsymbol{\varepsilon}$,

$$
\begin{aligned}
\langle \mathbf{R}X, \mathbf{R}Y\rangle &= \sum_{k=0}^{+\infty} \frac{a_{\lfloor \frac{k}{2}\rfloor}}{\sqrt{2}} \frac{b_{\lfloor \frac{k}{2}\rfloor}}{\sqrt{2}} = \frac{1}{2}\sum_{p=0}^{+\infty} a_{\lfloor \frac{2p}{2}\rfloor} b_{\lfloor \frac{2p}{2}\rfloor} + \frac{1}{2}\sum_{p=0}^{+\infty} a_{\lfloor \frac{2p+1}{2}\rfloor} b_{\lfloor \frac{2p+1}{2}\rfloor} \\
&= \sum_{p=0}^{+\infty} a_p b_p = \langle X, Y\rangle.
\end{aligned}
$$

As a result, $\mathbf{R}$ is an isometry on $\mathcal{H}_t(\boldsymbol{\varepsilon})$ and the abstract Wold theorem (i.e., Theorem 1.1 in Nagy, Foias, Bercovici, and Kérchy (2010)) applies.

The abstract Wold theorem supplies the orthogonal decomposition $\mathcal{H}_t(\boldsymbol{\varepsilon}) = \hat{\mathcal{H}}_t(\boldsymbol{\varepsilon}) \oplus \tilde{\mathcal{H}}_t(\boldsymbol{\varepsilon})$, where

$$
\hat{\mathcal{H}}_t(\boldsymbol{\varepsilon}) = \bigcap_{j=0}^{+\infty} \mathbf{R}^j \mathcal{H}_t(\boldsymbol{\varepsilon}), \qquad \tilde{\mathcal{H}}_t(\boldsymbol{\varepsilon}) = \bigoplus_{j=1}^{+\infty} \mathbf{R}^{j-1} \mathcal{L}_t^{\mathbf{R}}
$$

and $\mathcal{L}_t^{\mathbf{R}} = \mathcal{H}_t(\boldsymbol{\varepsilon}) \ominus \mathbf{R}\mathcal{H}_t(\boldsymbol{\varepsilon})$ is called *wandering subspace*.

In particular, we show that $\hat{\mathcal{H}}_t(\boldsymbol{\varepsilon})$ is the null subspace. Indeed, the subspaces $\mathbf{R}^j\mathcal{H}_t(\boldsymbol{\varepsilon})$ are made of linear combinations of innovations $\varepsilon_t$ with coefficients equal to each others $2^j$-by-$2^j$:

$$
\mathbf{R}^j \mathcal{H}_t(\boldsymbol{\varepsilon}) = \left\{ \sum_{k=0}^{+\infty} c_k^{(j)} \left( \sum_{i=0}^{2^j-1} \varepsilon_{t-k2^j-i} \right) \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : c_k^{(j)} \in \mathbb{R} \right\}.
$$

As a result, $\hat{\mathcal{H}}_t(\boldsymbol{\varepsilon})$ can just include variables as $\sum_{h=0}^{\infty} c\varepsilon_{t-h}$ with $c \in \mathbb{R}$. These elements belong to $\mathcal{H}_t(\boldsymbol{\varepsilon})$, hence $\sum_{k=0}^{\infty} c^2$ is finite and this is possible just in case $c=0$. As a result, $\hat{\mathcal{H}}_t(\boldsymbol{\varepsilon}) = \{0\}$ and $\mathcal{H}_t(\boldsymbol{\varepsilon}) = \tilde{\mathcal{H}}_t(\boldsymbol{\varepsilon})$.

We now focus on the subspace $\tilde{\mathcal{H}}_t(\boldsymbol{\varepsilon})$. As the orthogonal complement of $\mathbf{R}\mathcal{H}_t(\mathbf{x})$ is the kernel of the adjoint operator $\mathbf{R}^*$ (see, e.g., Theorem 1, Section 6.6 in Luenberger (1968)), we determine $\mathbf{R}^*$. Specifically, $\mathbf{R}^* : \mathcal{H}_t(\boldsymbol{\varepsilon}) \longrightarrow \mathcal{H}_t(\boldsymbol{\varepsilon})$ is such that

$$
\mathbf{R}^* : \sum_{k=0}^{+\infty} a_k \varepsilon_{t-k} \longmapsto \sum_{k=0}^{+\infty} \frac{a_{2k}+a_{2k+1}}{\sqrt{2}} \varepsilon_{t-k}.
$$

Indeed, $\mathbf{R}^*$ is well-defined and the relation $\langle \mathbf{R}X, Y\rangle = \langle X, \mathbf{R}^*Y\rangle$ holds for any $X=\sum_{h=0}^{\infty} b_h \varepsilon_{t-h}$, $Y=\sum_{k=0}^{\infty} a_k\varepsilon_{t-k}$ in $\mathcal{H}_t(\boldsymbol{\varepsilon})$, due to the white noise nature of $\boldsymbol{\varepsilon}$:

$$
\begin{aligned}
\langle \mathbf{R}X, Y\rangle &= \sum_{h=0}^{+\infty}\sum_{k=0}^{+\infty} \frac{b_{\lfloor \frac{h}{2}\rfloor}}{\sqrt{2}} a_k \langle \varepsilon_{t-h}, \varepsilon_{t-k}\rangle = \sum_{k=0}^{+\infty} b_{\lfloor \frac{k}{2}\rfloor} \frac{a_k}{\sqrt{2}} = \sum_{k=0}^{+\infty} b_k \frac{a_{2k}+a_{2k+1}}{\sqrt{2}} \\
&= \sum_{h=0}^{+\infty}\sum_{k=0}^{+\infty} b_h \frac{a_{2k}+a_{2k+1}}{\sqrt{2}} \langle \varepsilon_{t-h}, \varepsilon_{t-k}\rangle = \big\langle X, \mathbf{R}^* Y\big\rangle.
\end{aligned}
$$

As for the kernel of $\mathbf{R}^*$, we prove that

$$\ker\big(\mathbf{R}^*\big) = \left\{\sum_{k=0}^{+\infty} d_k^{(1)}(\varepsilon_{t-2k} - \varepsilon_{t-2k-1}) \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : d_k^{(1)} \in \mathbb{R}\right\}.$$

Take any element of $\mathcal{H}_t(\boldsymbol{\varepsilon})$ of the kind $X = \sum_{k=0}^{\infty} d_k^{(1)}(\varepsilon_{t-2k} - \varepsilon_{t-2k-1})$ for some square-summable sequence of real numbers $\{d_k^{(1)}\}_k$. Such $X$ can be rewritten as $X = \sum_{h=0}^{\infty} a_h \varepsilon_{t-h}$ with $a_{2k+1} = -a_{2k}$ for all $k \in \mathbb{N}_0$, that is $a_{2k} + a_{2k+1} = 0$. Therefore, by the expression of $\mathbf{R}^*$, we realize that $\mathbf{R}^* X = 0$. Thus,

$$\left\{\sum_{k=0}^{+\infty} d_k^{(1)}(\varepsilon_{t-2k} - \varepsilon_{t-2k-1}) \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : d_k^{(1)} \in \mathbb{R}\right\} \subset \ker\big(\mathbf{R}^*\big). \tag{23}$$

Conversely, consider $X = \sum_{h=0}^{\infty} a_h \varepsilon_{t-h}$ in $\ker(\mathbf{R}^*)$. Since $\mathbf{R}^* X = 0$ in the $L^2$-norm, $\sum_{k=0}^{\infty}(a_{2k} + a_{2k+1})^2 = 0$. As a consequence, $a_{2k+1} = -a_{2k}$ for any $k \in \mathbb{N}_0$ and we can write $X = \sum_{k=0}^{\infty} d_k^{(1)}(\varepsilon_{t-2k} - \varepsilon_{t-2k-1})$ with $d_k^{(1)} = a_{2k}$. As a result, also the converse inclusion in (23) holds and

$$\mathcal{L}_t^{\mathbf{R}} = \ker\big(\mathbf{R}^*\big) = \left\{\sum_{k=0}^{+\infty} b_k^{(1)} \varepsilon_{t-2k}^{(1)} \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : b_k^{(1)} \in \mathbb{R}\right\}.$$

In addition,

$$\mathbf{R}\mathcal{L}_t^{\mathbf{R}} = \left\{\sum_{k=0}^{+\infty} b_k^{(2)} \varepsilon_{t-4k}^{(2)} \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : b_k^{(2)} \in \mathbb{R}\right\}$$

and, for any $j \in \mathbb{N}$,

$$\mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}} = \left\{\sum_{k=0}^{+\infty} b_k^{(j)} \varepsilon_{t-k2^j}^{(j)} \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : b_k^{(j)} \in \mathbb{R}\right\}.$$

As the case with $j \in \mathbb{N}$ can be derived by induction, we focus on $\mathbf{R}\mathcal{L}_t^{\mathbf{R}}$ and show that

$$\mathbf{R}\mathcal{L}_t^{\mathbf{R}} = \left\{\sum_{k=0}^{+\infty} d_k^{(2)}(\varepsilon_{t-4k} + \varepsilon_{t-4k-1} - \varepsilon_{t-4k-2} - \varepsilon_{t-4k-3}) \in \mathcal{H}_t(\boldsymbol{\varepsilon}) : d_k^{(2)} \in \mathbb{R}\right\}. \tag{24}$$

Consider any $Y \in \mathbf{R}\mathcal{L}_t^{\mathbf{R}}$. Since $Y$ is the image of some $X \in \mathcal{L}_t^{\mathbf{R}}$, there exists a square-summable sequence of real numbers $\{d_k^{(1)}\}_k$ such that

$$X = \sum_{k=0}^{+\infty} d_k^{(1)}(\varepsilon_{t-2k} - \varepsilon_{t-2k-1}), \qquad Y = \sum_{k=0}^{+\infty} \frac{d_k^{(1)}}{\sqrt{2}}(\varepsilon_{t-4k} + \varepsilon_{t-4k-1} - \varepsilon_{t-4k-2} - \varepsilon_{t-4k-3}).$$

As a result, $\mathbf{R}\mathcal{L}_t^{\mathbf{R}}$ is included in the set in (24). Vice versa, take any $Y = \sum_{k=0}^{\infty} d_k^{(2)}(\varepsilon_{t-4k} + \varepsilon_{t-4k-1} - \varepsilon_{t-4k-2} - \varepsilon_{t-4k-3})$ for some square-summable sequence of coefficients $\{d_k^{(2)}\}_k$. Then $Y$ belongs to $\mathbf{R}\mathcal{L}_t^{\mathbf{R}}$, too, because it is the image of $X = \sum_{k=0}^{\infty} \sqrt{2} d_k^{(2)}(\varepsilon_{t-2k} - \varepsilon_{t-2k-1})$, which belongs to $\mathcal{L}_t^{\mathbf{R}}$. Therefore, the characterization in (24) is assessed. $\square$

*Proof of Theorem 1*

PROOF. By applying the classical Wold decomposition to the zero-mean, weakly stationary purely nondeterministic process $\mathbf{x}$, we find that $x_t$ belongs to the Hilbert space $\mathcal{H}_t(\boldsymbol{\varepsilon})$, where $\boldsymbol{\varepsilon}$ is the unit variance white noise of classical Wold innovations of $\mathbf{x}$. Importantly, $\mathcal{H}_t(\boldsymbol{\varepsilon})$ orthogonally decomposes as in Lemma A.1. By denoting $g_t^{(j)}$ the orthogonal projections of $x_t$ on the subspaces $\mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}}$, we find that $x_t = \sum_{j=1}^{\infty} g_t^{(j)}$, where the equality is in the $L^2$-norm. Then, by using the characterizations of subspaces $\mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}}$, for any scale $j \in \mathbb{N}$ we find a square-summable sequence of real coefficients $\{\beta_k^{(j)}\}_k$ such that equation (9) holds. As a result, we are allowed to decompose the variable $x_t$ as in equation (5).

We now show (i). As we can see in equation (6), the process $\boldsymbol{\varepsilon}^{(j)}$ is an MA$(2^j - 1)$ with respect to the fundamental innovations $\boldsymbol{\varepsilon}$. In addition, the subprocess $\{\varepsilon_{t-k2^j}^{(j)}\}_{k\in\mathbb{Z}}$ is weakly stationary. Indeed, since $\boldsymbol{\varepsilon}$ is a unit variance white noise, for any $k \in \mathbb{Z}$,

$$\mathbb{E}\left[\left(\varepsilon_{t-k2^j}^{(j)}\right)^2\right] = \frac{1}{2^j}\mathbb{E}\left[\left(\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-k2^j-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-k2^j-2^{j-1}-i}\right)^2\right] = \frac{1}{2^j}\sum_{i=0}^{2^j-1} \mathbb{E}\left[\varepsilon_t^2\right] = 1.$$

Thus, $\mathbb{E}[(\varepsilon_{t-k2^j}^{(j)})^2]$ is finite and it does not depend on $k$. Moreover, $\mathbb{E}[\varepsilon_{t-k2^j}^{(j)}] = 0$ for any $k \in \mathbb{Z}$ and the expectation does not depend on $k$. Finally, we analyze cross-moments in the support $S_t^{(j)} = \{t - k2^j : k \in \mathbb{N}_0\}$. By taking $h \neq k$,

$$\begin{aligned}
&\mathbb{E}\left[\varepsilon_{t-h2^j}^{(j)}\varepsilon_{t-k2^j}^{(j)}\right] \\
&= \frac{1}{2^j}\mathbb{E}\left[\left(\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-h2^j-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-h2^j-2^{j-1}-i}\right)\right. \\
&\quad \left.\cdot\left(\sum_{l=0}^{2^{j-1}-1} \varepsilon_{t-k2^j-l} - \sum_{l=0}^{2^{j-1}-1} \varepsilon_{t-k2^j-2^{j-1}-l}\right)\right] \\
&= \frac{1}{2^j}\left\{\sum_{i=0}^{2^{j-1}-1}\sum_{l=0}^{2^{j-1}-1} \mathbb{E}[\varepsilon_{t-h2^j-i}\varepsilon_{t-k2^j-l}] - \sum_{i=0}^{2^{j-1}-1}\sum_{l=0}^{2^{j-1}-1} \mathbb{E}[\varepsilon_{t-h2^j-i}\varepsilon_{t-k2^j-2^{j-1}-l}]\right. \\
&\quad \left. - \sum_{i=0}^{2^{j-1}-1}\sum_{l=0}^{2^{j-1}-1} \mathbb{E}[\varepsilon_{t-h2^j-2^{j-1}-i}\varepsilon_{t-k2^j-l}] + \sum_{i=0}^{2^{j-1}-1}\sum_{l=0}^{2^{j-1}-1} \mathbb{E}[\varepsilon_{t-h2^j-2^{j-1}-i}\varepsilon_{t-k2^j-2^{j-1}-l}]\right\}.
\end{aligned}$$

Since $h \neq k$, the sets of indices $\{h2^j, \ldots, h2^j + 2^j - 1\}$ and $\{k2^j, \ldots, k2^j + 2^j - 1\}$ are disjoint and so the last sums are null. Therefore, $\mathbb{E}[\varepsilon_{t-h2^j}^{(j)}\varepsilon_{t-k2^j}^{(j)}] = 0$ for all $h \neq k$.

As a result, $\{\varepsilon_{t-k2^j}^{(j)}\}_{k\in\mathbb{Z}}$ is weakly stationary on $S_t^{(j)}$ and it is a unit variance white noise.

We now turn to (ii). For any fixed scale $j \in \mathbb{N}$, since the variables $\varepsilon_{t-k2^j}^{(j)}$ are orthonormal when $k$ varies, the component $g_t^{(j)}$ has a unique representation as in equation (8). Thus, the coefficients $\beta_k^{(j)}$ are uniquely defined, and clearly, $\sum_{j=1}^{\infty}\sum_{k=0}^{\infty}(\beta_k^{(j)})^2$ is finite.

In order to find the explicit expression of coefficients $\beta_k^{(j)}$, we exploit the orthogonal decompositions of $\mathcal{H}_t(\boldsymbol{\varepsilon})$ at different scales $J \in \mathbb{N}$:

$$\mathcal{H}_t(\boldsymbol{\varepsilon}) = \mathbf{R}^J \mathcal{H}_t(\boldsymbol{\varepsilon}) \oplus \bigoplus_{j=1}^{J} \mathbf{R}^{j-1} \mathcal{L}_t^{\mathbf{R}}.$$

We call $\pi_t^{(j)}$ the orthogonal projection of $x_t$ on the subspace $\mathbf{R}^j \mathcal{H}_t(\boldsymbol{\varepsilon})$ and we proceed inductively.

We start by the first decomposition $x_t = \pi_t^{(1)} + g_t^{(1)}$ coming from scale $J = 1$, namely $\mathcal{H}_t(\boldsymbol{\varepsilon}) = \mathbf{R}\mathcal{H}_t(\boldsymbol{\varepsilon}) \oplus \mathcal{L}_t^{\mathbf{R}}$. By the definitions of elements in $\mathbf{R}\mathcal{H}_t(\boldsymbol{\varepsilon})$ and $\mathcal{L}_t^{\mathbf{R}}$ described in Lemma A.1, we set

$$\pi_t^{(1)} = \sum_{k=0}^{+\infty} \gamma_k^{(1)} \frac{\varepsilon_{t-2k} + \varepsilon_{t-(2k+1)}}{\sqrt{2}} = \sum_{k=0}^{+\infty} c_k^{(1)} (\varepsilon_{t-2k} + \varepsilon_{t-(2k+1)}),$$

$$g_t^{(1)} = \sum_{k=0}^{+\infty} \beta_k^{(1)} \varepsilon_{t-2k}^{(1)} = \sum_{k=0}^{+\infty} d_k^{(1)} (\varepsilon_{t-2k} - \varepsilon_{t-2k-1})$$

for some sequences of coefficients $\{c_k^{(1)}\}_k$ and $\{d_k^{(1)}\}_k$, or equivalently $\{\gamma_k^{(1)}\}_k$ and $\{\beta_k^{(1)}\}_k$, to determine in order to have $x_t = \pi_t^{(1)} + g_t^{(1)}$, where we set $\sqrt{2}c_k^{(1)} = \gamma_k^{(1)}$ and $\sqrt{2}d_k^{(1)} = \beta_k^{(1)}$. The expressions above may be rewritten as

$$x_t = \sum_{k=0}^{+\infty} \left\{ \left(c_k^{(1)} + d_k^{(1)}\right) \varepsilon_{t-2k} + \left(c_k^{(1)} - d_k^{(1)}\right) \varepsilon_{t-2k-1} \right\}.$$

However, from the classical Wold decomposition of $\mathbf{x}$,

$$x_t = \sum_{k=0}^{+\infty} \{\alpha_{2k} \varepsilon_{t-2k} + \alpha_{2k+1} \varepsilon_{t-2k-1}\}$$

with the same fundamental innovations $\varepsilon_t$. By the uniqueness of writing of the classical Wold decomposition, the two expressions for $x_t$ must coincide. As a result, $c_k^{(1)}$ and $d_k^{(1)}$ are the solutions of the linear system

$$\begin{cases} c_k^{(1)} + d_k^{(1)} = \alpha_{2k}, \\ c_k^{(1)} - d_k^{(1)} = \alpha_{2k+1}, \end{cases}$$

that is,

$$c_k^{(1)} = \frac{\alpha_{2k} + \alpha_{2k+1}}{2}, \qquad d_k^{(1)} = \frac{\alpha_{2k} - \alpha_{2k+1}}{2}$$

and, in particular, we find

$$\gamma_k^{(1)} = \frac{\alpha_{2k} + \alpha_{2k+1}}{\sqrt{2}}, \qquad \beta_k^{(1)} = \frac{\alpha_{2k} - \alpha_{2k+1}}{\sqrt{2}}.$$

Next, we focus on the scale $J = 2$. We exploit the decomposition of the space $\mathbf{R}\mathcal{H}_t(\boldsymbol{\varepsilon}) = \mathbf{R}^2\mathcal{H}_t(\boldsymbol{\varepsilon}) \oplus \mathbf{R}\mathcal{L}_t^{\mathbf{R}}$ that implies the relation $\pi_t^{(1)} = \pi_t^{(2)} + g_t^{(2)}$. We follow the same track as in the previous case, by using the features of elements in $\mathbf{R}^2\mathcal{H}_t(\boldsymbol{\varepsilon})$ and in $\mathbf{R}\mathcal{L}_t^{\mathbf{R}}$ and, finally, by comparing the expression of $\pi_t^{(2)} + g_t^{(2)}$ with the (unique) writing of $\pi_t^{(1)}$ that we found before. Since

$$\pi_t^{(2)} = \sum_{k=0}^{+\infty} \gamma_k^{(2)} \frac{\varepsilon_{t-4k} + \varepsilon_{t-(4k+1)} + \varepsilon_{t-(4k+2)} + \varepsilon_{t-(4k+3)}}{2}, \qquad g_t^{(2)} = \sum_{k=0}^{+\infty} \beta_k^{(2)} \varepsilon_{t-4k}^{(2)},$$

by solving a simple linear system we discover that

$$\gamma_k^{(2)} = \frac{\alpha_{4k} + \alpha_{4k+1} + \alpha_{4k+2} + \alpha_{4k+3}}{2}, \qquad \beta_k^{(2)} = \frac{\alpha_{4k} + \alpha_{4k+1} - \alpha_{4k+2} - \alpha_{4k+3}}{2}.$$

At the generic scale $J = j$, we retrieve the expressions of $\beta_k^{(j)}$ and $\gamma_k^{(j)}$ of equation (7) and (11), where $\pi_t^{(j)}$ is defined in equation (10).

Finally, we show (iii). First, when $t$ is fixed, $\mathbb{E}[g_t^{(j)} g_t^{(l)}] = 0$ for all $j \neq l$ because $g_t^{(j)}$ and $g_t^{(l)}$ are, respectively, the projections of $x_t$ on the subspaces $\mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}}$ and $\mathbf{R}^{l-1}\mathcal{L}_t^{\mathbf{R}}$ that are orthogonal by construction. Now, consider any $g_{t-m2^j}^{(j)}$ with $m \in \mathbb{N}_0$. Clearly, $g_{t-m2^j}^{(j)}$ belongs to $\mathbf{R}^{j-1}\mathcal{L}_{t-m2^j}^{\mathbf{R}}$ but, by the definition of $g_t^{(j)}$, we can write

$$g_{t-m2^j}^{(j)} = \sum_{k=0}^{+\infty} \beta_k^{(j)} \varepsilon_{t-(m+k)2^j}^{(j)} = \sum_{K=0}^{+\infty} \beta_K^{(j)} \varepsilon_{t-K2^j}^{(j)},$$

where $\beta_K^{(j)} = 0$ if $K = 0, \dots, m-1$ and $\beta_K^{(j)} = \beta_k^{(j)}$ if $K = m + k$ for some $k \in \mathbb{N}_0$. As a result, $g_{t-m2^j}^{(j)}$ belongs to $\mathbf{R}^{j-1}\mathcal{L}_t^{\mathbf{R}}$, too. Similarly, at scale $l$, taken any $n \in \mathbb{N}_0$, it is easy to see that $g_{t-n2^l}^{(l)}$ belongs to $\mathbf{R}^{l-1}\mathcal{L}_t^{\mathbf{R}}$. Hence, the orthogonality of such subspaces guarantees that $\mathbb{E}[g_{t-m2^j}^{(j)} g_{t-n2^l}^{(l)}] = 0$ for all $j \neq l$ and $m, n \in \mathbb{N}_0$.

As for the general requirement on $\mathbb{E}[g_{t-p}^{(j)} g_{t-q}^{(l)}]$ for any $j, l \in \mathbb{N}$ and $p, q, t \in \mathbb{Z}$,

$$\begin{aligned}
\mathbb{E}\big[g_{t-p}^{(j)} g_{t-q}^{(l)}\big] &= \sum_{k=0}^{+\infty}\sum_{h=0}^{+\infty} \beta_k^{(j)} \beta_h^{(l)} \mathbb{E}\big[\varepsilon_{t-p-k2^j}^{(j)} \varepsilon_{t-q-h2^l}^{(l)}\big] \\
&= \frac{1}{\sqrt{2^{j+l}}} \sum_{k=0}^{+\infty}\sum_{h=0}^{+\infty} \beta_k^{(j)} \beta_h^{(l)} \sum_{u=0}^{2^{j-1}-1} \sum_{v=0}^{2^{l-1}-1} \big\{ \mathbb{E}[\varepsilon_{t-p-k2^j-u} \varepsilon_{t-q-h2^l-v}] \\
&\quad - \mathbb{E}[\varepsilon_{t-p-k2^j-u} \varepsilon_{t-q-h2^l-2^{l-1}-v}] - \mathbb{E}[\varepsilon_{t-p-k2^j-2^{j-1}-u} \varepsilon_{t-q-h2^l-v}] \\
&\quad + \mathbb{E}[\varepsilon_{t-p-k2^j-2^{j-1}-u} \varepsilon_{t-q-h2^l-2^{l-1}-v}] \big\}
\end{aligned}$$

and so

$$\begin{aligned}
\mathbb{E}\big[g_{t-p}^{(j)} g_{t-q}^{(l)}\big] &= \frac{1}{\sqrt{2^{j+l}}} \sum_{k=0}^{+\infty}\sum_{h=0}^{+\infty} \beta_k^{(j)} \beta_h^{(l)} \sum_{u=0}^{2^{j-1}-1} \sum_{v=0}^{2^{l-1}-1} \big\{ \gamma\big(p - q + k2^j + u - h2^l - v\big) \\
&\quad - \gamma\big(p - q + k2^j + u - h2^l - 2^{l-1} - v\big)
\end{aligned}$$

$$
\begin{aligned}
&- \gamma(p-q+k2^j+2^{j-1}+u-h2^l-v)\\
&+ \gamma(p-q+k2^j+2^{j-1}+u-h2^l-2^{l-1}-v)\big\},
\end{aligned}
$$

where coefficients $\beta^{(j)}_k$, $\beta^{(l)}_h$ do not depend on $t$ and $\gamma$ denotes the autocovariance function of $\boldsymbol{\varepsilon}$. After the summations over $u$, $v$ and $k$, $h$, the one remaining variables are $j$, $l$, $p-q$. In other words, $\mathbb{E}[g^{(j)}_{t-p}g^{(l)}_{t-q}]$ depends at most on $j$, $l$, $p-q$. □

*Proof of Theorem 2*

PROOF. First, we show that any process $\mathbf{g^{(j)}}$ is well-defined. Indeed, by using the moving average representation of each $g^{(j)}_t$ with respect to the innovations on the support $S^{(j)}_t$ and the definition of detail processes $\boldsymbol{\varepsilon}^{(j)}$, we have

$$
g^{(j)}_t = \sum_{k=0}^{+\infty} \beta^{(j)}_k \varepsilon^{(j)}_{t-k2^j} = \sum_{k=0}^{+\infty}\sum_{i=0}^{2^j-1} \beta^{(j)}_k \delta^{(j)}_i \varepsilon_{t-k2^j-i} = \sum_{h=0}^{+\infty} \beta^{(j)}_{\lfloor \frac{h}{2^j}\rfloor} \delta^{(j)}_{h-2^j\lfloor \frac{h}{2^j}\rfloor} \varepsilon_{t-h}, \tag{25}
$$

where $h = k2^j + i$, $k = \lfloor \frac{h}{2^j}\rfloor$ and $i = h - 2^j\lfloor \frac{h}{2^j}\rfloor$. Condition (13) ensures the square-summability of the coefficients and so each $\mathbf{g^{(j)}}$ is well-defined.

In addition, the process $\mathbf{x}$ is well-defined because of Conditions (13) and (14). According to the latter, the components $g^{(j)}_t$ are orthogonal to each others at different scales for fixed $t \in \mathbb{Z}$. Therefore,

$$
\mathbb{E}[x_t^2] = \mathbb{E}\left[\left(\sum_{j=1}^{+\infty} g^{(j)}_t\right)^2\right] = \sum_{j=1}^{+\infty} \mathbb{E}[(g^{(j)}_t)^2] = \sum_{j=1}^{+\infty}\sum_{h=0}^{+\infty} \left(\beta^{(j)}_{\lfloor \frac{h}{2^j}\rfloor} \delta^{(j)}_{h-2^j\lfloor \frac{h}{2^j}\rfloor}\right)^2,
$$

which is finite because of (13). In consequence, the process $\mathbf{x}$ is well-defined.

Now we show that $\mathbf{x}$ is weakly stationary, with zero mean. We already observed that $\mathbb{E}[x_t^2]$ is finite and not dependent on $t$. In addition, since the processes $\mathbf{g^{(j)}}$ have zero mean, $\mathbb{E}[x_t] = 0$ for any $t \in \mathbb{Z}$. Finally, take $p \neq q$. Then

$$
\mathbb{E}[x_{t-p}x_{t-q}] = \mathbb{E}\left[\left(\sum_{j=1}^{+\infty} g^{(j)}_{t-p}\right)\left(\sum_{l=1}^{+\infty} g^{(l)}_{t-q}\right)\right] = \sum_{j=1}^{+\infty}\sum_{l=1}^{+\infty} \mathbb{E}[g^{(j)}_{t-p}g^{(l)}_{t-q}].
$$

As $\mathbb{E}[g^{(j)}_{t-p}g^{(l)}_{t-q}]$ are dependent at most on $j$, $l$ and $p-q$ (see, e.g., the computations in the proof of Theorem 1), $\mathbb{E}[x_{t-p}x_{t-q}]$ depends at most on the difference $p-q$. As a result, $\mathbf{x}$ is weakly stationary, with zero mean.

By taking the sum over scales $j \in \mathbb{N}$ in equation (25), we obtain the decomposition of $x_t$ with respect to the process $\boldsymbol{\varepsilon}$ stated in equation (16). Clearly, $\mathbf{x}$ is purely non-deterministic. □

PROPOSITION A.1. *The time series*

$$
\mathbf{R}x_t = \sum_{k=0}^{+\infty} \frac{\alpha_{\lfloor \frac{k}{2}\rfloor}}{\sqrt{2}} \varepsilon_{t-k} \quad \textit{and} \quad \mathbf{R_x}x_t = \frac{1}{\sqrt{2}}(x_t + x_{t-1})
$$

*have spectral density functions, respectively,*

$$f_{\mathbf{R}}(\lambda) = 2\cos^2\left(\frac{\lambda}{2}\right) f_x(2\lambda) \quad \textit{and} \quad f_{\mathbf{R_x}}(\lambda) = 2\cos^2\left(\frac{\lambda}{2}\right) f_x(\lambda),$$

*where $f_x(\lambda)$ is the spectral density function of $x_t$.*

PROOF. Define the time-invariant linear filter $A(\mathbf{L}) = \sum_{h=0}^{\infty} \alpha_h \mathbf{L}^h$, so that $x_t = A(\mathbf{L})\varepsilon_t$. Since $\sum_{h=0}^{\infty} |\alpha_h| < +\infty$ and the spectral density function of $\varepsilon_t$ is $f_\varepsilon(\lambda) = 1/2\pi$,

$$\begin{aligned} f_x(\lambda) &= \left|A\left(e^{-i\lambda}\right)\right|^2 f_\varepsilon(\lambda) = \left|\sum_{h=0}^{+\infty} \alpha_h e^{-ih\lambda}\right|^2 \frac{1}{2\pi} \\ &= \frac{1}{2\pi}\left\{\left(\sum_{h=0}^{+\infty} \alpha_h \cos(h\lambda)\right)^2 + \left(\sum_{h=0}^{+\infty} \alpha_h \sin(h\lambda)\right)^2\right\} \\ &= \frac{1}{2\pi}\sum_{h=0}^{+\infty}\sum_{k=0}^{+\infty} \alpha_h\alpha_k \cos\big(\lambda(k-h)\big). \end{aligned}$$

First, consider $\mathbf{R}x_t$. As $\sum_{k=0}^{\infty} |\alpha_{\lfloor \frac{k}{2} \rfloor}| = 2\sum_{h=0}^{\infty} |\alpha_h| < +\infty$, we have

$$\begin{aligned} f_{\mathbf{R}}(\lambda) &= \left|\sum_{k=0}^{+\infty} \frac{\alpha_{\lfloor \frac{k}{2} \rfloor}}{\sqrt{2}} e^{-ik\lambda}\right|^2 \frac{1}{2\pi} = \frac{1}{2\pi}\sum_{h=0}^{+\infty}\sum_{k=0}^{+\infty} \frac{\alpha_{\lfloor \frac{h}{2} \rfloor}\alpha_{\lfloor \frac{k}{2} \rfloor}}{2} \cos\big(\lambda(k-h)\big) \\ &= \frac{1}{2\pi}\sum_{h=0}^{+\infty}\sum_{k=0}^{+\infty} \alpha_h\alpha_k \left\{\cos\big(2\lambda(k-h)\big) + \frac{\cos\big(\lambda(2k-2h+1)\big) + \cos\big(\lambda(2k-2h-1)\big)}{2}\right\} \\ &= \frac{1}{2\pi}\sum_{h=0}^{+\infty}\sum_{k=0}^{+\infty} \alpha_h\alpha_k \cos\big(2\lambda(k-h)\big)\{1+\cos(\lambda)\} = 2\cos^2\left(\frac{\lambda}{2}\right) f_x(2\lambda). \end{aligned}$$

Now consider $\mathbf{R_x}x_t$. The spectral density function in the claim follows from

$$f_{\mathbf{R_x}}(\lambda) = \left|\frac{1}{\sqrt{2}}\left(e^0 + e^{-i\lambda}\right)\right|^2 f_x(\lambda) = \frac{1}{2}\left\{\big(1+\cos(\lambda)\big)^2 + \sin^2(\lambda)\right\} f_x(\lambda). \qquad \square$$

### A.2 *Forecasting from the persistence-based decomposition*

We provide the formulas for conditional expectation and variance of a process $\mathbf{x} = \{x_t\}_{t\in\mathbb{Z}}$ that has classical and extended Wold decompositions given by equations (4) and (5), respectively. We consider the filtration generated by the white noise $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t\in\mathbb{Z}}$ assuming that the innovations $\varepsilon_t$ are independent.

Fix $p \in \mathbb{N}$. The conditional expectation at time $t$ of $x_{t+p}$ is characterized by an offset of the classical Wold coefficients, namely $\mathbb{E}_t[x_{t+p}] = \sum_{h=0}^{\infty} \alpha_{h+p}\varepsilon_{t-h}$. Notably, such offset is inherited by the extended Wold decomposition of $\mathbb{E}_t[x_{t+p}]$:

$$\mathbb{E}_t[x_{t+p}] = \sum_{j=1}^{+\infty}\sum_{k=0}^{+\infty} \beta_{k,p}^{(j)} \varepsilon_{t-k2^j}^{(j)},$$

where, for any $j \in \mathbb{N}$ and $k \in \mathbb{N}_0$,

$$\beta_{k,p}^{(j)} = \frac{1}{\sqrt{2^j}} \left( \sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+i+p} - \sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+2^{j-1}+i+p} \right).$$

Therefore, once the extended Wold decomposition of $x_t$ is known, $p$-step ahead forecasts do not require a large additional effort because they are driven by the detail processes $\boldsymbol{\varepsilon}^{(j)}$, too, and coefficients $\beta_{k,p}^{(j)}$ are easily computed.

As to the conditional variance, the properties of the classical Wold decomposition ensure that $\mathrm{Var}_t(x_{t+p}) = \alpha_0^2 + \cdots + \alpha_{p-1}^2$. By Theorem 2 the coefficients $\alpha_h$ can be obtained from the scale-specific $\beta_k^{(j)}$ and so $\mathrm{Var}_t(x_{t+p})$ can be derived directly from them. For example, $\mathrm{Var}_t(x_{t+1}) = \alpha_0^2 = (\sum_{j=1}^{\infty} \beta_0^{(j)}/\sqrt{2^j})^2$.

## References


Luenberger, D. G. (1968), *Optimization by Vector Space Methods*. John Wiley & Sons. [2]

Nagy, B. S., C. Foias, H. Bercovici, and L. Kérchy (2010), *Harmonic Analysis of Operators on Hilbert Space*. Springer. [2]


---


Co-editor Frank Schorfheide handled this manuscript.

Manuscript received 3 October, 2017; final version accepted 4 August, 2019; available online 14 August, 2019.