IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 12, 2017, accepted September 5, 2017, date of publication September 19, 2017,
date of current version October 12, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2754346

Identifying Uncertainty Distributions and Confidence Regions of Power Plant Parameters

TETIANA BOGODOROVA[1], (Student Member, IEEE), LUIGI VANFRETTI[2], (Senior Member, IEEE), VEDRAN S. PERIĆ[3], (Member, IEEE), AND KONSTANTIN TURITSYN[4], (Member, IEEE)

[1]KTH Royal Institute of Technology, 114 28 Stockholm, Sweden
[2]Electrical, Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180 USA
[3]GE Energy Consulting, 80807 Munich, Germany
[4]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139-4307 USA

Corresponding author: Tetiana Bogodorova (tetianab@kth.se)

This work was supported by the EU through FP7 *iTesla* project.

ABSTRACT Power system operators, when obtaining a model's parameter estimates; require additional information to guide their decision on a model's acceptance. This information has to establish a relationship between the estimates and the chosen model in the parameter space. For this purpose, this paper proposes to extend the usage of the particle filter (PF) as a method for the identification of power plant parameters; and the parameters' confidence intervals, using measurements. Taking into consideration that the PF is based on the Bayesian filtering concept, the results returned by the filter contain more information about the model and its parameters than usually considered by power system operators. In this paper the samples from the multi-modal posterior distribution of the estimate are used to identify the distribution shape and associated confidence intervals of estimated parameters. Three methods [rule of thumb, least-squares cross validation, plug-in method (HSJM)] for standard deviation (bandwidth) selection of the Gaussian mixture distribution are compared with the uni-modal Gaussian distribution of the parameter estimate. The applicability of the proposed method is demonstrated using field measurements and synthetic data from simulations of a Greek power plant model. The distributions are observed for different system operation conditions that consider different types of noise. The method's applicability for model validation is also discussed.

INDEX TERMS Confidence intervals, measurements, parameter identification, particle filters, power plant models, power system identification, power system model validation, power systems.

I. INTRODUCTION

Accurate modeling of electric power system components is essential for power system operators. However, maintaining accurate dynamic models represents large technical and practical challenges for utilities, including those related to data management and analysis. The conventional approach to obtain dynamic models is through physical modeling, where values of parameters are obtained from name tables. This approach is relatively simple, but can lead to inaccurate model responses due to uncertain changes that parameters undergo through time. Model accuracy can be improved by periodically performing staged identification tests where parameters can be updated. However, the logistics and costs associated to these tests are non-negligible because staged tests require components to be disconnected from the network or other interventions that may hinder the normal operation of the power system. An approach to address these problems has been found in the application of modern system identification theory for parameter identification and calibration. Development of synchrophasor technology has made this approach even more appealing due to the increasing availability of synchronized high sample rate measurement data capturing the systems' dynamic response.

A. LITERATURE REVIEW

From the mathematical perspective, different methods for identification of the power system's parameters have been proposed [1]–[5]. Jones [1] focused on the use of the autoregressive moving average with integrator in noise model and exogenous inputs to represent a power system using a low-order dynamic model. Stefopoulos *et al.* [3] have shown the benefits of applying genetic algorithms for governor-turbine dynamic model identification in multi-machine power systems. In [4], researchers have successfully applied the

extended Kalman filter for generator dynamic model validation using phasor measurements. Zhou *et al.* in [5] extended a particle filter to estimate the dynamic states of generators.

The particle filter (PF) has a number of benefits that makes it attractive for parameter identification. A PF provides the parameter uncertainty distribution of estimated parameters by a set of samples (particles). In this way, this property makes a PF applicable for nonlinear system identification and robust to different types of noise, including non-Gaussian noise. As a result, this technique does not impose any constraints on the model structure used, meaning that the parameter set defining a power plant model can be identified entirely.

Parameter estimates are most meaningful when accompanied with their statistical properties, i.e. mean and confidence intervals. Such properties were received in the asymptotic variance evaluation of estimated modes that is extracted directly from auto regressive moving average model [6].

Meanwhile, in [5] the statistical properties were received for distribution of mean values evaluated using Monte Carlo simulations in order to verify the PF method. In this way, the estimated mean and standard error characterized the accuracy of the method, but not the internal dynamics of the nonlinear power system.

B. CONTRIBUTION

The main goal of this paper is to derive new knowledge for power system operators about the power system parameters that are estimated given measurements and a model. The ability to derive an estimated parameter uncertainty using an arbitrary distribution shape has not been analyzed in power systems. The current methodology used by power system operators limits their ability to determine parameter estimates relationships and consequently, design model validation experiments that consider the implications of these relationships. To address this gap, the reconstruction of the estimated parameter posterior distribution and confidence intervals is proposed.

The proposed methodology is based on particle filtering, it is applicable for non-linear systems, and robust to arbitrary noise characteristics. These features represent an enhancement in comparison to the previously proposed methods [4] that assume linear(-ized) models with Gaussian distributed noise.

The main benefit of the PF is that it can handle multimodal distribution of parameter estimates. Therefore, the proposed method calculates the uncertainty of parameter estimates in the form of multimodal (mixture) Gaussian distributions. The estimate distribution's reconstruction for Gaussian mixture distribution and its confidence regions calculation is presented in detail. Three methods (Rule of Thumb (ROT), Least-Squares Cross Validation (LCV), Plug-in method (HSJM)) for standard deviation (bandwidth) selection of the Gaussian mixture distribution are compared.

The performance of the Gaussian mixture distribution reconstruction is evaluated by comparing the confidence intervals estimated using the proposed method that exploits Gaussian multivariate mixture distribution and a Monte-Carlo method that approximates the uncertainty distribution with a uni-modal Gaussian distribution form. Finally, the estimates of the reconstructed distributions are compared to the actual parameter values. The proposed method can be parallelized to increase computational efficiency and applicable for online applications.

The remainder of this paper is organized as follows. Section II gives the problem formulation. The proposed particle filter identification method is formulated in Section III and expanded to evaluate confidence intervals in Section IV. The application of the proposed method is shown and discussed in Section V, whereas conclusions are drawn in Section VI.

II. PROBLEM FORMULATION

A. PROBLEM STATEMENT

To identify the model parameters and their confidence regions, assuming the model structure of a power plant that includes a generator with controls (Fig. 1), and having measurements from the tests and on the terminal bus.

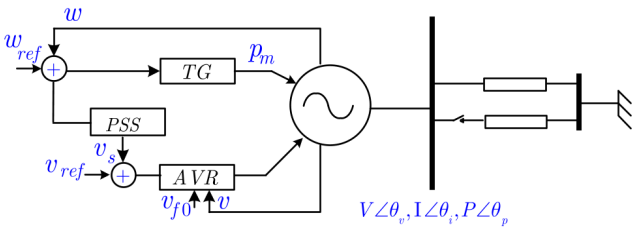

FIGURE 1. Generator with controllers connected to an infinite bus.

The formulated problem is illustrated on the model that consists of a synchronous machine GenVI [7], turbine governor TGI [7], automatic voltage regulator AVRIII [7] and power system stabilizer PSSIII [7]. The generator is connected to the infinite bus through transmission lines. The entire model[1] can be described in state space form, as follows:

$$\frac{dx}{dt} = f(x(t), \rho, u, \xi) \tag{1}$$

$$y = h(x(t), \rho, u, \sigma), \tag{2}$$

where $\{x, \rho\}$- states/parameters (some of them to be estimated); σ - measurement noise, ξ - process noise; and u - input (e.g. terminal voltage), y - measurements.

Given the discrete measurements (input and output) $u_n = \{u_i, i = 1..n\}$ and $y_n = \{y_i, i = 1..n\}$ obtained every $i\Delta t$, where $i = 1, 2...n$, the model can be discretized:

$$x_n = f(x_{n-1}, \rho_{n-1}, u_{n-1}\xi_{n-1}) \tag{3}$$

$$y_n = h(x_n, \rho_n, u_n, \sigma_n), \tag{4}$$

[1]The model that has been used for the case studies has been developed in Modelica. All the model components are included into OpeniPSL library and can be found in open access [8].

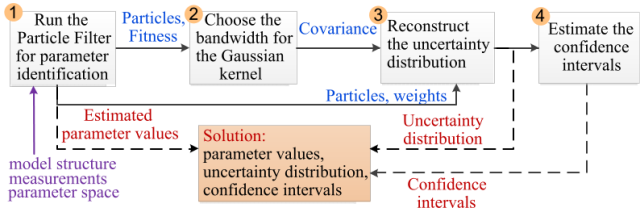

FIGURE 2. Global algorithm of the uncertainty distributions and confidence intervals identification.

III. IDENTIFICATION APPROACH

To address the problem formulated in Section II, the algorithm depicted in Fig. 2 have been developed:

1) Given measurements, the assumed model structure and the parameter space that defines the limits in which the parameters can vary, execute particle filter routines to identify parameter values. The output of the filter are particles that are coordinates in parameter space, their weights and fitness (Section III-A).
2) Choose the covariance for the selected Gaussian kernel[2] using one of the available algorithms (see Section IV-B) that exploit the particles and their weights, received as an output of the particle filter.
3) Using the selected covariance, particles that define mean values for Gaussian mixture, and weights, reconstruct the uncertainty distribution for the estimated set of parameters (Section IV-A).
4) Marginalize[3] the uncertainty distribution with respect to each parameter to receive a univariate pdf. The latter depends only on one parameter and defines the uncertainty interval of this parameter.

A. PARTICLE FILTER AS A BAYESIAN FILTERING IMPLEMENTATION

Bayesian filters, the class of filters to which PF belongs, recursively estimates a belief (pdf) in the parameters[4] $\{x_n\}$ [9] by using all available information about the system's structure and measurements (3), (4).

The algorithm starts by assuming that the initial pdf that is called prior, $p(x_0|y_0) = p(x_0)$, is given. Then the final pdf that is often referred as a posterior, $p(x_n|y_{1:n})$, is iteratively constructed. This iterative process is a stochastic Markovian process itself, where the next $(n+1)$ parameter set is determined by the pdf at point (n) given the model of the system (prediction step) and real system measurements (update step). PF proceeds in this manner through discrete approximations the posterior pdf $p(x_{n+1}|y_n)$ by a set of random samples (particles) of a parameter space $\{x_n^{(i)}\}$ drawn from the posterior pdf. Each particle is a concrete instantiation of the state/parameter value at time t. Let N be a number of particles the parameter space $x_n^{(i)}$ is defined with, so that the *prediction* step of discrete approximation is described by the Chapman-Kolmogorov equation, [10]:

$$p(x_n|y_{1:n-1}) = \sum_{i=1}^{N_{n-1}} p(x_n|x_{n-1}^{(i)}) p(x_{n-1}^{(i)}|y_{1:n-1}), \tag{5}$$

where $p(x_n|x_{n-1}^{(i)})$ - transition density in importance function. The result of prediction step at $(n+1)$ is improved by using Bayes' rule once the measurements (y_n) at point $(n+1)$ become available.

$$p(x_n|y_{1:n}) = \Omega_n^{-1} p(y_n|x_n) p(x_n|y_{1:n-1}) \tag{6}$$

The normalizing constant C can be evaluated using:

$$\Omega = p(y_n|y_{1:n-1}) = \sum_{i=1}^{N_{n-1}} p(y_n|x_n^{(i)}) p(x_n^{(i)}|y_{1:n-1}). \tag{7}$$

The likelihood function $p(y_n|x_n)$ is represented through the measurement equation (4), where the properties of the measurement noise are known.

In Algorithm 1, the action of assigning and updating weights $\omega_0^{(i)}$ for each particle is called weighting. The weighting is specific and based on the value of the fitness function that is defined as the relative squared difference between simulated and real measurements in Algorithm 1. Transition density $p(x_n|x_{n-1}^{(i)})$ can be obtained from model structure by passing samples trough equation (3).

IV. EXTENSIONS TO THE PARTICLE FILTER ALGORITHM

In this section the estimate distribution with definition of all the parameters to be described. Ability to reconstruct the full distribution gives benefit to engineer to observe peaks and hollows meaning probability mass allocation of the estimate, in other words, local optimums. In this case, for nonlinear model the distribution to be reconstructed has non-Gaussian form, therefore, the confidence intervals quantitatively can't be estimated only by one covariance matrix in the closed form.

A. ESTIMATE DISTRIBUTION RECONSTRUCTION

The posterior distribution that is an output of particle filter is described by particles with weights assigned according to a fitness function that is presented in Algorithm 1. The continuous pdf can be constructed from PF by assigning so called kernel function to each particle. The solution is to use each particle as the center of the kernel and the overall density will be composed as a mixture of the kernel densities. In this particular case, a Gaussian kernel was chosen, and thus, the resulting distribution can be estimated using Gaussian mixtures, as illustrated in Fig. 3:

$$p(x) = \sum_{n=1}^{N} w_n \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_n, \Sigma_n), \tag{8}$$

[2]Gaussian kernel meant to be a normalized sum of Gaussians centralized around each particle and scaled by particle weight.In that way it forms the final uncertainty probability density function (pdf).

[3]Marginalization is an act of retrieving the pdf for a parameter discarding other parameters from the joint parameter pdf.

[4]The states are variable parameters that can be included into estimation process. Therefore, for simplicity lets assign $x_n{}^{(i)}$ to a set of parameters that are aimed to be identified.

Algorithm 1 Particle Filter

procedure PF($p_{min}, p_{max}, N, \epsilon, n_{it}$)
  while $n < n_{it}$ and {stop criteria} **do**
    Step 1. Initialization (sampling from uniform prior pdf):
    for $i = 1$ to N **do**
      draw the samples $x_n^{(i)} \propto p(x_0)$
      $\omega_0^{(i)} \leftarrow 1/N$.
    Step 2. Importance Sampling:
    for $i = 1$ to N **do**
      draw samples $\hat{x}_n^{(i)} \propto p(x_n|x_{n-1}^{(i)})$,
      $\hat{x}_n^{(i)} \leftarrow \{x_{0:n-1}^{(i)}, \hat{x}_n^{(i)}\}$
    Step 3. Weight update with normalization:
    for $i = 1$ to N **do**
      $RSSD(y, \hat{y}) \leftarrow \frac{1}{L}\sum_i (\frac{y_i - \hat{y}_i}{y_i})^2$
      $fitness \leftarrow RSSD(y, \hat{y})$
      $\omega_n^{(i)} \leftarrow 1 - \frac{fitness}{\sum_i fitness}$
    Step 4. Resampling:
    Generate/prune particles $x_n^{(i)}$ from $\{\hat{x}_n^{(i)}\}$
    according to $\omega_n^{(i)}$/(prune if $\omega_n^{(i)} < \epsilon$)
    to obtain N random samples

where $[p_{min}, p_{max}]$ - range of parameters' values which define parameter space, N - number of particles used to fill the space of parameter values; Y - real measurements; $\hat{Y}$ - estimate; ϵ - prune threshold (defines which percentage of particles with lowest weight will not survive); {stop criteria} - set of conditions to finish the main procedure cycle; L - number of measurement instances.

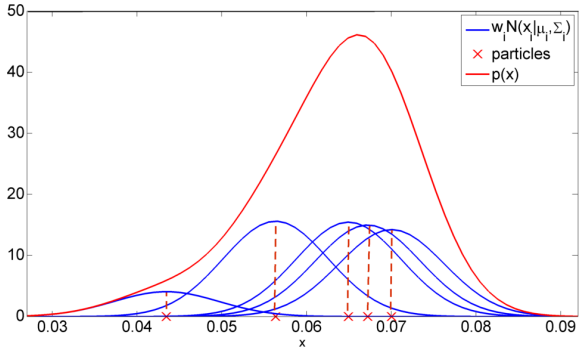

FIGURE 3. Gaussian mixture distribution reconstruction [illustration for equation (8)].

where $\sum_{n=1}^{N} w_n = 1$, $0 \leqslant w_n \leqslant 1$ - normalization of N individual M-dimentional Gaussian components $\mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_n, \Sigma_n)$, where M - number of parameters.

Each component is a multivariate Gaussian distribution given by:

$$\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \Sigma) \stackrel{def}{=} \frac{1}{(2\pi)^{M/2}|\Sigma|^{1/2}} \times \exp\Big(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big) \quad (9)$$

where $\boldsymbol{x} = [x_1 .. x_M]$ - vector of parameters.

In (9), $\Delta^2 = (\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ represents the squared Mahalanobis distance [11] between x and μ. When Δ^2 is equal to a constant, it will define ellipsoids that result in locus of the equal density level of each Gaussian component. This means that it defines the confidence contours on the required, user-predefined confidence region percentage (e.g. 95%).

Decomposing the covariance in (9) to eigenvalues and eigenvectors using the Cholesky decomposition [12] ($\Sigma = U\Lambda U^T$) gives the following:

$$\Sigma^{-1} = U^{-T}\Lambda^{-1}U^{-1} = U\Lambda^{-1}U^T = \sum_{m=1}^{M}\frac{1}{\lambda_m}\boldsymbol{u}_m\boldsymbol{u}_m^T \quad (10)$$

where U - eigenvectors, Λ - diagonal matrix of eigenvalues $[\lambda_1..\lambda_m]_{m=1:M}$.

Substituting (10) in the squared Mahalanobis distance of (9), it can be shown that:

$$(\boldsymbol{x}-\boldsymbol{\mu})^T\Big(\sum_{m=1}^{M}\frac{1}{\lambda_m}\boldsymbol{u}_m\boldsymbol{u}_m^T\Big)(\boldsymbol{x}-\boldsymbol{\mu}) = \sum_{m=1}^{M}\frac{y_m^2}{\lambda_m} \quad (11)$$

where $y_m \stackrel{def}{=} \boldsymbol{u}_m^T(\boldsymbol{x}-\boldsymbol{\mu})$.

The confidence space[5] of a parameter vector $\boldsymbol{x}$ can be defined as the union of M-dimentional ellipsoids (see Appendix A). The number of the ellipsoids corresponds to the number of Gaussians in a Gaussian Mixture model, which is created locating the means of Gaussian components in each PF particle.

From (11), the M-dimensional ellipsoid described by a constant Mahalanobis distance, can be presented as follows:

$$S_i = \sum_{m=1}^{M}\frac{y_m^2}{\lambda_m} = a \quad (12)$$

where a - defines the scale of the ellipse.

The squared Mahalanobis distance Δ^2 is a sum of squared M Gaussian data samples that are distributed according to the Chi-Square distribution (χ^2) with M degrees of freedom. Therefore, one has to find the probability $P(x \leq a_t) = I$ where x is less than or equal to a specific value, which can easily be obtained using the cumulative Chi-Square distribution.

Note that the union of the ellipses is given by:

$$S = \bigcup_{i=1}^{N} S_i \quad (13)$$

where N is the number of particles.

B. COVARIANCE SELECTION

Covariance in the Gaussian kernel (8) plays the role of a smoothing parameter that defines the shape of the reconstructed estimate distribution. Small covariance values Σ_n of each M-dimensional Gaussian component increase the asymptotic variance, meaning that the resulting distribution

[5]The confidence space is referred as confidence interval when the space is of 1-dimension.

shape will be "wavy" with many extra graphical features, such as peaks and cavities. On the other hand, big values of the covariances reduce the (asymptotic) variance but increase the (asymptotic) bias, thus smoothing away the true distribution peaks [13].

In this paper three methods for bandwidth selection are applied and compared: Rule of Thumb, Least-Squares Cross Validation, and the Plug-in method [14].

1) RULE OF THUMB (ROT)

This method is based on the asymptotic mean squared error estimator $AMISE(h)$. The optimal bandwidth $\hat{h}_{rot}$ obtained from differentiating $AMISE(h)$ w.r.t. standard deviation (bandwidth)[6] h and calculating root of the derivative. The result is simplified substituting the unknown true density function $f(x)$ with a standard normal distribution rescaled to have variance equal to the sample variance $\hat{\sigma}^2$:

$$\begin{aligned} \hat{h}_{rot} &= \underset{h}{\operatorname{argmin}}\, AMISE(h) = 1.06\hat{\sigma}N^{-\frac{1}{5}} \\ AMISE(h) &= (Nh)^{-1}\int K^2(x)dx \\ &\quad + \frac{h^4}{4}\Big(\int x^2K(x)dx\Big)^2\int\Big(\frac{d^2f(x)}{dx}\Big)^2dx \end{aligned} \tag{14}$$

where $K_i = N\omega_i\mathcal{N}(x_i|\mu_i, h_i)$- the kernel function[7]; $\hat{f}_h(x) = \frac{1}{N}\sum_{i=1}^{N} K_i$ - density function estimator. Note: Integration is performed over the x parameter space.

2) LEAST-SQUARES CROSS VALIDATION (LCV)

This method is based on the integrated squared error function $ISE(h)$.

$$\begin{aligned} ISE(h) &= \int \big(\hat{f}_h(x) - f(x)\big)^2dx \\ &= \int \hat{f}_h^2(x)dx - 2\int \hat{f}_h(x)f(x)dx + \int f^2(x)dx \end{aligned}$$

Due to the third term of $ISE(h)$ doesn't depend on h, one can use Least-Squares Cross-Validation function $LSCV(h)$ as an estimator for $ISE(h) - \int f^2(x)dx$:

$$\hat{h}_{LSCV} = \underset{h}{\operatorname{argmin}}\, LSCV(h) \tag{15}$$

$$= \underset{h}{\operatorname{argmin}}\Big(\int \hat{f}_h^2(x)dx - 2\sum_{i=1}^{N}\hat{f}_h\Big) \tag{16}$$

3) PLUG-IN METHOD (HSJM)

The main idea of the HSJM[8] method is to take one further term Ψ in the asymptotic expansion of the integrated squared bias:

$$\hat{h} = \underset{h}{\operatorname{argmin}}\, AMISE_2(h) = \underset{h}{\operatorname{argmin}}(AMISE(h) - \Psi) \tag{17}$$

where

$$\Psi = \frac{1}{24}h^6\int x^2K(x)dx\int x^4K(x)dx\int\Big(\frac{d^3f(x)}{dx}\Big)^2dx \tag{18}$$

The minimizer (17) is not easy to calculate, therefore, the asymptotically equivalent $\hat{h}_{HSJM}$ is used:

$$\hat{h}_{HSJM} = \Big(\frac{\int K^2(x)dx}{N\Big(\int x^2K(x)dx\Big)^2\hat{I}_2}\Big)^{\frac{1}{5}} \tag{19}$$

$$+ \frac{\int x^4K(x)dx\hat{I}_3}{20\int x^2K(x)dx\hat{I}_2}\Big(\frac{\int K^2(x)dx}{N\Big(\int x^2K(x)dx\Big)^2\hat{I}_2}\Big)^{\frac{3}{5}} \tag{20}$$

where $\hat{I}_2$ and $\hat{I}_3$ are functionals that depend on true density function derivatives of 2 and 3 order respectively and should be estimated (for more details refer to of [16, Sec. 3]).

C. CONFIDENCE REGION (INTERVALS) IDENTIFICATION

The problem of the mixture of multivariate Gaussians lies in the fact that the maximum likelihood solution for the parameters no longer has a closed-form analytical solution. This means that the moments of the distribution (mean, variance, peaks) need to be found numerically as a superposition of weighted multivariate Gaussians.

In this context, the task of a required confidence space allocation, and consequence confidence intervals allocation, is transformed into the following problem:

Find the cutting surface $C : p(x) = b$ that intersects with the Gaussian mixtures pdf $p(x)$ given by (8) and gives a projection contour area on $(x_1, .. , x_M)$ coordinates plane equal to S, where S - union of ellipses.

To find the confidence interval for each estimated parameter, the Gaussian mixture pdf has to be marginalized, thus defining the uncertainty distribution for each parameter x_i.

V. CASE STUDIES

The following case studies were performed to assess the proposed algorithm, which includes posterior distribution and confidence intervals estimation. For this purpose the RaPId Toolbox [17] was extended by implementing the proposed theory in Section IV. The simulation data was generated using the FMI Toolbox for Matlab allowing for Simulink usage.

The single-machine infinite-bus power system (Fig. 1) was modeled using Modelica and used to simulate the system's response. RaPId [18] allows the parameters estimation of the generator and its controls. To excite the system a probing signal was given to reference inputs and a tree-phase fault was applied. The parameter values used to initialize the algorithm were chosen according to known ranges that are available for different components in the literature [19]. Hence, the prior pdf was chosen as uniform in the parameter space defined for initialization.

[6] In statistical research, the standard deviation that has to be estimated is referred as bandwidth [13].

[7] For simplicity, the kernel and the estimators are presented for 1-d space kernel. For multidimentional representation, it is recommended to refer to [15].

[8] The abbreviation refers to Hall, Sheather, Jones, and Marron's work in [16].

Each case study subsection is divided into parts:

- **Model**: It describes the model setup for the particular case study. The parameters that are estimated and their range of values. Also, the excitation input signal is presented if any is applied. For each case the model in Fig. 1 with parameters in Appendix B is used.
- **Measurements**: The measurements that are used as output signals used in evaluation of fitness in Algorithm 1.
- **Choice of the number of particles, Choice of Σ, etc.**. Those contain the results analysis of the global algorithm steps corresponding to Fig. 2

The results for all cases are summarized in Table 4.

A. CASE 1: ILLUSTRATIVE EXAMPLE USING MEASUREMENT DATA FROM STAGED TESTS

This case aims to illustrate the proposed method showing: a) the reconstruction of the uncertainty pdf of the estimated parameters, b) the identification of the confidence regions of a model estimated parameters. This case targets have been illustrated using the identification of confidence regions in a 2-dimensional form. In this case the field measurements have been used.

Measurements: The measurements used in this example come from staged tests that were performed on a turbine-governor set to determine the response of the turbine to system frequency deviations, thus the measured output data is the turbine mechanical power output p_m.

Model: The turbine parameters of the power plant system model that most influence the power output are the droop (R) and governor time constant (T_s). Hence, they were chosen for estimation. For this case study the power plant was operated at different power dispatch levels by varying the initial power p_{in}^* [20]. The initial power was measured in percentage of the so-called Maximum Continuous Rating (MCR) that refers to the gas turbine output at which it enters into the temperature limit control regime under normal air (temperature/humidity) ambient conditions, which is 95% of MCR for this particular case. An incremental signal ($\Delta\omega = +0.2\,Hz\,(0.004\,p.u)$) was injected to the frequency reference input of the governor to mimic the effect of a variation in the system's frequency.

Initial parameter values were sampled from a uniform distribution on the space $x_{min}(R, T_s) = [0.03\ 0.6]$, $x_{max}(R, T_s) = [0.08\ 1.5]$. To define the appropriate number of particles for estimation, experiments that compare the error value dependency on the number of particles were performed. The results from the experiments are shown in Figs. 4-7.

1) CHOICE OF THE PARTICLES NUMBER

The computational complexity of the Particle Filter highly increases with the increase of chosen number of particles (see Table 1) [21]. In addition, complexity of the inverse computation of Σ rises with the increase of dimension. For this reasons, it is important to choose an optimal number of particles for the selected case. In Fig. 4(a) the estimation error of the algorithm steeply drops with the increase of the number of particles. After $N = 10$, the estimation error (fitness), which is presented in Fig. 4(b), remains approximately the same, i.e. equal $6 \cdot 10^{-6}$. Therefore, $N = 10$ is chosen.

TABLE 1. Computational complexity evaluation (Case 1).

	N=5	N=10	N=20	N=30	N=50	N=70	N=100
Time, s	60.03	78.76	143.67	176.67	274.79	391.45	588.04

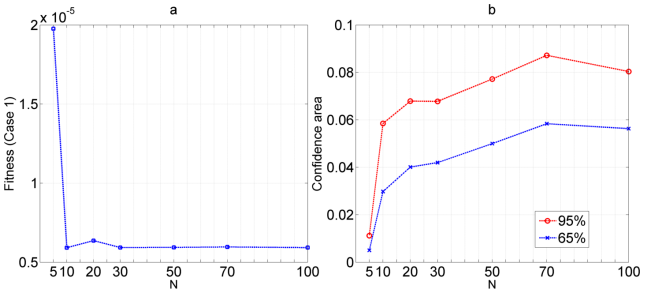

FIGURE 4. (a) Estimation error (fitness) and (b) Confidence area dependency on number of particles (Case 1 in Table 4).

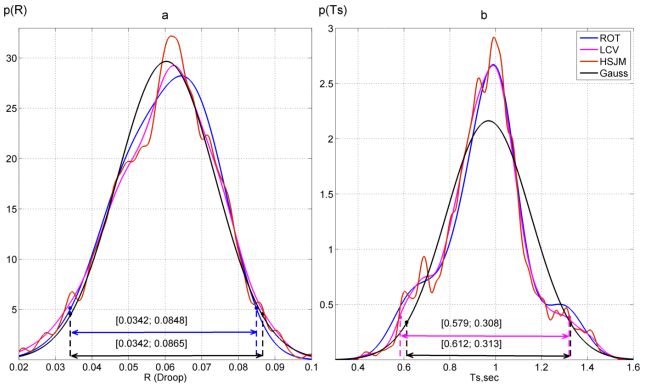

FIGURE 5. Marginalized estimate uncertainty pdf(s) and σ choice (Case 1 when N=10 in Table 4).

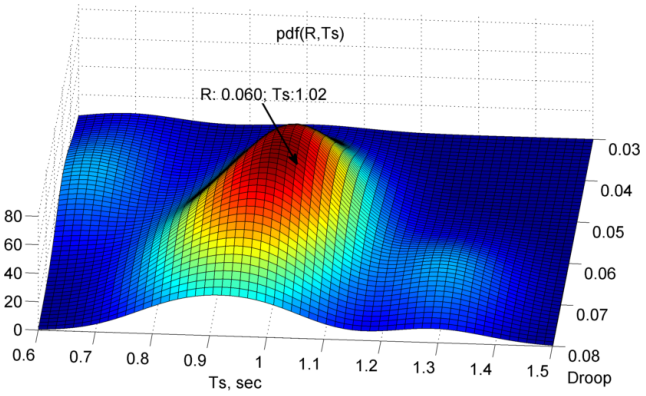

FIGURE 6. Reconstructed estimate uncertainty probability density function (Case 1 when N=10 in Table 4).

2) CHOICE OF Σ

To reconstruct the estimate uncertainty distribution (given particles and their weights after the Particle Filter run), the bandwidth of the kernel has to be estimated (see Fig. 2).

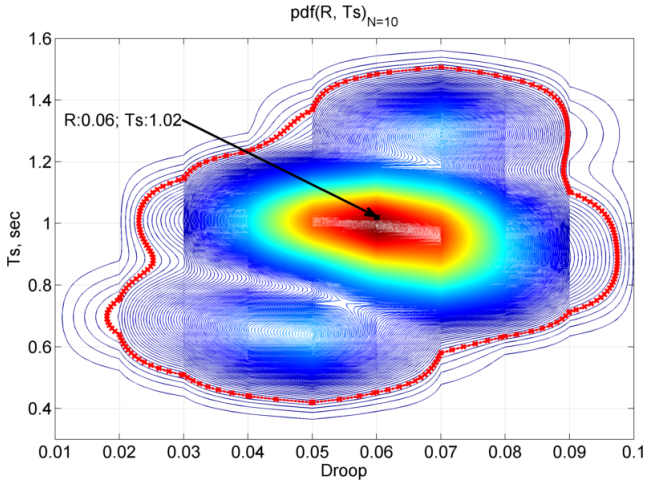

FIGURE 7. 95% confidence region of [R, Ts] parameters uncertainty of estimate (corresponds to Case 1 when N=10 in Table 4).

TABLE 2. Estimates of reconstructed pdf (Fig. 5).

Method	$mean(p(R))$	$mean(p(Ts))$
ROT	0.064	0.99
LCV	0.062	0.99
HSJM	0.061	0.99
Gauss	0.060	0.97
Real value	0.060	1

In the case of the Gaussian kernel, the standard deviation is such so-called ''bandwidth'', that has to be estimated. It is assumed that each component of the Gaussian mixture has the same bandwidth and that all the estimated parameters are independent. In other words, Σ (8) is a M xM diagonal matrix, where each diagonal value corresponds to the dimension of each parameter. For this purpose three methods for data-based bandwidth selection were applied using the KDE Toolbox for MATLAB [22].

The methods for droop estimation of the same performance in terms of confidence in estimates are the Rule of Thumb method and the Least-Squares Cross Validation method. The Rule of Thumb method minimizes the sample variance while maximizing confidence. Least-Squares Cross Validation provides a more prominent peak at the parameter value. As it can be observed in Fig. 5, ROT smooths out the distribution features more than other methods, and therefore it is not appropriate for estimation of ''wavy'' distributions with two peaks close to each other, which is not the case in this example. The LCV method is more sensitive to variance than ROT, but less sensitive than HSJM (Fig. 5). The best method for the time constant uncertainty pdf is LCV, as it gives maximal confidence (Fig. 5).

Another way to evaluate the performance of the proposed methods is to compare their estimates to the real (known) parameter value. Thus, after reconstructing the distributions (5), the mean values have been compared. The result shows that all the values are close to the real value, but in the case of the case of droop R estimate, the best performance is shown by Gauss method, while the ROT gives the worst. The opposite behavior of Gauss and ROT has been observed in the case of time constant estimate Ts. The closest estimates to the real values have been given by HSJM method.

The effectiveness of the algorithm can be assessed by comparing the confidence regions estimated using a Gaussian distribution and the proposed method. Under the assumption that the samples of the posterior distribution are from the Gaussian distribution, the confidence region can be evaluated. For each parameter, i.e. droop [Fig. 5(a)] and time constant [Fig. 5(b)], a comparison between the best estimated bandwidth using the proposed methods and the bandwidth estimated using a Gaussian distribution is shown. It can be concluded that the Gaussian distribution simplifies the shape and it is valid only for rough approximations, while the proposed methods provide increased shape resolution and an increase of confidence for smaller percentages of probability. For example, in the case of 95% probability, the right value lies in the estimated interval, the difference in the intervals estimated from the compared methods is negligible; however, for 50% probability, the intervals shrink considerably.

3) ESTIMATE UNCERTAINTY DISTRIBUTION AND CONFIDENCE REGION RECONSTRUCTION

For the estimated characteristics (w, μ, Σ, N), the distribution (8) is reconstructed (Fig. 6) and the confidence region ((13) and Section IV-C) is defined. The estimated and the real parameters values lay in the identified confidence region. It is worth to notice that the real values of the identified turbine governor parameters lay at the peak of the reconstructed distribution. Considering the high precision or, namely, small estimation error of the Particle Filter, an engineer can use smaller than 95% confidence interval when defining the parameter's variance.

4) VERIFICATION OF THE PARTICLE NUMBER CHOICE

To verify the choice of the number of particles, the confidence regions of 95% and 65% were estimated for particle numbers in the range of [5..100]. Results are presented in Fig. 4. The shape of the distribution, created by increasing the number of particles, causes an increase in the area of the confidence regions as shown in Fig. 4. An adequate number of particles must correspond to a minimal fitness and a small uncertainty or confidence area. Therefore, $N = 10$ is an adequate number for this case study.

To analyze the dependency of confidence intervals to the number of particles for each estimated parameter, 95% confidence intervals were extracted by marginalizing the reconstructed distributions. Fig. 8 shows the estimates with respect to the real value of the parameters and the 95% confidence region limits for each parameter. The 95% confidence region for the T_s estimate grows as the number of particles increases from $N = 5$ to $N = 70$, while there is no clear dependency for the 95% confidence region of R estimate.

These results imply that the droop R can be precisely estimated when $N > 5$, while the time constant's T_s value is

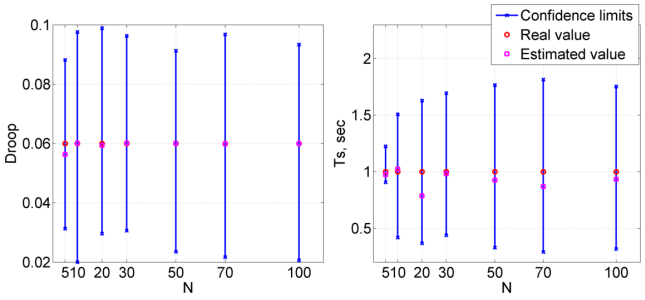

FIGURE 8. Confidence limits of the 95% probability for each parameter vs the number of particles (N) (Case 1 in Table 4).

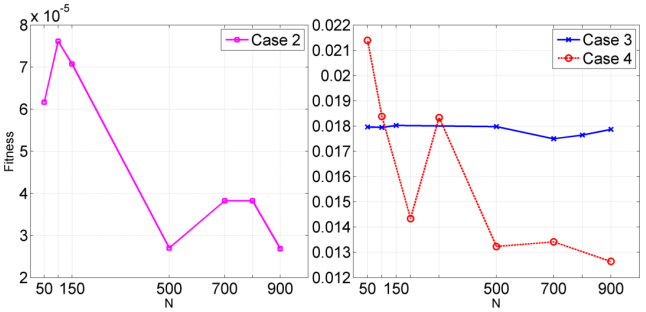

FIGURE 9. Estimation error (fitness) dependency on number of particles (Cases 2-4 in Table 4).

underestimated for most of the cases. This can be explained by considering the fact that the tests performed in the plant did not heavily excite transient dynamics, and thus, these dynamics are difficult to observe in the measurement data, in addition to the presence of measurement noise. This causes a comparatively large divergence in the estimates of T_s with respect to those of R.

B. CASE 2: POWER PLANT PARAMETER ESTIMATION UNDER STEADY STATE OPERATION USING SIMULATED MEASUREMENTS

In order to illustrate the methodology on a large set of estimated parameters, simulations of a power plant connected to a grid were performed.

Model: The power plant model used in this case is shown in Fig. 1. The detailed list of parameters of each component is available in [20]. The parameters to be estimated are the generator's inertia M_0 (initial value range [6..7]), governor droop R (initial value range [0.04..0.065]), governor time constant T_s (initial value range [0.5..1.5]), automatic voltage regulator (AVR) gain K_0 (initial value range [4..6]), the AVR's field circuit time constant T_e (initial value range [0.005..0.03]), washout gain of the power system stabilizer (PSS) K_ω (initial value range [20..90]), and the PSS's washout time constant T_ω (initial value range [5..40]).

Measurements: The following synthetic measurements have been used: voltage phasor ($v.i$, $v.r$) and current phasor ($i.i$, $i.r$) from bus (Fig. 1), mechanical power (p_m) (output of the turbine governor), speed (ω) (input to the turbine governor), voltage (v_s) as the PSS output, and the field voltage (v_f) as the AVR's output. The sampling rate of the simulation was set to 50 samples/s. A 1% Gaussian white noise was added to the voltage and current phasors.

1) CHOICE OF NUMBER OF PARTICLES

To keep the estimation error low, the number of particles needed for the set of 7 parameters to estimate is bigger than in the case for 2 parameters. Fig. 9 shows that for Case 2 the estimation error is minimized when $N = 500$.

2) CHOICE OF Σ

Taking into consideration the analysis for the choice of Σ in Section V-A, the ROT method was chosen for Cases 2-4.

Results are summarized in Table 3.

TABLE 3. Bandwidth estimation by ROT method ($N = 500$) (Case 2-4).

Param. name	Case 2	Case 3	Case 4
Generator inertia GEN.M0	0.1606	0.1622	0.1644
Governor droop TG.R	0.0040	0.0041	0.0032
Governor time constant TG.Ts	0.1608	0.1695	0.1706
Regulator gain AVR.K0	0.3254	0.3295	0.3301
Field circuit time constant AVR.Te	0.0041	0.0040	0.0041
Washout gain PSS.Kp	11.4467	11.0949	11.3064
Washout time constant PSS.Tw	5.4861	5.6393	5.5079

C. CASE 3: POWER PLANT PARAMETERS ESTIMATION UNDER TRANSIENT GRID DISTURBANCES USING SIMULATED MEASUREMENTS

Model: This case study was performed for a case when the system is excited with a 3 phase fault occurred at $t = 10-10.2\,s$. Such a test should enable engineer to obtain a more precise estimation of a certain time constants by having additional system excitation. In addition, it allows to test the method with a larger dynamic variation in the system states, which makes the identification problem more complex. All the other simulation and identification settings are the same as in Case 2.

Measurements: In this case, a 1% Gaussian process noise was added to the bus in order to imitate the natural power fluctuations observed in reality.

1) CHOICE OF NUMBER OF PARTICLES

When a large disturbance is applied to the model, the system dynamics are activated and reflected in the simulated data (measurements). The operating state may vary and non-linear components become active, which brings more uncertainty into identification process. For this reason, the estimation error is higher than in the steady state operation case (Case 2). The error is relatively unchanged with growth of the particles number (Fig. 9), so $N = 500$ was chosen for convenience and consistency with other cases.

D. CASE 4: POWER PLANT PARAMETERS ESTIMATION THAT UNDER A TRANSIENT GRID DISTURBANCE AND NON-GAUSSIAN NOISE USING SIMULATED MEASUREMENTS

There are several types of noise that are more prominent in power systems, which may corrupt measurements or influence the process. The largest influence is that of thermal

TABLE 4. Numerical experiment results.

	Exp. Name	Exp. Setup	Parameter	Estimate	Confidence Region's (95%) boundary	Est.Error(Fitness)
1	TG	$\Delta\omega = +0.2\,Hz$	$Droop\ (R)$	0.0601	[0.02..0.0976]	5.91e-06
	staged tests	$(0.004\,p.u)$	$Ts,\ sec$	1.0242	[0.4175..1.5088]	
2	PP id: Gen	Gaussian noise	M_0	6.4257	[5.9241..7.0676]	2.7003e-005
	Controls TG	ss operation	R, Ts	(0.051, 1.0905)	([0.0372..0.0669], [0.4143..1.6132])	
	AVR		K_0, Te	(4.1919, 0.0091)	([3.7889..6.2082], [0.0033..0.0321])	
	PSS		$K\omega, T\omega$	(67.4998, 22.0769)	([12.8073..97.4829], [3.6416..42.7645])	
3	PP id: Gen	Gaussian noise +	M_0	6.3806	[5.9329..7.0807]	0.01797
	Controls TG	3 phase fault	R, Ts	(0.0516, 0.6335)	([0.0378..0.0679], [0.3695..1.6176])	
	AVR		K_0, Te	(4.5100, 0.0286)	([3.7773..6.2155],[0.0031..0.0320])	
	PSS		$K\omega, T\omega$	(20.6461, 26.1437)	([12.4190..94.6687],[2.9388..42.5331])	
4	PP id: Gen +	Pink noise +	M_0	6.6001	[5.9257..7.0802]	0.01323
	Controls TG	3 ph fault	R, Ts	(0.0391, 0.5751)	([0.0282..0.0515], [0.3571..1.6078])	
	AVR		K_0, Te	(4.0983,0.0067)	([3.7487..6.1979], [0.0033..0.0322])	
	PSS		$K\omega, T\omega$	(50.4064, 15.4898)	([12.8326..96.0916], [4.1923..43.4088])	

Acronyms: PP id - power plant identification, TG - turbine governor, Gen - generator, Exp. - experiment, Est. - estimation, ss - steady state

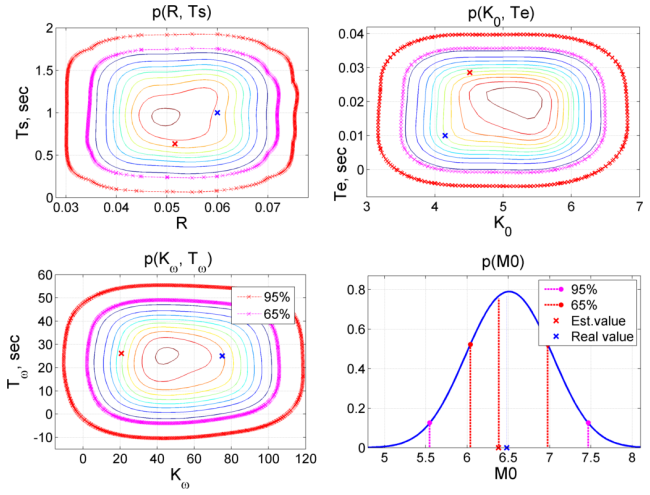

FIGURE 10. Marginalized estimate uncertainty pdf(s) (Case 3 (N=500) in Table 4).

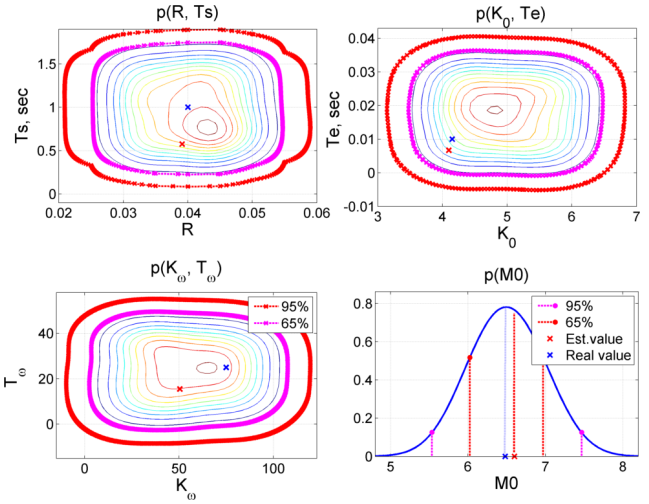

FIGURE 11. Marginalized estimate uncertainty pdf(s) (Case 4 (N=500) in Table 4).

noise (Johnson-Nyquist noise) present in every conductor, and Flicker noise (or 1/f noise) that is present in almost all electronic devices. The latter, contrary to thermal noise, has a non-white wide-band spectrum that is known as ''pink'' noise.

Model: In this case study the performance of the algorithm is observed in the case when a 3 phase fault disturbance and a 1% non-Gaussian ''pink'' noise as process noise is applied. The process noise has influence on each input-output signal measured in the system.

Measurements: The synthetic measurements have been taken and used in this case study are the same as for the Case 2. The difference with Case 3 is that 1% non-Gaussian ''pink'' noise is added to the voltage and current phasors at the bus (Fig. 1).

1) CHOICE OF NUMBER OF PARTICLES

Fig. 9 shows that the estimation error decreases with the increase of particles number. The optimal particle number can be considered equal to $N = 500$, due to a large drop of the error value from $N = 300$ to $N = 500$, and approximately the same value for the error after $N = 500$. In addition, it allows for comparison with Cases 2-4.

E. ESTIMATE UNCERTAINTY DISTRIBUTION AND CONFIDENCE REGION RECONSTRUCTION: COMPARISON OF CASES 2-4

For the estimated distribution characteristics (w, μ, Σ, N), the marginal distributions (8) are reconstructed in Figs. 10-11 for Case 3 and 4 respectively. The marginal confidence regions, that are formulated in (13) and Section IV-C, are estimated (Table 4). The estimated and real parameter values lay in a region that is smaller than 65% of confidence. It is worth to notice that the AVR parameter values are the most distant from the peak of the reconstructed distribution than other parameters. Considering the high precision or small estimation error of the Particle Filter, an engineer can use smaller confidence interval than 95% when defining the

parameter variance. The assumption is that the solution to the formulated problem is unique in the defined initial parameter space is confirmed. It is possible to observe one prominent peak for each marginalized distribution.

VI. DISCUSSION

The developed methodology of parameters identification with their uncertainty representation in form of the distribution can play important role in power system models validation and parameters calibration [23], [24]. The model validation regularly performed by transmission system operators (TSOs) to keep the models up to date. These models are used to test system operation in different conditions. Based on these tests TSOs derive guidelines that comprise the set of actions in order to preserve normal operation of the grid. Current practices mainly based on Kalman filters [4] estimating the uncertainties by uni-modal Gaussian distribution. This approximation is not always valid.

One of the possible cases in which the shape of estimated distribution is important become when bivariate ambiguity in parameter estimation takes place. Namely, when exchange of estimated parameters values won't change precision of system output. For example, several time constants of the controller (e.g. static var compensator) are laying in the same range of values, so the algorithm will be initiated in the same parameter estimation range. But having measured only input and output of the controller will not allow to define which estimated value belongs to which parameter. So the problem in such formulation will have two possible answers. The final decision on a proper estimate has to be taken by the operator. Advantage of the proposed method consists in ability to estimate two or more- peaked distribution instead of rough approximation of Gaussian estimate.

The shape of the uncertainty distribution in parameter space is crucial in decision of particularization of the assumed model. This can help the engineer to decide on another experiment set up for model validation or parameters calibration. For example, if the top of the distribution is flat, there is a big range of the parameters that give approximately the same fitness. The reason for that could be that the measured signals are not are not sensitive to the parameters. In this case the measured signals to use in identification workflow should be selected by considering the dynamics where the parameters of interest are involved [6].

Observing the estimated parameter uncertainties can help to verify the assumed model structure. Such information can be found in the estimation errors (fitness) calculated by the algorithm. In the case the model structure is erroneous, the fitness value is large. This indicated to the operator to change the model assumption.

This work can also be valuable for robust control design [25]. For this task it is of crucial importance to consider the uncertainties in parameters of real engineering systems that are vulnerable to external disturbances and measurement noise. In addition, the calibrated model can significantly decrease the requirements for robust control design.

The proposed method can be used for uncertainty assessment of state tracking. However, the precision of the state estimator can be improved employing the Interacting Multiple Model (IMM) algorithm [26]. IMM combines state hypotheses from multiple filter models (e.g. Particle Filter, Kalman Filters) to obtain better estimates. IMM has been applied in the aerospace field [27] for highly maneuvering target tracking and could be applied in complex nonlinear power systems for state tracking of generators and their controls.

The proposed methods for bandwidth selection are not the only that could solve the density reconstruction problem. The expectation maximization (EM) algorithm, that has been widely used for other applications in power systems [28], [29], could also be applied. This can be considered as a future extension of the work presented in this paper.

VII. IMPLEMENTATION CONSIDERATIONS

The proposed methodology is generic and can be applied to any physical system, not only the power systems. This covers the range of the applications in multi-domain modeling. In order to get a good estimate located closer to the peak of the uncertainty distribution, the assumption of the model structure has to be correct and the measurements being sensitive to the estimated parameters dynamics. As more measurements are used, more precise the parameters can be estimated (avoiding dynamics interference or superposition effect). The quality of the estimates depends also on the sampling frequency. Suppose an engineer wants to estimate the time constants, then the measurements has to be sampled following Nyquist-Shannon-Kotelnikov theorem.

Precision of the distribution reconstruction highly depends on number of particles chosen in the algorithm. The small number of particles can lead to the erroneous conclusions on the uncertainty and large estimation error. Therefore, it is recommended to calibrate the algorithm on the known model before applying to the real system parameters identification.

VIII. CONCLUSION

A new methodology to estimate model parameters and their uncertainty distributions, therefore, the parameter confidence intervals, has been presented and justified. The particularization of the particle filter as the nonparametric Bayesian filtering concept has been studied. The number of particles has been chosen considering computational time and fitness value. The bandwidth estimation algorithms has been compared by considering estimates' real values and the distributions' shape. In addition, the confidence area/intervals dependency on the number of particles, and the presence of noise, has been studied. The proposed and alternative methods were applied to estimate the parameters of a Greek power plant considering steady-state and transient operation conditions. The proposed methodology gives additional insight into power system properties when estimating the parameters of the model. This allows power system analysts to decide on the design of validation tests of the chosen model. Consequently, this technique is recommended for power system analysts

when performing power system model validation and parameter calibration tasks.

APPENDIX A

Theorem: The union of the ellipsoids S_i formed by the weighted mixture of Gaussians contains exactly or more than I percent of probability if each surface contains I of probability.

Proof: Let S_i be the M-dimensional ellipsoid surface such as $\mathbb{P}(X_i \in S_i) = I$, where I is a chosen percent of probability, which S_i contains, or the confidence that the estimated value located inside the ellipsoid with probability I. Let the union of the surfaces be defined as $\bigcup_{i=1}^{M} S_i$, then from the definition of probability and the probability product rule, the following equation can be defined:

$$\begin{aligned}\int_{\bigcup_{i=1}^{M} S_i} p(x)dx &= \mathbb{P}(X_j \in \bigcup_{i=1}^{M} S_i) \\ &= \sum_{i=1}^{M} \mathbb{P}(X_j \in \bigcup_{i=1}^{M} S_i | J = j)\mathbb{P}(J = j) \geq I,\end{aligned}$$

because

$$\mathbb{P}(X_j \in \bigcup_{i=1}^{M} S_i | J = j) \geq I, \quad \sum_{i=1}^{M} \mathbb{P}(J = j) = 1.$$

It is known that the confidence region of the parameter estimate (ellipsoids S_i) is proportional to square root of variance (standard deviation) of the parameter estimate. Therefore, if we show that the resulting variance of the weighted distribution is equal or bigger than sum of weighted variances of separate distributions the theorem will be proved. From equation (8) and by the definition of moments (i.e. mean, variance) of distributions $p(x) \sim \mathcal{N}(\mu^{(1)}, \mu^{(2)})$ find:

$$\mu^{(k)} = \mathbb{E}_p[x^k] = \sum_i \omega_i \mathbb{E}_{p_i}[x^k] = \sum_i \omega_i \mu_i^{(k)} \tag{21}$$

where k - the moment number

$$\begin{aligned}\mu_i^{(2)} &= \sigma_i^2 + (\mu_i^{(1)})^2 \\ Var(p) &= \mu^{(2)} - (\mu^{(1)})^2 = \sum_i \omega_i \mu_i^{(2)} - (\sum_i \omega_i \mu_i^{(1)})^2 \\ &= \sum_i \omega_i (\sigma_i^2 + (\mu_i^{(1)})^2) - (\sum_i \omega_i \mu_i^{(1)})^2 \\ &= \sum_i \omega_i \sigma_i^2 + \sum_i \omega_i (\mu_i^{(1)})^2 - (\sum_i \omega_i \mu_i^{(1)})^2 \end{aligned} \tag{22}$$

According to Jensen's inequality [11]:

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)], \tag{23}$$

where ϕ is a convex function. Thus, the average squared mean can be no less than the square of the average mean.

$$Var(p) = \sum_i \omega_i \sigma_i^2 + \varepsilon, \quad \varepsilon \geq 0 \tag{24}$$

The variance of the mixture is the mixture of the variances plus a non-negative term accounting for the (weighted) dispersion of the means.

APPENDIX B

Greek Power Plant parameters: Generator VI [7]: $S_n = 100$ MVA, $V_n = 19$ kV, $r_{a0} = 0.0028$ p.u., $x_{d0} = 2.08$ p.u., $x_{q0} = 2$ p.u., $x'_{d0} = 0.305$ p.u., $x'_{q0} = 0.49$ p.u., $x''_{d0} = 0.245$ p.u., $x''_{q0} = 0.245$ p.u., $t'_{d0} = 6.8$ s, $t'_{q0} = 0.62$ s, $t''_{d0} = 0.0402$ s, $t''_{q0} = 0.077$ s, $t_{aa} = 0$ s, $M_0 = 6.48$ kWs/kVA; TG TypeI [7]: $R = 0.04$, $T_s = 1$ s, $T_c = 0.3$ s, $T_3 = 0.04$ s, $T_4 = 5$ s, $T_5 = 4$ s, $p_{max} = 0.5$ p.u., $p_{min} = 0$; AVR TypeIII [7]: $v_0 = 1$ p.u., $K_0 = 4.15$, $T_2 = 1$, $T_1 = 1$, $T_e = 0.01$ s; PSS TypeII [7]: $K_p = 75$, $T_w = 25$ s, $T_1 = 0.15$ s, $T_2 = 0.01$ s, $T_3 = 0.15$ s, $T_4 = 0.01$ s, $Y_{min} = -0.1$, $Y_{max} = 0.1$; Input references: for PSS: $V = 1$ p.u.; for AVR: $V_0 = 1$ p.u.

The Modelica components of the model are included into the iPSL library and available in Github (see [8]).

ACKNOWLEDGMENT

The authors are thankful to George Antonopoulos of IPTO who kindly provided measurements and Greek power plant structure information.

REFERENCES

[1] D. Jones, "Estimation of power system parameters," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 1980–1989, Nov. 2004.
[2] M. Burth, G. C. Verghese, and M. Vélez-Reyes, "Subset selection for improved parameter estimation in on-line identification of a synchronous generator," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 218–225, Feb. 1999.
[3] G. K. Stefopoulos, P. S. Georgilakis, N. D. Hatziargyriou, and A. P. S. Meliopoulos, "A genetic algorithm solution to the governor-turbine dynamic model identification in multi-machine power systems," in *Proc. 44th IEEE Conf. Decision Control*, Dec. 2005, pp. 1288–1294.
[4] Z. Huang, P. Du, D. Kosterev, and S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1939–1949, May 2013.
[5] N. Zhou, D. Meng, and S. Lu, "Estimation of the dynamic states of synchronous machines using an extended particle filter," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4152–4161, Nov. 2013.
[6] V. S. Perić, X. Bombois, and L. Vanfretti, "Optimal signal selection for power system ambient mode estimation using a prediction error criterion," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2621–2633, Jul. 2016.
[7] F. Milano. (May 2010). *Power System Analysis Toolbox Documentation for PSAT Version 2.1.6*. [Online]. Available: http://faraday1.ucd.ie/psat.html
[8] L. Vanfretti, T. Rabuzin, M. Baudette, and M. Murad, "iTesla power systems library (iPSL): A Modelica library for phasor time-domain simulations," *SoftwareX*, vol. 5, pp. 84–88, May 2016. [Online]. Available: http://dx.doi.org/10.1016/j.softx.2016.05.001
[9] S. C. Kramer and H. W. Sorenson, "Bayesian parameter estimation," *IEEE Trans. Autom. Control*, vol. AC-33, no. 2, pp. 217–222, Feb. 1988.
[10] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer, 2001.
[11] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
[12] G. Strang, *Introduction to Linear Algebra*. Wellesley Hills, MA, USA: Cambridge Press, 1986.
[13] B. A. Turlach, "Bandwidth selection in kernel density estimation: A review," Inst. Statist. Ökonometrie, Humboldt-Univ. Berlin, Berlin, Germany, Tech. Rep., Jan. 1993, pp. 23–493.
[14] M. C. Jones, J. S. Marron, and S. J. Sheather, "A brief survey of bandwidth selection for density estimation," *J. Amer. Statist. Assoc.*, vol. 91, no. 433, pp. 401–407, 1996.
[15] T. Duong and M. L. Hazelton, "Cross-validation bandwidth matrices for multivariate kernel density estimation," *Scandin. J. Statist.*, vol. 32, no. 3, pp. 485–506, 2005.
[16] P. Hall, S. J. Sheather, M. C. Jones, and J. S. Marron, "On optimal data-based bandwidth selection in kernel density estimation," *Biometrika*, vol. 78, no. 2, pp. 263–269, 1991.

[17] L. Vanfretti, T. Bogodorova, and M. Baudette, "Power system model identification exploiting the Modelica language and FMI technologies," in *Proc. IEEE Int. Conf. Intell. Energy Power Syst.*, Jun. 2014, pp. 127–132.
[18] L. Vanfretti *et al.*, "RaPId: A modular and extensible toolbox for parameter estimation of Modelica and FMI compliant models," *SoftwareX*, vol. 5, pp. 144–149, Aug. 2016. [Online]. Available: http://dx.doi.org/10.1016/j.softx.2016.07.004
[19] P. M. Anderson and A. A. Fouad, *Power System Control and Stability*. New York, NY, USA: Wiley, 2008.
[20] T. Bogodorova, L. Vanfretti, and K. Turitsyn, "Bayesian parameter estimation of power system primary frequency controls under modeling uncertainties," *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 461–465, Oct. 2015.
[21] N. Zhou, D. Meng, Z. Huang, and G. Welch, "Dynamic state estimation of a synchronous machine using PMU data: A comparative study," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 450–460, Jan. 2015.
[22] A. Ihler and M. Mandel. (2003). *Kernel Density Estimation Toolbox for MATLAB*. [Online]. Available: http://www.ics.uci.edu/ihler/code/kde.html
[23] D. N. Kosterev, W. A. Mittelstadt, and C. W. Taylor, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 967–979, Aug. 1999.
[24] E. Allen, D. Kosterev, and P. Pourbeik, "Validation of power system models," in *Proc. IEEE Power Energy Gen. Meet.*, Jul. 2010, pp. 1–7.
[25] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 742–753, May 2006.
[26] A. F. Genovese, "The interacting multiple model algorithm for accurate state estimation of maneuvering targets," *Johns Hopkins APL Tech. Dig.*, vol. 22, no. 4, pp. 614–623, 2001.
[27] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: A survey," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 1, pp. 103–123, Jan. 1998.
[28] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, Feb. 2010.
[29] F. M. Gonzalez-Longatt, J. L. Rueda, I. Erlich, D. Bogdanov, and W. Villa, "Identification of Gaussian mixture model using mean variance mapping optimization: Venezuelan case," in *Proc. 3rd IEEE PES Int. Conf. Exhib. Innov. Smart Grid Technol. (ISGT Europe)*, Oct. 2012, pp. 1–6.

TETIANA BOGODOROVA (S'12) received the B.S. degree in computerized systems, automatics and control, the M.Sc. degree in automatic and control systems from the National Technical University of Ukraine–Kyiv Polytechnic Institute. She is currently pursuing the Ph.D. degree with Electrical Power Systems Division, School of Electric Engineering, KTH Royal Institute of Technology, Stockholm. Her experience includes development of the operations support system for telecommunication industry as a System Engineer with System Analytics Group, Research and Development, NetCracker Technology Corp. Her current research interests lie on intersection of system identification theory and power systems analysis, modelling, and validation.

LUIGI VANFRETTI (SM'15) is currently an Associate Professor (Tenured) and Docent with the Electric Power Systems Department, KTH Royal Institute of Technology, Stockholm, Sweden. He was conferred the Swedish Title of Docent in 2012 and was an Assistant Professor with the Electric Power Systems Department from 2010 to 2013. Since 2011, he has served as an Advisor to the Research and Development Division of Statnett SF, Oslo, Norway, where he is a Special Advisor in strategy and public affairs (SPA–Strategi og Samfunnskontakt). His major research funded projects are IDE4L, Ideal Grid for All; FP7-Energy-2013-7-1-1 Call. PI for KTH funded by the European Commission; iTesla, Innovative Tools for Electric Power System Security within Large Areas; FP7-Energy-2011-1 Call. PI for KTH funded by the European Commission; STRONg2rid, Smart Transmission Grids Operation and Control; Funded by Nordic Energy Research and Svenska Kraftnät, Sustainable Energy Systems 2050 call. PI for KTH. He is mainly active in the Power and Energy Society, where he contributes to several working groups, task forces, and committees. He served, since 2009, in the IEEE Power and Energy Society PSDP Working Group on Power System Dynamic measurements, where he is currently the Vice-Chair. In addition, since 2009, he has served as the Vice-Chair of the IEEE PES CAMS Task Force on Open Source Software.

VEDRAN S. PERIĆ (M'17) received the M.S. degrees in power systems and power electronics from the University of Novi Sad, Serbia, the joint Ph.D. degree from the KTH Royal Institute of Technology, Stockholm, Sweden (primary institution), Delft University of Technology, Delft, Netherlands and Comillas Pontifical University, Madrid, Spain in 2016. He held positions of Research and Teaching Assistant with the University of Novi Sad and Visiting Researcher with the Delft University of Technology. He was a Senior Power System Engineer with GE Grid Solutions Research and Development Department, GE WAMS Center of Excellence and as a Senior Business Analyst with Regional Security Coordinator, TSCNET Services GmbH. He is currently a Senior Power System Consultant with GE Consulting in Munich, Germany. His research interests include a wide range of topics related to power system operation and control, with the focus on commercial application of innovative technologies.

KONSTANTIN TURITSYN (M'09) received the M.Sc. degree in physics from the Moscow Institute of Physics and Technology, Moscow, Russia, and the Ph.D. degree in physics from the Landau Institute for Theoretical Physics, Moscow, Russia, in 2007. He was an Oppenheimer Fellow with the Los Alamos National Laboratory, Los Alamos, NM, USA, and a Kadanoff-Rice Post-Doctoral Scholar with the University of Chicago, Chicago, IL, USA. He is currently an Assistant Professor with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. His current research interests include nonlinear and stochastic dynamics of complex systems, energy-related fields such as, stability and security assessment, integration of distributed, and renewable generation.