

VALIDATING MARKOV SWITCHING VAR THROUGH SPECTRAL REPRESENTATIONS

Monica Billio* and Maddalena Cavicchioli**

Abstract. *We develop a method to validate the use of Markov switching models in modelling time series subject to structural changes. Particularly, we consider multivariate autoregressive models subject to Markov switching and derive close-form formulae for the spectral density of such models, based on their autocovariance functions and stable representations. Within this framework, we check the capability of the model to capture the relative importance of high- and low-frequency variability of the series. Numerical examples and applications to U.S. macroeconomic and financial data illustrate the behaviour at different frequencies.*

Keywords: Markov Switching VAR, Stable representation, Spectral density, Long memory, Frequency variability.

JEL Classification: C01, C32, C50.

* Department of Economics - Università Cà Foscari - Cannaregio 873, 30121 Venezia, Italy. E-mail: <billio@unive.it>

** Department of Economics "Marco Biagi", Viale Berengario 51, 41121, Modena (Italy). Office Phone No: +39 059 2056846. E-mail: <maddalena.cavicchioli@unimore.it>

1. Introduction

Issues about persistence and frequency variability of time series are often raised in macroeconomics and finance. In particular, a large literature in econometrics has developed parametric tools to capture the low-frequency behaviour of some time series, which exhibit strong persistence in time. Models such as local-to-unity, fractional or "long memory" have been particularly applied to capture low-frequency behaviour of the stochastic component of those series. As an example, take the long debate starting from Ding, Engle and Granger (1993) and Ding and Granger (1996) on the persistency characteristic of asset returns. A common finding is that returns themselves contain little serial correlation, while absolute returns and their power transformations are highly correlated. In Ding et al. (1993) a long memory property for the absolute returns of S&P500 daily stock market is established empirically. However, recent literature have focused on the possibility of confusing long memory and structural change. Mikosch and Starica (2000) find structural change in asset return dynamics and argue that it could be responsible for evidence of long memory. Moreover, Diebold and Inoue (2001) show analytically that stochastic regime switching is easily confused with long memory, even asymptotically, so long as a small amount of regime switches occurs. Granger and Hyung (2004) study occasional structural breaks and their empirical results suggest the possibility such at least part of the long memory can be caused by the presence of neglected breaks in the series. According to these recent findings, the persistency characterizing some time series should be taken into account when modelling the series as a non-linear model. In this paper we propose a method to validate the use of Markov switching models through the use of their spectral density functions. We first apply some new tools recently proposed in Cavicchioli (2014a) to detect the presence of structural changes in the data. Then we derive close-form formulae for the spectral representations of Markov switching VAR processes which are necessary to evaluate high- and low-frequency variability of time series. The aim is twofold: from one side, we investigate non-linear features of the

data to correctly specify the parametric model, from the other, we check the correct specification analyzing their frequency contents through the spectral density function. If the empirically detected persistency is captured by the chosen parametrization, then we can be more confident in the application of our parametric model. Our results are related to the work of Krolzig (1997) in terms of state space representation and stable representation and to the paper of Pataracchia (2011) where a different Markovian representation has been considered. However, note that in the latter paper, it is assumed that the constant term (which is also governed by Markov chain) is zero. Here we find more general and useful expressions. Thus our primary interest is to test non-linearity in the data, study their behaviour at different frequencies through spectral functions and validate the chosen model in relation with its empirical counterpart. The plan of the paper is the following. In Section 2 we study the spectral density functions of Markov-switching (MS) VAR(0) and VAR(p) processes in close-form, both from their switching state-space representations and from stable VARMA representations. Section 3 shows some numerical examples of the spectral densities for MS models. In Section 4 we investigate the presence of structural changes in real data. Then we check the ability of the chosen models to capture high- and low-frequency variability, using arguments from Section 2. Section 5 concludes. Derivation of some formulae can be found in the Appendix.

2. Spectra of Markov-switching VAR

2.1 Spectra of hidden Markov process

Let us consider the model

$$(2.1) \quad \mathbf{y}_t = \boldsymbol{\nu}_{s_t} + \boldsymbol{\Sigma}_{s_t} \mathbf{u}_t$$

where $\mathbf{u}_t \sim IID(\mathbf{0}, \mathbf{I}_k)$, \mathbf{y}_t , $\boldsymbol{\nu}_{s_t}$ and \mathbf{u}_t are $K \times 1$, $\boldsymbol{\Sigma}_{s_t}$ is $K \times K$ and (s_t) follows an M -state (irreducible and ergodic) Markov chain. Let $\mathbf{P} = (p_{ij})_{i,j=1,\dots,M}$ be the transition matrix of the chain, where $p_{ij} = Pr(s_t = j | s_{t-1} = i)$. Ergodicity implies the existence of a stationary vector of probabilities $\boldsymbol{\pi} =$

$(\pi_1 \dots \pi_M)'$ satisfying $\boldsymbol{\pi} = \mathbf{P}'\boldsymbol{\pi}$ and $\mathbf{i}'_M\boldsymbol{\pi} = 1$, where \mathbf{i}_M denotes the $(M \times 1)$ vector of ones. Irreducibility implies that $\pi_m > 0$ for $m = 1, \dots, M$, meaning that all unobservable states are possible. An useful representation for (s_t) is obtained by letting $\boldsymbol{\xi}_t$ denote a random $(M \times 1)$ vector whose m th element is equal to unity if $s_t = m$ and zero otherwise. Then the Markov chain follows a VAR(1) process

$$\boldsymbol{\xi}_t = \mathbf{P}'\boldsymbol{\xi}_{t-1} + \mathbf{v}_t$$

where $\mathbf{v}_t = \boldsymbol{\xi}_t - E(\boldsymbol{\xi}_t | \boldsymbol{\xi}_{t-1})$ is a zero mean martingale difference sequence. Consequently, we have the following standard properties ($h > 0$):

$$\begin{aligned} E(\boldsymbol{\xi}_t) &= \boldsymbol{\pi} & E(\boldsymbol{\xi}_t \boldsymbol{\xi}'_t) &= \mathbf{D} = \text{diag}(\pi_1 \dots \pi_M) \\ E(\boldsymbol{\xi}_t \boldsymbol{\xi}'_{t+h}) &= \mathbf{D}\mathbf{P}^h & \mathbf{v}_t &\sim IID(\mathbf{0}, \mathbf{D} - \mathbf{P}'\mathbf{D}\mathbf{P}) \end{aligned}$$

Define $\boldsymbol{\Lambda} = (\boldsymbol{\nu}_1 \dots \boldsymbol{\nu}_M)$ and $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_1 \dots \boldsymbol{\Sigma}_M)$. We get a first state space representation of (2.1)

$$(2.2) \quad \begin{cases} \mathbf{y}_t = \boldsymbol{\Lambda}\boldsymbol{\xi}_t + \boldsymbol{\Sigma}(\boldsymbol{\xi}_t \otimes \mathbf{I}_K)\mathbf{u}_t \\ \boldsymbol{\xi}_t = \mathbf{P}'\boldsymbol{\xi}_{t-1} + \mathbf{v}_t \end{cases}$$

In fact, for $s_t = m$, $\boldsymbol{\xi}_t = \mathbf{e}_m$ the m th column of the identity matrix \mathbf{I}_M . So we get

$$\begin{aligned} \mathbf{y}_t &= \begin{pmatrix} \boldsymbol{\nu}_1 & \dots & \boldsymbol{\nu}_M \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_M \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{I}_K \\ \vdots \\ \mathbf{0} \end{pmatrix} \mathbf{u}_t \\ &= \boldsymbol{\nu}_m + \boldsymbol{\Sigma}_m \mathbf{I}_K \mathbf{u}_t = \boldsymbol{\nu}_m + \boldsymbol{\Sigma}_m \mathbf{u}_t. \end{aligned}$$

The transition equation in (2.2) differs from a stable linear VAR(1) process by the fact that one eigenvalue of \mathbf{P}' is equal to one, and the covariance matrix is singular due to the adding-up restriction. For analytical purposes, a slightly different formulation of the transition equation in (2.2) is more

useful, where the identity $\mathbf{i}'_M \boldsymbol{\xi}_t = 1$ is eliminated. See Krolzig (1997), Chp.3. This procedure alters the state-space representation by using a new $(M-1)$ -dimensional state vector

$$\boldsymbol{\delta}_t = \begin{pmatrix} \xi_{1,t} - \pi_1 \\ \vdots \\ \xi_{M-1,t} - \pi_{M-1} \end{pmatrix}.$$

The transition matrix \mathbf{F} associated with $\boldsymbol{\delta}_t$ is given by

$$\mathbf{F} = \begin{pmatrix} p_{1,1} - p_{M,1} & \cdots & p_{M-1,1} - p_{M,1} \\ \vdots & & \vdots \\ p_{1,M-1} - p_{M,M-1} & \cdots & p_{M-1,M-1} - p_{M,M-1} \end{pmatrix}.$$

The eigenvalues of \mathbf{F} are less than 1 in absolute value. Here the relations

$$\xi_{M,t} = 1 - \sum_{m=1}^{M-1} \xi_{mt} \qquad \pi_M = 1 - \sum_{m=1}^{M-1} \pi_m$$

have been used. Then we have

$$\boldsymbol{\xi}_t - \boldsymbol{\pi} = P'(\boldsymbol{\xi}_{t-1} - \boldsymbol{\pi}) + \mathbf{v}_t$$

hence

$$\boldsymbol{\delta}_t = \mathbf{F} \boldsymbol{\delta}_{t-1} + \mathbf{w}_t$$

where

$$\mathbf{w}_t = (\mathbf{I}_{M-1} \quad -\mathbf{i}_{M-1}) \mathbf{v}_t.$$

This gives a second (unrestricted) state-space representation

$$\mathbf{y}_t = \boldsymbol{\Lambda} \boldsymbol{\pi} + \boldsymbol{\Lambda}(\boldsymbol{\xi}_t - \boldsymbol{\pi}) + \boldsymbol{\Sigma}((\boldsymbol{\xi}_t - \boldsymbol{\pi}) \otimes \mathbf{I}_K) \mathbf{u}_t + \boldsymbol{\Sigma}(\boldsymbol{\pi} \otimes \mathbf{I}_K) \mathbf{u}_t$$

hence

$$(2.3) \quad \begin{cases} \mathbf{y}_t = \boldsymbol{\Lambda} \boldsymbol{\pi} + \tilde{\boldsymbol{\Lambda}} \boldsymbol{\delta}_t + \tilde{\boldsymbol{\Sigma}}(\boldsymbol{\delta}_t \otimes \mathbf{I}_K) \mathbf{u}_t + \boldsymbol{\Sigma}(\boldsymbol{\pi} \otimes \mathbf{I}_K) \mathbf{u}_t \\ \boldsymbol{\delta}_t = \mathbf{F} \boldsymbol{\delta}_{t-1} + \mathbf{w}_t \end{cases}$$

where

$$\tilde{\Lambda} = (\boldsymbol{\nu}_1 - \boldsymbol{\nu}_M \dots \boldsymbol{\nu}_{M-1} - \boldsymbol{\nu}_M) \quad \tilde{\Sigma} = (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_M \dots \boldsymbol{\Sigma}_{M-1} - \boldsymbol{\Sigma}_M).$$

We then have the following standard properties:

$$\begin{aligned} E(\boldsymbol{\delta}_t) &= 0 & E(\boldsymbol{\delta}_t \boldsymbol{\delta}_t') &= \tilde{\mathbf{D}} \\ E(\boldsymbol{\delta}_t \boldsymbol{\delta}_{t+h}') &= \tilde{\mathbf{D}}(\mathbf{F}')^h, \quad h > 0 & \mathbf{w}_t &\sim IID(\mathbf{0}, \tilde{\mathbf{D}} - \mathbf{F}\tilde{\mathbf{D}}\mathbf{F}') \end{aligned}$$

where

$$\tilde{\mathbf{D}} = \begin{pmatrix} \pi_1(1 - \pi_1) & \dots & -\pi_1\pi_{M-1} \\ \vdots & & \vdots \\ -\pi_{M-1}\pi_1 & \dots & \pi_{M-1}(1 - \pi_{M-1}) \end{pmatrix}.$$

The autocovariance function of the process (\mathbf{y}_t) in (2.3) is given by

$$\begin{aligned} \Gamma_{\mathbf{y}}(0) &= \tilde{\Lambda}\tilde{\mathbf{D}}\tilde{\Lambda}' + \tilde{\Sigma}(\tilde{\mathbf{D}} \otimes \mathbf{I}_K)\tilde{\Sigma}' + \boldsymbol{\Sigma}((\mathbf{D}\mathbf{P}_\infty) \otimes \mathbf{I}_K)\boldsymbol{\Sigma}' \\ \Gamma_{\mathbf{y}}(h) &= \tilde{\Lambda}\mathbf{F}^h\tilde{\mathbf{D}}\tilde{\Lambda}', \quad h > 0 \end{aligned}$$

where $\mathbf{D}\mathbf{P}_\infty = \boldsymbol{\pi}\boldsymbol{\pi}'$ and $\mathbf{P}_\infty = \lim_n \mathbf{P}^n = \mathbf{i}_M\boldsymbol{\pi}'$. The multivariate spectral matrix describes the spectral density functions of each element of the state vector in the diagonal terms. The off-diagonal terms are defined cross spectral density functions and they are typically complex numbers. Here we are only interested in the diagonal terms. Therefore, we can compute them, without loss of generality, considering the summation

$$F_{\mathbf{y}}(\omega) = \sum_{h=-\infty}^{+\infty} \Gamma_{\mathbf{y}}(|h|)e^{-i\omega h}$$

where the frequency ω belongs to $[-\pi, \pi]$. See also Pataracchia (2011) where a different spectral representation was obtained. Since the spectral radius $\rho(\mathbf{F})$ of \mathbf{F} is less than 1, the spectral density matrix of the process (\mathbf{y}_t) in (2.3) is given by

$$(2.4) \quad F_{\mathbf{y}}(\omega) = Q + 2\tilde{\Lambda}\mathbf{F}\mathcal{R}e\{(\mathbf{I}_{M-1}e^{i\omega} - \mathbf{F})^{-1}\}\tilde{\mathbf{D}}\tilde{\Lambda}'$$

where \mathcal{Re} denotes the real part of the complex matrix $(\mathbf{I}_{M-1}e^{i\omega} - \mathbf{F})^{-1}$, and

$$Q = \tilde{\mathbf{\Lambda}}\tilde{\mathbf{D}}\tilde{\mathbf{\Lambda}}' + \tilde{\mathbf{\Sigma}}(\tilde{\mathbf{D}} \otimes \mathbf{I}_K)\tilde{\mathbf{\Sigma}}' + \mathbf{\Sigma}((\mathbf{D}\mathbf{P}_\infty) \otimes \mathbf{I}_K)\mathbf{\Sigma}'.$$

Complete derivation of Formula (2.4) is given in the Appendix. An alternative approach to the same problem is based on a stable representation of (2.3). Set $\mu_{\mathbf{y}} = \mathbf{\Lambda}\boldsymbol{\pi}$. From (2.3) we get

$$\boldsymbol{\delta}_t = F(L)^{-1}\mathbf{w}_t$$

where $F(L) = \mathbf{I}_{M-1} - \mathbf{F}L$ (here L is the lag operator). Substituting this relation into the measurement equation in (2.3) yields

$$|F(L)|(\mathbf{y}_t - \mu_{\mathbf{y}}) = \tilde{\mathbf{\Lambda}}F(L)^*\mathbf{w}_t + \tilde{\mathbf{\Sigma}}(F(L)^*\mathbf{w}_t \otimes \mathbf{I}_K)\mathbf{u}_t + |F(L)|\mathbf{\Sigma}(\boldsymbol{\pi} \otimes \mathbf{I}_K)\mathbf{u}_t$$

where $F(L)^*$ denotes the adjoint matrix of $F(L)$ and $|F(L)|$ is the determinant of $F(L)$. Thus we get a stable VARMA(p^*, q^*) representation of the process (\mathbf{y}_t) in (2.3)

$$(2.5) \quad \phi(L)(\mathbf{y}_t - \mu_{\mathbf{y}}) = \theta(L)\boldsymbol{\epsilon}_t$$

where $p^* = q^* \leq M - 1$, $\phi(L) = |F(L)|$ is scalar and

$$\theta(L) = (\tilde{\mathbf{\Lambda}}F(L)^* \quad \tilde{\mathbf{\Sigma}}(F(L)^* \otimes \mathbf{I}_K) \quad |F(L)|\mathbf{I}_K).$$

See also Cavicchioli (2014a), Th.6. The error term is also given by

$$\boldsymbol{\epsilon}_t = (\mathbf{w}_t' \quad \mathbf{u}_t'(\mathbf{w}_t' \otimes \mathbf{I}_K) \quad \mathbf{u}_t'(\boldsymbol{\pi}' \otimes \mathbf{I}_K)\mathbf{\Sigma}')'$$

with variance matrix

$$\Xi = \text{Var}(\boldsymbol{\epsilon}_t) = \text{diag}(\tilde{\mathbf{D}} - \mathbf{F}\tilde{\mathbf{D}}\mathbf{F}', (\tilde{\mathbf{D}} - \mathbf{F}\tilde{\mathbf{D}}\mathbf{F}') \otimes \mathbf{I}_K, \mathbf{\Sigma}((\mathbf{D}\mathbf{P}_\infty) \otimes \mathbf{I}_K)\mathbf{\Sigma}').$$

Using (2.5) the spectral density matrix of the process (\mathbf{y}_t) in (2.3) is also given by

$$F_{\mathbf{y}}(\omega) = \frac{\theta(e^{i\omega})\Xi\theta'(e^{-i\omega})}{|\phi(e^{i\omega})|^2}.$$

In fact, we can apply a well-known result (see, for example, Gouriéroux and Monfort (1997), Chp.8, Formula 8.3, p.257). The spectral density of a VARMA process

$$\Phi(L)\mathbf{y}_t = \Theta(L)\boldsymbol{\epsilon}_t,$$

with $Var(\boldsymbol{\epsilon}) = \boldsymbol{\Omega}$, is given by

$$(2.6) \quad F_{\mathbf{y}}(\omega) = \frac{1}{2\pi} \Phi^{-1}(\exp(i\omega)) \Theta(\exp(i\omega)) \boldsymbol{\Omega} \overline{\Theta(\exp(i\omega))}' \overline{\Phi^{-1}(\exp(i\omega))}'$$

This formula can be applied when $\det \Phi(z)$ has all its roots outside the unit circle. Moreover, we can also write $F_{\mathbf{y}}(\omega)$ as

$$F_{\mathbf{y}}(\omega) = \frac{1}{2\pi} \frac{\Phi^*(\exp(i\omega)) \Theta(\exp(i\omega)) \boldsymbol{\Omega} \overline{\Theta(\exp(i\omega))}' \overline{\Phi^*(\exp(i\omega))}'}{|\det \Phi(\exp(i\omega))|^2}$$

where Φ^* denotes the adjoint matrix of Φ . Here, we apply these formulae ignoring the coefficient. Written in this form $F_{\mathbf{y}}(\omega)$ is a matrix whose elements are rational functions of $\exp(i\omega)$. This property is a characteristic of the VARMA process.

2.2 Spectra of MS-VAR(p)

Let us consider the MS-VAR(p), $p > 0$, process

$$(2.7) \quad A(L)\mathbf{y}_t = \boldsymbol{\nu}_{st} + \boldsymbol{\Sigma}_{st} \mathbf{u}_t$$

where $A(L) = \mathbf{I}_K - \mathbf{A}_1 L - \dots - \mathbf{A}_p L^p$ is a $(K \times K)$ -dimensional lag polynomial. Assume that there are no roots on or inside the unit circle of the complex plane, i.e., $|A(z)| \neq 0$ for $|z| \leq 1$. Reasoning as above, the process (\mathbf{y}_t) in (2.7) admits a stable VARMA(p^*, q^*) with $p^* \leq M + p - 1$ and $q^* \leq M - 1$:

$$(2.8) \quad \Psi(L)(\mathbf{y}_t - \mu_{\mathbf{y}}) = \theta(L)\boldsymbol{\epsilon}_t$$

where $\Psi(L) = |F(L)|A(L) = \phi(L)A(L)$ and $\theta(L)\boldsymbol{\epsilon}_t$ is as in (2.4). If we want the autoregressive part of the stable VARMA in (2.8) to be scalar, we have to multiply (2.8) on the left with the adjoint $A(L)^*$ to give a stable VARMA(p', q') representation, where the bounds satisfy $p' \leq M + Kp - 1$

and $q' \leq M + (K - 1)p - 1$. Thus the spectral density matrix of the process (\mathbf{y}_t) in (2.8) is given by

$$\begin{aligned} F_{\mathbf{y}}(\omega) &= \frac{A^{-1}(e^{i\omega})\theta(e^{i\omega})\Xi\theta'(e^{-i\omega})[A'(e^{-i\omega})]^{-1}}{|\phi(e^{i\omega})|^2} \\ &= \frac{A^*(e^{i\omega})\theta(e^{i\omega})\Xi\theta'(e^{-i\omega})A^{*'}(e^{-i\omega})}{|\phi(e^{i\omega})|^2|\det A(e^{i\omega})|^2}. \end{aligned}$$

From the above section we can also obtain the matrix expression

$$(2.9) \quad \begin{aligned} F_{\mathbf{y}}(\omega) &= A^{-1}(e^{i\omega})Q[A'(e^{-i\omega})]^{-1} + 2A^{-1}(e^{i\omega})\tilde{\mathbf{\Lambda}}\mathbf{F} \\ &\quad \times \mathcal{R}e\{(\mathbf{I}_{M-1}e^{i\omega} - \mathbf{F})^{-1}\}\tilde{\mathbf{D}}\tilde{\mathbf{\Lambda}}'[A'(e^{-i\omega})]^{-1}. \end{aligned}$$

A similar result can be obtained for a Markov switching VAR(p), $p > 0$, process

$$(2.10) \quad A_{s_t}(L)\mathbf{y}_t = \boldsymbol{\nu}_{s_t} + \boldsymbol{\Sigma}_{s_t}\mathbf{u}_t$$

where we assume that the state variable is independent of the observables. Define

$$\mathbf{A}(L) = (A_1(L) \dots A_M(L))$$

where

$$A_m(L) = \mathbf{I}_K - \mathbf{A}_{1,m}L - \dots - \mathbf{A}_{p,m}L^p$$

for $m = 1, \dots, M$. Recall that $s_t \in \{1, \dots, M\}$. Then (2.10) can be written in the form

$$\mathbf{A}(L)(\boldsymbol{\xi}_t \otimes \mathbf{I}_K)\mathbf{y}_t = \boldsymbol{\Lambda}\boldsymbol{\xi}_t + \boldsymbol{\Sigma}(\boldsymbol{\xi}_t \otimes \mathbf{I}_K)\mathbf{u}_t.$$

Assume that $B(L) = \mathbf{A}(L)(\boldsymbol{\pi} \otimes \mathbf{I}_K)$ is invertible. Then the spectral density matrix of the process (\mathbf{y}_t) in (2.10) is given by

$$(2.11) \quad F_{\mathbf{y}}(\omega) = \frac{B^{-1}(e^{i\omega})\theta(e^{i\omega})\Xi\theta'(e^{-i\omega})[B'(e^{-i\omega})]^{-1}}{|\phi(e^{i\omega})|^2}.$$

Finally, we can also obtain the matrix expression

$$(2.12) \quad \begin{aligned} F_{\mathbf{y}}(\omega) &= B^{-1}(e^{i\omega})Q[B'(e^{-i\omega})]^{-1} + 2B^{-1}(e^{i\omega})\tilde{\mathbf{\Lambda}}\mathbf{F} \\ &\quad \times \mathcal{R}e\{(\mathbf{I}_{M-1}e^{i\omega} - \mathbf{F})^{-1}\}\tilde{\mathbf{D}}\tilde{\mathbf{\Lambda}}'[B'(e^{-i\omega})]^{-1}. \end{aligned}$$

3. Simulation exercises

As a first example, let us consider an univariate MS(2)-AR(1) model defined as

$$(3.1) \quad \begin{cases} y_t = a_1 y_{t-1} + \sigma_1 u_t & s_t = 1 \\ y_t = a_2 y_{t-1} + \sigma_2 u_t & s_t = 2. \end{cases}$$

In Figure 1 we plot the spectral density of the model (3.1), where the error term is $u_t \sim IIDN(0, 1)$ and $\sigma_1 = \sigma_2 = 1$. The autoregressive coefficients and transition probabilities are as follows: $a_1 = .8, a_2 = .3, p_{11} = .6, p_{22} = .4$ (top left panel); $a_1 = .8, a_2 = .3, p_{11} = .9, p_{22} = .1$ (top right panel); $a_1 = -.8, a_2 = -.3, p_{11} = .6, p_{22} = .4$ (bottom left panel); $a_1 = .8, a_2 = -.3, p_{11} = .6, p_{22} = .4$ (bottom right panel). In general, spectra are a sort of weighted average of the two underlying linear models. Moreover, when the autoregressive coefficients are both positive ($a_1 = 0.8$ and $a_2 = 0.3$) or negative ($a_1 = -0.8$ and $a_2 = -0.3$), the shape is similar to the typical spectral representation of an AR(1) with positive/negative coefficients. When the sign is opposite ($a_1 = 0.8$ and $a_2 = -0.3$) the prevailing shape depends on which model dominates in terms of absolute value of the coefficients and underlying probabilities. Finally, if the transition probabilities are very different ($p = 0.9$ and $q = 0.1$), the more persistent regime becomes more likely and the peak of the spectrum increases.

As a second example, let us consider a simulation experiment from Diebold and Inoue (2001) (Example 4.3, page 149) in which they study the finite-sample property of some MS(2)-AR(0) models. We analyze the same model from the frequency domain perspective, taking advantage of our results in Section 2. The simulated processes are the following:

$$(3.2) \quad y_t = \mu_{s_t} + \epsilon_t$$

where $\epsilon_t \sim IIDN(0, \sigma^2)$, and s_t and ϵ_τ are independent for all t and τ . The intercept term takes values $\mu_0 = 0$ in the first regime and $\mu_1 = 1$ in the

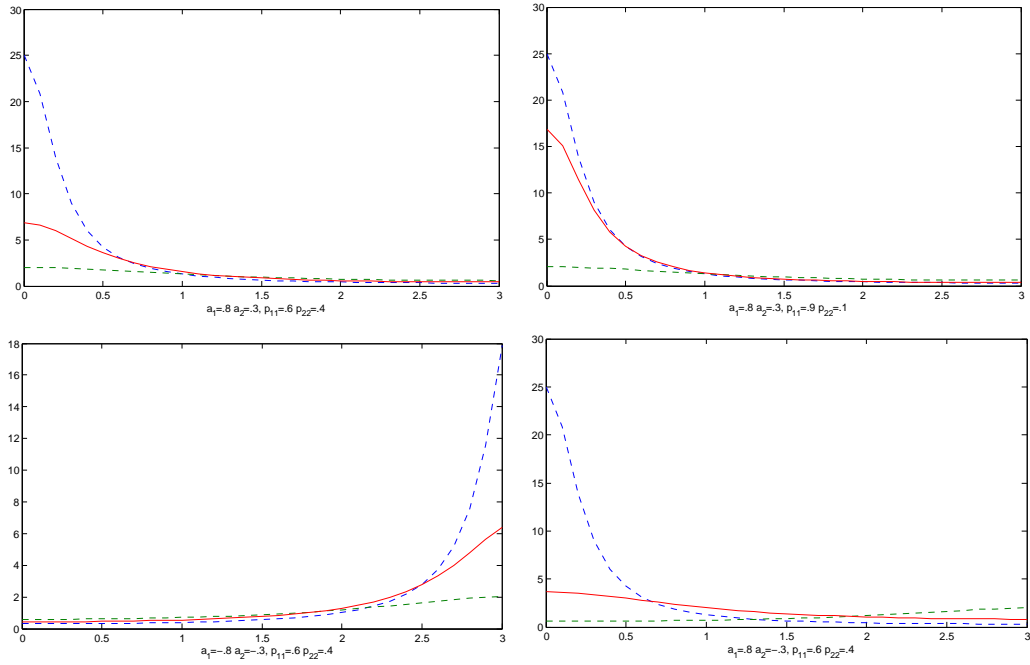


Figure 1: Plot of spectral density functions from the simulated MS(2)-AR(1) model in (3.1), where $\sigma_1 = \sigma_2 = 1$ and autoregressive coefficients and transition probabilities are as follows: $a_1 = .8, a_2 = .3, p_{11} = .6, p_{22} = .4$ (top left panel); $a_1 = .8, a_2 = .3, p_{11} = .9, p_{22} = .1$ (top right panel); $a_1 = -.8, a_2 = -.3, p_{11} = .6, p_{22} = .4$ (bottom left panel); $a_1 = .8, a_2 = -.3, p_{11} = .6, p_{22} = .4$ (bottom right panel). Spectral densities (solid lines) are depicted together with the two underlying linear models (dashed lines).

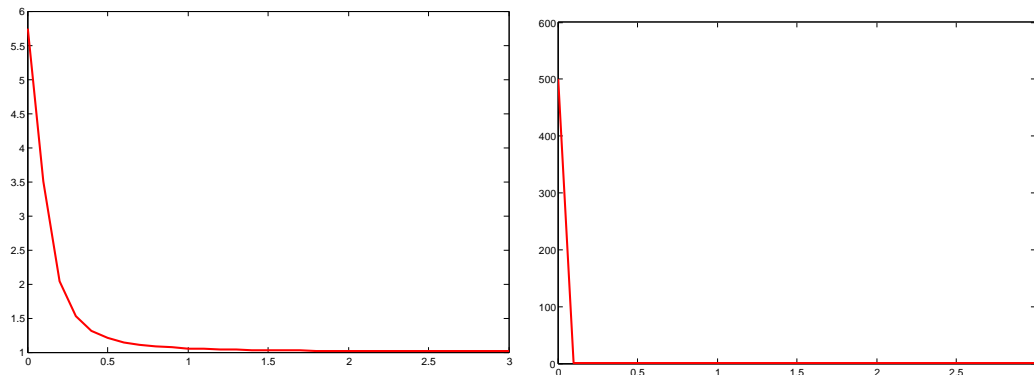


Figure 2: Plot of spectral density functions from simulated MS(2)-AR(0) models in (3.2) with the intercept term equal to 0 in the first regime and 1 in the second regime. Transition probabilities are $p_{00} = p_{11} = 0.95$ (left panel) and $p_{00} = p_{11} = 0.9995$ (right panel). This example is taken from Diebold and Inoue (2001) (4.3, page 149).

second regime. For a short time series ($T = 400$) or a long one ($T = 10,000$) the following results remain the same. Diebold and Inoue (2001) observe the behaviour of the data for different values of probabilities (p_{00} and p_{11}) and show that for equal probabilities close to one (e.g., 0.9995), the regime does not change with positive probability, so that it does a good job of mimicking long memory. On the contrary, when we observe equal transition probabilities but well away from unity (e.g., 0.95), the long memory feature is less marked in the data. Note at this point that traditionally long memory has been defined in the time domain in terms of decay rate of long-lag autocorrelations, or in the frequency domain in terms of rates of explosion of low-frequency spectra. Here we are able to confirm Diebold and Inoue (2001) conclusions through the analysis of the spectra of the two described cases. Figure 2 shows the spectral densities of the MS(2)-AR(0) model with equal probabilities which are away from unity (left panel) and with probabilities very close to one (right panel). In the first case the spectrum shows a smoother behaviour which vanishes only at frequency equal to 1, while in the second we witness an explosion at low-frequency in the spectra. Therefore, we conclude that stochastic regime switching is intimately related to long memory and could

be easily confused with it, as long as only a small amount of regime switches occurs in an observed sample path.

4. Frequency variability in real data

A recent paper by Müller and Watson (2008) has proposed a framework to study how successful are time series models in explaining low-frequency variability. In fact, some econometric models (local-to-unity or fractional) were specifically designed to capture low-frequency variability of the data. However, they provide reliable guidance for empirical analysis only if they are able to accurately describe not only low-frequency behaviour of the time series, but also high-frequency. In particular, the authors focus on lower frequencies than the business cycle, that is a period greater than eight years, and some inference is proposed on the low-frequency component of the series of interest by computing weighted averages of the data, where the weights are low-frequency trigonometric series. We propose to look at the relative importance of low- and high-frequencies in a time series from a different perspective. We assume that a suitable parametric model should be able to capture the relative importance of the different frequencies which characterize the behaviour of the series. Our aim is to study some empirical questions of interest from Müller and Watson (2008) with a different approach. In particular, we firstly use recent test from Cavicchioli (2014a) to correctly parametrize the process we are considering. Then, by using simple Maximum Likelihood Estimation (MLE) expressions from Cavicchioli (2014b) we proceed to estimate the model. Finally, using spectral density results presented in Section 2, we check if the chosen model is able to extract frequency variability of the initial process. Following Müller and Watson (2008), we investigate the following questions: (1) after accounting for a deterministic linear trend, is real gross domestic product (GDP) consistent with a $I(1)$ model? (2) is the term spread consistent with the $I(0)$ model, that is are long term and short term interest rates cointegrated? (3) are absolute daily returns, which are characterized by "slow decay of autocorrelations" consistent with an $I(1)$ or $I(0)$ model?

We study those questions allowing the possibility that those series may be affected by structural changes and, if it is the case, they should be taken into account when fitting a model on the data. We take postwar quarterly U.S. data and focus on a period greater than 32 quarters, that is frequencies lower than the business cycle, as in Müller and Watson (2008). In particular, we consider quarterly values (1952:Q1 - 2005:Q3) of the logarithm of de-trended real GDP and de-measured term spread - difference between interest rates for 10 years and 1 year U.S Treasury bonds. Moreover, we observe daily absolute returns (January 2nd, 1957 - September 30th, 2013) computed as the logarithm of the ratio between consecutive closing prices from S&P500. Data are taken from the FRED database. Before proceeding with our analysis, we plot sample periodograms of the data in order to have a preliminary idea on the different behaviour of the series. In Figure 3 we recognize a mixed pattern of low- and high-frequency cycles for real GDP and bond spread which produces uncertainty on the relative importance of the two components. On the contrary, we recognize the explosion at the low-frequency in the periodogram of absolute returns, as we expected from "long-memory" considerations. To correctly estimate the process, the first step is to test linearity or non-linearity of the model and, if it is case, the number of regimes which characterizes the time series. For the determination of regime number, we use results from Cavicchioli (2014a).

With regard to the real GDP, we select a linear model, that is one regime is sufficient to describe the data. We include one lag for the autoregressive model (as suggested by standard information criteria for the AR model) and plot its spectral density in Figure 4 (upper panel). The spectra of this model is typical of an autoregressive model; here low-frequencies are the most important, giving credit to an $I(1)$ model. However, if we take the first differences of the series, we somehow depurate the process from the stochastic trend (not only from the linear deterministic one, as before). Here the test suggests a $MS(2)$ - $AR(1)$ model and the spectra in Figure 4 (lower panel) retains only high-frequency movements. It suggests that the long-

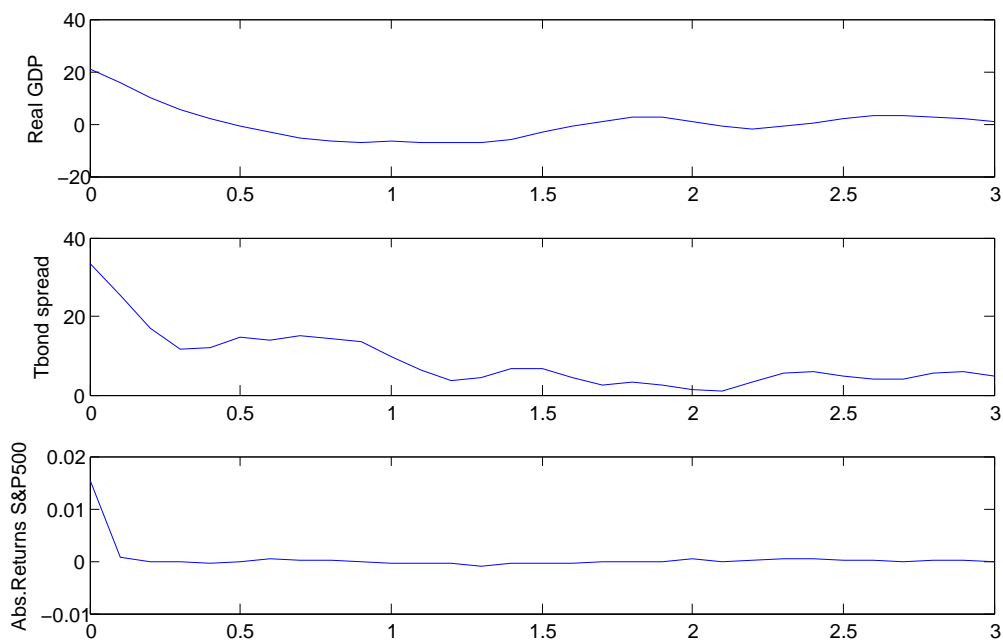


Figure 3: Periodograms of logarithm of detrended U.S. real GDP (top panel), demeaned Treasury bond spread as the difference between interest rates for 10 years and 1 year U.S Treasury bonds (middle panel); those two series have quarterly frequency and the period is from 1952:Q1 to 2005:Q3. Absolute daily returns (bottom panel) as the logarithm of the ratio between consecutive closing prices from S&P500 (January 2nd, 1957 - September 30th, 2013). Data are taken from FRED database.

run pattern characterized by two phases of the economy is captured by the switching model.

When considering the Treasury bond spread, a 2-state switching model is selected. Thus, we estimate a MS(2)-AR(1) model which turns to be as follows

$$\begin{aligned}\hat{\mu} &= [0.4033 \quad 0.0047] \\ sd(\hat{\mu}) &= [1.2356 \quad 0.4711] \\ \hat{\phi} &= [-1.1343 \quad -0.8474] \\ sd(\hat{\phi}) &= [1.3273 \quad 0.2711] \\ \hat{\sigma} &= [0.8737 \quad 0.3331] \\ sd(\hat{\sigma}) &= [0.6178 \quad 0.2356] \\ \hat{P} &= \begin{bmatrix} 0.76 & 0.24 \\ 0.24 & 0.76 \end{bmatrix}\end{aligned}$$

Then we use the estimated values in the spectral formulae of Section 2 to depict the spectral representation of the data, which is in Figure 5. The spectrum suggests that only high-frequencies of the series matter, even if with relatively uncertainty given by the 95% confidence bands. This seems to be consistent with an I(0) model, where high-frequency variability dominates the process.

Finally, we evaluate the behaviour of absolute returns from S&P500, which typically suffer of "long memory". Here a 3-state switching model suits the data and we estimate a MS(3)-AR(1). The estimated parameters are the following

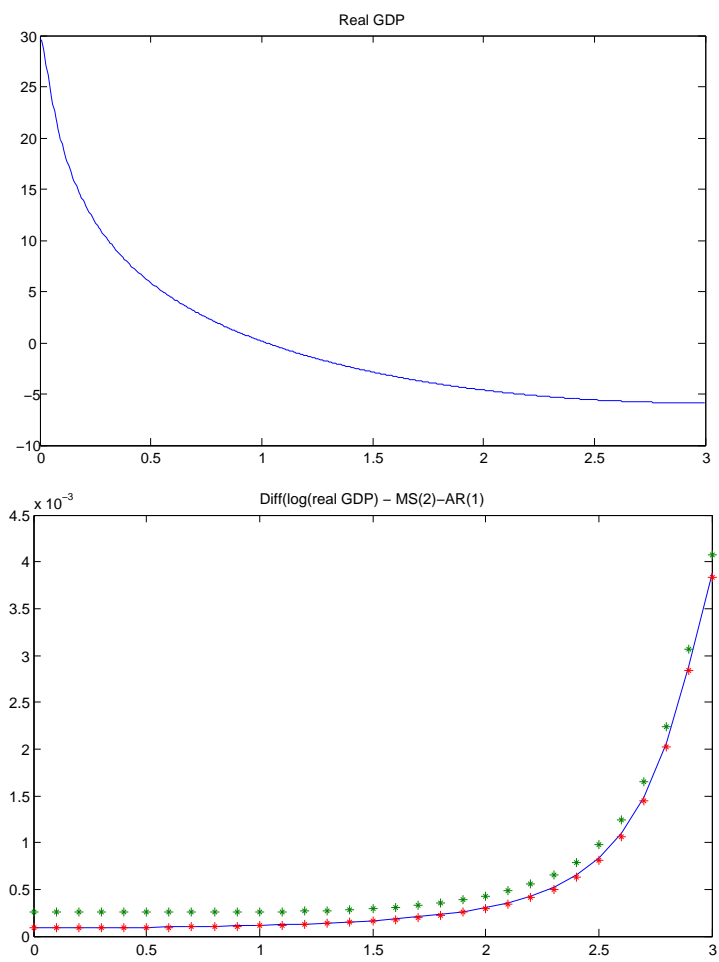


Figure 4: Spectral density functions (solid lines) for the logarithm of detrended U.S. real GDP modelled as a linear AR(1) (upper panel) and for the logarithm of differenced U.S. real GDP modelled as MS(2)-AR(1) (lower panel) along with 95% confidence interval bands (starred lines). Both series have quarterly frequency for the period 1952:Q1 to 2005:Q3. Data are taken from FRED database.

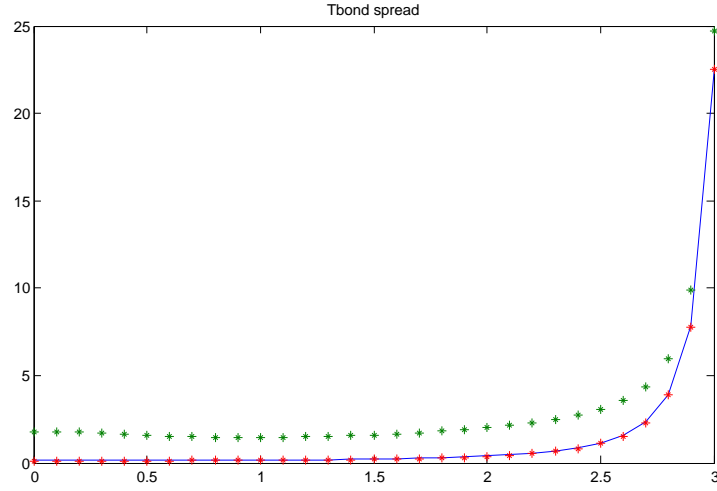


Figure 5: Spectral density (solid line) of demeaned Treasury bond spread (difference between interest rates for 10 years and 1 year) at quarterly frequency (1952:Q1 to 2005:Q3) modelled as a MS(2)-AR(1), along with 95% confidence interval bands (starred lines). Data are taken from FRED database.

$$\begin{aligned}
 \hat{\mu} &= [-0.0408 \quad 0.0004 \quad 0.0049] \\
 sd(\hat{\mu}) &= [0.1276 \quad 0.0009 \quad 0.0084] \\
 \hat{\phi} &= [0.2013 \quad -0.0001 \quad -0.0819] \\
 sd(\hat{\phi}) &= [4.8271 \quad 0.2609 \quad 0.5717] \\
 \hat{\sigma} &= [0.0737 \quad 0.0005 \quad 0.0048] \\
 sd(\hat{\sigma}) &= [0.0471 \quad 0.0003 \quad 0.0031] \\
 \hat{P} &= \begin{bmatrix} 0.98 & 0.00 & 0.02 \\ 0.02 & 0.98 & 0.00 \\ 0.00 & 0.04 & 0.96 \end{bmatrix}
 \end{aligned}$$

Using Formula (2.12), we construct the spectral density of the process having the above estimated parameters. This is plotted in Figure 6 and it is very close to the sample periodogram, with very tight confidence bands, opening up a room for considering structural change rather than long memory attributes of the process. Moreover, from the estimated values, we recog-

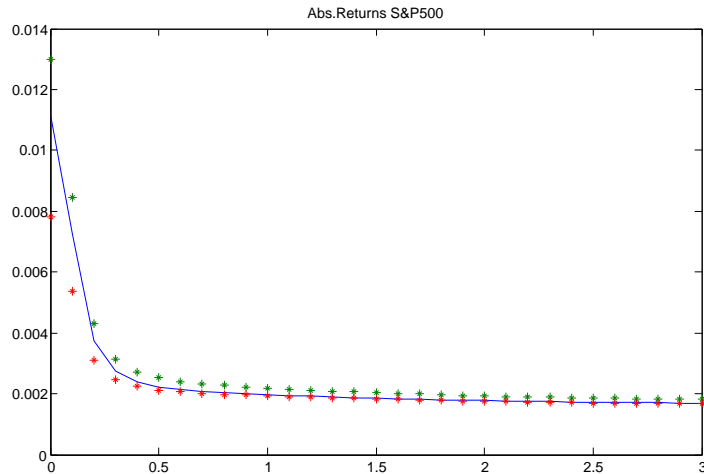


Figure 6: Spectral density (solid line) of absolute daily returns (logarithm of the ratio between consecutive closing prices from S&P500) for the period January 2nd, 1957 - September 30th, 2013, modelled as a MS(3)-AR(1), along with 95% confidence interval bands (starred lines). Data are taken from FRED database.

nize a first regime of high-volatility and negative returns, a second regime of low volatility and high returns and a third state of moderate volatility and average returns. In particular, estimated transition probabilities show that regimes are very persistent, which is also in line with the conclusion given by Diebold and Inoue (2001).

6. Conclusion

In this work we study multivariate AR models subject to Markov Switching in the most general form and derive close-form formulae for the spectral density functions of such processes. The spectral densities of these models can be very useful as a tool to infer information on the persistency characterizing the series and to check the correct parametrization of the process. In particular, after having assessed linearity or non-linearity of the series, spectral analysis gives some insights on the relative importance of high- and low- frequency variability and help to validate the assumed model. We applied the method to some macroeconomic and financial data to evaluate their

frequency variability and to investigate the problem of "structural change vs long memory" of returns via spectral analysis.

References

- [1] Cavicchioli, M. (2014a) *Determining the number of regimes in Markov-switching VAR and VMA models*, Journal of Time Series Analysis **35**(2), 173-186.
- [2] Cavicchioli, M. (2014b) *Analysis of the likelihood function for Markov switching VAR(CH) models*, Journal of Time Series Analysis **35**(6), 624-639.
- [3] Diebold, F.X. and Inoue, A. (2001) *Long memory and regime switching*, Journal of Econometrics **105**, 131-159.
- [4] Ding, Z., Granger, C.W.J. and Engle, R.F. (1993) *A long memory property of stock market returns and a new model*, Journal of Empirical Finance **1**, 83-106.
- [5] Ding, Z. and Granger, C.W.J. (1996) *Modeling volatility persistence of speculative returns: a new approach*, Journal of Econometrics **73**, 185-215.
- [6] Gouriéroux, C. and Monfort, A. (1997) *Time Series and Dynamic Models*, Cambridge University Press.
- [7] Granger, C.W.J. and Hyung, N. (2004) *Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns*, Journal of Empirical Finance **11**, 399-421.
- [8] Mikosch, T. and Stărică, C. (2000) *Is it really long memory we see in financial returns?*, Extremes and Integrated Risk Management **12**, 149-168.

- [9] Müller, U. K. and Watson, M. W. (2008) *Testing Models of Low Frequency Variability*, *Econometrica* **76**(5), 979-1016.
- [10] Krolzig, H.M. (1997) *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference and Application to Business Cycle Analysis*, Springer Verlag, Berlin-Heidelberg-New York.
- [11] Pataracchia, B., (2011) *The Spectral Representation of Markov Switching ARMA Models*, *Economics Letters* **112**, 11-15.

Appendix

Derivation of Formula (2.4):

The spectral density of the process (\mathbf{y}_t) in (3) is given by

$$\begin{aligned}
 F_{\mathbf{y}}(\omega) &= \sum_{h=-\infty}^{+\infty} \Gamma_{\mathbf{y}}(|h|)e^{-i\omega h} = \Gamma_{\mathbf{y}}(0) + \sum_{h=1}^{+\infty} \Gamma_{\mathbf{y}}(h)e^{-i\omega h} + \sum_{k=-\infty}^{-1} \Gamma_{\mathbf{y}}(k)e^{-i\omega k} \\
 &= \Gamma_{\mathbf{y}}(0) + \sum_{h=1}^{+\infty} \Gamma_{\mathbf{y}}(h)e^{-i\omega h} + \sum_{h=1}^{+\infty} \Gamma_{\mathbf{y}}(h)e^{i\omega h}.
 \end{aligned}$$

Note that

$$\left(\sum_{h=1}^n A^h \right) (I - A) = (A + A^2 + \dots + A^n)(I - A) = A - A^{n+1}$$

which is equal to A when n goes to infinity with the spectral radius of A less than 1. Hence

$$\left(\lim_{n \rightarrow +\infty} \sum_{h=1}^n A^h \right) (I - A) = A \quad \text{and} \quad \sum_{h=1}^{+\infty} A^h = A(I - A)^{-1}.$$

It turns out that spectral density of the process in (3) is given by

$$\begin{aligned}
F_{\mathbf{y}}(\omega) &= \tilde{\Lambda}\tilde{\mathbf{D}}\tilde{\Lambda}' + \tilde{\Sigma}(\tilde{\mathbf{D}} \otimes \mathbf{I}_K)\tilde{\Sigma}' + \Sigma((\mathbf{D}\mathbf{P}_\infty) \otimes \mathbf{I}_K)\Sigma' \\
&\quad + \sum_{h=1}^{+\infty} \tilde{\Lambda}F^h\tilde{\mathbf{D}}\tilde{\Lambda}'e^{-i\omega h} + \sum_{h=1}^{+\infty} \tilde{\Lambda}F^h\tilde{\mathbf{D}}\tilde{\Lambda}'e^{i\omega h} \\
&= Q + \tilde{\Lambda} \sum_{h=1}^{+\infty} (Fe^{-i\omega})^h \tilde{\mathbf{D}}\tilde{\Lambda}' + \tilde{\Lambda} \sum_{h=1}^{+\infty} (Fe^{i\omega})^h \tilde{\mathbf{D}}\tilde{\Lambda}' \\
&= Q + \tilde{\Lambda}(Fe^{-i\omega})(I - Fe^{-i\omega})^{-1}\tilde{\mathbf{D}}\tilde{\Lambda}' + \tilde{\Lambda}(Fe^{i\omega})(I - Fe^{i\omega})^{-1}\tilde{\mathbf{D}}\tilde{\Lambda}' \\
&= Q + 2\tilde{\Lambda}\mathbf{F}\mathcal{R}e\{(\mathbf{I}_{M-1}e^{i\omega} - \mathbf{F})^{-1}\}\tilde{\mathbf{D}}\tilde{\Lambda}'
\end{aligned}$$

where $\mathcal{R}e$ denotes the real part of the complex matrix $(\mathbf{I}_{M-1}e^{i\omega} - \mathbf{F})^{-1}$, and

$$Q = \tilde{\Lambda}\tilde{\mathbf{D}}\tilde{\Lambda}' + \tilde{\Sigma}(\tilde{\mathbf{D}} \otimes \mathbf{I}_K)\tilde{\Sigma}' + \Sigma((\mathbf{D}\mathbf{P}_\infty) \otimes \mathbf{I}_K)\Sigma'. \quad \square$$