

# A Prediction-Error Covariance Estimator for Adaptive Kalman Filtering in Step-Varying Processes: Application to Power-System State Estimation

Lorenzo Zanni, *Member, IEEE*, Jean-Yves Le Boudec, *Fellow, IEEE*, Rachid Cherkaoui, *Senior Member, IEEE*, and Mario Paolone, *Senior Member, IEEE*

**Abstract**—In this paper, we present a new method for the estimation of the prediction-error covariances of a Kalman filter (KF), which is suitable for step-varying processes. The method uses a series of past innovations (i.e., the difference between the upcoming measurement set and the KF predicted state) to estimate the prediction-error covariance matrix by means of a constrained convex optimization problem. The latter is designed to ensure the symmetry and the positive semidefiniteness of the estimated covariance matrix, so that the KF numerical stability is guaranteed. Our proposed method is straightforward to implement and requires the setting of one parameter only, i.e., the number of past innovations to be considered. It relies on the knowledge of a linear and stationary measurement model. The ability of the method to track state step-variations is validated in ideal conditions for a random-walk process model and for the case of power-system state estimation. The proposed approach is also compared with other methods that estimate the KF stochastic parameters and with the well-known linear weighted least squares. The comparison is given in terms of both accuracy and computational time.

**Index Terms**—Adaptive Kalman filter (AKF), covariance estimation, phasor measurement unit (PMU), power systems, state estimation, step processes.

## I. INTRODUCTION

A WELL-KNOWN challenge in the application of the Kalman filter (KF) to real systems is the identification of the parameters of the stochastic error distributions. Indeed, the correctness of the KF is subject to their proper assessment.

Let us consider a system described by a process model consisting in the following linear time-variant discrete-time equation ( $k$  is the discrete-time index):

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{w}_k \quad (1)$$

Manuscript received May 14, 2016; revised August 31, 2016; accepted November 2, 2016. Manuscript received in final form November 9, 2016. This work was supported in part by the Swiss Competence Center for Energy Research (FURIES) and in part by the NanoTera Swiss National Science Foundation Project S3-Grids. Recommended by Associate Editor A. Alessandri.

The authors are with the Swiss Federal Institute of Technology of Lausanne, 1015 Lausanne, Switzerland (e-mail: lorenzo.zanni@epfl.ch; jean-yves.leboudec@epfl.ch; rachid.cherkaoui@epfl.ch; mario.paolone@epfl.ch).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCST.2016.2628716

in which  $\mathbf{x}_k \in \mathbb{R}^n$  is the system state and  $\mathbf{w}_k \in \mathbb{R}^n$  is the process noise that is assumed to be a white Gaussian sequence with covariance matrix  $\mathbf{Q}_k$ . This autoregressive integrated moving average (ARIMA) (0,1,0) model, adopted for instance in [1], is particularly suitable in case high-resolution measurements are available as the state does not change substantially from one time-step to the other. It also assumes there are no controllable inputs. The process noise should be able to account for occasional steps. To this end, we enable the covariance matrix  $\mathbf{Q}_k$  to be time-varying and unknown. In contrast, the measurement model is assumed to be linear, known, and time-invariant. Hence, it can be defined as

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k \quad (2)$$

in which  $\mathbf{z}_k \in \mathbb{R}^m$  is the measurement set,  $\mathbf{H} \in \mathbb{R}^{m \times n}$  is the measurement matrix, and  $\mathbf{v}_k \in \mathbb{R}^m$  is the measurement noise that is assumed to be a white Gaussian sequence with covariance matrix  $\mathbf{R}$ . We also assume that the system is observable, i.e.,  $\mathbf{H}$  is of full rank.

Given the process and measurement model (1) and (2), the well-known linear KF equations are as follows:

1) *Prediction*:

$$\widehat{\mathbf{x}}_{k|k-1} = \widehat{\mathbf{x}}_{k-1|k-1} \quad (3)$$

$$\mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-1} + \mathbf{Q}_k. \quad (4)$$

2) *Measurement Update*:

$$\mathbf{L}_k = \mathbf{P}_{k|k-1}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R})^{-1} \quad (5)$$

$$\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{z}_k - \mathbf{H}\widehat{\mathbf{x}}_{k|k-1}) \quad (6)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{H})\mathbf{P}_{k|k-1} \quad (7)$$

in which  $\widehat{\mathbf{x}}_{k|k-1}$  is the predicted state given the knowledge of the process prior to time-step  $k$ ,  $\mathbf{L}_k$  is the Kalman gain,  $\widehat{\mathbf{x}}_{k|k}$  is the estimated state given  $\mathbf{z}_k$ , and  $\mathbf{I}$  is the identity matrix.  $\mathbf{P}_{k|k-1}$  and  $\mathbf{P}_{k|k}$  are the covariance matrices of the prediction and estimation errors  $\varepsilon_{k|k-1}$  and  $\varepsilon_{k|k}$ , respectively

$$\varepsilon_{k|k-1} = \mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1} \quad (8)$$

$$\mathbf{P}_{k|k-1} = E[\varepsilon_{k|k-1}\varepsilon_{k|k-1}^T] \quad (9)$$

$$\varepsilon_{k|k} = \mathbf{x}_k - \widehat{\mathbf{x}}_{k|k} \quad (10)$$

$$\mathbf{P}_{k|k} = E[\varepsilon_{k|k}\varepsilon_{k|k}^T] \quad (11)$$

where  $E$  is the expected value operator. After the prediction step, we can also define the innovation  $\mathbf{y}_k$  and its covariance matrix  $\mathbf{S}_k$  as

$$\mathbf{y}_k = \mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1} \quad (12)$$

$$\mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R}. \quad (13)$$

As it is known, the KF estimate is optimal if we use the optimal Kalman gain  $\mathbf{L}$  that depends on matrices  $\mathbf{P}_{k|k-1}$ ,  $\mathbf{H}$ , and  $\mathbf{R}$ . The matrix  $\mathbf{P}_{k|k-1}$  contains the uncertainty of the process noise  $\mathbf{Q}$ . Assuming the knowledge of  $\mathbf{H}$ , the literature dealing with KF-parameters estimation proposes several methods to infer the correct  $\mathbf{R}$  and  $\mathbf{Q}$  (a detailed literature review on the subject is presented in Section II). Most of these methods are suitable for time-invariant or slow-varying systems. Indeed, the KF was conceived to filter the measurement noise of linear systems by modeling their time behavior. However, some systems are characterized by fast state variations, such as unpredictable and sudden state steps. In these cases, given a certain process model, the Kalman gain has to be updated. To this end, Myers and Tapley [2] use a direct estimation of both  $\mathbf{R}$  and  $\mathbf{Q}$  from their sample covariances at every sampling step. Recently, Zanni *et al.* [3] presented a heuristic method that, given the knowledge of  $\mathbf{R}$ , assesses  $\mathbf{Q}$  from the state estimates. The tests are carried out using a power system in quasi-steady-state conditions and show the better performance of the filter compared with the linear weighted least-squares (LWLS) algorithm. This technique is further tested and validated in [4]. However, this method is not proved to be optimal as it is heuristic, and state step-variations are not considered. The aforementioned methods are unable to track a state step-variation as soon as it occurs, because they update the KF parameters based on the state estimates.

In this paper, we present an adaptive KF (AKF) that filters effectively the measurement noise during quasi-steady-state conditions and is also able to track occasional state step-variations. We consider a process that can be modeled as (1) with constant or slow-varying  $\mathbf{Q}$  and that exhibits state step-variations that violate the process model. Assuming  $\mathbf{H}$  and  $\mathbf{R}$  are known and time-invariant, the goal is to keep tracking the state in spite of these violations. We exploit the knowledge of the sample covariance matrix of the past and current innovations (12) in order to estimate the prediction-error covariance matrix  $\mathbf{P}_{k|k-1}$ . Hence, we automatically update the Kalman gain before the measurement update. Our proposed method for inferring  $\mathbf{P}_{k|k-1}$  is called *prediction-error covariance estimation* (PECE), and it consists in a constrained convex optimization problem based on maximum-likelihood (ML) estimation that ensures the positive semidefiniteness of  $\mathbf{P}_{k|k-1}$ . The accuracy of the PECE method is demonstrated, considering power-system state estimation (SE) as application example. State step-variations are common in power systems as they are associated, for instance, with the connection/disconnection of different components (mainly loads, generators, and transmission lines). It is important to note that our proposed method can be used for any other application, as long as the Assumptions 1–3 introduced in Section III hold. We also provide

a comparison with the methods presented in [2] and [3], and with the LWLS widely used in power-system SE.

This paper is structured as follows. Section II consists of a literature review on the KF-covariance-estimation methods that have been proposed since the end of the 1960s. In Section III, we focus on the description of the proposed PECE method and provide the accuracy assessment for ideal cases. In Section IV, we apply the PECE-based AKF to a power-system SE example. In Section V, we present a computational-time assessment. In Sections III–V, we compare the PECE method with other two methods presented in [2] and [3]. The conclusions are given in Section VI.

## II. LITERATURE REVIEW

In the early 1970s, Mehra [5] classified into four categories the methods for the estimation of the KF covariances: Bayesian [6], [7], ML [8]–[10], correlation [11]–[20], and covariance matching [2], [3].

The Bayesian and ML estimation methods are characterized by a high level of complexity and are usually employed for the case of time-invariant systems. As observed in [10], a relevant advantage of these approaches is that they can be applied also when measurements are available at irregular intervals, which is quite common in real applications.

Correlation techniques estimate the measurement and process-noise covariance matrices by exploiting the sample autocorrelation functions of the innovations at different lags. Only time-invariant systems are considered in [11], [12], and [14]–[19], whereas [13] extends the problem formulation to time-variant systems. However, as remarked in [5] and shown in [13], the correlation methods provide sound results mainly for time-invariant systems. Odelson *et al.* [15] propose a least-squares optimization problem based on the correlations between the innovations; it is able to ensure the positive semidefiniteness of both  $\mathbf{R}$  and  $\mathbf{Q}$ . Then, a follow-up of this paper is presented by Rajamani and Rawlings [18], where they present new conditions for the uniqueness of the covariance estimates as well as an optimal weighting to be used in the least-squares objective to ensure minimum variance in the estimates. The work presented in [20] deals with adaptive estimation of both  $\mathbf{R}$  and  $\mathbf{Q}$  for time-variant models. The method is based on the correlation-innovations approach and includes an approach to ensure the positive definiteness of the covariance matrices.

Covariance-matching methods consist in adaptive algorithms that, at every sampling step, assess the KF parameters directly from the past state estimates. They are particularly effective in case either  $\mathbf{R}$  or  $\mathbf{Q}$  is known, as stated in [5]. The covariance-matching methods presented in [2] and [3] have already been discussed in Section I.

Correlation and covariance-matching techniques need *ad hoc* procedures in order to ensure the positive semidefiniteness of the estimated covariance matrices, which is not guaranteed by the method itself. Myers and Tapley [2] use an approximated countermeasure to ensure positive diagonal elements of the covariance matrices, which consists in replacing the estimated diagonal elements with their absolute values.

Similarly, in [20], the positive definiteness of the estimated  $\mathbf{R}$  and  $\mathbf{Q}$  is ensured with further processing of the two matrices after their estimation; therefore, also this method is approximated. An effective way to ensure the positive semidefiniteness of the estimated covariance matrices is to solve a constrained optimization problem, as in [15].

In the context of power systems, very few papers deal with the KF-parameter evaluation, usually limited to the estimation of  $\mathbf{Q}$ . Indeed, the elements of  $\mathbf{R}$  are associated with the measurement device characteristics that are known. However, in the literature dealing with power-systems SE using KF, the value of  $\mathbf{Q}$  is often arbitrarily selected and is assumed to be constant [21]–[23]. This approach leads to a nonadaptive filter that could have very poor estimation performance, with respect to a filter whose stochastic parameters are frequently updated. Zhang *et al.* [24] formulate an optimization problem that is only able to inflate the process-noise variances when a step-variation in the power-system state occurs. Then, the authors just mention that the inflated  $\mathbf{Q}$  could be decreased by employing an exponential decay with time constant to be defined by the user. As already mentioned in Section I, a  $\mathbf{Q}$  assessment method for a KF applied to power-system SE is also proposed in [3], although it is heuristic and suboptimal. Finally, Chowdhury *et al.* [25] propose the use of KF to detect the presence of a fault in power networks. Their goal is to detect a fault by analyzing the measurement residuals. The method exploits the fact that when a step in the state variables occurs, the process model does not match the real process, thus resulting in a sudden increase of the estimation residuals. This is flagged as the occurrence of a fault. Whereas, in this paper, we track the state even when there is a state step-variation. To do so, we adapt the process model uncertainty when the step occurs.

The methods presented in the aforementioned papers are conceived to estimate the stochastic parameters of time-invariant or slow-varying processes. Still needed is a KF able to filter the measurement noise and to track sudden state step-variations with minimum delay. In this paper, we propose an AKF that addresses this problem and uses power-system SE as an application example. Nevertheless, our proposed AKF is not specific for power-system SE only, actually it can be applied to every process where Assumptions 1–3 introduced in Section III hold.

### III. NEW METHOD FOR THE PREDICTION-ERROR COVARIANCE ESTIMATION

The PECE method we propose in this paper applies to cases where the following assumptions hold.

*Assumption 1:* The process model is the linear and time-variant ARIMA (0,1,0) model described by (1). In particular, we consider the general case in which the process-noise covariance matrix  $\mathbf{Q}_k$  is unknown and changes as a function of time. Furthermore, we are interested in the case where the system state is characterized by occasional step-variations that violate the process model.

*Assumption 2:* The measurement model is linear, known, and time-invariant. We assume to have the perfect knowledge

of both the measurement matrix  $\mathbf{H}$  and the measurement-noise covariance matrix  $\mathbf{R}$ .

*Assumption 3:* The system is assumed to be fully observable by using a number of measurements equal to or higher than the number of states:  $m \geq n$  and matrix  $\mathbf{H}$  has full rank.

If  $\mathbf{Q}_k$  were known, the prediction-error covariance matrix  $\mathbf{P}_{k|k-1}$  would be computed iteratively from (4) and all the other quantities of interest would derive from (5)–(7). However, this is not possible because of Assumption 1. The objective of PECE is to provide an estimate of  $\mathbf{P}_{k|k-1}$  from the measurements, without direct estimation of  $\mathbf{Q}_k$ .

#### A. PECE Method

Given Assumption 1, the KF-prediction equation is the ARIMA (0,1,0) model of (3), recalled here as follows:

$$\widehat{\mathbf{x}}_{k|k-1} = \widehat{\mathbf{x}}_{k-1|k-1}. \quad (14)$$

In order to quickly react to step-variations of the system state, the PECE method takes advantage of the innovation  $\mathbf{y}$  that contains the new information brought by the measurements at each time-step. In what follows, the PECE method's algorithm is presented in four steps.

- 1) *Step 1:* At time-step  $k$ , after the computation of the predicted state  $\widehat{\mathbf{x}}_{k|k-1}$  by means of (14) and when the new measurement set  $\mathbf{z}_k$  is available, the innovation  $\mathbf{y}_k$  is calculated by using (12).
- 2) *Step 2:* In steady-state conditions, the innovations represent a white Gaussian sequence with covariance matrix  $\mathbf{S}_\infty$ , defined as

$$\mathbf{S}_\infty = \mathbf{H}\mathbf{P}_\infty\mathbf{H}^T + \mathbf{R} \quad (15)$$

where  $\mathbf{P}_\infty$  is the value of  $\mathbf{P}_{k|k-1}$  as  $k \rightarrow \infty$ . An approximation of  $\mathbf{S}_\infty$  is the sample innovation-covariance matrix  $\widehat{\mathbf{C}}_k$ , which is calculated at time-step  $k$  by considering a moving-window composed of  $N$  time-steps as

$$\widehat{\mathbf{C}}_k = \text{cov}(\mathbf{y}_k, \mathbf{y}_{k-1}, \dots, \mathbf{y}_{k-N+1}). \quad (16)$$

As it is known, for stationary processes, the sample covariance tends to the true one as  $N$  increases. In non-stationary conditions, the true innovation-covariance matrix varies at each time-step and we denote it as  $\mathbf{S}_k$ . Thus, given the knowledge of the matrices  $\mathbf{H}$  and  $\mathbf{R}$  (Assumption 2), we can rewrite (15) as (13), recalled here as follows:

$$\mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R}. \quad (17)$$

Given the KF-prediction equation (14), when the state has a sudden change,  $\mathbf{S}_k$  changes as well. The sample matrix  $\widehat{\mathbf{C}}_k$  follows the variations of the true matrix  $\mathbf{S}_k$ , because the innovations incorporate the information of the measurements. The PECE method exploits this feature in order to quickly react to the state changes, as explained in the following steps.

- 3) *Step 3:* Estimate  $\mathbf{P}_{k|k-1}$  from the innovation samples, using ML estimation, as described in Theorem 2. This provides the estimate  $\widehat{\mathbf{P}}_{k|k-1}$ , which is symmetric and

positive semidefinite; using (5), this estimate is used to calculate the Kalman gain  $\mathbf{L}$  that is updated at every time-step. In quasi-steady-state conditions,  $\mathbf{L}$  remains fairly constant. When a state step-variation occurs, the KF-prediction equation (14) is inaccurate, causing an increase of  $\widehat{\mathbf{P}}_{k|k-1}$  and consequently of  $\mathbf{L}$ . Therefore, the KF trusts the measurements more than the predicted state, which is the right action to take in order to quickly react to the state variation.

*Estimation of  $\mathbf{P}_{k|k-1}$ :* A natural method for estimating  $\mathbf{P}_{k|k-1}$  would consist in replacing in (17) the matrix  $\mathbf{S}_k$  by the sample innovation-covariance  $\widehat{\mathbf{C}}_k$  computed in Step 2 and solving for  $\mathbf{P}_{k|k-1}$  (note that matrices  $\mathbf{H}$  and  $\mathbf{R}$  are known, by Assumption 2). However, this process, in general, will not produce a semidefinite matrix. A more adequate and more generally applied method is an ML estimation that can guarantee the positive semidefiniteness of  $\mathbf{P}_{k|k-1}$ .

The estimation procedure uses the following optimization problem:

$$\begin{aligned} \min_{\Sigma} \quad & \{-\log[\det(\Sigma)] + \text{trace}(\Sigma\mathbf{E})\} \\ \text{s.t.} \quad & \Sigma \text{ real symmetric and } \Sigma > \mathbf{0} \\ & \mathbf{I}_n - \Sigma \geq \mathbf{0}. \end{aligned} \quad (18)$$

In this optimization problem: 1) the optimization variable is the  $n \times n$ , real matrix  $\Sigma$ , assumed to be symmetric; 2)  $\mathbf{I}_n$  is the identity matrix of size  $n$ ; 3)  $\mathbf{E}$  is a known and fixed matrix; and 4) the notation  $\Sigma > \mathbf{0}$  means that  $\Sigma$  is positive definite and the notation  $\mathbf{I}_n - \Sigma \geq \mathbf{0}$  means that  $\mathbf{I}_n - \Sigma$  is positive semidefinite.

*Theorem 1:* The optimization problem (18) is a convex problem; it has one unique optimal solution  $\widehat{\Sigma}$ .

The proof is given in the Appendix. Note that problem (18) is a MAXDET problem for which there exists efficient software. In this paper, problem (18) is modeled and solved with YALMIP employing the *sdpt3* solver [26], [27].

We can now describe the estimation procedure. First, perform a QR-decomposition of the matrix  $\mathbf{R}^{-\frac{1}{2}}\mathbf{H}$  and obtain

$$\mathbf{R}^{-\frac{1}{2}}\mathbf{H} = \mathbf{V} \begin{pmatrix} \mathbf{U} \\ \mathbf{0}_{m-n,n} \end{pmatrix} \quad (19)$$

where  $\mathbf{U}$  is an upper triangular, real  $n \times n$  matrix,  $\mathbf{0}_{m-n,n}$  is the null rectangular matrix of dimensions  $(m-n) \times n$ , and  $\mathbf{V}$  is an orthogonal matrix of dimensions  $m \times m$ . Note that  $\mathbf{U}$  is invertible, because  $\mathbf{H}$  and  $\mathbf{R}$  are full rank.

Second, let  $\widehat{\mathbf{E}}_k$  be the square matrix of dimensions  $n \times n$  made of the first  $n$  rows and columns of  $\mathbf{V}^T \mathbf{R}^{-\frac{1}{2}} \widehat{\mathbf{C}}_k \mathbf{R}^{-\frac{1}{2}} \mathbf{V}$

$$\widehat{\mathbf{E}}_k = (\mathbf{V}^T \mathbf{R}^{-\frac{1}{2}} \widehat{\mathbf{C}}_k \mathbf{R}^{-\frac{1}{2}} \mathbf{V})(1:n, 1:n). \quad (20)$$

*Theorem 2:* The ML estimate of  $\mathbf{P}_{k|k-1}$  based on the observation of the innovations is equal to

$$\widehat{\mathbf{P}}_{k|k-1} = \mathbf{U}^{-1}(\widehat{\Sigma}^{-1} - \mathbf{I}_n)\mathbf{U}^{-T} \quad (21)$$

where  $\widehat{\Sigma}$  is the optimal solution of problem (18) with  $\mathbf{E}$  replaced by  $\widehat{\mathbf{E}}_k$ . The matrix  $\widehat{\mathbf{P}}_{k|k-1}$  is symmetric and positive semidefinite.

The proof is in the Appendix.

## B. Numerical Simulations

In this section, we evaluate the accuracy of the PECE method by considering ideal processes where the time evolution of the system state is controlled, so that the true value of every parameter of the process is known. The purpose is to verify whether the ability of the PECE method to track step-variations of the system state is not at the expense of its accuracy when there are no steps.<sup>1</sup> For this purpose, we first consider a process without step-variations in the system state, and then study a process where we add a step to the state variables. For all the cases treated in this section, we assume  $\mathbf{R}$  to be diagonal and constant

$$\mathbf{R} = r\mathbf{I}_m \quad (22)$$

where  $r$  is a scalar and  $\mathbf{I}_m \in \mathbb{R}^{m \times m}$  is the identity matrix.

The two case studies are described in detail here as follows.

1) *Base Case:* The process consists in a random walk as (1) with  $\mathbf{Q}$  that is diagonal and constant

$$\mathbf{Q} = q\mathbf{I}_n \quad (23)$$

where  $q$  is a scalar and  $\mathbf{I}_n \in \mathbb{R}^{n \times n}$  is the identity matrix. We have tested the two possible conditions that can occur, i.e.,  $q/r < 1$  and  $q/r > 1$ .

2) *Base Case Plus Steps:* It consists in the same random-walk processes of the *base case* plus occasional state step-variations of intentionally large amplitudes that violate the process model.

It is worth observing that this section proves the effectiveness of the PECE method for generic physical processes, as it can be applied to every process, as long as Assumptions 1–3 hold.

The PECE method is compared with two other covariance-estimation methods (their formulation is recalled in the Appendix).

- 1) The method of Myers and Tapley [2], henceforth in this paper, will be called *Myers*. In [2], the assessment of both  $\mathbf{Q}$  and  $\mathbf{R}$  is discussed. Whereas, we assume  $\mathbf{R}$  is known, so that only  $\mathbf{Q}$  has to be estimated.
- 2) The method of Zanni *et al.* [3], henceforth in this paper, will be called *Zanni*.

Three KFs that use the three considered methods (PECE, *Myers*, and *Zanni*) are run in parallel. We assume that the KFs have the knowledge of  $\mathbf{R}$ , but they do not know  $\mathbf{Q}$ .

The PECE and *Myers* methods use a moving window composed of  $N = 5000$  time-steps, as it has been found to be an effective tradeoff between filtering performance and fast tracking of state step-variations (the influence of  $N$  on the SE accuracy is given in Section III-B3). The *Zanni* method uses  $N = 30$ , because it does not need a large moving-window length to estimate  $\mathbf{Q}$  (see the Appendix for further details).

<sup>1</sup>Indeed, a simple way to track state step-variations would be to always overestimate the value of  $\mathbf{Q}$ , as the method proposed in [3] does (see Fig. 1). Obviously, the filtering effectiveness is compromised when no steps are present. Another way would be to inflate  $\mathbf{Q}$  only when the step occurs, as it is proposed in [24]. The problem is what value of  $\mathbf{Q}$  should be set before and after the step.

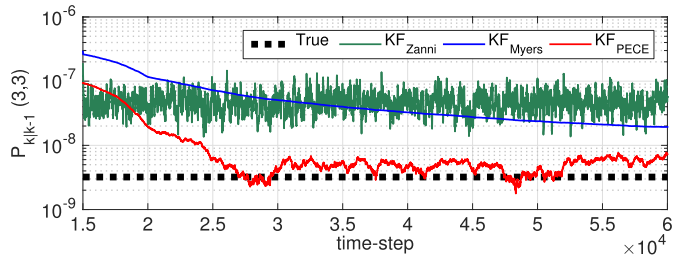


Fig. 1. *Base case*,  $q = 10^{-10}$ . Time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  (the other diagonal elements exhibit an analogous behavior): true and estimated values.

1) *Base Case*: Let us consider a multidimensional state  $\mathbf{x} \in \mathbb{R}^n$ , with  $n = 6$ . The process and measurement equations are

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{z}_k &= \mathbf{x}_k + \mathbf{v}_k \\ \mathbf{w} &\sim N(0, \mathbf{Q}) \\ \mathbf{v} &\sim N(0, \mathbf{R}) \end{aligned}$$

in which  $\mathbf{R}$  and  $\mathbf{Q}$  are defined as in (22) and (23), respectively. First, we consider the case where  $q/r < 1$ , with  $r = 10^{-7}$  and  $q = 10^{-10}$ . As a consequence, the true value of the KF-prediction-error covariance matrix is  $\mathbf{P}_\infty = 3.21 \cdot 10^{-9} \mathbf{I}_n$ . Note that the important parameter is the ratio  $q/r$ . Indeed, different values of  $r$  and  $q$  resulting in the same ratio lead to similar results.

The estimated state  $\hat{\mathbf{x}}$  and the matrix  $\mathbf{Q}$ , which are used in the three KFs, are initialized by using the initial conditions listed here as follows.

- 1) The state vector is initialized to  $\mathbf{x}_0 = \mathbf{1}$  (where  $\mathbf{1}$  is a vector of ones), which is different from the true state.
- 2) For the first 15 000 time-steps, we use  $\mathbf{Q} = q_0 \mathbf{I}_n$ . To help the convergence of the initial state  $\mathbf{x}_0$  toward the true one, we set a high value of  $q_0 = 10^{-5}$ .

Fig. 1 shows the time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  estimated by the three methods and its true value (the other diagonal elements exhibit an analogous behavior). The time evolution of the norm of the estimation errors is shown in Fig. 2 for these three KFs plus a KF that uses the exact value of  $\mathbf{Q}$  denoted as “KF<sub>Qexact</sub>”. At time-step  $k$ , the estimation error vector  $\mathbf{e}_k$  is defined as the difference between the estimated and the true state

$$\mathbf{e}_k = \hat{\mathbf{x}}_k^{\text{est}} - \mathbf{x}_k^{\text{true}}. \quad (24)$$

Note that the norm of  $\mathbf{e}_k$  accounts for the errors of all the six state variables. Until the 15 000th time-step, the three KFs use a large value of  $\mathbf{Q}$  equal to  $10^{-5} \mathbf{I}_n$ , so that the estimation errors are the same for every KF and remain large. Afterward, the process covariances start to be assessed and the errors of the three KFs decrease, thus reflecting the behavior of the  $\mathbf{P}_{k|k-1}$  estimates. As visible in Fig. 1,  $\mathbf{P}_{k|k-1}$  inferred by the PECE method converges to the true value of  $\mathbf{P}_{k|k-1}$  in about 10 000 time-steps, which corresponds to twice the moving-window length  $N = 5000$ . Afterward, the  $\mathbf{P}_{k|k-1}$  estimates oscillate around the true value of  $\mathbf{P}_{k|k-1}$  (see the

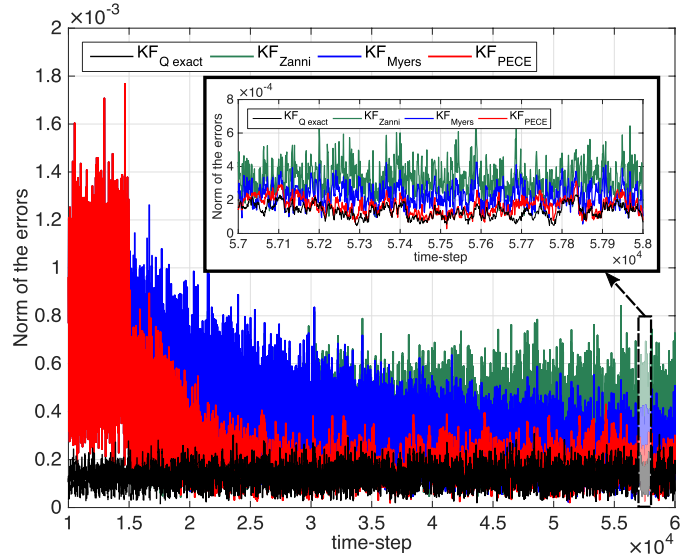


Fig. 2. *Base case*,  $q = 10^{-10}$ . Time evolution of the norm of the estimation errors.

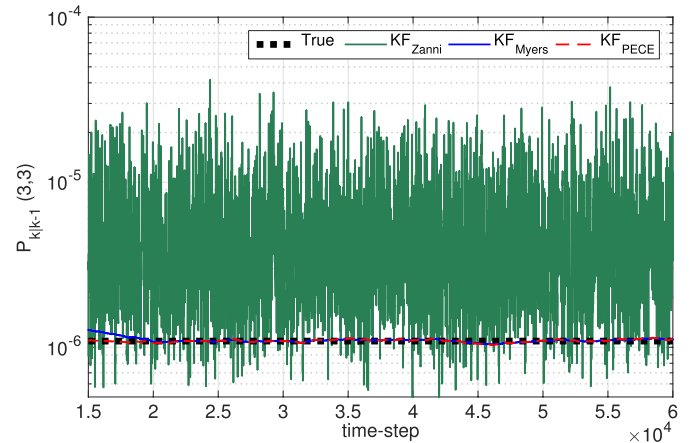


Fig. 3. *Base case*,  $q = 10^{-6}$ . Time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  (the other diagonal elements exhibit an analogous behavior): true and estimated values.

enlarged part of Fig. 2) and the estimation errors become similar to those of the KF that uses the exact value of  $\mathbf{Q}$ . The convergence of the *Myers* method is much slower and the  $\mathbf{P}_{k|k-1}$  estimates take a longer time to converge to the true values, i.e., hundreds of thousands of time-steps. However, the estimates are characterized by smaller variations compared with the PECE method (see Fig. 1). The *Zanni* method’s convergence phase lasts few time-steps, because it uses a small moving-window length  $N = 30$ . Then, it tends always to overestimate the value of  $\mathbf{P}_{k|k-1}$ . As a consequence, the measurement noise is not filtered effectively, which leads to significant estimation errors.

Let us now consider the case in which  $q/r > 1$ , with  $r = 10^{-7}$  and  $q = 10^{-6}$ . The corresponding true value of  $\mathbf{P}_\infty$  is  $1.09 \cdot 10^{-6} \mathbf{I}_n$ . Fig. 3 shows the time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  for this case (the other diagonal elements exhibit an analogous behavior). The *Myers* and PECE methods estimate its precise true value, whereas the *Zanni* method provides again larger variances. It can be

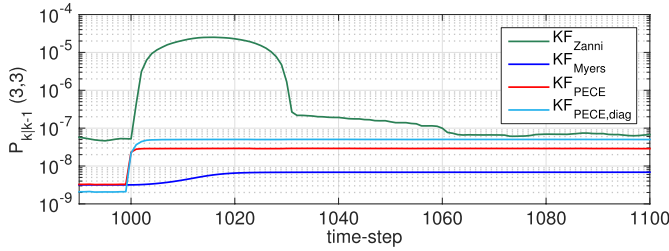


Fig. 4. *Base case plus steps*,  $q = 10^{-10}$ . Time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  (the other diagonal elements exhibit an analogous behavior). A time-window of 110 time-steps, including the step-variation, is shown.

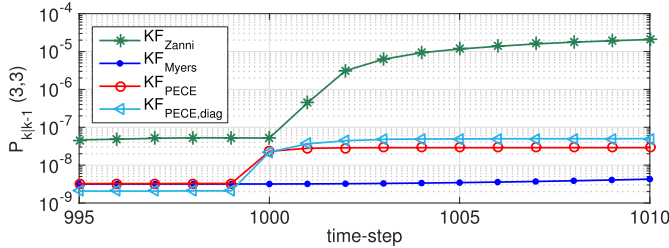


Fig. 5. *Base case plus steps*,  $q = 10^{-10}$ . Zoomed-in view of Fig. 4 close to the step-variation that occurs at the 1000th time-step. It can be seen that only the PECE and PECE<sub>diag</sub> methods are able to react exactly by the time the step occurs.

noted that the PECE method estimates immediately the correct value of  $\mathbf{P}_{k|k-1}$ , whereas the *Myers* method takes 5000 time-steps to approach it. However, the estimation errors of these three KFs are almost equal to those of the KF that uses the exact value of  $\mathbf{Q}$ . We do not show the norm of the errors for this case, because they are indistinguishable and oscillate between  $0.2 \cdot 10^{-3}$  and  $1.6 \cdot 10^{-3}$ . These results are caused by the predominance of the process noise, making simple its estimation from the samples. This case is not of high interest, because the process model has a limited influence on the KF solution. Indeed, the simple processing of the measurements leads to state estimates very close to the optimal ones.

2) *Base Case Plus Steps*: Let us consider the same process of the *base case*, with  $r = 10^{-7}$ ,  $q = 10^{-10}$ , and  $\mathbf{x}_0 = \mathbf{1}$ . For this case, we wait until the convergence phase is finished (time instant 0), and 1000 time-steps later, we simulate a state step-variation of amplitude  $10^{-2}$ , which is significantly larger than the process-noise standard deviation. By the time the state change occurs, we expect  $\mathbf{P}_{k|k-1}$  to increase, because the assumed ARIMA (0,1,0) in (14) is no longer accurate. Note that, as the step is applied to every state variable, they become correlated. In order to show the effect of the estimated off-diagonal elements of  $\mathbf{P}_{k|k-1}$  on the state estimates, we add a further KF, called PECE<sub>diag</sub>. The latter uses the PECE method in which the variable  $\Sigma$  of the optimization problem (18) is diagonal.

Fig. 4 shows the time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  estimated by the four KFs when the step occurs, and Fig. 5 shows a zoomed-in view of Fig. 4 (the other diagonal elements exhibit an analogous behavior). The trend of the off-diagonal elements of  $\mathbf{P}_{k|k-1}$  is highly dependent on the specific simulation parameters. Their effect is evident by

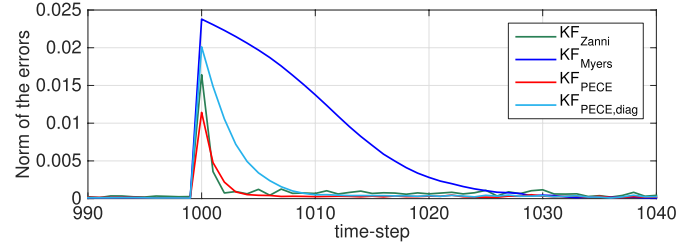


Fig. 6. *Base case plus steps*,  $q = 10^{-10}$ . Time evolution of the norm of the estimation errors. A time-window of 50 time-steps, including the step-variation, is shown.

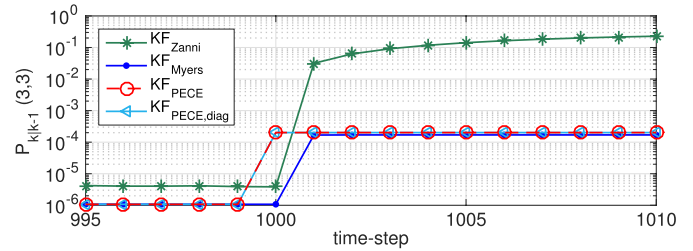


Fig. 7. *Base case plus steps*,  $q = 10^{-6}$ . Time evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  (the other diagonal elements exhibit an analogous behavior). A time-window of 15 time-steps, including the step-variation, is shown.

comparing the behavior of the KF estimates given by the PECE and PECE<sub>diag</sub> methods. The KF that uses the *Myers* method has a remarkable delay in inflating  $\mathbf{P}_{k|k-1}$  and consequently in tracking the system state during and after the step, as is visible from the norm of the estimation errors in Fig. 6. This behavior is typical of the covariance-estimation methods that infer  $\mathbf{Q}$  by exploiting only the knowledge of the past state estimates. It is known that  $\mathbf{Q}$  influences the state estimates, and vice versa, as a closed loop; therefore, both take several time-steps to increase. At the time of the step, the *Zanni* method had smaller errors, because it overestimates the value of  $\mathbf{Q}$ , as we have seen for the *base case*. Shortly, after the step, the *Zanni* method inflates  $\mathbf{Q}$  rapidly, so that the estimation errors decrease in only a few time-steps. Unlike the other two methods, the PECE method uses the past innovations and the innovation at the current time-step that already contains the information brought by the upcoming set of measurements. Then, it updates the value of  $\mathbf{P}_{k|k-1}$  before the measurement update. Indeed, the PECE method is the only one able to increase  $\mathbf{P}_{k|k-1}$  exactly by the time the step occurs (i.e., with no delay), as shown in Fig. 5. This characteristic enables the PECE method to immediately react to state step-variations. Fig. 6 shows that the estimation errors of the KF that uses the PECE method are the smallest, both exactly when the step occurs and shortly afterward. In Fig. 6, it is also evident that the KF that uses the PECE method reacts faster than the KF that uses the PECE<sub>diag</sub> method. Therefore, the estimated off-diagonal terms of  $\hat{\mathbf{P}}_{k|k-1}$  improve the state-tracking capability.

Finally, we consider a case in which  $q/r > 1$  ( $r = 10^{-7}$  and  $q = 10^{-6}$ ), and a state step-variation is present. The step-variation is of amplitude 1 and occurs at the 1000th time-step. Note that the step amplitude is larger than the case where  $q/r < 1$ , because a step amplitude of  $10^{-2}$  would be comparable to the process noise. Fig. 7 shows the time

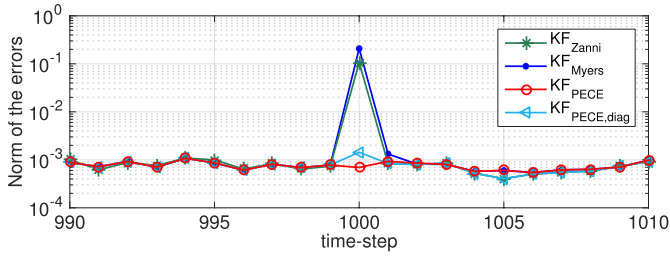


Fig. 8. *Base case plus steps*,  $q = 10^{-6}$ . Time evolution of the norm of the estimation errors. A time-window of 20 time-steps, including the step-variation, is shown.

evolution of the third diagonal element of  $\mathbf{P}_{k|k-1}$  estimated by the *Zanni*, *Myers*, PECE, and PECE<sub>diag</sub> methods (the other diagonal elements exhibit an analogous behavior). As in the case where  $q = 10^{-10}$ , the PECE and PECE<sub>diag</sub> methods increase  $\mathbf{P}_{k|k-1}$  as soon as the step occurs, whereas the *Zanni* and *Myers* methods react with a delay of one time-step. Indeed, only the PECE and PECE<sub>diag</sub> methods are characterized by limited errors during the step, as shown in Fig. 8. Note that, again, the accuracy of the PECE method is better than that of PECE<sub>diag</sub>, due to the effect of the estimated off-diagonal elements of  $\mathbf{P}_{k|k-1}$ .

3) *Influence of Parameter  $N$  on the SE Accuracy*: The only parameter that has to be set in the PECE method is  $N$ , specifically the number of previous innovations used to calculate the sample innovation-covariance matrix in (16). In this section, we present the influence of  $N$  on the SE accuracy considering both the *base case* and the *base case plus steps* in which  $q/r < 1$ . As explained in Section III-B1,  $q/r < 1$  is the case of interest where the noise filtering can be effective. The investigation involves the PECE and *Myers* methods. Unlike in Sections III-B1 and III-B2, where the norm of the estimation errors was calculated at each time-step, here the norm is computed by considering of all the errors of the  $M$  considered time-steps. Besides, we show the median value of the norms obtained in ten different simulations. It is worth mentioning that we wait until the convergence phase is finished before performing this assessment.

Fig. 9(a) shows the norm of the estimation errors as a function of  $N$  for the *base case*. For each value of  $N$ , the norm is computed by considering  $M = 2000$  time-steps. The PECE method degrades the estimation accuracy as we consider a smaller number of innovations. The innovations are affected by the measurement noise that is not effectively filtered if a small number of innovations are used. The norm of the errors of the *Myers* method remains quite stable and smaller than that of the PECE method for  $N > 500$ . Note that the use of a smaller value of  $N$  also leads to a faster convergence of the *Myers* method to the true value of  $\mathbf{P}_{k|k-1}$  compared with Fig. 1. However, the accuracy of the *Myers* method drops for small values of  $N$  (i.e.,  $N = 20$  and  $N = 100$ ) due to the fact that the positive semidefiniteness of  $\mathbf{P}_{k|k-1}$  is not guaranteed. As a consequence, the KF numerical stability might be compromised and the KF solution sometimes diverges from the true state. This phenomenon is visible in Fig. 10 where  $N = 100$ : the state estimated by the KF that uses the *Myers*

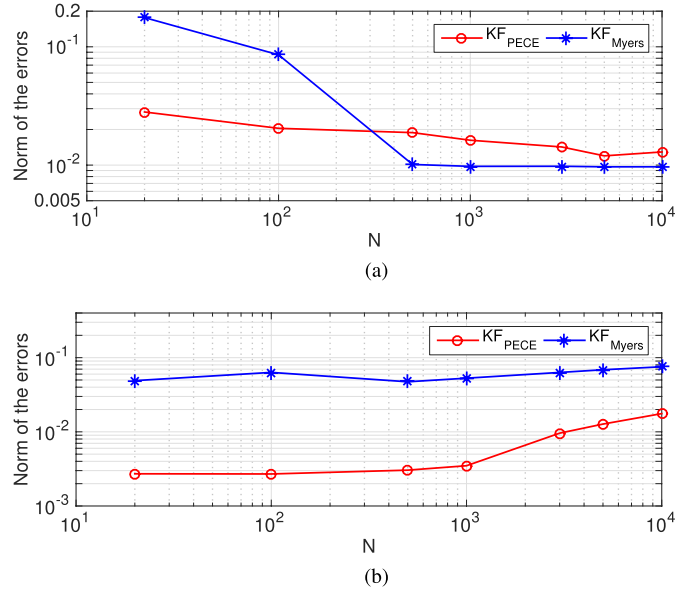


Fig. 9. Influence of  $N$  on the estimation errors of the KFs that use the PECE and *Myers* methods. In the *base case*, the norm of the errors is computed considering 2000 time-steps. In the *base case plus steps*, the norm of the errors is computed considering the 30 time-steps after the state step-variation. (a) *Base case* ( $q = 10^{-10}$ ,  $r = 10^{-7}$ ). (b) *Base case plus steps* ( $q = 10^{-10}$ ,  $r = 10^{-7}$ ).

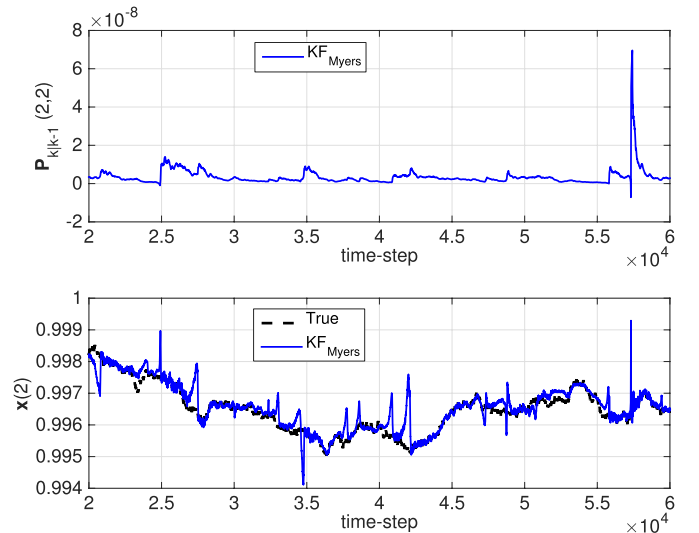


Fig. 10. *Base case*,  $q = 10^{-10}$ ,  $r = 10^{-7}$ , and  $N = 100$ . Time evolution of the second diagonal element of  $\mathbf{P}_{k|k-1}$  and of the associated second element of the state vector: true state and state estimated by the KF that uses the *Myers* method. This figure shows the numerical stability problems of the *Myers* method.

method loses track of the true state in many occasions, e.g., when negative diagonal elements of  $\mathbf{P}_{k|k-1}$  are estimated at time-step  $5.7 \cdot 10^4$ . For the ideal case considered in this section, significant estimation errors related to this problem occur only for  $N < 300$ .

Fig. 9(b) shows the norm of the estimation errors as a function of  $N$  for the *base case plus steps*. For each value of  $N$ , the norm is computed by considering the  $M = 30$  time-steps after the state step-variation. As expected, the PECE method outperforms the *Myers* method in the period

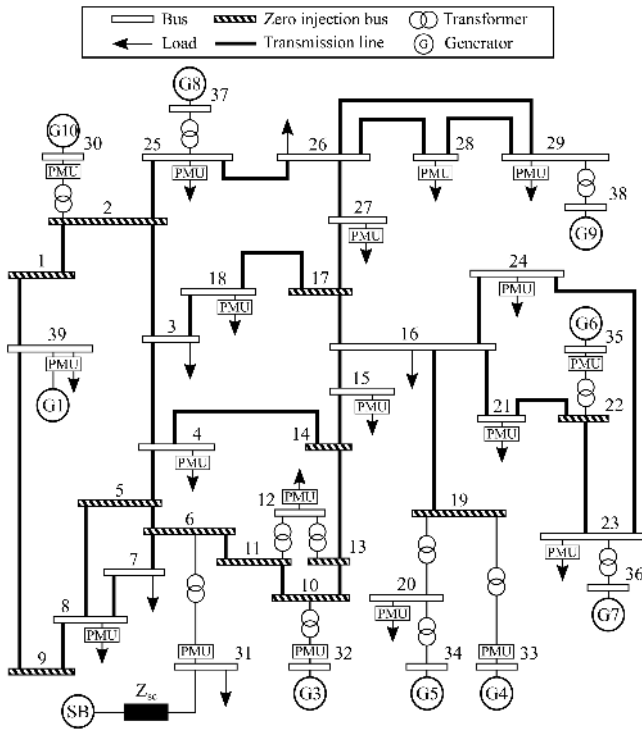


Fig. 11. Network topology of the IEEE 39-bus test system [28], together with the adopted PMU placement. We assume that bus #31 is the connection point of the system to an external network that is represented by a voltage source called SB in series with the short-circuit impedance  $Z_{sc}$ .

following the step. In contrast with the *base case*, the smaller the  $N$  value is, the smaller the errors are, because the last innovations have a higher weight in the computation of the sample covariance (16).

#### IV. APPLICATION TO POWER-SYSTEM STATE ESTIMATION

The effectiveness of the PECE method is here proved in the context of power-system SE. We simulate the behavior of a power network via the procedure described here as follows.

- 1) We consider the IEEE 39-bus test system, a widely used transmission network benchmark [28]. A schematic of the network is shown in Fig. 11.
- 2) The network boundary conditions are represented by the nodal power injections/absorptions of generators/loads.
- 3) In order to show the behavior of the PECE method when a state step-variation occurs, we simulate a sudden drop of the reactive power absorbed by the load at bus #4.<sup>2</sup>
- 4) The network true state and the true values of the measurements are computed via a *power flow* procedure that uses as inputs the network characteristics and operating conditions described in *steps* #1–3 (a detailed description of the *power flow* procedure is contained in [29]).

<sup>2</sup>This can be caused by the disconnection of a shunt reactor, which is a common operation in transmission networks. Note that a step-variation of the active power or of both active and reactive powers simultaneously might produce step-variations of different state variables. However, the covariance assessment methods are independent of the particular state variables that experience the step-variation.

TABLE I  
VOLTAGE LEVELS CHOSEN FOR THE IEEE 39-BUS TEST SYSTEM

Rated voltage	Bus #
380 kV	31
230 kV	1–11, 13–19, 21–29, 39
125 kV	12, 20
15 kV	30, 32–38

- 5) The measurements used in SE are composed of nodal voltage and nodal injected/absorbed current phasors at the network buses marked as phasor measurement unit “(PMU)” in Fig. 11. They are generated by adding noise to the true quantities computed in *step* #4.
- 6) In order to evaluate the accuracy of each state estimator, the estimated state returned by each state estimator is compared with the true state provided by the *power flow* procedure.

Further technical details and explanations are given in the following. The voltage levels at every bus of the 39-bus network of Fig. 11 are given in Table I. The zero-injection buses shown in Fig. 11 are buses where no load or generator is connected. We assume that bus #31 is the connection point of the system to an external network. The latter is characterized by a short-circuit power of 50 GVA and a rated voltage of 380 kV that result in a short-circuit impedance  $Z_{sc} = i 2.89 \Omega$  ( $i$  denotes the imaginary unit of complex numbers). The external network equivalent circuit is represented by a fixed voltage source called *slack bus* (SB) in series with  $Z_{sc}$ . The SB is the reference bus used in the *power flow* procedure, where we impose the phase angle of the voltage phasor to be zero.

The state of a power network composed of  $\ell$  buses is usually represented by the voltage phasors at every network bus, which can be expressed in rectangular coordinates as

$$\mathbf{x} = [V_{1re}, \dots, V_{jre}, \dots, V_{\ell re}, V_{1im}, \dots, V_{jim}, \dots, V_{\ell im}]$$

$$V_j = V_{jre} + i V_{jim} \quad (25)$$

in which  $V_j$  is the voltage phasor at the network bus # $j$ . In our case,  $\ell = 39$  and the state dimension  $n = 78$ .

The *power flow* procedure computes the system state imposing: 1) voltage magnitude and phase angle at the SB; 2) active power and voltage magnitude at the generator buses; and 3) active and reactive powers at the load buses. The active powers and the voltage magnitudes at the generator buses are the ones specified in [28]. In order to use realistic time series, the active and reactive powers absorbed by the loads come from real power measurements provided at 50 frames/s by PMUs<sup>3</sup> installed in the 125-kV subtransmission network of Lausanne, Switzerland. These power measurements are adapted to the values specified in [28].<sup>4</sup>

<sup>3</sup>According to the definition in [30], a PMU is a device that produces synchronized phasor, frequency, and rate-of-change of frequency estimates from voltage and/or current signals and a time-synchronizing signal.

<sup>4</sup>As we do not use transformer tap changers, the powers at buses #7, #8, and #12 have been changed in order to get the voltage magnitude at every bus within the range  $\pm 5\%$  of the rated voltage, which is a required operating condition in transmission networks. In particular, the absorbed active  $P_a$  and reactive  $P_r$  powers at the aforementioned buses are:  $(P_{7a}, P_{7r}) = (23.38, 8.4)$  MVA,  $(P_{8a}, P_{8r}) = (52.2, 17.6)$  MVA, and  $(P_{12a}, P_{12r}) = (7.5, 30)$  MVA.



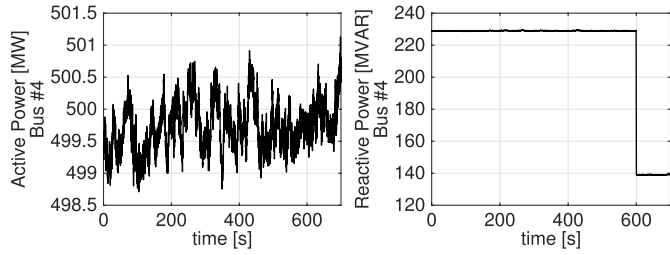


Fig. 12. Time profiles of the active and reactive power at bus #4. A drop of about 90 MVAR of reactive power occurs at 600 s.

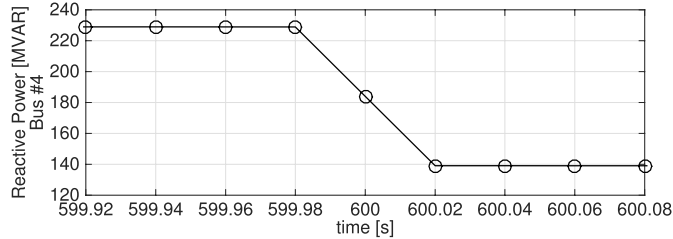


Fig. 13. Time profile of the reactive power at bus #4: zoomed-in view close to the 90-MVAR step-variation.

At 600 s, we simulate a sudden drop of 90 MVAR of the reactive power at bus #4, which is shown in Fig. 12. This produces step-wise perturbations in the state variables; these perturbations are similar to those of Section III-B2. Fig. 13 highlights the fact that the drop affects two time-steps as a result of the response to transients of the PMU phasor estimation algorithm [31].

The PMU locations are chosen in order to have the system fully observable, i.e., matrix  $\mathbf{H}$  is of full rank (Assumption 3 holds). The measured phasors are expressed in rectangular coordinates and are obtained by perturbing the true quantities with randomly generated Gaussian white noise characterized by a variance related to the PMU accuracy. The latter is assumed to be known and the measurement uncertainties are independent; these are common assumptions in power systems. Thus,  $\mathbf{R}$  is known and diagonal. In particular, we assume the use of PMUs characterized by a total vector error (TVE) of 0.14%, which is equally distributed between real and imaginary parts, i.e., an error of 0.1% each (see [30] for the definition of TVE). The standard deviations of the measurement noises are assumed to be one third of it, specifically 0.33%. They are written in percentages of the full scales of the sensors [32, p. 159]. The voltage sensors' full scales are assumed to be equal to the rated voltage, and the current sensors' full scales are assumed to be equal to twice the largest value of the injected-current magnitude at the respective bus.

Expressing both states and measurements in rectangular coordinates leads to a linear measurement model characterized by an exact measurement matrix  $\mathbf{H}$ . The latter is not approximated and its elements are known and constant along the time, as explained in detail in [4]. The part of  $\mathbf{H}$  related to the voltage measurements is composed of zeros and ones. The part of  $\mathbf{H}$  related to the injected-current measurements is derived from the admittance matrix of the network [4].

The latter is constructed from the network topology and the electrical parameters of the transmission lines that are commonly assumed to be known and constant in time (see [29] for the details on the construction of the admittance matrix).

The above-mentioned hypotheses on  $\mathbf{H}$  and  $\mathbf{R}$  lead to a linear, known, and stationary measurement model (Assumption 2 holds).

Moreover, we use the process model given in (14), because Assumption 1 holds for the following reasons: 1) the true state of a power system is always time-variant and hidden; 2) the power-system inputs are usually not controllable from the SE perspective; and 3) we assume the use of high-resolution measurements at 50 frames/s.

#### A. Accuracy Performance Analysis

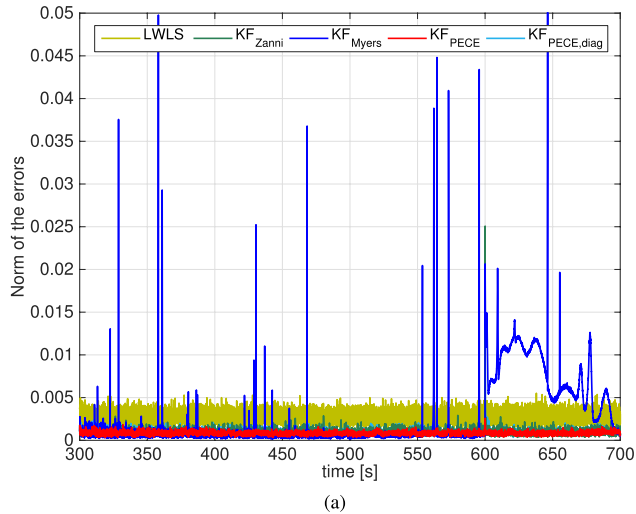
In this section, we analyze the accuracy performance of five state estimators, i.e., four KFs that use the *Myers*, *Zanni*, PECE and PECE<sub>diag</sub> methods, and the LWLS. The LWLS formulation is described in detail in [4] and briefly recalled in the Appendix. As in Section III-B, the *Zanni* method uses  $N = 30$ , whereas the PECE, *Myers*, and PECE<sub>diag</sub> methods use  $N = 5000$ . The influence of parameter  $N$  on the SE accuracy for the specific application of power-system SE is presented in Section IV-B.

The first part of the simulation consists in an initial phase of 300 s (corresponding to 15000 time-steps as in Section III-B1) in which we keep a constant value of  $\mathbf{Q}$  for the considered KFs and their estimated states converge from a *flat-start initialization*<sup>5</sup> toward the true state. For brevity, the initial phase is not shown in this section.

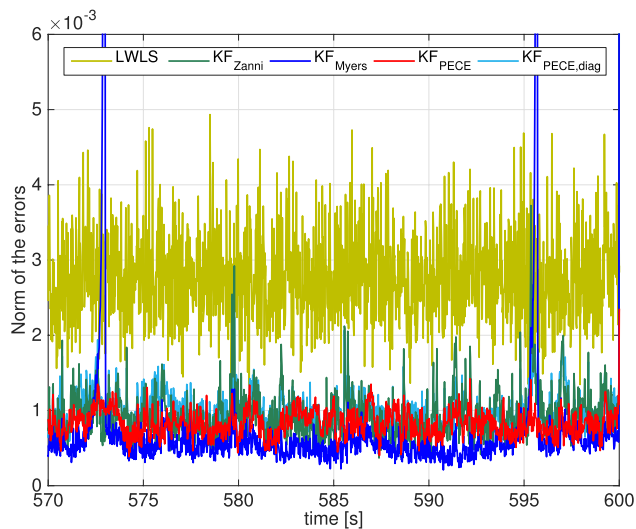
The time evolution of the norm of the estimation errors of the five considered state estimators is given in Fig. 14; it is composed of three subfigures: Fig. 14(a) shows the entire simulation; Fig. 14(b) shows a portion of the simulation where the network is in quasi-steady-state conditions, and Fig. 14(c) shows the portion of the simulation close to the state step-variation. The norm of the estimation errors of the *Myers* method exhibits high sporadic spikes that are clearly visible in Fig. 14(a). This numerical stability issue of the *Myers* method has already been described in Section III-B3 (in particular, see Fig. 10), and it is due to the fact that the *Myers* method does not ensure the positive semidefiniteness of  $\mathbf{P}_{k|k-1}$ . Note that the smaller the  $N$  value, the worse the numerical stability issue.

Fig. 14(b) shows the norm of the estimation errors when the network is in quasi-steady-state conditions, i.e., the load powers are varying smoothly with no steps of significant amplitude. The errors of the four KFs are comparable except for the occasional spikes of the *Myers* method, whereas the LWLS errors are about three times larger on average. We can observe that, unlike in Section III-B1, the accuracy of the *Zanni* method is just slightly worse than that of the other KFs.

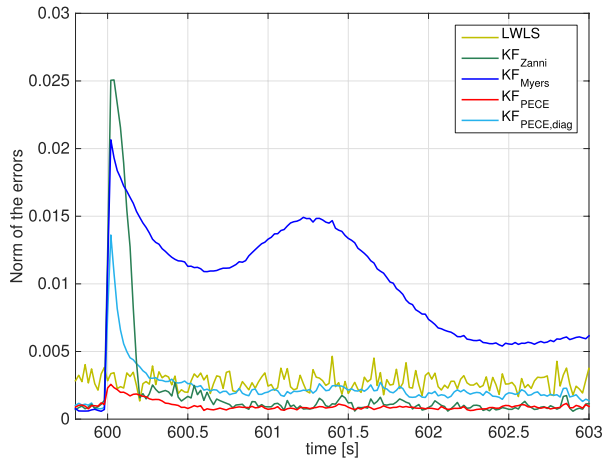
<sup>5</sup>In the context of power systems, the *flat-start initialization* refers to a state vector with the real parts of the voltage phasors equal to 1 per unit of the rated voltage and the imaginary parts equal to 0. This is a common practice in power systems as, in normal operating conditions, the voltage magnitudes are close to the rated values and the voltage phase angles are small.



(a)



(b)



(c)

Fig. 14. Time evolution of the norm of the estimation errors. (a) Entire simulation. (b) Zoomed-in view of Fig. 14(a) during quasi-steady-state conditions. (c) Zoomed-in view of Fig. 14(b) close to the state step-variation.

Fig. 14(c) shows the norm of the estimation errors during and after the step-variation. The LWLS keeps the same estimation accuracy; because at every time-step, it relies only on

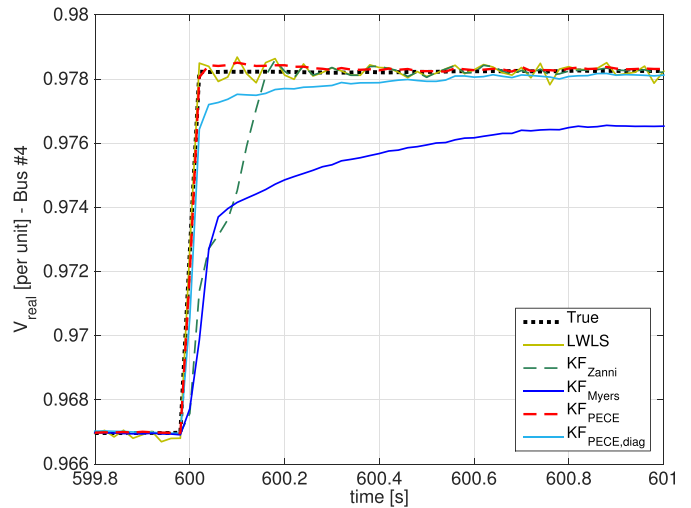


Fig. 15. Time evolution of the real part of the voltage phasor (in per unit of the rated voltage) at bus #4: true and estimated values. Bus #4 is the one where the reactive power step-variation occurs; therefore, the voltage at this bus is the most affected by the perturbation. Only the LWLS and PECE methods are capable of tracking the true state during the step-variation.

the measurements taken at that time-step. The KF, instead, also uses the past information to compute the system state. As soon as the step occurs,  $\mathbf{P}_{k|k-1}$  should be adapted in order to attribute a lower weight to the predicted state with respect to the measurements. In other words, the KF behavior should approach that of the LWLS. The *Myers* and *Zanni* methods detect the occurrence of the step with a time-step delay, because they use only the state estimates. Therefore, they are characterized by large estimation errors when the step occurs. After the step, the *Zanni* method recovers the correct state tracking faster than the *Myers* method. The PECE method is the only one that has a peak of the errors comparable to the one of LWLS. Indeed, it estimates the new  $\mathbf{P}_{k|k-1}$  by using the innovations before the measurement update, so that the information brought by the measurements at the time of the step is already considered in the Kalman gain. The accuracy of the PECE<sub>diag</sub> method is significantly worse than that of the PECE method, proving the importance of using a full matrix in the optimization problem (18). These results confirm those obtained for the ideal case of Section III-B2.

The reactive power step-variation affects mainly the real part of the voltage phasors at bus #4 and at the neighbor buses. Fig. 15 shows the time evolution of the real part of the voltage phasor (in per unit of the rated voltage) at bus #4: the true value and the value estimated by the considered state estimators. It reflects the estimation errors shown in Fig. 14(c). It is worth mentioning that some of the state variables estimated by the *Myers*, *Zanni*, and PECE<sub>diag</sub> methods go even on the opposite direction of the true state variation. For instance, this undesired behavior occurs with the *Myers* and *Zanni* methods for the real part of the voltage phasor at bus #31, as shown in Fig. 16. This is due to the incorrect value of  $\mathbf{P}_{k|k-1}$  estimated by these methods when the step occurs. These wrong estimates are other possible equilibrium points of the power system far from the true ones.

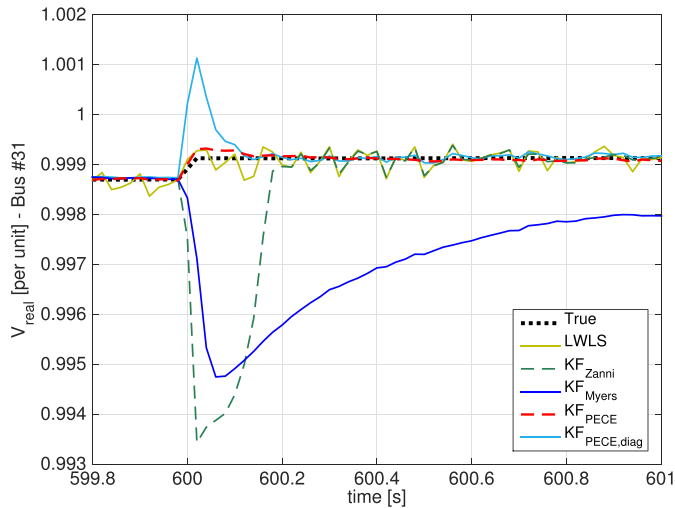


Fig. 16. Time evolution of the real part of the voltage phasor (in per unit of the rated voltage) at bus #31: true and estimated values. The quantities estimated by the *Myers*, *Zanni*, and *PECE<sub>diag</sub>* methods exhibit a peculiar and undesired behavior. Only the *LWLS* and *PECE* methods track the true state correctly.

### B. Influence of Parameter $N$ on the SE Accuracy

This section shows the influence of parameter  $N$  on the SE accuracy. This study is carried out separately for the network in quasi-steady-state conditions and during the step-variation. In the first case, we consider the 100 s prior to the step. In the second case, we consider the 2 s following the step-variation. For both cases, we compute the norm of all the estimation errors contained in the considered time-window and we show the median value of the norms obtained in ten different simulations. The norm values as a function of  $N$  are shown in Fig. 17(a) and (b) for the system in quasi-steady-state conditions and during the step-variation, respectively. The accuracy of both the *PECE* and *Myers* methods is examined.

Concerning the system in quasi-steady-state conditions, the estimation accuracy of both the *PECE* and *Myers* methods degrades as  $N$  diminishes. However, it can be seen that the accuracy of the *PECE* method is much better than that of the *Myers* method for all the considered values of  $N$ . This is mainly due to the numerical stability problem of the *Myers* method that we have discussed in Section IV-A.

As expected, also the study on the step-variation reveals that the *PECE* method outperforms the *Myers* method for all the considered values of  $N$ . The accuracies of both methods decrease as  $N$  decreases, which indicates an opposite trend with respect to the results of Section III-B3. Regarding the *Myers* method, the reason lies in more significant numerical stability problems. Instead, the worsening of the *PECE* method's accuracy as  $N$  decreases is due to the increasing effect of the measurement noise on the sample covariance matrix (16). Indeed, we have observed that the small peak of the errors of the *PECE* method close to the step-variation has the same amplitude irrespectively of the value of  $N$ , whereas the errors after the step are highly affected by the measurement

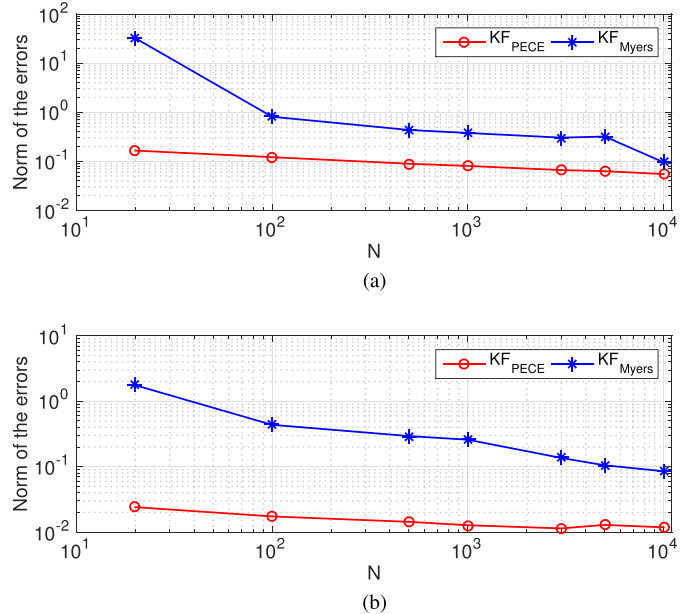


Fig. 17. (a) Quasi-steady-state conditions. (b) State step-variation. Influence of parameter  $N$  on the SE accuracy for both the *PECE* and *Myers* methods. For case (a) where the network is in quasi-steady-state conditions, the norm of the errors is computed considering the 100 s prior to the step. For case (b) where the network state has a step-variation, the norm of the errors is computed considering the 2 s following the step.

noise. The latter is not effectively filtered if a small value of  $N$  is set.

In conclusion, the *PECE* method guarantees a better accuracy than the *Myers* method for all the considered values of  $N$  in both quasi-steady-state conditions and during the state step-variation.

### C. Computational-Time Performance Assessment

The remarkable estimation accuracy of the *PECE* method comes at the expense of the computational time. The latter is almost entirely devoted to solving the convex optimization problem of (18) that depends only on the state dimension  $n$ . We can instead neglect the time used to compute the sample covariance matrix (16), so that the computational time is not function of  $N$ . Fig. 18(a) shows the computational time of the *PECE* method as a function of  $n$ . Each value is an average over 100 simulations. These results are obtained by solving the optimization problem (18) with YALMIP employing the *sdpt3* solver implemented in the MATLAB 2014b [26], [27]. The laptop is an Apple MacBook Pro with a 2.5-GHz CPU, 16-GB RAM. The increase of the computational time shown in Fig. 18(a) is exponential as the y-axis is in logarithmic scale.

On the contrary, the computational time of the *Myers* and *Zanni* methods is significantly affected by the parameter  $N$ . Fig. 18(b) and (c) shows the computational time of these methods, as a function of the state dimension  $n$ ; the three curves refer to three values of  $N$ , i.e., 100, 1000, and 10000. Both methods are considerably faster than the *PECE* method. In particular, the *Myers* method is characterized by a computational time that increases exponentially as a function of  $n$ , and it remains below 100 ms for the considered values of  $n$ .

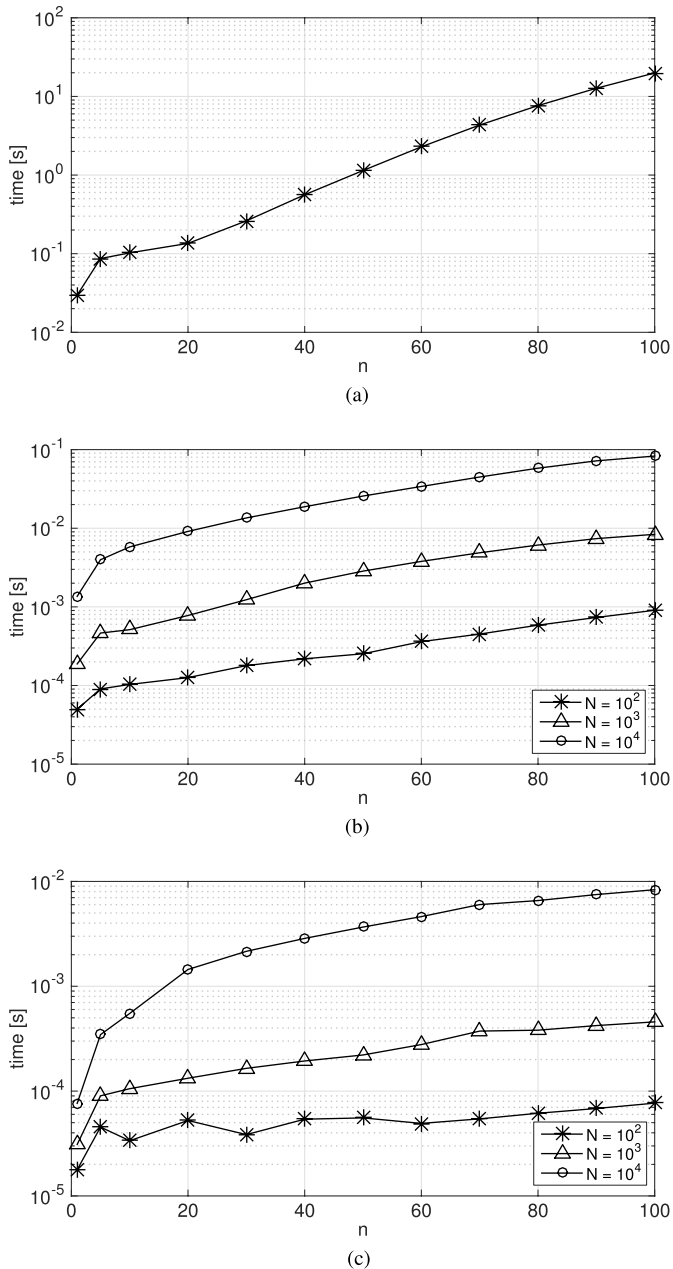


Fig. 18. Computational time as a function of  $n$ . For the *Myers* and *Zanni* methods, multiple curves refer to different values of  $N$ . (a) *PECE* method. (b) *Myers* method. (c) *Zanni* method.

The *Zanni* method is the fastest one with a computation time always below 10 ms.

## V. CONCLUSION

In this paper, we have proposed a new method, called *PECE*, for the optimal assessment of the KF-prediction-error covariance matrix  $\mathbf{P}_{k|k-1}$  that is suitable for SE of step-varying processes where the process-noise covariances are time-varying and unknown. To the best of our knowledge, for the first time in the literature, it is proposed a method that: 1) correctly estimates the value of  $\mathbf{P}_{k|k-1}$  for a process characterized by constant noise covariances and 2) rapidly tracks the system state subject to large step-variations. We have tested the *PECE* method for an ideal case and for a realistic

application as power-system SE. On one hand, in tracking state step-variations, the *PECE* method is faster than the two covariance-estimation methods that, in this paper, are called *Myers* and *Zanni*. On the other hand, the proposed method is computationally expensive for high-dimensional systems, whereas *Myers* and *Zanni* methods are more suitable for a real-time implementation. The *PECE* method relies on the knowledge of a linear and stationary measurement model, and it has been validated for a random-walk process model. It makes use of a constrained convex optimization problem that computes  $\mathbf{P}_{k|k-1}$  from the innovations and ensures the symmetry and positive semidefiniteness of  $\mathbf{P}_{k|k-1}$ . It requires the setting of a single parameter  $N$  (the number of past innovations used to compute  $\mathbf{P}_{k|k-1}$ ), so that the parameter tuning for specific applications is simplified. A study regarding the influence of parameter  $N$  on the *PECE* accuracy has shown its robustness compared with the *Myers* method. It is also important to highlight that the *PECE* method estimates the full matrix  $\mathbf{P}_{k|k-1}$ , and we have shown that the correct assessment of the off-diagonal entries of  $\mathbf{P}_{k|k-1}$  plays an important role in the proper tracking of state step-variations. Future research will focus on a more general process model, which is a not trivial contribution that requires a thorough validation.

## APPENDIX

Here, we give the proofs of Theorems 1 and 2. We also briefly recall the covariance-estimation methods proposed by *Myers* and *Tapley* [2] and *Zanni et al.* [3], as well as the *LWLS* formulation.

### A. Proof of Theorem 1

The objective function is convex and the set of feasible  $\Sigma$  is convex [33, Ch. 7.1.1, pp. 355–357]. Furthermore, the infimum of the objective function cannot occur when  $\lambda_{\min}(\Sigma) \rightarrow 0$  (where  $\lambda_{\min}$  denotes the smallest eigenvalue) because the objective function becomes infinite when  $\lambda_{\min}(\Sigma) \rightarrow 0$ . As the feasible set is bounded (because of the condition  $\mathbf{I}_n - \Sigma \geq \mathbf{0}$ ), it follows that the optimization problem has a finite minimum, which is attained for one or several values of  $\Sigma$ . Furthermore,  $\log[\det(\Sigma)]$  is strictly concave [34, Lemma 6.2.2, p. 101] and  $\text{trace}(\Sigma\mathbf{E})$  is linear in  $\Sigma$ ; therefore, the objective function is strictly convex. It follows that the minimum is reached at one unique value of  $\Sigma$ . (QED)

### B. Proof of Theorem 2

First, note that the innovations form a Gaussian random vector with zero mean and covariance matrix given by (17). By [33, Ch. 7.1.1, pp. 355–357], it follows that the ML estimation of the covariance matrix of the innovation is obtained as the optimal value of  $\mathbf{S}$  in the following optimization problem<sup>6</sup> (recall that  $\mathbf{H}$  and  $\mathbf{R}$  are fixed and known):

$$\begin{aligned} & \min_{\mathbf{S}, \mathbf{P}} \{ \log[\det(\mathbf{S})] + \text{trace}(\mathbf{S}^{-1}\mathbf{C}_k) \} \\ & \text{s.t. } \mathbf{S} = \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R} \\ & \mathbf{P} \text{ real symmetric and } \mathbf{P} \geq \mathbf{0}. \end{aligned} \quad (26)$$

<sup>6</sup>Note that (26) is a convex problem as (18).

Second, we show that the feasible sets of the optimization problems (18) and (26) are equivalent. More precisely, there is a one-to-one mapping between any feasible  $\Sigma$  of problem (18) and a feasible  $(\mathbf{S}, \mathbf{P})$  pair of problem (26). In one direction, given  $\Sigma$ ,  $\mathbf{S}$  and  $\mathbf{P}$  are obtained by

$$\mathbf{P} = \mathbf{U}^{-1}(\Sigma^{-1} - \mathbf{I}_n)\mathbf{U}^{-T} \quad (27)$$

$$\mathbf{S} = \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}. \quad (28)$$

We need to prove that if  $\Sigma$  satisfies the conditions of problem (18), then  $\mathbf{P} \geq \mathbf{0}$ . To show this, observe that  $\Sigma \succ \mathbf{0}$  and  $\mathbf{I}_n - \Sigma \geq \mathbf{0}$ ; therefore,  $\Sigma^{-1} - \mathbf{I}_n \geq \mathbf{0}$  [35, Appendix C]. This in turn implies that  $\mathbf{P} = \mathbf{U}^{-1}(\Sigma^{-1} - \mathbf{I}_n)\mathbf{U}^{-T} \geq \mathbf{0}$ .

In the reverse direction, given  $\mathbf{S}$  and  $\mathbf{P}$ ,  $\Sigma$  is obtained by

$$\Sigma = (\mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n)^{-1}. \quad (29)$$

Similarly, we need to show that if  $\mathbf{P}$  satisfies the condition of problem (26), then  $\Sigma$  is well defined,  $\Sigma \succ \mathbf{0}$ , and  $\mathbf{I}_n - \Sigma \geq \mathbf{0}$ . First, observe that  $\mathbf{P} \geq \mathbf{0}$ ; therefore,  $\mathbf{U}\mathbf{P}\mathbf{U}^T \geq \mathbf{0}$  and  $\mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n \geq \mathbf{I}_n$ . Thus,  $\mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n \succ \mathbf{0}$  and is invertible. It follows that  $\Sigma$  is well defined and  $\Sigma \succ \mathbf{0}$ . Furthermore,  $\Sigma^{-1} = \mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n \geq \mathbf{I}_n$  and thus  $\Sigma^{-1} \leq \mathbf{I}_n$  [35, Appendix C].

Third, we show that the values of the objective functions differ by a constant when  $\Sigma$  and  $(\mathbf{S}, \mathbf{P})$  are mapped by the above correspondence. Indeed, by (28)

$$\begin{aligned} \mathbf{R}^{-\frac{1}{2}}\mathbf{S}\mathbf{R}^{-\frac{1}{2}} &= \mathbf{R}^{-\frac{1}{2}}\mathbf{H}\mathbf{P}\mathbf{H}^T\mathbf{R}^{-\frac{1}{2}} + \mathbf{I}_m \\ &= \mathbf{V} \begin{pmatrix} \mathbf{U} \\ \mathbf{0}_{m-n,n} \end{pmatrix} \mathbf{P} (\mathbf{U}^T \ \mathbf{0}_{m-n,n}) \mathbf{V}^T + \mathbf{I}_m \\ &= \mathbf{V} \begin{pmatrix} \mathbf{U}\mathbf{P}\mathbf{U}^T & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{0}_{m-n,m-n} \end{pmatrix} \mathbf{V}^T + \mathbf{I}_m \\ &= \mathbf{V} \begin{pmatrix} \mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{I}_{m-n} \end{pmatrix} \mathbf{V}^T. \end{aligned} \quad (30)$$

Therefore, noticing that  $\det(\mathbf{V})\det(\mathbf{V}^T) = 1$  and using rules for block-diagonal matrices

$$\det(\mathbf{R})^{-1}\det(\mathbf{S}) = \det(\mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n) = \det(\Sigma)^{-1}. \quad (31)$$

Taking the inverse of (30) gives

$$\begin{aligned} \mathbf{R}^{\frac{1}{2}}\mathbf{S}^{-1}\mathbf{R}^{\frac{1}{2}} &= \mathbf{V} \begin{pmatrix} \mathbf{U}\mathbf{P}\mathbf{U}^T + \mathbf{I}_n & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{I}_{m-n} \end{pmatrix}^{-1} \mathbf{V}^T \\ &= \mathbf{V} \begin{pmatrix} \Sigma & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{I}_{m-n} \end{pmatrix} \mathbf{V}^T \end{aligned} \quad (32)$$

and thus

$$\mathbf{S}^{-1}\mathbf{C}_k = \mathbf{R}^{-\frac{1}{2}}\mathbf{V} \begin{pmatrix} \Sigma & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{I}_{m-n} \end{pmatrix} \mathbf{V}^T \mathbf{R}^{-\frac{1}{2}}\mathbf{C}_k. \quad (33)$$

Using the property that  $\text{trace}(\mathbf{A}\mathbf{B}) = \text{trace}(\mathbf{B}\mathbf{A})$ , it comes

$$\begin{aligned} \text{trace}(\mathbf{S}^{-1}\mathbf{C}_k) &= \text{trace} \left( \begin{pmatrix} \Sigma & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{I}_{m-n} \end{pmatrix} \mathbf{V}^T \mathbf{R}^{-\frac{1}{2}}\mathbf{C}_k \mathbf{R}^{-\frac{1}{2}} \mathbf{V} \right) \\ &= \text{trace} \left( \begin{pmatrix} \Sigma & \mathbf{0}_{n,m-n} \\ \mathbf{0}_{m-n,n} & \mathbf{I}_{m-n} \end{pmatrix} \begin{pmatrix} \mathbf{E}'_k & \mathbf{E}''_k \\ \mathbf{E}'_k & \mathbf{E}'''_k \end{pmatrix} \right) \\ &= \text{trace}(\Sigma\mathbf{E}'_k) + \text{trace}(\mathbf{E}'''_k) \end{aligned} \quad (34)$$

where  $\mathbf{E}'_k$ ,  $\mathbf{E}''_k$ , and  $\mathbf{E}'''_k$  are subblocks of appropriate sizes and  $\mathbf{E}_k$  is defined in (20).

Putting together (31) and (34) gives the following relation between the objective functions:

$$\log \det(\mathbf{S}) + \text{trace}(\mathbf{S}^{-1}\mathbf{C}_k) = -\log \det(\Sigma) + \text{trace}(\Sigma\mathbf{E}'_k) + \alpha \quad (35)$$

where  $\alpha = \log \det(\mathbf{R}) + \text{trace}(\mathbf{E}'''_k)$  is a constant.

It follows from all the above that  $\Sigma$  is optimal for problem (18) if and only if  $(\mathbf{S}, \mathbf{P})$  is optimal for problem (26), when  $\Sigma$  and  $(\mathbf{S}, \mathbf{P})$  are mapped by (27)–(29). (QED)

### C. Covariance-Estimation Method of Myers and Tapley [2]

Myers and Tapley [2] have proposed a simple covariance-matching method to assess both  $\mathbf{Q}$  and  $\mathbf{R}$ . In this paper, we assume  $\mathbf{R}$  is known, so that only  $\mathbf{Q}$  has to be estimated. At time-step  $k$  and denoting the state estimate with  $\hat{\mathbf{x}}$ , the value of  $\mathbf{Q}_k$  is inferred by using the following procedure.

- 1) Compute  $N$  residual vectors ( $j = 1, \dots, N$ )

$$\mathbf{r}_j = \hat{\mathbf{x}}_{k-j|k-j} - \mathbf{A}_{k-j-1}\hat{\mathbf{x}}_{k-j-1|k-j-1}. \quad (36)$$

- 2) Then, compute the sample covariance matrix

$$\hat{\mathbf{M}}_k = \text{cov}(\mathbf{r}_1, \dots, \mathbf{r}_N). \quad (37)$$

- 3) Compute the estimated process-noise covariance matrix

$$\begin{aligned} \hat{\mathbf{Q}}_k &= \hat{\mathbf{M}}_k - \frac{1}{N} \sum_{j=2}^{N+1} (\mathbf{A}_{k-j}\mathbf{P}_{k-j|k-j}\mathbf{A}_{k-j}^T \\ &\quad - \mathbf{P}_{k-j+1|k-j+1}). \end{aligned} \quad (38)$$

In stationary conditions,  $\hat{\mathbf{Q}}$  approaches the true value as the parameter  $N$  increases. In case of variations of the process stochastic parameters, the smaller the  $N$  value, the faster the KF reaction.

### D. Covariance-Estimation Method of Zanni *et al.* [3]

The analytical formulation of the heuristic method for the assessment of the process-noise covariance matrix  $\mathbf{Q}$  proposed by Zanni *et al.* [3] is recalled here as follows. At time-step  $k$  and denoting the state estimate with  $\hat{\mathbf{x}}$ , the procedure to estimate  $\mathbf{Q}_k$  is the following.

- 1) Defining a vector  $\mathbf{g} \in \mathbb{R}^n$ , the  $j$ th element of  $\mathbf{g}$  is computed as the sample variance of a vector composed of the  $j$ th elements of the last  $N$  estimated states

$$\mathbf{g}(j) = \text{var}[\hat{\mathbf{x}}_{k-1|k-1}(j), \dots, \hat{\mathbf{x}}_{k-N|k-N}(j)]. \quad (39)$$

- 2) Then, the elements of  $\mathbf{g}$  constitute the diagonal of the estimated  $\mathbf{Q}_k$

$$\hat{\mathbf{Q}}_k = \text{diag}(\mathbf{g}). \quad (40)$$

Note that  $\hat{\mathbf{Q}}$  is diagonal but its elements are different from each other. The elements of  $\mathbf{g}$  increase when the state changes monotonically (increasing or decreasing). Hence, the method is able to rapidly increase  $\hat{\mathbf{Q}}$  in case of state variations. Note that the larger the  $N$  value, the larger the  $\hat{\mathbf{Q}}$  value. A small value of  $N$  can be set if a slowly-varying process has to be estimated. However, it is suggested to use few tens of state estimates in order to adapt  $\hat{\mathbf{Q}}$  in case of quick state variations.

### E. Linear Weighted Least-Squares State Estimator

The LWLS is a static state estimator as it uses only the information contained in the current measurement set  $\mathbf{z}$ . The procedure is the following.

- 1) First, compute the so-called Gain matrix

$$\mathbf{G} = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}. \quad (41)$$

- 2) Then, the estimated state is computed as

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{z}. \quad (42)$$

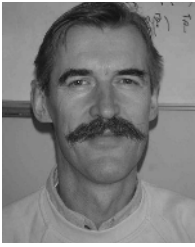
### REFERENCES

- [1] A. S. Debs and R. E. Larson, "A dynamic estimator for tracking the state of a power system," *IEEE Trans. Power Appl. Syst.*, vol. PAS-89, no. 7, pp. 1670–1678, Sep. 1970.
- [2] K. Myers and B. D. Tapley, "Adaptive sequential estimation with unknown noise statistics," *IEEE Trans. Autom. Control*, vol. 21, no. 4, pp. 520–523, Aug. 1976.
- [3] L. Zanni, S. Sarri, M. Pignati, R. Cherkaoui, and M. Paolone, "Probabilistic assessment of the process-noise covariance matrix of discrete Kalman filter state estimation of active distribution networks," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Jul. 2014, pp. 1–6.
- [4] S. Sarri, L. Zanni, M. Popovic, J.-Y. Le Boudec, and M. Paolone, "Performance assessment of linear state estimators using synchrophasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 3, pp. 535–548, Mar. 2016.
- [5] R. K. Mehra, "Approaches to adaptive filtering," *IEEE Trans. Autom. Control*, vol. 17, no. 5, pp. 693–698, Oct. 1972.
- [6] D. L. Alspach, "A parallel filtering algorithm for linear systems with unknown time varying noise statistics," *IEEE Trans. Autom. Control*, vol. 19, no. 5, pp. 552–556, Oct. 1974.
- [7] C. G. Hilborn and D. G. Lainiotis, "Optimal estimation in the presence of unknown parameters," *IEEE Trans. Syst. Sci.*, vol. 5, no. 1, pp. 38–43, Jan. 1969.
- [8] R. L. Kashyap, "Maximum likelihood identification of stochastic linear systems," *IEEE Trans. Autom. Control*, vol. 15, no. 1, pp. 25–34, Feb. 1970.
- [9] T. Bohlín, "Four cases of identification of changing systems," in *System Identification Advances and Case Studies*, vol. 126, R. K. Mehra and D. G. Lainiotis, Eds. New York, NY, USA: Academic, 1976.
- [10] V. A. Bavdekar, A. P. Deshpande, and S. C. Patwardhan, "Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter," *J. Process Control*, vol. 21, no. 4, pp. 585–601, Apr. 2011.
- [11] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Trans. Autom. Control*, vol. 15, no. 2, pp. 175–184, Apr. 1970.
- [12] B. Carew and P. Belanger, "Identification of optimum filter steady-state gain for systems with unknown noise covariances," *IEEE Trans. Autom. Control*, vol. 18, no. 6, pp. 582–587, Dec. 1973.
- [13] P. R. Bélanger, "Estimation of noise covariance matrices for a linear time-varying stochastic process," *Automatica*, vol. 10, no. 3, pp. 267–275, May 1974.
- [14] R. G. Reynolds, "Robust estimation of covariance matrices," *IEEE Trans. Autom. Control*, vol. 35, no. 9, pp. 1047–1051, Sep. 1990.
- [15] B. J. Odelson, M. R. Rajamani, and J. B. Rawlings, "A new autocovariance least-squares method for estimating noise covariances," *Automatica*, vol. 42, no. 2, pp. 303–308, 2006.
- [16] B. J. Odelson, A. Lutz, and J. B. Rawlings, "The autocovariance least-squares method for estimating covariances: Application to model-based control of chemical reactors," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 3, pp. 532–540, May 2006.
- [17] B. M. Åkesson, J. B. Jørgensen, N. K. Poulsen, and S. B. Jørgensen, "A generalized autocovariance least-squares method for Kalman filter tuning," *J. Process Control*, vol. 18, nos. 7–8, pp. 769–779, Aug./Sep. 2008.
- [18] M. R. Rajamani and J. B. Rawlings, "Estimation of the disturbance structure from data using semidefinite programming and optimal weighting," *Automatica*, vol. 45, no. 1, pp. 142–148, Jan. 2009.
- [19] Y. Zhen and J. Harlim, "Adaptive error covariances estimation methods for ensemble Kalman filters," *J. Comput. Phys.*, vol. 294, pp. 619–638, Aug. 2015.
- [20] G. Noriega and S. Pasupathy, "Adaptive estimation of noise covariance matrices in real-time preprocessing of geophysical data," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 5, pp. 1146–1159, Sep. 1997.
- [21] A. M. L. D. Silva, M. B. D. C. Filho, and J. M. C. Cantera, "An efficient dynamic state estimation algorithm including bad data processing," *IEEE Trans. Power Syst.*, vol. 2, no. 4, pp. 1050–1058, Nov. 1987.
- [22] S. Sarri, M. Paolone, R. Cherkaoui, A. Borghetti, F. Napolitano, and C. A. Nucci, "State estimation of active distribution networks: Comparison between WLS and iterated Kalman-filter algorithm integrating PMUs," in *Proc. 3rd IEEE PES Int. Conf. Exhibit. Innov. Smart Grid Technol. (ISGT Europe)*, Oct. 2012, pp. 1–8.
- [23] A. Sharma, S. C. Srivastava, and S. Chakrabarti, "Testing and validation of power system dynamic state estimators using real time digital simulator (RTDS)," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2338–2347, May 2016.
- [24] J. Zhang, G. Welch, G. Bishop, and Z. Huang, "A two-stage Kalman filter approach for robust and real-time power system state estimation," *IEEE Trans. Sustainable Energy*, vol. 5, no. 2, pp. 629–636, Apr. 2014.
- [25] F. N. Chowdhury, J. P. Christensen, and J. L. Aravena, "Power system fault detection and state estimation using Kalman filter with hypothesis testing," *IEEE Trans. Power Del.*, vol. 6, no. 3, pp. 1025–1030, Jul. 1991.
- [26] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. Comput. Aided Control Syst. Design*, Sep. 2004, pp. 284–289.
- [27] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3—A MATLAB software package for semidefinite programming, version 1.3," *Optim. Methods Softw.*, vol. 11, nos. 1–4, pp. 545–581, 1999.
- [28] A. Pai, *Energy Function Analysis for Power System Stability*. Norwell, MA, USA: Kluwer, 1989.
- [29] J. J. Grainger and W. D. Stevenson, *Power System Analysis*. New York, NY, USA: McGraw-Hill, 1994.
- [30] *IEEE Standard for Synchrophasor Measurements for Power Systems*, IEEE Standard C37.118.1-2011, (Revision IEEE Standard C37.118-2005), Dec. 2011, pp. 1–61.
- [31] P. Romano and M. Paolone, "Enhanced interpolated-DFT for synchrophasor estimation in FPGAs: Theory, implementation, and validation of a PMU prototype," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 2824–2836, Dec. 2014.
- [32] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York, NY, USA: Marcel Dekker, 2004.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [34] B. Gärtner and J. Matousek, *Approximation Algorithms Semidefinite Programming*. New York, NY, USA: Springer, 2012.
- [35] A. Van den Bos, *Parameter Estimation for Scientists Engineers*. New York, NY, USA: Wiley, 2007.



**Lorenzo Zanni** (M'13) was born in Italy in 1988. He received the B.Sc. and M.Sc. (Hons.) degrees in electrical engineering from the University of Bologna, Bologna, Italy, in 2010 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Power-System Group and the Distributed Electrical System Laboratory, Swiss Federal Institute of Technology of Lausanne, Switzerland.

His current research interests include real-time power-system monitoring, control and protections by using PMUs with particular focus on state estimation, and fault location.



**Jean-Yves Le Boudec** (F'04) received the Aggregation degree in mathematics from the École Normale Supérieure de Saint-Cloud, Paris, France, in 1980, and the Ph.D. degree from the University of Rennes 1, Rennes, France, in 1984.

From 1984 to 1987, he was with the Institut National Des Sciences Appliquées/Institut de Recherche en Informatique et Systèmes Aléatoires, Rennes. In 1987, he joined the Bell Northern Research, Ottawa, ON, Canada, as a member of scientific staff with the Network and Product Traffic Design Department. In 1988, he joined the IBM Zurich Research Laboratory, Rüschlikon, Switzerland, where he was the Manager of the Customer Premises Network Department. In 1994, he became an Associate Professor with the Swiss Federal Institute of Technology of Lausanne, where he is currently a Professor. He has co-authored a book on network calculus, which forms a foundation to many traffic control concepts in the internet, an introductory textbook on information sciences, and is the author of the book *Performance Evaluation*. His interests are in the performance and architecture of communication systems and smart grids.



**Mario Paolone** (M'07–SM'10) received the M.Sc. (Hons.) and Ph.D. degrees in electrical engineering from the University of Bologna, Bologna, Italy, in 1998 and 2002, respectively.

In 2005, he was an Assistant Professor in power systems with the University of Bologna, where he was with the Power Systems Laboratory until 2011. He is currently an Associate Professor with the Swiss Federal Institute of Technology, Lausanne, Switzerland, where he accepted the Chair of the Distributed Electrical Systems Laboratory. He has authored or co-authored over 230 scientific papers published in reviewed journals and international conferences. His current research interests include power systems with particular reference to real-time monitoring and operation of active distribution networks, integration of distributed energy-storage systems, power-system protections, and power-system transients.

Dr. Paolone received the IEEE EMC Society Technical Achievement Award in 2013. He is the Editor-in-Chief of the journal *Sustainable Energy, Grids and Networks*.



**Rachid Cherkaoui** (M'05–SM'07) received the M.Sc. and Ph.D. degrees in electrical engineering from the Swiss Federal Institute of Technology of Lausanne (EPFL), Lausanne, Switzerland, in 1983 and 1992, respectively.

He is currently a Senior Scientist with EPFL, leading the Power Systems Group. He was the IEEE Swiss chapter officer from 2005 to 2011. He has authored and co-authored over 100 scientific publications. His current research interests include electricity market deregulation, distributed generation and storage, and power system vulnerability mitigation.

Dr. Cherkaoui was a member of International Council on Large Electric Systems task force's and working group's. He is a member of technical program committees of various conferences and