RESEARCH

Root tracking using time-varying autoregressive moving average models and sigma-point Kalman filters

Kyriaki Kostoglou[*] and Michael Lunglmayr

Abstract

Root tracking is a powerful technique that provides insight into the mechanisms of various time-varying processes. The poles and the zeros of a signal-generating system determine the spectral characteristics of the signal under consideration. In this work, time-frequency analysis is achieved by tracking the roots of time-varying processes using autoregressive moving average (ARMA) models in cascade form. A cascade ARMA model is essentially a high-order infinite impulse response (IIR) filter decomposed into a series of first- and second-order sections. Each section is characterized by real or conjugate pole/zero pairs. This filter topology allows individual root tracking as well as immediate stability monitoring and correction. Also, it does not suffer from high round-off error sensitivity, as is the case with the filter coefficients of the direct-form ARMA structure. Instead of using conventional gradient-based recursive methods, we investigate the performance of derivative-free sigma-point Kalman filters for root trajectory tracking over time. Based on simulations, the sigma-point estimators provide more accurate estimates, especially in the case of tightly clustered poles and zeros. The proposed framework is applied to real data, and more specifically, it is used to examine the time-frequency characteristics of raw ultrasonic signals from medical ultrasound images.

Keywords: ARMA models, Time-varying, Cascade structure, Sigma-point Kalman filter, Root tracking, Genetic algorithm, Ultrasound imaging

1 Introduction

Most of the signals generated by natural or artificial systems exhibit nonstationary characteristics, i.e., their properties change over time. Nonstationarity could be the result of unobservable interactions/trends or an inherent feature of the signal itself. Examples of time-varying (TV) signals are abundant, including biological measurements such as cardiac and brain signals, music, speech, seismic waves, as well as financial time series, and climate data. A useful tool for gaining insight into the nonstationary nature of a signal is time-frequency (TF) analysis [1]. TF analysis is a collection of mathematical formulations and signal processing/modeling techniques used to describe the spectral and temporal variations of a signal. TF analysis has been extensively used in both industrial and academic environments in a wide range of applications, from engineering to ecology and meteorology. Furthermore, TF features and maps

have been exploited successfully in various pattern-recognition and machine-learning classification problems [2–5]. Another important feature of TF analysis is the instantaneous frequency concept [6], which has been proven valuable in the field of telecommunications as well as in radar and sonar applications.

Existing TF estimation approaches can be grossly classified into two categories: parametric and nonparametric. The basic idea behind nonparametric approaches is to devise a joint function or else a distribution that describes the energy density of a signal in both the time and frequency domains. Examples of such distributions are the short-time Fourier transform (STFT) [1, 7], the Gabor transform [1, 8], the Wigner–Ville distribution (WVD) [1], Cohen's class time-frequency distributions [1], and the continuous wavelet transform (CWT) [9, 10]. These methods are simple, fast, and do not require an explicit model or prior knowledge of the signal characteristics. However, they exhibit important limitations, for example, in STFT good frequency localization

* Correspondence: kyriaki.kostoglou@jku.at
Institute of Signal Processing, Johannes Kepler University, Linz, Austria

depends heavily on the length of the signal being processed, in CWT the selected mother wavelet may affect significantly the TF representation, and in WVD cross terms and negative distribution values deteriorate the TF resolution. Parametric approaches, on the other hand, assume that the observed signal is a realization of a stochastic process that can be described by deterministic or probabilistic models. The estimated TV model coefficients are then used to create a TF representation of the signal. If the model is correctly chosen, parametric methods can achieve higher TF resolution than the nonparametric techniques, even in relatively short datasets [11]. However, the choice of an appropriate model, as well as the tuning of all necessary hyperparameters, may increase the computational complexity and runtime [12].

In this paper, we focus on parametric approaches and specifically TF representations based on TV autoregressive moving average (TV-ARMA) models [13–15]. TV-ARMA models provide parsimonious descriptions of various nonstationary processes. From a digital signal processing view, they are equivalent to the well-known infinite impulse response (IIR) filters with a finite number of feedforward (MA) and feedback (AR) coefficients that vary over time. The TV coefficients are usually estimated using quasi-stationary or recursive techniques and then translated in both the time and frequency domains, generating fine TF representations. The main issue with the conventional TV-ARMA structure (also known as the direct-form ARMA model) is the difficulty of maintaining stability, especially if the underlying process is narrowband or adaptation becomes too rapid.

A more natural and less sensitive representation that provides robust stability monitoring is the TV-ARMA model in cascade form. Essentially, the direct form TV-ARMA model is reparametrized in terms of its roots and expressed in a cascade structure of first- and second-order sections. This enables independent tracking of the location of the poles and the zeros of each filter. Root tracking has been proposed earlier for various applications such as speech [16, 17], biosignal analysis [18–20], fault diagnosis and condition monitoring [21], and channel prediction in telecommunications [22]. Even though the aforementioned representation provides a more natural description of the process, the transition from the conventional TV-ARMA form to the cascade form results in a model that is nonlinear in its parameters. Various recursive techniques have been proposed; however, they exhibit high sensitivity to initial conditions and noise and require gradient calculations. To this end, we propose a gradient-free approach based on sigma-point Kalman filters (KF). Specifically, we employ a Rao–Blackwellized sigma-point KF structure that uses the Stirling approximation method of the central difference KF. The performance of the estimator is validated using

simulations as well as real data from the field of ultrasound imaging.

2 Methods

2.1 Time-varying autoregressive moving average (TV-ARMA) model in direct form

Wold's decomposition theorem [23, 24] states that any discrete-time stationary process can be described as the sum of two components: one deterministic and one stochastic. The deterministic counterpart of the process can be predicted with no error using its entire past. The stochastic component, on the other hand, can be expressed as a linear combination of lagged white noise process values (the so-called infinite-order moving average, MA(∞)). Wold's theorem constitutes the basis of the well-known ARMA models [15]. ARMA models approximate the infinite lag polynomial of the Wold representation using the ratio of two finite-lag polynomials. Extending Wold's theorem to nonstationary processes, Cramer–Wold's decomposition [25] allows the representation of nonstationary signals as ARMA processes with TV coefficients. A TV-ARMA model of order (p, q) is expressed as follows [13–15]:

$$
\begin{aligned}
y(n) = &-\sum_{k=1}^{p} a_k(n)y(n-k) \\
&+ \sum_{k=1}^{q} b_k(n)e(n-k) + e(n),
\end{aligned} \tag{1}
$$

where $y \in R^{N \times 1}$ is the output (i.e., the signal under consideration in our case), $e \in R^{N \times 1}$ is the driving white noise signal with zero mean and variance σ_e^2, and $a(n) = [a_1(n)...a_p(n)]^T \in R^{p \times 1}$ and $b(n) = [b_1(n)...b_q(n)]^T \in R^{q \times 1}$ are the TV autoregressive (AR) and moving average (MA) model coefficients, respectively, at time point n. In practical applications, e is not known, and therefore, the purpose of estimating a TV-ARMA is twofold: (a) identify the TV coefficients of the model and (b) extract the underlying driving noise of the process. Note that herein we investigate only real signals, and therefore, y is real. In the z-domain, Eq. (1) can be written as:

$$
\begin{aligned}
Y(z) &= H(z,n)E(z) \\
&= \frac{B(z,n)}{A(z,n)} E(z) = \frac{1+\sum_{k=1}^{q} b_k(n)z^{-k}}{1+\sum_{k=1}^{p} a_k(n)z^{-k}} E(z),
\end{aligned} \tag{2}
$$

where $H(z,n) = \frac{1+\sum_{k=1}^{q} b_k(n)z^{-k}}{1+\sum_{k=1}^{p} a_k(n)z^{-k}}$ is the transfer function of the signal generating system in direct form. Based on Eq. (2), the nonstationary process y can be described as the output of a causal linear TV filter driven by a white noise input sequence e. The poles and the zeros of $H(z, n)$ (i.e., the roots of $A(z, n)$ and $B(z, n)$, respectively) should reside inside the unit circle $|z| \leq 1$ to guarantee stability and invertibility.

2.1.1 TV-ARMA estimation in direct form

Eq. (1) can be reformulated as:

$$
y(n) = \begin{bmatrix} \boldsymbol{a}^T(n) & \boldsymbol{b}^T(n) \end{bmatrix} \cdot \begin{bmatrix} -y(n-1) \\ \vdots \\ -y(n-p) \\ e(n-1) \\ \vdots \\ e(n-q) \end{bmatrix} + e(n)
$$

$$
= \boldsymbol{c}^T(n)\boldsymbol{\phi}(n) + e(n), \tag{3}
$$

where $\boldsymbol{c}^T(n) = [\boldsymbol{a}^T(n) \quad \boldsymbol{b}^T(n)] \in \boldsymbol{R}^{1 \times d}$ are the $d = p + q$ model coefficients at time n, and $\boldsymbol{\phi}(n) \in \boldsymbol{R}^{d \times 1}$ is the regressor vector that consists of the p and q past lags of $\boldsymbol{y}$ and $\boldsymbol{e}$, respectively. One common method used to estimate the trajectory of the coefficient vector $\boldsymbol{c}(n)$ over time is the recursive estimation technique. At each time step, $\boldsymbol{c}(n)$ is updated, enabling the adaptation of the model to possible variations in the process. In this work, we focus on the Kalman filtering technique [15]. Use of a KF assumes that the model coefficients follow a random walk driven by Gaussian white noise (GWN) with covariance $\boldsymbol{R_1} = R_1 \boldsymbol{I}_{d \times d}$, where $R_1 = \sigma_\xi^2$ (see Eq. 4a). R_1 essentially dictates the expected magnitude of the coefficient fluctuations. In case of large variations, R_1 is assigned with a large value and vice versa. The measurement noise is also assumed to be GWN with variance R_2. In state-space form this can be expressed as:

$$
\boldsymbol{c}(n) = \boldsymbol{c}(n-1) + \xi(n), \quad \boldsymbol{\xi} \sim N(0, R_1), \tag{4a}
$$

$$
y(n) = \boldsymbol{c}^T(n)\boldsymbol{\phi}(n) + e(n), \quad \boldsymbol{e} \sim N(0, R_2), \tag{4b}
$$

where the coefficient vector $\boldsymbol{c}(n)$ is the unknown state, and $y(n)$ is the observation at time point n. The KF is thus described by the following set of recursive equations [15]:

$$
\hat{e}(n) = y(n) - \hat{\boldsymbol{\phi}}^T(n)\hat{\boldsymbol{c}}(n-1), \tag{5a}
$$

$$
r(n) = \hat{\boldsymbol{\phi}}^T(n)\boldsymbol{P}(n-1)\hat{\boldsymbol{\phi}}(n), \tag{5b}
$$

$$
\boldsymbol{K}(n) = \frac{\boldsymbol{P}(n-1)\hat{\boldsymbol{\phi}}(n)}{R_2 + r(n)}, \tag{5c}
$$

$$
\boldsymbol{P}(n) = \boldsymbol{P}(n-1) + \boldsymbol{R_1} - \boldsymbol{K}(n)\hat{\boldsymbol{\phi}}^T(n)\boldsymbol{P}(n-1), \tag{5d}
$$

$$
\hat{\boldsymbol{c}}(n) = \hat{\boldsymbol{c}}(n-1) + \boldsymbol{K}(\boldsymbol{n})\hat{e}(n), \tag{5e}
$$

where $\hat{e}(n)$ is the a priori prediction error, $\hat{\boldsymbol{c}}(n)$ are the estimated coefficients, $\boldsymbol{K}(n) \in \boldsymbol{R}^{d \times 1}$ is the Kalman gain matrix, and $\boldsymbol{P}(n)$ is the a posteriori error covariance. The initial value for $\boldsymbol{P}$ is a diagonal matrix $\boldsymbol{P}(0) = P_0 \boldsymbol{I}_{p \times p}$, whereas the initial value for the coefficient vector is $\hat{\boldsymbol{c}}(0)$. Note that $\hat{\boldsymbol{\phi}}(n)$ is given as:

$$
\hat{\boldsymbol{\phi}}^T(n) = [-y(n-1) \ldots -y(n-k) \quad \hat{e}(n-1) \ldots \hat{e}(n-q)], \tag{6}
$$

with $\hat{e}(n)$ being the a priori prediction error of Eq. (5a). The rationale is that in cases where $\hat{\boldsymbol{c}}(n-1) \approx \hat{\boldsymbol{c}}(n)$, then $\hat{e}(n)$ is a good approximation of $e(n)$. The described approach is known in the literature as recursive pseudo-linear least squares (RPLS) [14]; however, the updating of the coefficients is realized using a recursive least squares (RLS) strategy. Therefore, we refer to the combination of RPLS with KF as KF-RPLS.

Recursive estimation may produce noisy estimates, depending on the initialization and the adaptation speed of the KF algorithm. Optionally, smoothing can be applied to the extracted TV coefficients using, for example, the Rauch–Tung–Striebel [26, 27] fixed-interval equations,

$$
\boldsymbol{W}(n) = \boldsymbol{P}(n)[\boldsymbol{P}(n+1) + \boldsymbol{R_1})]^{-1}, \tag{7}
$$

$$
\hat{\boldsymbol{c}}^s(n) = \hat{\boldsymbol{c}}(n) + \boldsymbol{W}(n)[\hat{\boldsymbol{c}}^s(n+1) - \hat{\boldsymbol{c}}(n)], \tag{8}
$$

where $\hat{\boldsymbol{c}}^s(n)$ is the smooth estimate of the coefficient vector. $\hat{\boldsymbol{c}}^s(n)$ is updated in a backward fashion starting from time point $n = N - 1$ up to $n = 1$ setting $\hat{\boldsymbol{c}}^s(N) = \hat{\boldsymbol{c}}(N)$.

2.1.2 TV-ARMA direct form spectrum

A smoothed version of the power spectral density (PSD) of the signal under consideration can be generated using the TV-ARMA model as an interpolating function. Once the model coefficients are computed, the TV PSD of $\boldsymbol{y}$ can be estimated by taking, at each time instant, the square of the magnitude of the transfer function $H(z, n)$ [13],

$$
S_{\boldsymbol{y}}(z, n) = S_{\boldsymbol{e}}(z, n)|H(z, n)|^2
$$

$$
= \sigma_e^2 \cdot \frac{\left|1 + \sum_{k=1}^{q} b_k(n)z^{-k}\right|^2_{z=e^{j2\pi f \Delta T}}}{\left|1 + \sum_{k=1}^{p} a_k(n)z^{-k}\right|^2_{z=e^{j2\pi f \Delta T}}}, \tag{9}
$$

where $\Delta T = \frac{1}{f_s}$, f_s is the sampling rate, $f \leq \frac{f_s}{2}$ is the frequency of interest, and σ_e^2 is the variance of the driving noise. Note that $S_e(z, n) = \sigma_e^2$, since $\boldsymbol{e}$ is assumed to be GWN. The variance σ_e^2 is unknown, and therefore the variance of the TV-ARMA residuals (Eq. 5a) is used as an estimate of σ_e^2.

2.2 Time-varying autoregressive moving average (TV-ARMA) model in cascade form

The main drawback of the direct-form TV-ARMA model is its temporal instability, especially if the underlying process is narrowband, that is, at least one of the roots is located very close to the unit circle [28, 29]. Large estimation errors either due to abrupt variations or excessive noise may temporarily force the poles

outside the unit circle. One way to mitigate this issue is to transfer all unstable poles inside the unit circle by factoring the characteristic polynomial of the denominator of Eq. (2) to roots and dividing each unstable root by its squared radius [15]. However, this implies extra computational power. In addition, even if this method works for low-order filters, it may fail for higher-order models due to root sensitivity to round-off errors in the coefficient of the polynomials [30]. A more convenient approach would be to directly adapt and control the location of each pole and zero, instead of adjusting the model coefficients. A representation that enables root tracking is the cascade form TV-ARMA [16–20, 31–35], where the transfer function of Eq. (2) is expressed as:

$$H(z,n) = \frac{\prod_{k=1}^{q_c} B_k{}^{(2)} \cdot \prod_{k=1}^{q_r} B_k{}^{(1)}}{\prod_{k=1}^{p_c} A_k{}^{(2)} \cdot \prod_{k=1}^{p_r} A_k{}^{(1)}} = \frac{\prod_{k=1}^{q_c}\{[1-\lambda_k(n)z^{-1}][1-\lambda_k^*(n)z^{-1}]\} \cdot \prod_{k=1}^{q_r}[1-\mu_k(n)z^{-1}]}{\prod_{k=1}^{p_c}\{[1-\rho_k(n)z^{-1}][1-\rho_k^*(n)z^{-1}]\} \cdot \prod_{k=1}^{p_r}[1-\xi_k(n)z^{-1}]}.$$
(10)

Equation (10) can be described as a cascade of first- and second-order filter sections. $B_k{}^{(2)}$, $A_k{}^{(2)}$ are the kth second-order sections and $B_k{}^{(1)}$, $A_k{}^{(1)}$ are the kth first-order sections of the numerator and denominator, respectively; $\{\lambda_k(n), \lambda_k^*(n)\}$ is the kth pair of complex-conjugate zeros at time point n; $\{\rho_k(n), \rho_k^*(n)\}$ is the kth pair of complex-conjugate poles; and $\mu_k(n)$ and $\xi_k(n)$ are the kth real zero and pole, respectively. The parameters p_r and p_c represent the total number of real and complex-conjugate pairs of poles. Similarly, q_r and q_c are the total number of real and complex-conjugate pairs of zeros. It is assumed that the number of poles and zeros remains constant throughout time. A complex pair of poles describes oscillatory signal behavior (i.e., spectral peaks), whereas real poles contribute to the high-pass or low-pass characteristics of the process. The effect of zeros is exactly the opposite than that of poles. Note that a TV-ARMA of order (p, q) in direct form has $p/2$ complex-conjugate pairs of poles if p is even, or $(p + 1)/2$ complex-conjugate pairs and one real pole if p is odd. The same applies to zeros. If q is even, then the model consists of $q/2$ complex-conjugate pairs of zeros. If q is odd, then the model has $(q + 1)/2$ complex-conjugate pairs and one real zero.

2.2.1 TV-ARMA estimation in cascade form
The complex-conjugate poles and the zeros of Eq. (10) can be represented in various forms (i.e., polar or rectangular representation) however, herein, we select the Cartesian coordinate representation [35, 36],

$$\rho_k(n) = \rho_{Rk}(n) + j\rho_{Ik}(n), \quad \rho_k^*(n)$$
$$= \rho_{Rk}(n) - j\rho_{Ik}(n),$$
(11a)

$$\lambda_k(n) = \lambda_{Rk}(n) + j\lambda_{Ik}(n), \quad \lambda_k^*(n)$$
$$= \lambda_{Rk}(n) - j\lambda_{Ik}(n),$$
(11b)

where $-1 \leq \rho_{Rk}(n)$, $\lambda_{Rk}(n) \leq 1|$ and $0 <= \{\rho_I\}\_k(n)$, $\{\lambda_I\}\_k(n) <= 1$. Since we are dealing with real signals, the poles and zeros should represent only positive frequencies. The central frequencies mapped by the poles and the zeros are obtained from the angle of their respective complex representations, i.e., $f_k(n) = \frac{\arg[\rho_k(n)]}{2\pi} f_s$ and $f_k(n) = \frac{\arg[\lambda_k(n)]}{2\pi} f_s$, where f_s is the sampling rate. Based on the Cartesian coordinate representation of Eqs. (11a) and (11b), the new model coefficient vector is defined as:

$$\boldsymbol{c}(n) = [\boldsymbol{\rho}(n) \; \boldsymbol{\lambda}(n) \; \boldsymbol{\xi}(n) \; \boldsymbol{\mu}(n)]^T \in \boldsymbol{R}^{d\times1},$$
(12a)

$$\boldsymbol{\rho}(n) = \left[\rho_{R1}(n)...\rho_{Rp_c}(n) \; \rho_{I1}(n)...\rho_{Ip_c}(n)\right]^T,$$
(12b)

$$\boldsymbol{\lambda}(n) = \left[\lambda_{R1}(n)...\lambda_{Rq_c}(n) \; \lambda_{I1}(n)...\lambda_{Iq_c}(n)\right]^T,$$
(12c)

$$\boldsymbol{\xi}(n) = \left[\xi_1(n)...\xi_{p_r}(n)\right]^T,$$
(12d)

$$\boldsymbol{\mu}(n) = \left[\mu_1(n)...\mu_{q_r}(n)\right]^T,$$
(12e)

where the total number of coefficients d is now $d = 2(p_c + q_c) + p_r + q_r$. The transfer function of the model (Eq. 10) then becomes:

$$H(z,n) = \frac{\prod_{k=1}^{q_c} B_k{}^{(2)}(n) \cdot \prod_{k=1}^{q_r} B_k{}^{(1)}(n)}{\prod_{k=1}^{p_c} A_k{}^{(2)}(n) \cdot \prod_{k=1}^{p_r} A_k{}^{(1)}(n)}$$

$$= \frac{\prod_{k=1}^{q_c}\{1-[\lambda_k(n)+\lambda_k^*(n)]z^{-1} + [\lambda_k(n)\lambda_k^*(n)]z^{-2}\} \cdot \prod_{k=1}^{q_r}[1-\mu_k(n)z^{-1}]}{\prod_{k=1}^{p_c}\{1-[\rho_k(n)+\rho_k^*(n)]z^{-1} + [\rho_k(n)\rho_k^*(n)]z^{-2}\} \cdot \prod_{k=1}^{p_r}[1-\xi_k(n)z^{-1}]}$$

$$= \frac{\prod_{k=1}^{q_c}\{1-2\lambda_{Rk}(n)z^{-1} + [\lambda_{Rk}^2(n)+\lambda_{Ik}^2(n)]z^{-2}\} \cdot \prod_{k=1}^{q_r}[1-\mu_k(n)z^{-1}]}{\prod_{k=1}^{p_c}\{1-2\rho_{Rk}(n)z^{-1} + [\rho_{Rk}^2(n)+\rho_{Ik}^2(n)]z^{-2}\} \cdot \prod_{k=1}^{p_r}[1-\xi_k(n)z^{-1}]}.$$
(13)

The main reason for selecting the Cartesian coordinate representation is its robustness to noise. Random perturbations of the real or imaginary part of a complex root will not affect the location of the corresponding pole or zero in the z-plane [37] significantly.

Based on the reparameterization of the TV-ARMA model in terms of its roots, the estimation problem can now be expressed in state-space form as:

$$\boldsymbol{c}(n) = \boldsymbol{c}(n-1) + \xi(n), \quad \boldsymbol{\xi} \sim N(0, R_1),$$
(14a)

$$y(n) = f\{\boldsymbol{c}(n)\} + e(n), \quad \boldsymbol{e} \sim N(0, R_2).$$
(14b)

In contrast to the direct-form case, the problem now is nonlinear in the coefficients, i.e., the observation $y(n)$ is a nonlinear function of the states/coefficients. In order to minimize $E(e^2)$, various recursive techniques have been developed. The most popular method in root

tracking is the recursive prediction error method (RPEM), which is basically a stochastic gradient algorithm. As in the case of the RPLS method, the RPEM follows the same structure as the conventional RLS technique. The only difference is that the main regressor vector, $\hat{\boldsymbol{\phi}}(n)$, is substituted with the negative gradient vector, defined as:

$$\boldsymbol{\psi}^T(n) = -\frac{\partial e(n)}{\partial \boldsymbol{c}(n)}. \tag{15}$$

In this work, we incorporate the RPEM method in the KF technique (Eq. 5a a–e). The combination of these two methods will be referred to as the KF-RPEM algorithm, which is summarized as follows:

$$\hat{e}(n) = y(n) - f\{\hat{\boldsymbol{c}}(n-1)\}, \tag{16a}$$

$$r(n) = \boldsymbol{\psi}^T(n)\boldsymbol{P}(n-1)\boldsymbol{\psi}(n), \tag{16b}$$

$$\boldsymbol{K}(n) = \frac{\boldsymbol{P}(n-1)\boldsymbol{\psi}(n)}{R_2 + r(n)}, \tag{16c}$$

$$\boldsymbol{P}(n) = \boldsymbol{P}(n-1) + \boldsymbol{R_1} - \boldsymbol{K}(n)\boldsymbol{\psi}^T(n)\boldsymbol{P}(n-1), \tag{16d}$$

$$\hat{\boldsymbol{c}}(n) = \hat{\boldsymbol{c}}(n-1) + \boldsymbol{K}(n)\hat{e}(n). \tag{16e}$$

The KF-RPEM is very similar to the well-known extended Kalman filter (EKF). However, there are some subtle differences, also described in [37, 38]. In order to estimate the gradient of e in terms of the coefficient vector $\boldsymbol{c}(n)$ (Eq. (15)), we applied the methodology described in [34]. The authors assume polar complex root representation; however, the extension to the Cartesian coordinates description of Eqs. (11a, 11b) is straightforward. The nonlinearity $f\{\hat{\boldsymbol{c}}(n-1)\}$ can be estimated indirectly, by computing first $\hat{e}(n)$ and then using Eq. (16a). Based on Eq. (2),

$$Y(z) = H(z,n)E(z) \Rightarrow E(z) = \frac{1}{H(z,n)} Y(z)$$
$$= \frac{\prod_{k=1}^{p_c} A_k^{(2)}(n) \cdot \prod_{k=1}^{p_r} A_k^{(1)}(n)}{\prod_{k=1}^{q_c} B_k^{(2)}(n) \cdot \prod_{k=1}^{q_r} B_k^{(1)}(n)} Y(z). \tag{17}$$

Equation (17) provides the a priori error $\hat{e}(n)$ by feeding $y(n)$ through the inverse filter $\frac{1}{H(z,n)}$ (Fig. 1). Following Eq. (16a), one can then estimate $f\{\hat{\boldsymbol{c}}(n-1)\}$ as: $f\{\hat{\boldsymbol{c}}(n-1)\} = y(n) - \hat{e}(n)$.

As an alternative estimation technique, we propose sigma-point Kalman filters [39] such as the unscented Kalman filter (UKF) [40–42] or the central difference Kalman filter (CDKF) [43], explicitly developed for nonlinear estimation problems. The main idea of UKF is the generation of several sampling points, also known as sigma points, around the current state estimate (i.e., the coefficient vector in our case) and the propagation of

these points through the true nonlinearity. This leads to a collection of transformed points that enables accurate estimation of the mean and covariance of the transformed distribution (up to third order under any type of nonlinearity, assuming the prior distribution of the state is Gaussian). The CDKF is very similar to the UKF; however, it approximates the nonlinearity using Stirling's polynomial interpolation. In contrast to the commonly used KF or EKF filters, the sigma-point filters are not limited to Gaussian and linear modeling assumptions and do not require explicit Jacobian or Hessian calculations. They constitute attractive alternatives when analytical expressions of the system dynamics cannot be easily formulated or linearized. Furthermore, their computational complexity is rather moderate and comparable with that of the EKF and the KF-RPLS.

The conventional sigma-point KFs assume nonlinearities in both state and measurement equations. However, in our case (Eq. (14)), the model coefficients, which are the unknown states, follow simple random walks that are described by first-order difference equations. Therefore, herein, we propose the fusion of Rao–Blackwellized UKFs (RB-UKF) [42, 44] with CDKFs. The RB-UKF is essentially a simplified UKF that assumes that the state dynamics are linear and the measurement dynamics are nonlinear, reducing the UKF sampling requirements. Instead of the unscented transformation, we apply Stirling's polynomial interpolation method in the measurement update stage. The latter requires the optimal tuning of only one hyperparameter, whereas the unscented transformation requires the selection of three scaling parameters (related to the spread of the sigma points and prior information about the distribution characteristics). The combination of the two methods is referred to as RB-CDKF.

The conventional RB-UKF can be described by the following equations:

Time update

$$\hat{\boldsymbol{c}}^-(n) = \hat{\boldsymbol{c}}(n-1), \tag{18a}$$

$$\boldsymbol{P}^-(n) = \boldsymbol{P}(n-1) + \boldsymbol{R_1}, \tag{18b}$$

Sigma points generation

$$\boldsymbol{\mathcal{X}}(n) = [\boldsymbol{c}^-(n) \dots \boldsymbol{c}^-(n)] + \gamma[\boldsymbol{0} \quad \sqrt{\boldsymbol{P}^-(n)} \quad -\sqrt{\boldsymbol{P}^-(n)}], \tag{18c}$$

Measurement update

$$\boldsymbol{\mathcal{Y}}(n) = f(\boldsymbol{\mathcal{X}}(n)) = [\mathcal{Y}_0(n) \dots \mathcal{Y}_{2d}(n)], \tag{18d}$$

$$\hat{y}(n) = \boldsymbol{\mathcal{Y}}(n)\boldsymbol{w}_m, \tag{18e}$$

$$P_{yy}(n) = [\boldsymbol{\mathcal{Y}}(n) - \hat{y}(n)]\boldsymbol{W}_c[\boldsymbol{\mathcal{Y}}(n) - \hat{y}(n)]^T + R_2, \tag{18f}$$

$$\boldsymbol{P_{xy}}(n) = \{\boldsymbol{\mathcal{X}}(n) - [\boldsymbol{c}^-(n) \dots \boldsymbol{c}^-(n)]\}\boldsymbol{W}\hat{y}(n), \tag{18g}$$

Fig. 1 Proposed procedure for estimating the nonlinearity $f\{\hat{\mathbf{c}}(n-1)\}$ in the case of the cascade TV-ARMA structure (Eq. 16a). $y(n)$ is fed into the inverse filter $\frac{1}{H(z,n)}$ providing a priori estimate $\hat{e}(n)$. $f\{\hat{\mathbf{c}}(n-1)\}$ can thus be computed as $f\{\hat{\mathbf{c}}(n-1)\} = y(n)-\hat{e}(n)$
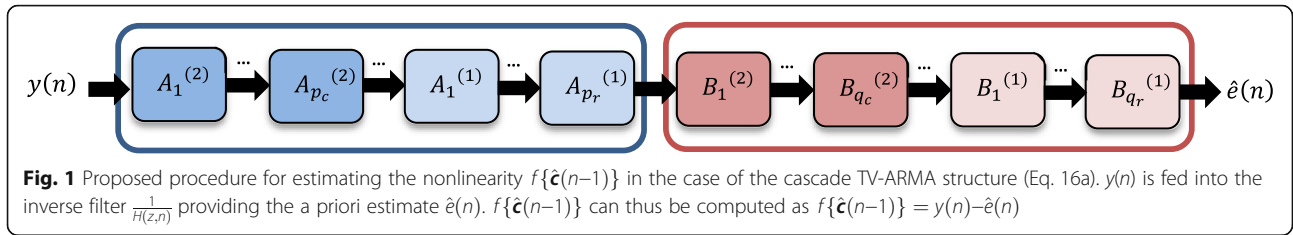
$$K(n) = \frac{P_{xy}(n)}{P_{yy}(n)}, \tag{18h}$$

$$\hat{e}(n) = y(n)-\hat{y}(n), \tag{18i}$$

$$\hat{c}(n) = c^-(n) + K(n)\hat{e}(n), \tag{18j}$$

$$P(n) = P^-(n)-K(n)S_{yy}(n)K^T(n), \tag{18k}$$

where $\mathcal{X}(n) \in \mathbf{R}^{d \times (2d+1)}$ are the $2d+1$ sigma points generated at time point n, γ is a scaling parameter (see Eq. (19 g)), $c^-(n)$ and $P^-(n)$ are the predicted state (i.e., model coefficients) mean and covariance, $\hat{y}(n)$ and $P_{yy}(n)$ are the predicted mean and covariance of the measurement, and $P_{xy}(n)$ is the predicted cross-covariance of the state and the measurement. Note that P is assumed to be positive definite. In order to compute the square root of P, the lower triangular Cholesky factorization is used [45]. In Eq. (18c), the sigma points are generated based on the predicted state mean and covariance. Each sigma point is then propagated through the nonlinearity f (Eq. 18d). The mean and covariance of the measurement output y are approximated using a weighted sample mean and covariance of the posterior sigma points (Eqs. 18e, 18f). The weights w_m and W_c are given as:

$$w_m = [w_{m0}...w_{m2d}]^T \in \mathbf{R}^{(2d+1) \times 1}, \tag{19a}$$

$$w_{m0} = \frac{\delta}{d+\delta}, \tag{19b}$$

$$w_{mj} = \frac{1}{2(d+\delta)}, \quad j = 1,...,2d, \tag{19c}$$

$$W_c = \begin{bmatrix} w_{c0} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{c2d} \end{bmatrix}^T \in \mathbf{R}^{(2d+1) \times (2d+1)}, \tag{19d}$$

$$w_{c0} = \frac{\delta}{d+\delta} + (1-\alpha^2 + \beta), \tag{19e}$$

$$w_{cj} = \frac{1}{2(d+\delta)}, \quad j = 1,...,2d, \tag{19f}$$

$$\gamma = \sqrt{d+\delta}, \tag{19g}$$

$$\delta = \alpha^2(d + \kappa), \quad \alpha > 0, \kappa \geq 0, \tag{19h}$$

where d is the total number of coefficients, and δ is a scaling parameter. The parameters α and κ determine the spread of the sigma points around the state, while β incorporates knowledge regarding the distribution characteristics of the state (i.e., $\beta = 2$ for a Gaussian distribution).

As mentioned earlier, instead of the unscented transformation, we use Stirling's polynomial interpolation since it requires the tuning of only one parameter. The proposed RB-CDKF follows the same structure as the RB-UKF. However, the measurement update Eqs. (18f, 18g) are expressed as:

$$P_{yy}(n) = \sum_{j=1}^{d} w_{cj}^{(1)} \left[\mathcal{Y}_j(n)-\mathcal{Y}_{d+j}(n) \right]^2 + w_{cj}^{(2)} \left[\mathcal{Y}_j(n)+\mathcal{Y}_{d+j}(n)-2\mathcal{Y}_0(n) \right]^2 + R_2, \tag{20a}$$

$$P_{xy}(n) = \sqrt{w_{c1}^{(1)} P^-(n)} \cdot \left[\mathcal{Y}_{1:d}(n)-\mathcal{Y}_{d+1:2d}(n) \right]^T. \tag{20b}$$

The weights w_m, $w_c^{(1)}$, and $w_c^{(2)}$ are given as:

$$w_m = [w_{m0}...w_{m2d}]^T \in \mathbf{R}^{(2d+1) \times 1}, \tag{21a}$$

$$w_{m0} = \frac{\gamma^2-d}{\gamma^2}, \tag{21b}$$

$$w_{mj} = \frac{1}{2\gamma^2}, \quad j = 1,...,2d, \tag{21c}$$

$$w_c^{(1)} = \left[w_{c1}^{(1)}...w_{c2d}^{(1)} \right]^T \in \mathbf{R}^{2d \times 1}, \tag{21d}$$

$$w_{cj}^{(1)} = \frac{1}{4\gamma^4}, \quad for \ j = 1,...,2d, \tag{21e}$$

$$w_c^{(2)} = \left[w_{c1}^{(2)}...w_{c2d}^{(2)} \right]^T \in \mathbf{R}^{2d \times 1}, \tag{21f}$$

$$w_{cj}^{(2)} = \frac{\gamma^2-1}{4\gamma^4}, \quad j = 1,...,2d. \tag{21g}$$

The parameter γ ($\gamma \geq 1$), here, is the central difference interval size and is equal to the square root of the kurtosis of the state's distribution. Assuming a Gaussian

distribution, γ can be optimally set to $\gamma = \sqrt{3}$ [43]. Compared with the RB-UKF, which requires the optimal tuning of three scaling parameters (namely α, β, κ), the RB-CDKF is only dependent on γ. Smoothing may also be applied once the TV coefficients are extracted. We chose the Rauch–Tung–Striebel fixed-interval smoothing algorithm, summarized as [46]:

$$\boldsymbol{G}(n) = \boldsymbol{P}(n)[\boldsymbol{P}^-(n+1)]^{-1}, \tag{22}$$

$$\hat{\boldsymbol{c}}^s(n) = \hat{\boldsymbol{c}}(n) + \boldsymbol{G}(n)[\hat{\boldsymbol{c}}^s(n+1) - \boldsymbol{c}^-(n+1)]. \tag{23}$$

The smoother runs backward starting from time point N with initial conditions $\hat{\boldsymbol{c}}^s(N) = \hat{\boldsymbol{c}}(N)$.

2.2.2 TV-ARMA cascade-form spectrum
Similarly, with the direct-form case, the PSD of the TV-ARMA process $\boldsymbol{y}$ in cascade form can be estimated as:

$$S_y(z,n) = \sigma_e^2 \cdot \frac{\prod_{k=1}^{q_c} |[1-\lambda_k(n)z^{-1}][1-\lambda_k^*(n)z^{-1}]|^2_{z=e^{j2\pi f\Delta T}} \cdot \prod_{k=1}^{q_r} |1-\mu_k(n)z^{-1}|^2_{z=e^{j2\pi f\Delta T}}}{\prod_{k=1}^{p_c} |[1-\rho_k(n)z^{-1}][1-\rho_k^*(n)z^{-1}]|^2_{z=e^{j2\pi f\Delta T}} \cdot \prod_{k=1}^{p_r} |1-\xi_k(n)z^{-1}|^2_{z=e^{j2\pi f\Delta T}}}, \tag{24}$$

where $\Delta T = \frac{1}{f_s}$, f_s is the sampling rate, $f \leq \frac{f_s}{2}$ is the frequency of interest, and σ_e^2 is the variance of the driving noise. Note that the closer a pole or a zero is to the unit circle, the higher its contribution to the total PSD. For poles, this can be translated as more prominent peaks, and for zeros deeper spectral valleys.

2.3 Model order selection and hyperparameter optimization
A particularly crucial step in the estimation of the TV-ARMA models in both direct and cascade form is the model order selection procedure. In the direct-form case, model order is defined as the number of AR and MA coefficients (p, q). In the cascade structure, the model order is linked to the number of poles and zeros of the filter, i.e., (p_r, p_c, q_r, q_c). A high model order leads to overly complex models that overfit the data. In terms of TF representation, this may lead to spurious spectral peaks. On the other hand, an underdetermined model may produce over-smoothed spectrums, lacking detail, and important spectral information. The final TF distribution is also affected by the hyperparameters of the recursive estimators. Suboptimal tuning of the estimators may lead to noisy or biased spectral representations. In this work, we are interested in exploring the capabilities of the models in both forms and compare them fairly under the same grounds.

To achieve this, optimal tuning of the models is necessary in order to extract the maximum possible performance. Grid search or exhaustive search procedures would require days of computations due to the large number of hyperparameters. To this end, we apply mixed integer

genetic algorithms (GA) [47] (ga function from Math-Works MATLAB) to select the best possible combination of model orders and hyperparameters in a fast and efficient manner. In Table 1, we summarize all the variables that have to be optimized for each case. Note that from this point and on, when we refer to KF-RPLS, we automatically assume that the estimation is realized using direct-form TV-ARMA models, whereas KF-RPEM, RB-UKF, and RB-CDKF refer to the cascade TV-ARMA structure. For a fitness function, we used the Akaike information criterion (AIC). The AIC takes into account both the predictive accuracy of the models (affected by both model order and estimator hyperparameters), as well as the model complexity (affected by the model order only). For each candidate solution $\boldsymbol{Cs_j}$, the GA provides a model order and a set of hyperparameters to the estimators and evaluates their performance on all the available data based on the AIC score defined as [48, 49]:

$$AIC(\boldsymbol{Cs_j}) = \frac{N}{2} \log\left(\frac{J}{N}\right) + d, \tag{25}$$

where N is the length of the dataset, d is the total number of model coefficients, and $J = \sum_{n=1}^{N} \hat{e}^2(n)$ is the sum of squares of the a priori prediction error. The optimal solution is the one with the lowest AIC score [50, 51].

3 Results
3.1 Time-invariant ARMA processes
The performance of the algorithms was tested on a time-invariant ARMA (6, 4) process of length 1200 samples, driven by 30 different realizations of zero-mean white noise $(\sigma_e^2 = 1)$. The process had three complex pairs of poles and two complex pairs of zeros. In the frequency domain, this can be translated as three spectral peaks and two spectral valleys. No additive measurement noise was considered. Note here that we selected an ARMA process in order to have a defined ground truth of the exact PSD. This PSD can then be used to compare the estimation capabilities of each algorithm. We investigated three different scenarios:

- Scenario I: The poles and zeros of the process were positioned far apart from each other. The contribution of each root was distinguishable in the PSD (Figs. 2 and 4a).

  $Poles: \rho_1 = -0.9e^{\pm j2.5}, \rho_2 = 0.9e^{\pm j2.5}, \rho_3 = 0.9e^{\pm j1.5}$

  $Zeros: \lambda_1 = -0.9e^{\pm j1.1}, \lambda_2 = 0.9e^{\pm j1.1}$

- Scenario II: The process was narrowband and consisted of closely spaced poles and zeros (Figs. 2 and 4c).

Table 1 Hyperparameters tuned by the GA for each recursive estimator

Model	Recursive estimator	Candidate solution form
TV-ARMA direct-form	KF-RPLS	$[p, q, R_1, R_2, P_0, \hat{\mathbf{c}}^T(0)]$
TV-ARMA cascade-form	KF-RPEM	$[p_r, p_c, q_r, q_c, R_1, R_2, P_0, \hat{\mathbf{c}}^T(0)]$
	RB-UKF	$[p_r, p_c, q_r, q_c, \alpha, \beta, \kappa, R_1, R_2, P_0, \hat{\mathbf{c}}^T(0)]$
	RB-CDKF	$[p_r, p_c, q_r, q_c, \gamma, R_1, R_2, P_0, \hat{\mathbf{c}}^T(0)]$

The upper and lower bounds ([LB, UB]) were set to [0,1] for R_1, R_2, and P_0 [1, 10]; for p and q [1, 5]; for p_c and q_c; [0, 2] for p_r and q_r [1, 2]; for γ; [0, 1] for α; [0, 3] for β; [0, 3] for κ; and [− 1, 1] for each element of $\hat{\mathbf{c}}^T(0)$

$$Poles: \quad \rho_1 = 0.9e^{\pm j1.6}, \rho_2 = 0.9e^{\pm j1.5}, \rho_3 = 0.9e^{\pm j1.4}$$

$$Zeros: \lambda_1 = 0.9e^{\pm j1.1}, \lambda_2 = 0.9e^{\pm j0.8}$$

- Scenario III: The process was narrowband and consisted of closely spaced poles and zeros as in the case of Scenario II; however, the roots were closer to the unit circle (Figs. 2 and 4e)

$$Poles: \quad \rho_1 = 0.98e^{\pm j1.6}, \rho_2 = 0.98e^{\pm j1.5}, \rho_3 = 0.98e^{\pm j1.4}$$

$$Zeros: \quad \lambda_1 = 0.98e^{\pm j1.1}, \lambda_2 = 0.98e^{\pm j0.8}$$
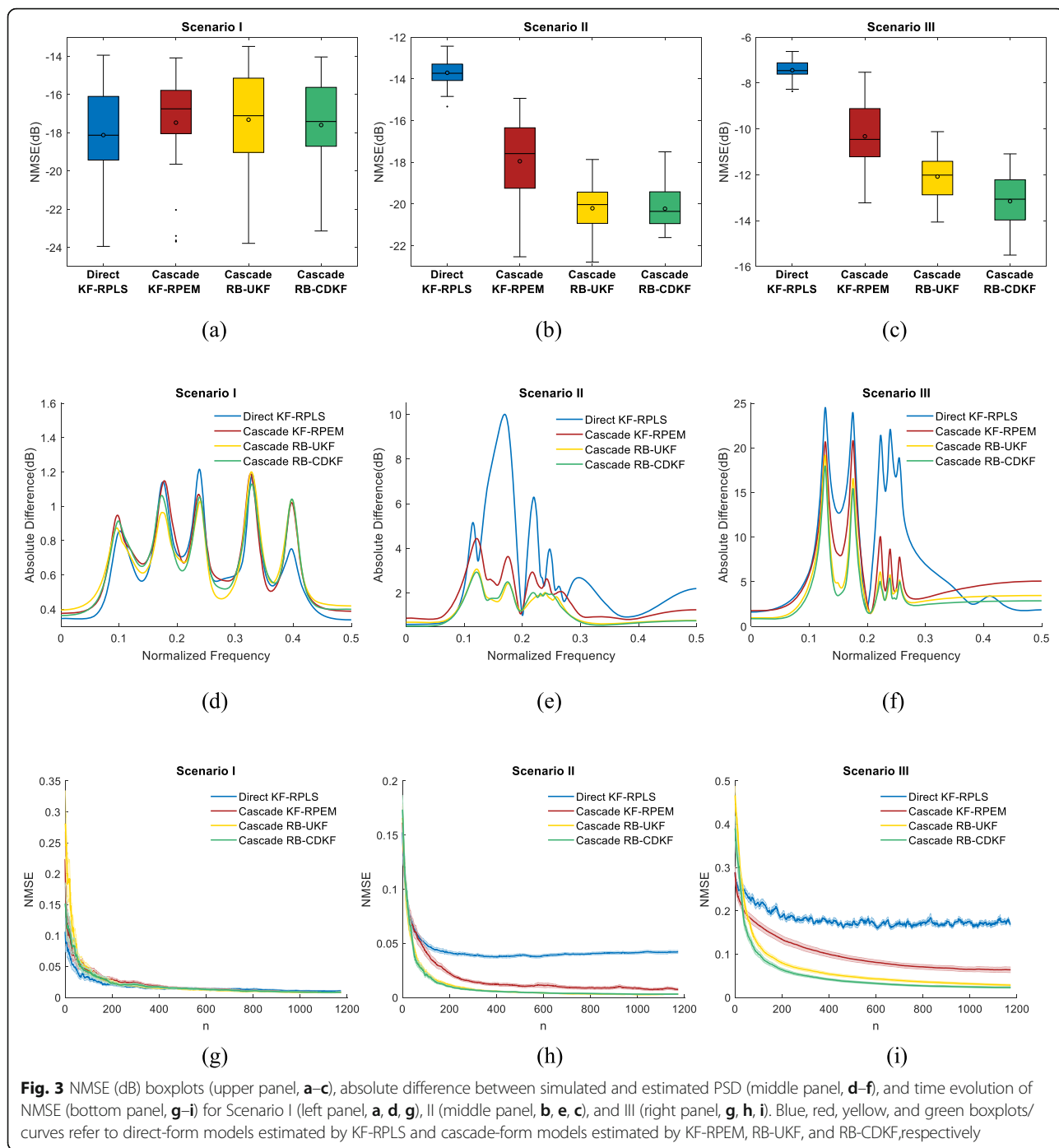
The number of poles and zeros was assumed to be known to avoid possible biases due to erroneous model order selection. During the adaptive process, stability and invertibility were enforced in all estimators, and therefore, all the roots were restricted to the interior of the unit circle. For the cascade structure, unstable roots were divided by their respective conjugate complex [52], which did not affect the magnitude of the final obtained PSD. On the other hand, in the direct-form TV-ARMA models, no stability monitoring/correction was applied due to the increased computational load required to
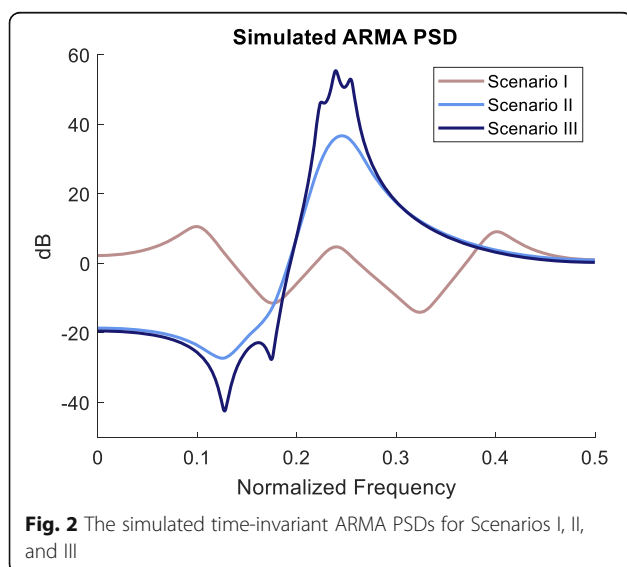
Fig. 2 The simulated time-invariant ARMA PSDs for Scenarios I, II, and III

factorize the polynomials to roots and re-express the roots back to polynomial coefficients. To quantify the overall performance of each algorithm, we used the normalized mean squared error (NMSE) between the simulated and estimated TV spectrum, expressed either with no units or in dB (NMSE(dB) = $10\log_{10}$NMSE).

Figure 3 a–f depicts boxplots of the obtained NMSE values as well as absolute differences between simulated and estimated PSDs for all three scenarios and the different TV-ARMA estimation algorithms. In Scenario I, where the root contributions were distinguishable, both direct and cascade models were able to approximate the simulated PSD accurately (Fig. 3a, d). In Fig. 4, we provide the zero-pole plots created based on all the realizations, using only the final root/coefficient estimates. All methods detected the exact location of the true roots of the system (Fig. 4b). However, in Scenarios II and III, the direct-form models were unable to capture all the PSD components correctly, leading to increased NMSE values (Figs. 3b, c, 4d, f).
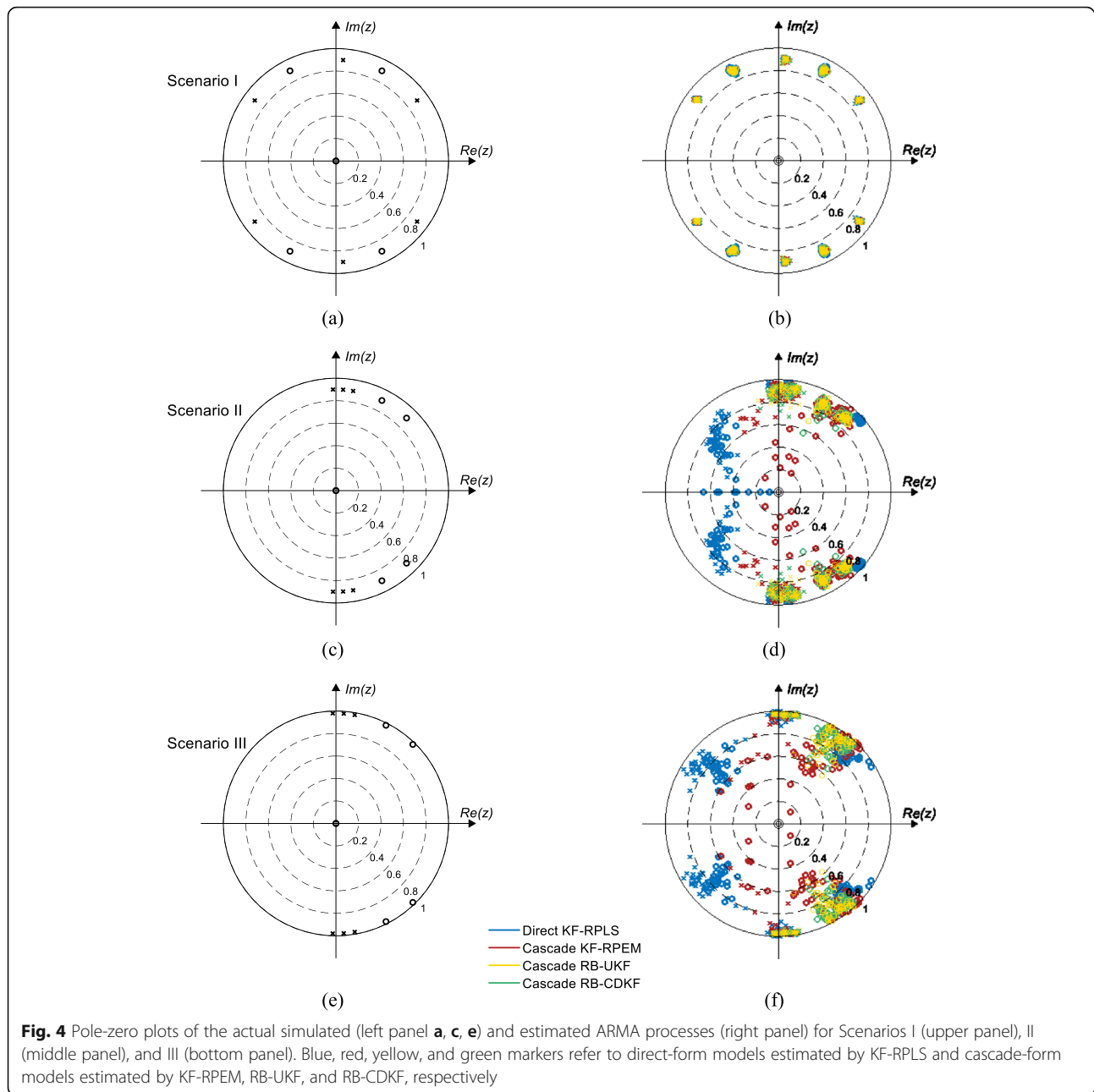
The cascade-form models, on the other side, exhibited superior performance compared with that of the direct-form models in all cases (Fig. 3a–f), and this is in line with previous work [53]. In [53], it was shown that cascade models converge faster, especially when the poles are tightly clustered. The author's main argument was that the condition number of each section of the cascade structure is not influenced by the roots of the other sections, as is the case with the direct form models. Therefore, the convergence rate of the cascade structure is higher compared to the direct form (Fig. 3h, i). Regarding the recursive estimators, the sigma-point KFs provided more accurate estimates than the KF-RPEM for the case of closely spaced poles (Fig. 3b, c, e, f). Optimizing the RB-UKF and RB-CDKF hyperparameters was rather straightforward. One should be aware, though, that the upper bound for the initial covariance matrices should not exceed one and, overall, should be relatively small. Taking into account the fact that the real and imaginary parts of the roots lie between − 1 and 1, the initialization should not force the roots outside the unit circle. Initializing the filters with unstable roots does not guarantee convergence. Optimization of the KF-RPEM hyperparameters, on the other hand, was not trivial for

Fig. 3 NMSE (dB) boxplots (upper panel, **a**–**c**), absolute difference between simulated and estimated PSD (middle panel, **d**–**f**), and time evolution of NMSE (bottom panel, **g**–**i**) for Scenario I (left panel, **a**, **d**, **g**), II (middle panel, **b**, **e**, **c**), and III (right panel, **g**, **h**, **i**). Blue, red, yellow, and green boxplots/curves refer to direct-form models estimated by KF-RPLS and cascade-form models estimated by KF-RPEM, RB-UKF, and RB-CDKF,respectively

Scenarios II and III. Two or three GA repetitions were required, with different initial solutions to obtain a good solution. The KF-RPEM is known to be sensitive to the initial conditions, and the high variance boxplots in Fig. 3 b and c indicate this. Between the two sigma-point filters, the RB-CDKF exhibited improved performance and convergence characteristics compared with the RB-UKF, especially in the case of the closely spaced pole scenarios (Fig. 3c, f, i).
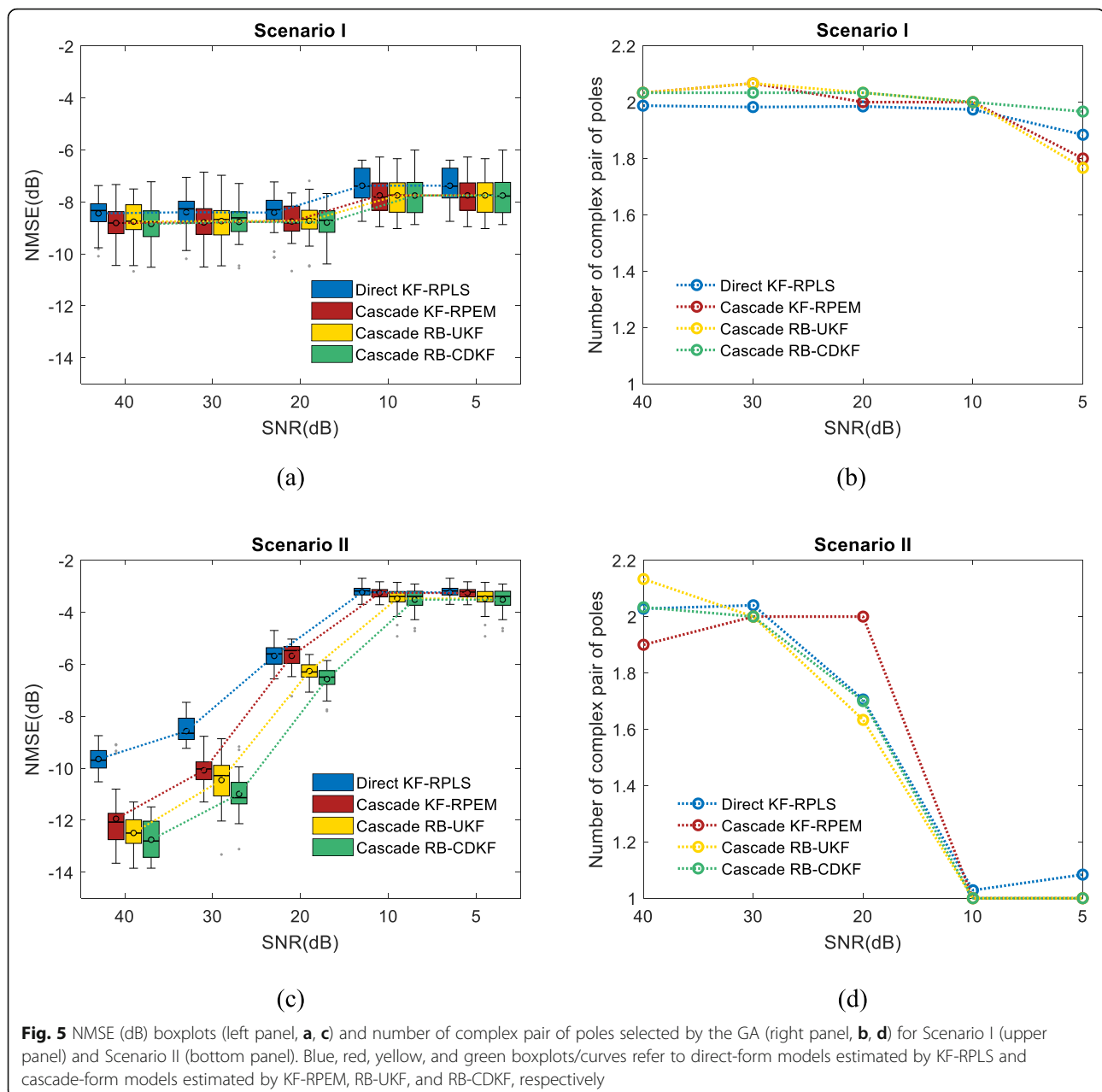
3.2 TV-ARMA processes

In order to investigate the tracking capabilities of the algorithms under noisy and TV environments, we generated 30 realizations of a TV-AR (4) process with additive GWN of 40, 30, 20, 10, and 5 dB signal-to-noise ratios (SNR). The process had two complex pair of poles. An AR (p) process in additive noise can be modeled either by a high-order AR or by an ARMA(p, p) model [54]. Two scenarios were examined:

Fig. 4 Pole-zero plots of the actual simulated (left panel **a**, **c**, **e**) and estimated ARMA processes (right panel) for Scenarios I (upper panel), II (middle panel), and III (bottom panel). Blue, red, yellow, and green markers refer to direct-form models estimated by KF-RPLS and cascade-form models estimated by KF-RPEM, RB-UKF, and RB-CDKF, respectively

- *Scenario I:* One pole remained constant throughout time, while the other varied sinusoidally (in terms of their angle). The poles were placed further apart from each other (Fig. 6a).
- *Scenario II:* The two poles of the process varied sinusoidally (in terms of their angle) and in close proximity to one another (Fig. 6d).

Herein, the model order/number of roots was not assumed to be known a priori and was optimized by the GA, along with all the rest of the hyperparameters (Table 1). The obtained NMSE values for both scenarios

can be seen in Fig. 5. In Scenario I, as in the time-invariant case, all algorithms performed equivalently, and no significant changes were observed with increased SNR (Fig. 5a). The GA was able to select the correct number of roots in almost all cases (Fig. 5b). On the other hand, in Scenario II, we observe the same pattern again as in the time-invariant case. The direct-form models faced difficulties with the closely spaced poles, resulting in a degraded performance. The sigma-point KFs provided more accurate estimates overall, with RB-CDKF being slightly better than the RB-UKF (Fig. 5c). We expect that in more complex processes, the

Fig. 5 NMSE (dB) boxplots (left panel, **a**, **c**) and number of complex pair of poles selected by the GA (right panel, **b**, **d**) for Scenario I (upper panel) and Scenario II (bottom panel). Blue, red, yellow, and green boxplots/curves refer to direct-form models estimated by KF-RPLS and cascade-form models estimated by KF-RPEM, RB-UKF, and RB-CDKF, respectively

differences between the algorithms will be more pronounced. Based on Fig. 5d, the two closely spaced TV peaks were not easily distinguishable for low SNR levels (5 and 10 dB). Thus, the GA was able to resolve only one spectral peak. Figure 6 illustrates the simulated and estimated TV PSDs (based on cascade-form models and RB-CDKF), averaged over all realizations and for SNR of 40 dB. Smoothed and non-smoothed versions of the TV PSDs are also provided. The average runtime over all 30 realizations was 4.54 ± 0.31, 4.63 ± 0.35, 6.33 ± 0.17, and 6.26 ± 0.20 s for KF-RPLS, KFR-RPEM, RB-UKF, and RB-CDKF, respectively (on an Intel Core™ i7-6700 @ 3.4 GHz, 32 GB using parallel computing and MATLAB MEX functions).

4 Applications in medical ultrasound imaging

One possibly interesting application of the described TV-ARMA methodology is the investigation of the propagation characteristics of ultrasonic waves in medical ultrasound (US) imaging. US imaging is a non-invasive technique that produces real time images of the human body anatomy. During a US scan, short acoustic pulses are emitted towards targeted areas, giving rise to reflections from internal structures. The echoes that return to the ultrasonic transducer are recorded and then combined to produce 2D images of the interrogated area. It is well known that the emitted pulse undergoes gradual waveform distortions and amplitude fluctuations as it propagates through the tissue.

Fig. 6 Simulated (**a**, **d**) and estimated (**b**, **c**, **e**, **f**) TV PSDs (averaged over all 30 realizations) for Scenario I (upper panel) and Scenario II (bottom panel), under 40 dB SNR and either with (right panel) or without (middle panel) smoothing. The estimation technique used was RB-CDKF combined with the cascade-form ARMA models
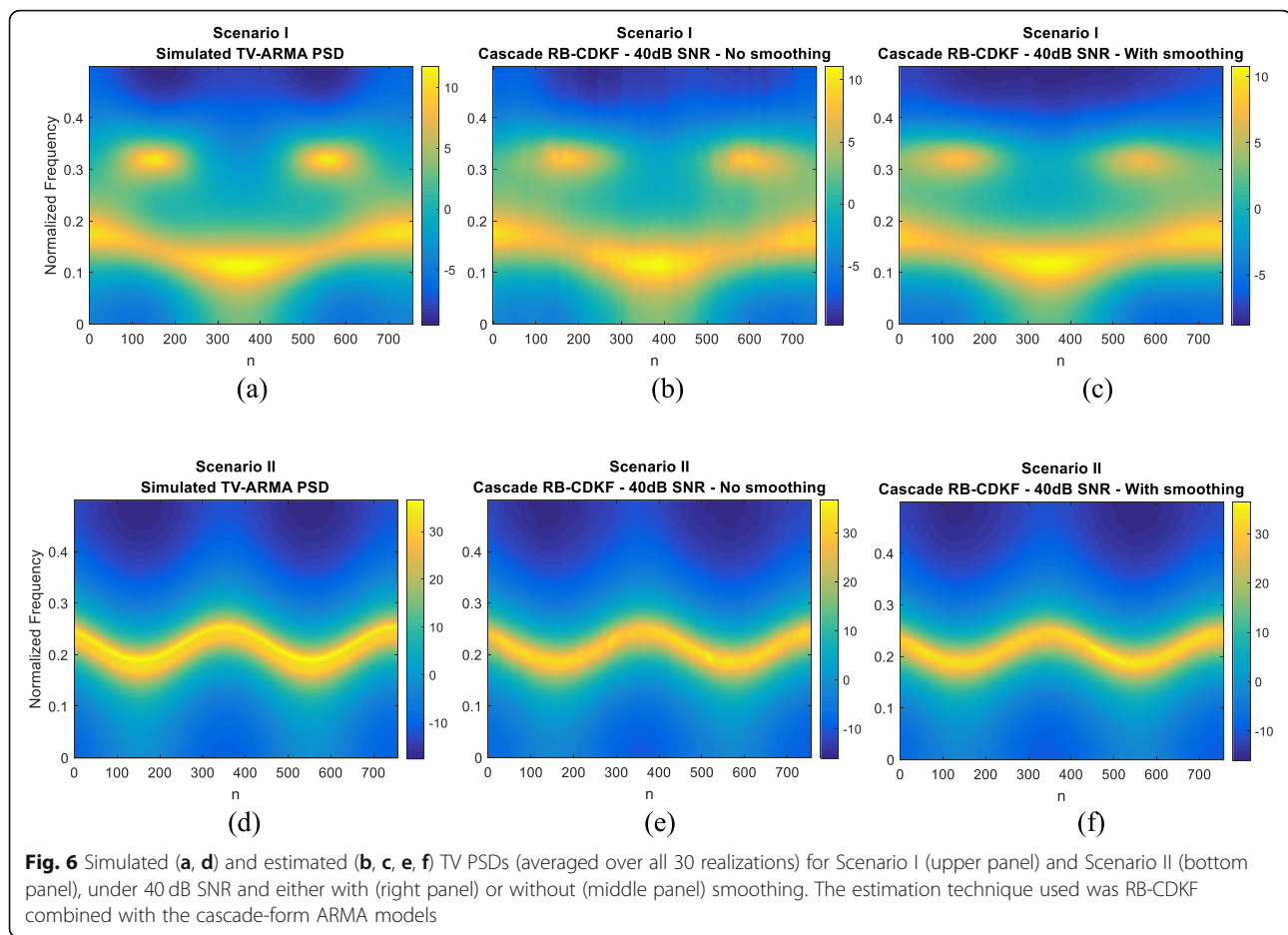
Our main aim is to use the proposed TV methodology to detect these changes in time. This may, first of all, give further insight into the underlying mechanisms of pulse propagation. Second, information extracted by the estimated TV models can also be used for segmentation, detection, classification (e.g., breast, prostate, lymph node lesion classification), and even tissue characterization. However, herein, we focus only on extracting the TV spectral characteristics of the ultrasonic pulse in order to validate the method described in this paper. The main reason for selecting this specific application is the nature of the problem and its close relation to the ARMA models. It has been shown that the echoes that are reflected back to the transducer can be modeled as ARMA processes [55], i.e., the output of an IIR filter driven by noise. The filter, in this case, is the TV ultrasonic pulse, and the noise is the so-called reflectivity function or else the strength of the acoustic reflection and scattering of the tissue as a function of its spatial coordinates. The recorded echoes, therefore, are the result of the convolution between the TV pulse and the underlying reflectivity signal. The tissue reflectivity function along a US image line can be grossly

described as a Gaussian–Bernoulli sequence; hence, it fits the description of the driving process noise.

For this study, we used the open-access database of raw ultrasonic signals acquired from malignant and benign breast lesions [56]. A representative example of an US image and a raw US signal can be found in Fig. 7. We applied our proposed methodology, and specifically, the cascade-form ARMA models along with the RB-CDKF estimator, on signals extracted from different lines of one of the US images of the database. Here, we present results obtained from one image line. The number of poles and zeros, as well as all the estimator hyper-parameters, was optimized using the GA. For this specific example, the GA selected an optimal number of ARMA roots, three complex pairs of poles and three complex pairs of zeros. To validate this selection, we manually increased the number of complex roots from one to six and optimized the rest of the hyperparameters with the help of the GA. Then, we compared the AIC values as well as the NMSE between the original and the predicted (a posteriori) time series, for different number of roots. The results are presented in Fig. 8. For both

Fig. 7 A representative US image [56] and a raw US signal extracted from one of the image lines. The US image is formed using the log-compressed envelope of the raw US signals of each line
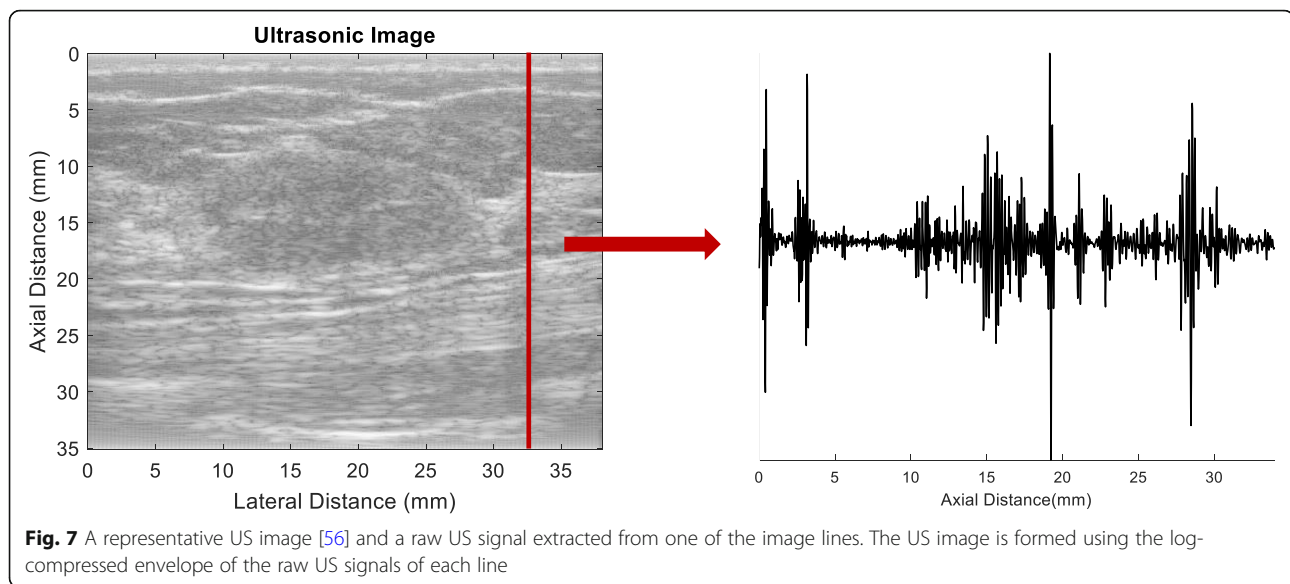
AIC and NMSE, the minimum was achieved, indeed, for three complex poles and three complex zeros.

The obtained TV-ARMA PSD, along with the time evolution of the instantaneous frequency estimates and the radiuses of the complex poles, is depicted in Figs. 9 a and 10 respectively. For comparison purposes, we also estimated the STFT, the CWT, and the smoothed pseudo-WVD (Fig. 9b–d). Compared with the rest of the methods, the TV-ARMA PSD was smoother and less noisy. This was expected since the used models take into account the stochastic nature of the signal. In the obtained PSDs, one can see a gradual, almost linear shift of the spectral characteristics of the pulse towards lower frequencies, which is in accordance with what has been

Fig. 8 AIC score and NMSE between original and predicted (a posteriori) time series for an increasing number of complex pair of poles and zeros. The smallest AIC and NMSE value was achieved for three complex poles and three complex zeros

previously observed [57]. During the propagation of the pulse through the tissue, high-frequency components undergo higher attenuation than the low-frequency components, giving rise to this spectral downshift observed with depth.

The fundamental frequency of the pulse was represented by the complex pole with the largest radius value and, as seen in Fig. 10 b, was pole ρ_1. Its contribution remained dominant up to almost 20 mm and then started to subside. The same pattern was observed for the ρ_3 pole, which probably describes a second harmonic. On the other hand, the spectral peak represented by the pole ρ_2 became more prominent with depth. Of course, one may select a larger number of poles and observe their time evolution, however, based on the model order selection procedure, three poles were adequate to describe the basic time variations of the process. Interpreting the exact mechanisms of pulse propagation is beyond the scope of this paper. Nonetheless, this section presents one possible direction for future work.
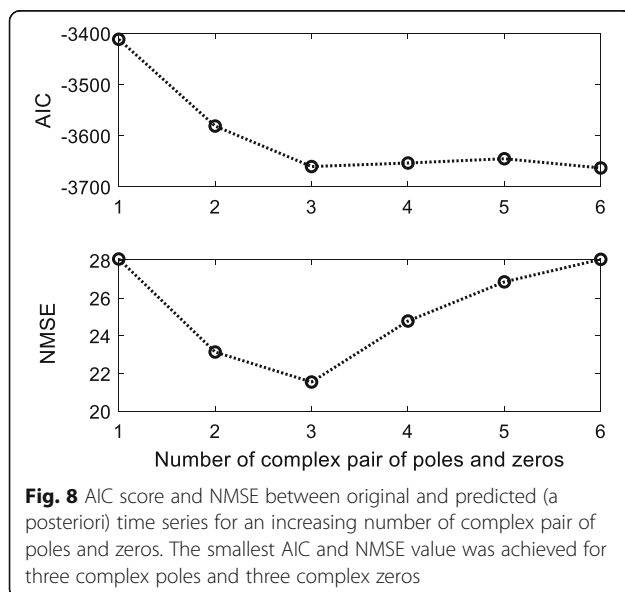
5 Discussion and conclusions

This work has provided a complete framework for estimating TV-ARMA processes and other types of nonstationary processes that are not necessarily stochastic in nature. The presented framework is based on the notion of root tracking, a powerful technique that provides insight into the core mechanisms of a process. Systems are usually identified by estimating the coefficients associated with the characteristic polynomials of their transfer functions. However, all the relevant information concerning the system dynamics is encompassed within the roots of these polynomials. In addition, transfer function coefficients can either become overly sensitive

TV-ARMA PSD

Spectrogram

Continuous Wavelet Transform

Smoothed Pseudo Wigner-Ville Distribution

(a)

(b)

(c)

(d)

Fig. 9 **a** Estimated TV-ARMA PSD. **b** STFT spectrogram using a window of 150 samples and an overlap of 149 samples. **c** CWT using the Bump wavelet. **d** Smoothed pseudo-WVD

(a)

(b)

Fig. 10 **a** Time evolution of the instantaneous frequency estimates that correspond to the three identified complex poles. **b** Time evolution of the radius of each complex pole

to numerical round-off errors or produce temporal instabilities that are computationally expensive to control and correct. To this end, the direct-form ARMA models are reparameterized in terms of their roots and expressed as a cascade of first- and second-order sections. Each section is therefore related to one of the roots of the process, allowing independent tracking and robust stability monitoring with minimal computational effort.

Tracking the roots of the system requires the use of recursive estimation techniques. However, classical methods, such as the KF or the RLS, cannot be directly applied since the cascaded ARMA structure is nonlinear in its coefficients. The most commonly used adaptive method is a type of gradient-based RLS algorithm. Herein, we explored, for the first time to our knowledge, the capabilities of sigma-point KFs. Sigma-point filters are gradient-free estimators that apply deterministic sampling approaches to deal with nonlinearities in the system. We combined the UKF and the CDKF in one algorithm and compared its performance with the conventional gradient-based technique.

Based on simulations, we concluded that sigma-point filters are less sensitive to the initial conditions and more easily tuneable. We also observed that when the process consists of tightly clustered roots, sigma-point filters provide more accurate results and may converge faster. To ease the identification procedure, instead of tuning in an ad hoc manner, all the relevant hyperparameters or resorting to exhaustive/grid search methods, we allowed a GA to optimize the models. Of course, the bottleneck in terms of speed compared with other methods is the GA itself (4 to 5 s for a signal of length 1200 and a maximum of 5 desired roots, on an Intel Core™ i7-6700 at 3.4 GHz, 32 GB using parallel computing, and MATLAB MEX functions); however, this gave us the opportunity to explore in a greater extent the capabilities of both direct- and cascade-form models, as well as the estimation performance of the different adaptive filtering techniques. For the sake of demonstration, the proposed framework was applied to medical ultrasound images to explore the TV characteristics of raw ultrasonic signals. Future work will revolve around minimizing the computational complexity related to the optimization of the models, as well as developing data-driven tuning algorithms. We will extend our research to methods that are robust to different types of noise (e.g., impulsive, heteroskedastic, or heavy-tailed noise). Once the algorithmic development is completed, the next challenging and interesting step is the implementation of the root tracking framework in the hardware. This will require refinements in order to balance the need for increased estimation accuracy with efficient hardware design.

Authors' contributions
KK and ML declare that they contributed to the manuscript. Both authors read and approved the final manuscript.

References
1. L. Cohen, Time-frequency analysis, vol. 778. Prentice hall, 1995
2. F.B. Vialatte et al., A machine learning approach to the analysis of time-frequency maps, and its application to neural dynamics. Neural Networks **20**(2), 194–209 (2007)
3. H.U. Amin et al., Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australasian physical & engineering sciences in medicine **38**(1), 139–149 (2015)
4. R. He et al., Automatic detection of atrial fibrillation based on continuous wavelet transform and 2D convolutional neural networks. Frontiers in physiology **9**, 2016 (2018)
5. B. Boashash, N.A. Khan, T. Ben-Jabeur, Time-frequency features for pattern recognition using high-resolution TFDs: a tutorial review. Digital Signal Processing **40**, 1–30 (2015)
6. B. Boashash, Estimating and interpreting the instantaneous frequency of a signal-part 2. Proceedings of the IEEE **80**(4), 540–568 (1992)
7. M.R. Portnoff, Time-frequency representation of digital signals. IEEE Transactions on Acoustics, Speech, and Signal Processing **28**(1), 55–69 (1980)
8. S. Farkash, S. Raz, Time-variant filtering via the Gabor expansion. Signal processing V: Theories and applications, 509–512 (1990)
9. S. Mallat, *A wavelet tour of signal processing* (Elsevier, 1999)
10. I. Daubechies, T. Paul, Time-frequency localisation operators-a geometric phase space approach: II. The use of dilations. Inverse Probl **4**(3), 661 (1988)
11. G. Wang, Z. Luo, X. Qin, Y. Leng, T. Wang, Fault identification and classification of rolling element bearing based on time-varying autoregressive spectrum. Mechanical Systems and Signal Processing **22**(4), 934–947 (2008)

12. A.M. Bianchi, L.T. Mainardi, S. Cerutti, Time-frequency analysis of biomedical signals. Transactions of the Institute of Measurement and Control **22**(3), 215–230 (2000)

13. Y. Grenier, Time-dependent ARMA modeling of non-stationary signals. IEEE Transactions on Acoustics, Speech, and Signal Processing **31**(4), 899–911 (1983)

14. L. Ljung and T. Söderström, Theory and practice of recursive identification. MIT press, 1983

15. L. Ljung, System identification: theory for the user. Englewood Cliffs, 1987

16. P.R. Scalassara et al., Autoregressive decomposition and pole tracking applied to vocal fold nodule signals. Pattern recognition letters **28**(11), 1360–1367 (2007)

17. N. Ouaaline, L. Radouane, Pole-zero estimation of speech signal based on zero-tracking algorithm. International Journal of Adaptive Control and Signal Processing **12**(1), 1–12 (1998)

18. L. Patomaki, J.P. Kaipio, P.A. Karjalainen, Tracking of nonstationary EEG with the roots of ARMA models. Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society **2**, 877–878 (1995)

19. L.T. Mainardi, A.M. Bianchi, G. Baselli, S. Cerutti, Pole-tracking algorithms for the extraction of time-variant heart rate variability spectral parameters. IEEE Transactions on Biomedical Engineering **42**(3), 250–259 (1995)

20. S. Cazares, M. Moulden, C.W.G. Redman, L. Tarassenko, Tracking poles with an autoregressive model: a confidence index for the analysis of the intrapartum cardiotocogram. Medical engineering & Physics **23**(9), 603–614 (2001)

21. S. Thanagasundram, S. Spurgeon, F. Soares Schlindwein, A fault detection tool using analysis from an autoregressive model pole trajectory. Journal of Sound and Vibration **317**(3–5), 975–993 (2008)

22. Y. Lee, Channel prediction with cascade AR modeling. In Advanced Int'l Conference on Telecommunications and Int'l Conference on Internet and Web Applications and Services, 40-40 (2006)

23. H. Wold, A study in the analysis of stationary time series. PhD diss., Almqvist & Wiksell, 1938

24. S. S. Haykin, Adaptive filter theory. Pearson Education India, 2005

25. H. Cramér, On some classes of nonstationary stochastic processes. Proceedings of the Fourth Berkeley symposium on mathematical statistics and probability **2**, 57–78 (1961)

26. H.E. Rauch, C.T. Striebel, F. Tung, Maximum likelihood estimates of linear dynamic systems. AIAA J. **3**(8), 1445–1450 (1965)

27. M.P. Tarvainen, J.K. Hiltunen, P.O. Ranta-aho, P.A. Karjalainen, Estimation of nonstationary EEG with Kalman smoother approach: an application to event-related synchronization (ERS). IEEE Transactions on Biomedical Engineering **51**(3), 516–524 (2004)

28. M.G. Hall, A.V. Oppenheim, A.S. Willsky, Time-varying parametric modeling of speech. Signal Processing **5**(3), 267–285 (1983)

29. M. Juntunen, J. Tervo, and J. P. Kaipio, Stabilization of stationary and time-varying autoregressive models. in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing **4**, 2173–2176 (1998)

30. P. Guillaume, J. Schoukens, R. Pintelon, Sensitivity of roots to errors in the coefficient of polynomials obtained by frequency-domain estimation methods. IEEE Transactions on Instrumentation and Measurement **38**(6), 1050–1056 (1989)

31. L.B. Jackson, S.L. Wood, Linear prediction in cascade form. IEEE Transactions on Acoustics, Speech, and Signal Processing **26**(6), 518–528 (1978)

32. M. Nayeri, W.K. Jenkins, Alternate realizations to adaptive IIR filters and properties of their performance surfaces. IEEE Transactions on Circuits and Systems **36**(4), 485–496 (1989)

33. A. Nehorai, D. Starer, Adaptive pole estimation. IEEE Transactions on Acoustics, Speech, and Signal Processing **38**(5), 825–838 (1990)

34. B. D. Rao, Adaptive IIR filtering using cascade structures. in Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, 194-198 (1993)

35. A. Al Zaman, X. Luo, M. Ferdjallah, and A. Khamayseh, A new TVAR modeling in cascaded form for nonstationary signals. in 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings **3**, III-III (2006)

36. Y.H. Tam, P.C. Ching, Y.T. Chan, Adaptive recursive filters in cascade form. IEE Proceedings F (Communications, Radar and Signal Processing) **134**(3), 245–252 (1987)

37. L. Ljung, Asymptotic behaviour of extended Kalman filter as a parameter estimator for linear systems. IEEE Transactions on Automatic Control **24**(1), 36–50 (1979)

38. J.B. Moore, H. Weiss, Recursive prediction error methods for adaptive estimation. IEEE Transactions on Systems, Man, and Cybernetics **9**(4), 197–205 (1979)

39. R. Van Der Merwe, E. Wan, and S. Julier, Sigma-point Kalman filters for nonlinear estimation and sensor-fusion: applications to integrated navigation. in AIAA Guidance, Navigation, and Control Conference and Exhibit, 5120 (2004)

40. S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, New approach for filtering nonlinear systems. in Proceedings of 1995 American Control Conference-ACC'95, **3**, 1628-1632 (1995)

41. S. Haykin, Kalman filtering and neural networks, vol. 5, no. 3. 2001.

42. E.A. Wan, R. van der Merwe, The unscented Kalman filter. Kalman Filter. Neural Networks, 221–280 (2003)

43. M. Nørgaard, N.K. Poulsen, O. Ravn, New developments in state estimation for nonlinear systems. Automatica **36**(11), 1627–1638 (2000)

44. M. Briers, S. R. Maskell, and R. Wright. A Rao-Blackwellised unscented Kalman filter. in Proceedings of the Sixth International Conference of Information Fusion, **1**, 55-61 (2003)

45. A. Iserles, Matrix computations. by GH Golub and CF Van Loan. Pp 642.£ 38. 1989. ISBN 0-8018-3772-3 (John Hopkins Press)." The Mathematical Gazette **74** (469), 322-324, (1990)

46. S. Särkkä, Unscented Rauch-Tung-Striebel smoother. IEEE Transactions on Automatic Control **53**(3), 845–849 (2008)

47. K. Deep, K.P. Singh, M.L. Kansal, C. Mohan, A real coded genetic algorithm for solving integer and mixed integer optimization problems. Applied Mathematics and Computation **212**(2), 505–518 (2009)

48. S. De Waele, P.M.T. Broersen, Order selection for vector autoregressive models. IEEE Transactions on Signal Processing **51**(2), 427–433 (2003)

49. L. Faes, S. Erla, and G. Nollo, Measuring connectivity in linear multivariate processes: definitions, interpretation, and practical analysis. Computational and Mathematical Methods in Medicine **2012**, (2012)

50. K. Kostoglou, G.D. Mitsis, Modelling of multiple-input, time-varying systems with recursively estimated basis expansions. Signal Processing **155**, 287–300 (2019)

51. K. Kostoglou, A.D. Robertson, B. MacIntosh, G.D. Mitsis, A novel framework for estimating time-varying multivariate autoregressive models and application to cardiovascular responses to acute exercise. IEEE Transactions on Biomedical Engineering **66**(11), 3257–3266 (2019)

52. J. Pardey, S. Roberts, L. Tarassenko, A review of parametric modelling techniques for EEG analysis. Medical Engineering & Physics **18**(1), 2–11 (1996)

53. G. Zakaria and A. A. Beex, Relative convergence of the cascade RLS with subsection adaptation algorithm. in Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat. No. CH37020), **1**, 810–814 (1999)

54. S. M. Kay, Modern spectral estimation. Pearson Education India, 1988.

55. J.A. Jensen, Deconvolution of ultrasound images. Ultrason. Imaging **14**(1), 1–15 (1992)

56. H. Piotrzkowska-Wróblewska et al., Open access database of raw ultrasonic signals acquired from malignant and benign breast lesions. Medical Physics **44**(11), 6105–6109 (2017)https://doi.org/10.5281/zenodo.545928.

57. P.A. Narayana, J. Ophir, Spectral shifts of ultrasonic propagation: a study of theoretical and experimental models. Ultrasonic Imaging **5**(1), 22–29 (1983)

Publisher's Note