

# An efficient method to take into account forecast uncertainties in large scale Probabilistic Power Flow

E. Ciapessoni, D. Cirio, A. Pitto  
Energy System Development Dept.  
Ricerca sul Sistema Energetico - RSE S.p.A.  
Milan, Italy  
emanuele.ciapessoni@rse-web.it

S. Massucco, F. Silvestro  
Intelligent Electric Energy Systems  
University of Genoa  
Genoa, Italy  
stefano.massucco@unige.it

**Abstract**— The simulation of uncertainties due to renewable and load forecasts is becoming more and more important in security assessment analyses performed on large scale networks. This paper presents an efficient method to account for forecast uncertainties in probabilistic power flow (PPF) applications, based on the combination of PCA (Principal Component Analysis) and PEM (Point Estimate Method), in the context of operational planning studies applied to large scale AC grids. The benchmark against the conventional PEM method applied to large power system models shows that the proposed method assures high speed up ratios, preserving a good accuracy of the marginal distributions of the outputs.

**Index Terms**-- uncertainty, risk, security, power systems, Point Estimate Method, Principal Component Analysis.

## I. INTRODUCTION

The increasing penetration of non-programmable Renewable Energy Sources (RES) increases the uncertainty of power system operation. An urgent requirement arises, to include RES and load forecast uncertainties in security assessment studies, both in operational planning and in nearly real time operation [1], [2]. Probabilistic Power Flow (PPF) can help to achieve this goal [2]-[7], [10].

Different approaches have been proposed for PPF. Monte Carlo (MC) sampling technique is accurate in modeling the uncertainties but it is time-consuming [2]. To improve its efficiency different variance of reduction techniques, such as importance sampling, have been proposed [3]; even though they reduce the number of individual deterministic power flow runs, still thousands of computations are necessary to solve large scale systems. Though less accurate with respect to MC sampling, analytical methods, like the cumulants' methods [4]-[5], are very efficient in case of independent variables, but accounting for dependence requires the computation of cumbersome integrals which slow down the overall computation, especially in case of thousands of stochastic variables. The Point Estimate Method (PEM) [6]-[8] is a hybrid method that represents a good tradeoff of accuracy and computational burden. This method requires a

number of “deterministic” calculations which grows linearly with the number of stochastic variables: in case of large scale grids, the very high number of stochastic variables may make the PEM not convenient with respect to a conventional MC sampling approach.

The original contribution of the present paper is to propose a flexible method to account for uncertainties in power system applications, which can calibrate the level of accuracy and computational efficiency according to the specific required application. In practice, the method can assure a strong reduction of the computational time with a precise quantification of the accuracy loss, which must be compared to the needs required by the specific application. Even though it's easy to parallelize the PEM runs, it must be considered that in the operational planning stage the need to assess large sets of contingencies with different features (fault location, type and duration) over a large set of plausible operating scenarios may greatly benefit by an approach which speeds up the evaluation of uncertainties. In this sense, the proposed approach might bring even more valuable benefits in applications like probabilistic dynamic security assessment, where time consuming domain simulations are required.

The paper is organized as follows: Section II proposes the novel methodology. Section III describes the uncertainty models adopted in the simulations, the benchmark method, the comparison metrics and the set-up of the simulation scenarios on two test systems of medium-large size. Section IV presents and discusses the simulation results. Section V draws some conclusions.

## II. METHODOLOGY

The proposed method combines a dimensionality reduction technique like the PCA (Principal Component Analysis) [9] with the Point Estimate Method which is a well-known hybrid method for the treatment of stochastic variables. The workflow of the proposed method is given in Figure 1. The inputs of the workflow consist in a set of dependent non-Gaussian stochastic variables  $X$  s.t.  $\dim(X) = N$ .

The main steps are the following:

1. Use Third Order Polynomial normal Transformation (TPNT) [10] by applying the expression  $x_i = a_{0,i} + a_{1,i} \cdot Z_i + a_{2,i} \cdot Z_i^2 + a_{3,i} \cdot Z_i^3$  for all variables  $i=1, \dots, N$  and evaluate the  $2N+1$  points in terms of original dependent non-Gaussian variables  $X$ . This step is convenient for two reasons:

- a. the original variables under study (i.e. the forecast errors of loads and RES injections) are usually far from being symmetric.
- b. The gaussianity of the variables is a convenient property for subsequent steps of the workflow

The outputs of this transformation consist in a set of  $N$  dependent normal transformed variables  $Z$  characterized by normal marginals and a correlation matrix  $R_z$  and in a set of parameters, called  $L$ -moments  $\lambda_i$  ( $i=1, \dots, 4$ ), used to build coefficients  $a_{0,i}-a_{3,i}$ .

2. Two alternative PCA decomposition schemes are applied: (1) application to correlation matrix  $R_z$  (correlation-based PCA), (2) application to the covariance matrix  $\text{diag}(a_{1,i}) * R_z * \text{diag}(a_{1,i})$  related to variables  $a_{0,i} * Z_i$  which represent the first order normal approximations of  $X$ 's without means (the PCA works better in case of Gaussian variables). Both schemes allow to model a defined fraction of explained variance equal to  $r$ . The outcomes consist in a smaller set of  $n$  (with  $n \ll N$  depending on parameter  $r$ ) retained normal and independent variables called Principal Components that are linear combinations of variables  $a_{0,i} * Z_i$ , and in the transformation matrix  $Q$  linking the principal components (PCs) to variables  $Z$ .
3. As the PCs are independent variables, one can apply the PEM (Point Estimate Method) to them, getting  $2n+1$  vectors of PCs, and the relevant  $2n+1$  weights  $W$
4. Using matrix  $Q$  it's possible to backproject the  $2n+1$  vectors of PCs onto  $2n+1$  vectors of the dependent transformed variables  $Z$
5. In the end, the application of the  $L$ -moments to the  $2n+1$  vectors of variables  $Z$  leads to the  $2n+1$  vectors of original variables  $X$ .
6. The generated vectors of variables  $X$  are then applied to the specific "power system analysis" tool (in the present study a load flow tool).
7. The  $2n+1$  results of the quantity of interest  $V$  (a node voltage or a branch power flow) are combined with the weights  $W$  computed at step 3, to obtain the raw moments of the marginal distribution of the quantity itself.

### III. TEST SYSTEMS AND SCENARIO SET-UP

The test systems to validate the present method are the IEEE 118 buses test system [11] and a 9241 bus model of the pan-European power system provided by the FP7 EU project Pegase [12]. The proposed method has been implemented using MATPOWER [13].

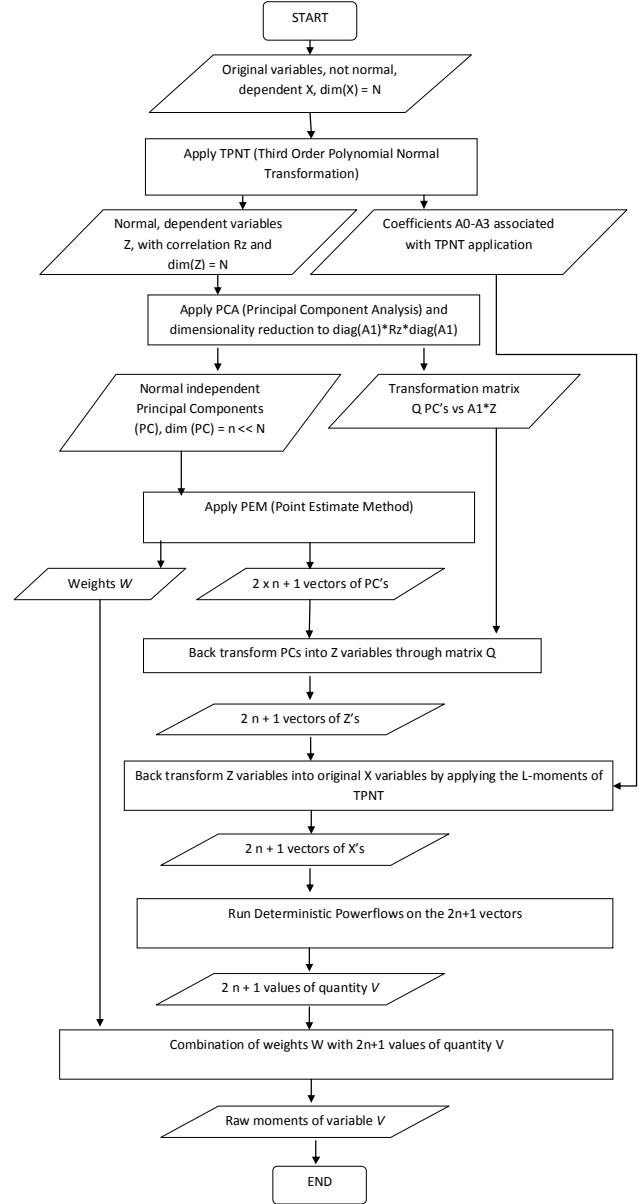


Figure 1. Flowchart of the proposed approach

#### A. Uncertainty modeling

The stochastic variables considered are the forecast errors for  $N$  loads and renewable injections. The approach presented in the simulations implements models drawn from different studies [14]-[17]. In particular, the non-symmetry of the forecast errors, derived from statistical analyses of historical data, suggests the use of non-symmetric distributions (like beta distributions) for wind and solar generation forecast errors. It is worth remarking that systematic errors in forecasts may determine a non-null mean value for the forecast errors, especially with regard to RES generation. The standard deviation of RES generation forecast errors depends on [15]-[17]:

- *Level of aggregation of RES*: the larger is the number of wind/solar farms aggregated into the same “equivalent” generation, the lower is the standard deviation
- *Geographic extension of the RES aggregation*: given the same number of RES sources aggregated into a single “equivalent” injection, the larger the area where they are distributed the higher the compensation effect among RES, thus the smaller is the standard deviation expressed in % of the total rating of the relevant injection
- *The forecast time horizon*: typically the larger the time horizon the larger the variance in forecast errors.

The standard deviation associated with load forecasts is usually very low (typically 1-4 % of the actual power).

Load forecast errors have been assumed Gaussian distributions as in [17], while renewable injection forecast errors depend on the current forecast value of the injections and – depending on this forecast value– may be represented by a beta distributions (for forecast values not to close to 0 or 1 p.u. of the rated power of the renewable plant) or with a truncated normal distributions (for more extreme values close to 0 or 1 p.u.) [6], [18]. In a general approach, the PV forecast error variance also depends on the clearness index [17]: however, the simulations performed in the present paper are not aimed to demonstrate the effect of a changing weather, thus they assume a clearness index corresponding to the worst weather conditions (i.e. maximum variance of the forecast error).

The spatial dependence model neglects potential non-linear dependencies and it is represented by a correlation matrix built as a Toeplitz matrix with a generating vector composed by linearly decreasing values from 1 to 0. The theorem in [19] assures the positive definiteness of the matrix built with a generating vector of linearly decreasing values – including negative ones - provided that the sum of the vector components is positive.

### B. Benchmarking method

The benchmark method used to validate the proposed approach consists in the well-established PEM method with Third Order Normal Transformation, already discussed and validated for PPF by the authors in [6]. This method consists in the following steps:

1. Apply the TPNT to the original variables  $X$  ( $\dim(X)=N$ ) to get normal dependent variables  $Z$  with correlation matrix  $R_z$ ,  $L$ -moments and coefficients  $a_{0,i}$ - $a_{3,i}$  to link  $Z$ 's with  $X$ 's.
2. Apply the Cholesky decomposition to  $R_z$  and get the Cholesky matrix  $G_z$
3. Perform the PEM procedure on normal independent variables  $Y$  getting the  $2N+1$  vectors for  $Y$  variables, and the corresponding  $2N+1$  vector of weights
4. Multiply  $G_z$  by the  $2N+1$  vectors of  $Y$ 's to derive the  $2N+1$  vectors of variables  $Z$
5. Apply the  $L$ -moments computed at step 1 to derive the  $2N+1$  vectors of original variables  $X$
6. Run the deterministic power flow and get the  $2N+1$  values of the quantity of interest  $V$

7. Combine the weights at step 3 and the values of  $V$  to get the raw moments of the marginal distribution of  $V$

The basic difference with the proposed method is that the benchmark method does not perform any dimensionality reduction. Thus the number of vectors to be evaluated at step 6 are  $2N+1 \gg 2n+1$ , depending on the chosen fraction of explained variance.

### C. Metrics for the validation

First of all, the quantities of interest investigated in the result comparison will be the node voltage magnitudes and phases, and the active and reactive flows along the branches.

The metrics chosen to compare the probability distributions of these variables obtained by the two methods (PEM and PROP in the following) can be divided into two groups:

- Component-oriented metrics, which compare the performance of the proposed method vs the benchmark at component level (local level), focusing on individual buses and branches.
- System-oriented metrics, which compare the performance at system level (aggregated level), considered the whole set of buses and branches. They consist in suitable combination of component-oriented metrics.

The local metrics adopted for the comparison are:

- The absolute errors on the first two statistical moments (mean and standard deviation) on  $j$ -th bus voltage or  $j$ -th branch power flow, as reported in (1).
- The Average Root Mean Square (ARMS) error of the CDF's, which is defined in (2)

$$|\Delta\mu_{V(P)j}| = \left| \mu_{V(P)j}^{PROP} - \mu_{V(P)j}^{PEM} \right| \quad |\Delta\sigma_{V(P)j}| = \left| \sigma_{V(P)j}^{PROP} - \sigma_{V(P)j}^{PEM} \right| \quad (1)$$

$$ARMS = \frac{\sqrt{\sum_{j=1}^{N_p} (CDF_j^{PEM} - CDF_j^{PROP})^2}}{N_p} \quad (2)$$

where  $N_p$  is the number of points at which the CDF's have been evaluated, while  $CDF_j^{PEM}$  and  $CDF_j^{PROP}$  are respectively the values of the CDF's computed with the two methods (PEM = benchmark method, PROP= proposed method) at  $j$ -th evaluation point,  $j = 1 \dots N_p$ .

The system oriented metrics are:

- 1, 5, 10, 50, 90, 95 99<sup>th</sup> quantiles of the distribution of  $|\Delta\mu|$  and  $|\Delta\sigma|$ .
- The weighted average of the relative errors as reported in (3), with weights corresponding to the statistical moments computed with the benchmark, i.e.

$$\begin{aligned} \mathcal{E}_{\mu V, \%} &= \frac{\sum_{j=1}^{N_{buses}} |\Delta\mu_{Vj}| \times |\mu_{Vj}|}{\sum_{j=1}^{N_{buses}} |\mu_{Vj}|} & \mathcal{E}_{\mu P, \%} &= \frac{\sum_{j=1}^{N_{branches}} |\Delta\mu_{Pj}| \times |\mu_{Pj}|}{\sum_{j=1}^{N_{branches}} |\mu_{Pj}|} \\ \mathcal{E}_{\sigma V, \%} &= \frac{\sum_{j=1}^{N_{buses}} |\Delta\sigma_{Vj}| \times \sigma_{Vj}}{\sum_{j=1}^{N_{buses}} \sigma_{Vj}} & \mathcal{E}_{\sigma P, \%} &= \frac{\sum_{j=1}^{N_{branches}} |\Delta\sigma_{Pj}| \times \sigma_{Pj}}{\sum_{j=1}^{N_{branches}} \sigma_{Pj}} \end{aligned} \quad (3)$$

where  $|\Delta\mu_{V(P)j}|_{\%}$  and  $|\Delta\sigma_{V(P)j}|_{\%}$  are the relative errors of the first two statistical moments on  $j$ -th bus voltage or  $j$ -th branch power flow and they are given in (4).

$$|\Delta\mu_{V(P)j}|_{\%} = \frac{|\mu_{V(P)j}^{PROP} - \mu_{V(P)j}^{PEM}|}{\mu_{V(P)j}^{PEM}}, \quad |\Delta\sigma_{V(P)j}|_{\%} = \frac{|\sigma_{V(P)j}^{PROP} - \sigma_{V(P)j}^{PEM}|}{\sigma_{V(P)j}^{PEM}} \quad (4)$$

In the present context, the main goal is to assure small errors between the benchmark and the proposed method: to this purpose, relative errors in (4) are not so useful as absolute errors in (1) because they could emphasize large errors on very low variance variables).

#### D. Scenario set up: IEEE 118 bus test system

This relatively small test system [11] is used to prove the concept of the proposed methodology. The grid contains 19 generators, 35 synchronous condensers, 177 lines, 9 transformers, and 91 loads. For simulation purposes,  $N = 10$  out of 19 synchronous generators (at buses 10, 12, 25, 26, 31, 46, 49, 59, 54 and 61) are replaced as equivalent renewable injections (5 wind parks and 5 solar parks) with the same rating as the original synchronous generator.

As for the uncertainty model, each wind (solar) park is composed by 2 (0.1) MW rated wind turbines (solar panels) on an equivalent area of 30 km. Unless differently specified, the 24 hour ahead forecast error standard deviation is 15% of the rated power for each wind turbine and solar panel. The forecast horizon adopted in the simulations is 6 h. The spatial correlation matrix among forecast errors is a Toeplitz matrix with off-diagonal coefficients decreasing linearly from 0.75 to 0 with 0.25 step. The simulation scenarios are the benchmark and the proposed method with  $r = 0.8$  and  $r = 0.6$ .

#### E. Scenario set up: pan-European grid

The pan European power system model in [12] includes 9241 Buses, 14044 Lines, 2234 Transformers, 80 Phase shifters, 5274 Loads, 289 Compensation banks, 1445 Generation Regulating Buses, 27 Areas. The total load is equal to 400 GW. 100 synchronous generators out of 1445 are replaced with equivalent renewable injections with the same rating as the original synchronous unit. The spatial correlation matrix is a Toeplitz matrix with off-diagonal elements linearly decreasing from 1 to 0 with 0.3 steps. As for the uncertainty model, the standard deviations for the 6 hour-ahead forecast errors of aggregated renewable injections are equal to 5% of the rated power of each injection.

Such a large system can better highlight the benefits coming from the proposed approach. In particular, the following simulation scenarios are run: the benchmark and the, proposed method with  $r = 0.95$ ,  $r = 0.9$ ,  $r = 0.8$  and  $r = 0.6$ .

### IV. SIMULATION RESULTS

The present section illustrates the main results of the application of the proposed methods to the two aforementioned test systems.

#### A. Validation of the proposed method

The first step consists in the validation of the proposed method against the benchmark one. To this aim, TABLE I and TABLE II compare the quantiles of the absolute errors on the

means and the standard deviations of the bus voltages and the active power flows obtained from (a) the benchmark method, (b) the proposed method with  $r = 1$ , which means that no dimensionality reduction is performed, respectively for the case of correlation-based PCA (TABLE I) and of covariance-based PCA (TABLE II).

TABLE I - VALIDATION OF THE PROPOSED METHOD AGAINST THE BENCHMARK, APPLICATION OF PCA TO CORRELATION MATRIX

	Quantiles	IEEE 118 bus	Pan European grid
$ \Delta\mu_{\text{on bus voltages, kV}}^{\text{PEM-PROP}} $	1%	0	$3.979 \times 10^{-13}$
	5%	0	$7.013 \times 10^{-8}$
	50%	$3.587 \times 10^{-6}$	$7.618 \times 10^{-6}$
	95%	$1.554 \times 10^{-4}$	$1.122 \times 10^{-4}$
	99%	$4.1328 \times 10^{-4}$	$2.820 \times 10^{-4}$
$ \Delta\sigma_{\text{on bus voltages, kV}}^{\text{PEM-PROP}} $	1%	$3.856 \times 10^{-9}$	0
	5%	$2.403 \times 10^{-7}$	$4.573 \times 10^{-7}$
	50%	$1.417 \times 10^{-4}$	$5.918 \times 10^{-5}$
	95%	$1.581 \times 10^{-2}$	$1.113 \times 10^{-3}$
	99%	$1.093 \times 10^{-1}$	$2.544 \times 10^{-3}$
$ \Delta\mu_{\text{on active power flows, MW}}^{\text{PEM-PROP}} $	1%	0	0
	5%	0	$1.977 \times 10^{-11}$
	50%	$4.430 \times 10^{-4}$	$1.809 \times 10^{-5}$
	95%	$3.561 \times 10^{-3}$	$1.299 \times 10^{-3}$
	99%	$1.067 \times 10^{-2}$	$3.787 \times 10^{-3}$
$ \Delta\sigma_{\text{on active power flows, MW}}^{\text{PEM-PROP}} $	1%	0	0
	5%	0	0
	50%	$4.713 \times 10^{-3}$	$3.520 \times 10^{-3}$
	95%	$1.152 \times 10^{-1}$	$1.138 \times 10^{-1}$
	99%	$3.582 \times 10^{-1}$	$3.354 \times 10^{-1}$

TABLE II - VALIDATION OF THE PROPOSED METHOD AGAINST THE BENCHMARK, APPLICATION OF PCA TO COVARIANCE MATRIX

	Quantiles	IEEE 118 bus	Pan European grid
$ \Delta\mu_{\text{on bus voltages, kV}}^{\text{PEM-PROP}} $	1%	0	$4.110 \times 10^{-6}$
	5%	0	$6.402 \times 10^{-8}$
	50%	$8.1556 \times 10^{-6}$	$7.673 \times 10^{-6}$
	90%	$1.2102 \times 10^{-4}$	$1.045 \times 10^{-4}$
	99%	$5.5832 \times 10^{-4}$	$4.094 \times 10^{-4}$
$ \Delta\sigma_{\text{on bus voltages, kV}}^{\text{PEM-PROP}} $	1%	0	0
	5%	$5.530 \times 10^{-7}$	$3.373 \times 10^{-7}$
	50%	$1.966 \times 10^{-4}$	$4.237 \times 10^{-5}$
	95%	$1.739 \times 10^{-2}$	$1.256 \times 10^{-3}$
	99%	$1.104 \times 10^{-1}$	$3.537 \times 10^{-3}$
$ \Delta\mu_{\text{on active power flows, MW}}^{\text{PEM-PROP}} $	1%	$3.000 \times 10^{-14}$	0
	5%	$8.600 \times 10^{-14}$	$2.217 \times 10^{-11}$
	50%	$2.235 \times 10^{-4}$	$1.407 \times 10^{-5}$
	95%	$2.624 \times 10^{-3}$	$1.578 \times 10^{-3}$
	99%	$6.133 \times 10^{-3}$	$5.263 \times 10^{-3}$
$ \Delta\sigma_{\text{on active power flows, MW}}^{\text{PEM-PROP}} $	1%	0	0
	5%	$2.035 \times 10^{-7}$	0
	50%	$5.582 \times 10^{-3}$	$2.530 \times 10^{-3}$
	95%	$1.263 \times 10^{-1}$	$1.406 \times 10^{-1}$
	99%	$2.232 \times 10^{-1}$	$6.416 \times 10^{-1}$

The validation tests performed show that the first two statistical moments obtained with the proposed approach, which performs an alternative sampling of  $2N+1$  PEM points with  $IR = 1$ , have a very good matching with the ones obtained from the benchmark method. The statement holds valid for both the PCA decomposition schemes.

For the IEEE 118 bus case, Figure 2 compares the  $2N+1$  points obtained by the two PCA decomposition schemes with respect to the  $2N+1$  points got from the benchmark method, for the machine with largest variance (i.e. G10) and one

stochastic injection with much smaller variance (G46). The two schemes are consistent with each other; however, it can be seen that the covariance based method proposes 2N+1 points closest to the “benchmark” points. The correlation-based PCA underestimates the contribution of injection G10 with larger variance. Similarly it can be verified that the contribution of the injection with the smallest variance (G31) is overestimated by the correlation-based PCA.

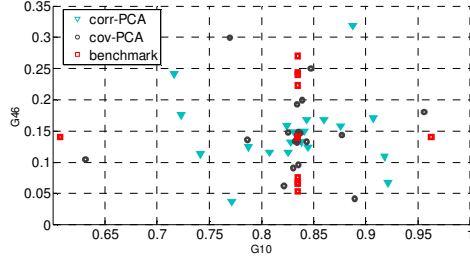


Figure 2. Scatterplot of 2N+1 points for G10 and G46 injections for the benchmark method (red squares), the correlation based PCA PEM (cyan triangles) and the covariance based PCA PEM (black circles)

This fact has an impact on the localization of the branches with largest errors on standard deviations: in the correlation based approach, the largest standard deviation errors (up to 0.39 MW) are detected on the branches with highest sensitivities (i.e. high PTDF’s - Power Transfer Distribution Factors) towards the stochastic injections with largest variances (e.g. lines 9-10 and 8-9), and with smallest variances (e.g. lines 17-31, 31-32).

Simulations performed considering different sizes of grid models and a fraction of explained variance lower than 1 show that the number of retained PC’s for the covariance-based PCA is much lower than the one for covariance based PCA (e.g. the speed up factor between the two PCA decompositions passes from 2.25 for  $r = 0.95$  to 4.20 for  $r = 0.6$  considering the European grid model described in section III), assuring an accuracy linearly increasing with fraction  $r$  (see subsection IV.C). For the present application context, the covariance-based PCA (henceforth named “cov-PCA”) shows an acceptable accuracy performance with a much smaller computational burden with respect to correlation-based PCA, thus cov-PCA is selected as the PCA decomposition for further analyses in the paper.

### B. IEEE 118 bus test system

The goal is to compare the statistical moments and the ARMS for bus voltage magnitudes and active power flows obtained from three different methods: (a) the benchmark approach, (b) the proposed approach with a fraction of explained variance  $r$  equal to 80%, (c) the proposed approach with  $r = 60\%$ . The simulations show that adopting a  $r = 0.8$  (0.6), the number of retained PCs is 3 (2).

TABLE III and TABLE IV report respectively the top ten bus voltages magnitude with the largest absolute errors on the first two statistical moments (mean and standard deviation), in case of  $r = 0.8$ .

TABLE V instead summarises quantiles  $Q_p$  ( $p=1, 5, 50, 90$  and  $99\%$ ) for the distribution of absolute errors  $|\Delta\mu^{\text{PEM-PROP}}|$  and  $|\Delta\sigma^{\text{PEM-PROP}}|$  over the sets of branches and nodes for the two

cases  $r = 0.8$  and  $r = 0.6$ . It’s worth noticing that using the proposed approach with  $r = 0.6$  allows to get still acceptably accurate results with a speed up ratio of 2.1 of the benchmark.

TABLE III – TOP TEN BUSES WITH THE LARGEST ABSOLUTE ERRORS ON THE MEANS AND THE STANDARD DEVIATIONS OF BUS VOLTAGES – IEEE 118 BUS TEST SYSTEM, COV-PCA

Bus ID	$ \Delta\mu^{\text{PEM-PROP}} $ , kV	Bus ID	$ \Delta\sigma^{\text{PEM-PROP}} $ , kV
38	0.0069636	38	0.11687
30	0.0029745	33	0.018775
64	0.0025073	9	0.018255
47	0.0020667	43	0.015339
63	0.0017572	44	0.013183
45	0.001347	37	0.01008
23	0.0012254	52	0.0083285
48	0.0012054	51	0.0078239
17	0.0011293	64	0.0068806
67	0.00098598	45	0.0068

TABLE IV - TOP TEN BRANCHES WITH THE LARGEST ABSOLUTE ERRORS ON THE MEANS AND THE STANDARD DEVIATIONS OF ACTIVE POWER FLOWS – IEEE 118 BUS TEST SYSTEM, COV-PCA

Branch ID	$ \Delta\mu^{\text{PEM-PROP}} $ , MW	Branch ID	$ \Delta\sigma^{\text{PEM-PROP}} $ , MW
17-31	0.61057	26-25	3.518
31-32	0.57804	60-61	2.3993
38-65	0.53379	8-5	2.1697
65-68	0.53333	64-61	2.1189
46-47	0.5184	11-12	1.9124
30-17	0.47942	49-66	1.7138
68-69	0.46559	49-66	1.7138
30-38	0.43931	54-59	1.691
23-32	0.41931	54-56	1.6611
25-27	0.35945	55-59	1.6143

TABLE V – QUANTILES  $Q_p$  FOR THE ABSOLUTE ERRORS FOR BUS VOLTAGES AND BRANCH ACTIVE POWER FLOWS – IEEE 118 BUS TEST CASE, COV-PCA

	$p$	$r=0.8$	$r=0.6$
$ \Delta\mu^{\text{PEM-PROP}} $ on bus voltages, kV	1%	$3.979 \times 10^{-15}$	$2.842 \times 10^{-14}$
	5%	$6.821 \times 10^{-14}$	$5.684 \times 10^{-14}$
	50%	$2.228 \times 10^{-4}$	$2.665 \times 10^{-4}$
	95%	$2.199 \times 10^{-3}$	$5.482 \times 10^{-3}$
	99%	$6.405 \times 10^{-3}$	$1.502 \times 10^{-2}$
$ \Delta\sigma^{\text{PEM-PROP}} $ on bus voltages, kV	1%	$1.274 \times 10^{-6}$	$1.955 \times 10^{-7}$
	5%	$2.697 \times 10^{-6}$	$1.907 \times 10^{-6}$
	50%	$8.700 \times 10^{-4}$	$1.285 \times 10^{-3}$
	95%	$1.621 \times 10^{-2}$	$6.924 \times 10^{-2}$
	99%	$1.031 \times 10^{-1}$	$1.230 \times 10^{-1}$
$ \Delta\mu^{\text{PEM-PROP}} $ on active power flows, MW	1%	$2.998 \times 10^{-14}$	$5.542 \times 10^{-14}$
	5%	$2.075 \times 10^{-13}$	$1.918 \times 10^{-13}$
	50%	$1.975 \times 10^{-2}$	$2.236 \times 10^{-2}$
	95%	$3.714 \times 10^{-1}$	$5.161 \times 10^{-1}$
	99%	$5.621 \times 10^{-1}$	$7.437 \times 10^{-1}$
$ \Delta\sigma^{\text{PEM-PROP}} $ on active power flows, MW	1%	0.000	0.000
	5%	$3.29 \times 10^{-7}$	0.000
	50%	$8.26 \times 10^{-2}$	$3.536 \times 10^{-1}$
	95%	1.624	3.706
	99%	2.317	9.195

In fact the standard deviation error of the power flow on branch 9-10 (the branch with the highest variance absolute error) corresponds to 2% of the initial branch power flow, and in both cases the median absolute error on both voltages and branch power flows is largely below the maximum absolute error. In terms of ARMS, the proposed approach assures good matchings in the shape of the CDF’s for both bus voltages and active power flows also for  $r = 0.6$ , as demonstrated by

the low ARMS values in TABLE VI reporting the quantiles of ARMS distribution with a number of points  $N_p = 2000$ .

TABLE VI – QUANTILES  $Q_p$  FOR ARMS DISTRIBUTION OVER THE NODE VOLTAGES AND ACTIVE POWER FLOWS – IEEE 118 BUS TEST CASE, COV-PCA

	$p$	$r = 0.8$	$r = 0.6$
ARMS on bus voltages, kV	1%	$1.578 \times 10^{-11}$	$3.660 \times 10^{-11}$
	5%	$1.061 \times 10^{-5}$	$1.484 \times 10^{-4}$
	50%	$3.913 \times 10^{-4}$	$6.334 \times 10^{-4}$
	95%	$2.126 \times 10^{-3}$	$2.904 \times 10^{-3}$
	99%	$6.474 \times 10^{-3}$	$4.729 \times 10^{-3}$
ARMS on active power flows, MW	1%	$6.764 \times 10^{-11}$	$6.383 \times 10^{-11}$
	5%	$6.719 \times 10^{-5}$	$1.997 \times 10^{-10}$
	50%	$3.176 \times 10^{-4}$	$6.126 \times 10^{-4}$
	90%	$9.695 \times 10^{-4}$	$2.172 \times 10^{-3}$
	99%	$2.273 \times 10^{-3}$	$2.482 \times 10^{-3}$

TABLE VII shows the system-oriented indicators (weighted averages of absolute errors on the means and standard deviations) for  $r = 0.8$  and  $r = 0.6$ . It is worth noticing that one can achieve a system-level percentage error lower within 10% over the means and standard deviations of  $|V|$  at PQ nodes and branch active power flows.

TABLE VII – WEIGHTED AVERAGE PERCENT ERRORS ON THE MEANS AND THE STANDARD DEVIATIONS – 118 BUS TEST CASE

	Weighted average percent error on bus voltage means $\epsilon_{\mu V, \%}$	Weighted average percent error on bus voltage standard deviations $\epsilon_{\sigma V, \%}$	Weighted average percent error on power flow means $\epsilon_{\mu P, \%}$	Weighted average percent error on power flow standard deviations $\epsilon_{\sigma P, \%}$
$r = 1$	$2.463 \times 10^{-3}$	5.675	$1.209 \times 10^{-3}$	0.3221
$r = 0.8$	$3.421 \times 10^{-4}$	6.549	$1.484 \times 10^{-1}$	4.125
$r = 0.6$	$9.496 \times 10^{-4}$	$1.825 \times 10$	$1.958 \times 10^{-1}$	$1.352 \times 10$

Figure 3 reports the pdf and the CDF of the active power flow along branch 26-25 (the one with largest absolute error on the standard deviations in case of cov-PCA) for the two fractions of explained variance. It can be noticed that there is not a significant improvement from  $r = 0.6$  to  $r = 0.8$ . The PPDFs relating to branch 26-25 show the strong sensitivity of this branch towards stochastic injections at buses 26 and 25.

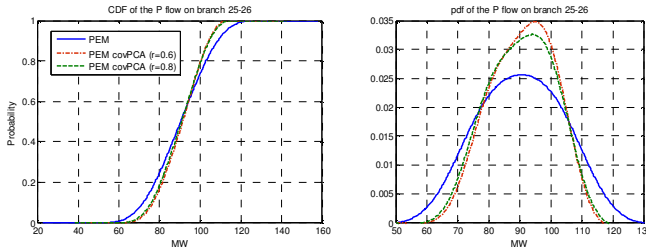


Figure 3. CDF (left) and pdf (right) of the active power flow along branch 26-25 for the two fractions of explained variance ( $r=0.8$  and  $0.6$ )

Figure 4 shows the loadings of the PC's on the injections at buses 26 and 25. Loadings are the coefficients of the linear combinations of PC's which provides the original variables and represent how much a variable is sensitive to a PC. From Figure 4(a) it can be found that the two injections are very sensitive to the fifth PC. Thus, a dimensionality reduction with  $r = 0.8$  (and only three retained PC's) does not allow a good reconstruction of the stochastic injection at bus 26.

Considering a 90% fraction of explained variance leads to 4 retained PC's, which are still not sufficient to improve the matching with the benchmark. Instead, a 95% fraction of explained variance leads to five retained PC's: this allows a reconstruction of the pdf which is much closer to the one obtained from the benchmark (see Figure 4(b) comparing the pdf's of the branch flow for  $r=90\%$  and  $r=95\%$ ).

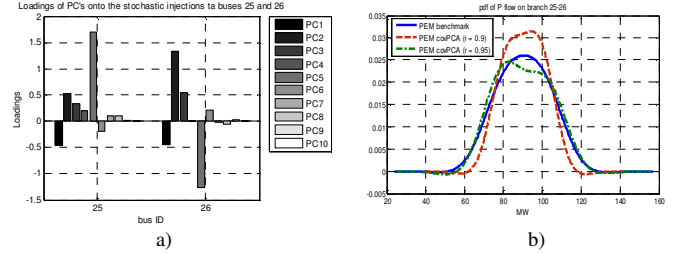


Figure 4. Loadings of PC's on injection at bus 26 (left diagram) and pdf (right diagram) of the P flow on branch 26-30 for  $r = 0.9$

### C. The 9241 bus model of the pan-European network

TABLE VIII reports quantiles  $Q_p$  with  $p = 1\%$ ,  $5\%$ ,  $50\%$ ,  $90\%$  and  $99\%$  for the ARMS distributions of the bus voltage magnitudes  $|V|$  and the active power flows "P", respectively over the set of PQ nodes and branches. The header of the table also reports the number of retained components in each case of application of the proposed method.

TABLE VIII – QUANTILES  $Q_p$  FOR THE ARMS DISTRIBUTION OF  $|V|$  ON PQ NODES AND OF ACTIVE POWER FLOWS – PAN EUROPEAN GRID MODEL

	$p$	$r=0.95$ (258 retained PC's)	$r=0.9$ (186 retained PC's)	$r = 0.8$ (118 retained PC's)	$r = 0.6$ (53 retained PC's)
ARMS on voltages, kV	1%	$7.793 \times 10^{-5}$	$1.567 \times 10^{-4}$	$3.273 \times 10^{-4}$	$6.357 \times 10^{-4}$
	5%	$2.309 \times 10^{-4}$	$3.780 \times 10^{-4}$	$6.461 \times 10^{-4}$	$1.154 \times 10^{-3}$
	50%	$1.763 \times 10^{-3}$	$2.144 \times 10^{-3}$	$2.642 \times 10^{-3}$	$3.447 \times 10^{-3}$
	95%	$1.609 \times 10^{-2}$	$1.547 \times 10^{-2}$	$1.480 \times 10^{-2}$	$1.449 \times 10^{-2}$
	99%	$1.927 \times 10^{-2}$	$1.867 \times 10^{-2}$	$1.744 \times 10^{-2}$	$1.670 \times 10^{-2}$
ARMS on active flows, MW	1%	$8.838 \times 10^{-8}$	$1.193 \times 10^{-7}$	$2.339 \times 10^{-7}$	$1.591 \times 10^{-7}$
	5%	$8.289 \times 10^{-5}$	$1.622 \times 10^{-4}$	$3.518 \times 10^{-4}$	$5.828 \times 10^{-4}$
	50%	$7.500 \times 10^{-4}$	$1.198 \times 10^{-3}$	$1.710 \times 10^{-3}$	$2.709 \times 10^{-3}$
	95%	$2.665 \times 10^{-3}$	$3.404 \times 10^{-3}$	$3.763 \times 10^{-3}$	$5.117 \times 10^{-3}$
	99%	$1.677 \times 10^{-2}$	$1.676 \times 10^{-2}$	$1.654 \times 10^{-2}$	$1.602 \times 10^{-2}$

TABLE IX reports the quantiles of the absolute errors of the means and the standard deviations of  $|V|$  and P respectively over the whole set of buses and of branches. Of course, neglecting a fraction of the total variance implies high errors in variance estimation for the power flows of the branches and the voltages at the nodes close to the stochastic injections with higher participation factors on the discarded PC's. This can be acceptable for a preliminary and fast investigation of the major sources of uncertainty in the grid. For more detailed probabilistic security evaluation, a higher IR must be considered. The proposed method allows to achieve speed up factors up to 8.3, 18.9 respectively for  $r = 0.8$  and  $r = 0.6$ . The system oriented metrics for the different fractions of explained variance (from 1 to 0.6) are reported in TABLE X. Thus, the percent errors on the standard deviations of individual branch power flow or voltage magnitudes can be higher than the percentage of the discarded variance of the



inputs depending on the sensitivities derived from the Jacobian, but the system-level amount of discarded "variance" in the outputs has a good matching with the amount of discarded "variance" on the inputs.

TABLE IX – QUANTILES  $Q_p$  FOR THE ABSOLUTE ERRORS ON THE MEANS AND THE VARIANCES – PAN EUROPEAN GRID MODEL

	$p$	$r = 0.95$	$r = 0.9$	$r = 0.8$	$r = 0.6$
$\Delta\mu_{PROP}^{EM}$ on bus voltages, kV	1%	$6.87 \times 10^{-12}$	$7.63 \times 10^{-12}$	$7.67 \times 10^{-12}$	$8.33 \times 10^{-12}$
	5%	$2.425 \times 10^{-5}$	$2.814 \times 10^{-5}$	$4.155 \times 10^{-5}$	$4.172 \times 10^{-5}$
	50%	$1.825 \times 10^{-3}$	$2.871 \times 10^{-3}$	$3.931 \times 10^{-3}$	$5.474 \times 10^{-3}$
	95%	$3.449 \times 10^{-2}$	$5.851 \times 10^{-2}$	$8.391 \times 10^{-2}$	$1.310 \times 10^{-1}$
	99%	$9.726 \times 10^{-2}$	$1.739 \times 10^{-1}$	$2.377 \times 10^{-1}$	$3.997 \times 10^{-1}$
$\Delta\sigma_{PROP}^{PEM}$ on bus voltages, kV	1%	0	0	0	0
	5%	$2.898 \times 10^{-5}$	$4.944 \times 10^{-5}$	$9.010 \times 10^{-5}$	$1.922 \times 10^{-4}$
	50%	$1.563 \times 10^{-3}$	$2.537 \times 10^{-3}$	$3.992 \times 10^{-3}$	$7.250 \times 10^{-3}$
	95%	$1.862 \times 10^{-2}$	$2.634 \times 10^{-2}$	$3.664 \times 10^{-2}$	$5.325 \times 10^{-2}$
	99%	$4.568 \times 10^{-2}$	$6.220 \times 10^{-2}$	$7.474 \times 10^{-2}$	$1.133 \times 10^{-1}$
$\Delta\mu_{PROP}^{PEM}$ active flows, MW	1%	0	0	0	0
	5%	$1.174 \times 10^{-9}$	$1.349 \times 10^{-9}$	$1.476 \times 10^{-9}$	$1.676 \times 10^{-9}$
	50%	$7.569 \times 10^{-2}$	$1.173 \times 10^{-1}$	$1.553 \times 10^{-1}$	$2.231 \times 10^{-1}$
	95%	3.423	5.713	7.810	$1.174 \times 10$
	99%	9.005	$1.461 \times 10^1$	$1.989 \times 10^1$	$3.175 \times 10^1$
$\Delta\sigma_{PROP}^{PEM}$ active flows, MW	1%	0	0	0	0
	5%	$2.106 \times 10^{-8}$	$4.151 \times 10^{-8}$	$1.907 \times 10^{-7}$	$1.460 \times 10^{-7}$
	50%	$1.146 \times 10^{-1}$	$1.858 \times 10^{-1}$	$3.059 \times 10^{-1}$	$5.533 \times 10^{-1}$
	95%	2.824	4.093	5.638	9.604
	99%	6.825	$1.052 \times 10$	$1.496 \times 10^1$	$2.062 \times 10^1$

TABLE X – WEIGHTED AVERAGE PERCENT ERRORS ON THE MEANS AND THE STANDARD DEVIATIONS OF BUS VOLTAGES AND BRANCH ACTIVE POWER FLOWS – PAN EUROPEAN GRID MODEL

	Weighted average percent error on bus voltage means $\epsilon_{\mu V, \%}$	Weighted average percent error on bus voltage standard deviations $\epsilon_{\sigma V, \%}$	Weighted average percent error on power flow means $\epsilon_{\mu P, \%}$	Weighted average percent error on power flow standard deviations $\epsilon_{\sigma P, \%}$
$r = 1$	$1.769 \times 10^{-5}$	0.5122	$2.850 \times 10^{-4}$	0.6661
$r = 0.95$	$3.419 \times 10^{-3}$	7.26	$5.772 \times 10^{-1}$	10.34
$r = 0.90$	$5.790 \times 10^{-3}$	10.360	$9.401 \times 10^{-1}$	$1.578 \times 10^1$
$r = 0.8$	$8.060 \times 10^{-3}$	14.2565	1.2940	22.4558
$r = 0.6$	$1.267 \times 10^{-2}$	23.6577	1.9973	35.9716

## V. CONCLUSIONS

This paper proposes an efficient method to speed up the probabilistic power flow in large power systems with hundreds of stochastic injections by combining PEM and PCA. The validation for  $r = 1$  against the standard PEM method on medium and large-size systems demonstrates the good matching between the two approaches in terms of absolute errors on the means and the standard deviations. Simulations also highlight the effect of dimensionality reduction on the accuracy and the execution time: even though the errors on standard deviations may be high close to the stochastic injections, the median errors on means and standard deviations are still acceptable, attaining speed up factors up to 19. Simulations also show that the discarded "variance" in the PPF outputs well matches the discarded "variance" on the inputs. The method can be effectively used in planning studies for fast estimation of the uncertainties

from a large number of stochastic injections on the security of large power systems. Future work will consist in an optimal selection of the PC's to attain a minimum target of "explained" variance in a defined subset of PPF outputs.

## ACKNOWLEDGMENT

This work has been financed by the Research Fund for the Italian Electrical System under the Contract Agreement between RSE S.p.A. and the Ministry of Economic Development - General Directorate for Nuclear Energy, Renewable Energy and Energy Efficiency in compliance with the Decree of March 8, 2006.

## REFERENCES

- [1] C. Lemaître, P. Panciatici, "iTesla: Innovative tools for electrical system security within large areas," 2014 IEEE PES General Meeting, National Harbor, July 27-31, 2014.
- [2] M. de Jong et al., "Impact of correlated infeeds on risk-based power system security assessment", PSCC, Wrocław Aug. 18-22, 2014.
- [3] M. Perninge, F. Lindskog and L. Soder, "Importance Sampling of Injected Powers for Electric Power System Security Analysis," *IEEE Transactions on Power Systems*, vol. 27, no. 1, Feb. 2012.
- [4] G. Li, X-P Zhang, "Comparison between two probabilistic load flow methods for reliability assessment", IEEE PES General Meeting, '09.
- [5] D.D. Le et al. "A detailed comparison of cumulant-based probabilistic power flow methods", *Int. Rev. El. Energy*, Vol.7, No.1, 2012.
- [6] E. Ciapessoni, D. Cirio, A. Pitto, "Effect of renewable and load uncertainties on the assessment of power system operational risk", PMAPS Conference, Durham, UK, July 2014.
- [7] E. Ciapessoni, D. Cirio, A. Pitto, S. Massucco, F. Silvestro, "A Novel Approach to Account for Uncertainty and Correlations in Probabilistic Power Flow", IEEE PES ISGT, Istanbul, 12-15 Oct. '14.
- [8] E. Rosenblueth, "Point estimates for probability moments", *Proc. Nat. Acad. of Sci.*, Vol 72, Issue 10, pp. 3812-3814, 1975.
- [9] I. T. Jolliffe, *Principal Component Analysis*, Second Edition, Springer Series in Statistics, 2002.
- [10] H. Yang, B. Zou, "The Point Estimate Method using third order polynomial normal transformation technique to solve probabilistic power flow with correlated wind source and load", APPEEC, 2012.
- [11] R. Christie, "Power Systems Test Case Archive," Aug. 1993. [Online]. Available: <http://www.ee.washington.edu/research/pstca/>
- [12] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, "Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive and corrective actions", *IEEE Trans. on Power Systems*, Vol 28, Issue 4, 2013.
- [13] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," *IEEE Tr. PS*, Vol. 26, No. 1, Feb. 2011.
- [14] S. Alessandrini, S. Sperati, P. Pinson, "The influence of new ECMWF Ensemble Prediction System resolution on wind power forecast accuracy and uncertainty estimation", *Adv. Sci. Res.*, Vol. 8, 2012, pp. 143-147.
- [15] U. Focken et al., "A statistical analysis of the reduction of the wind power prediction error by spatial smoothing effects", *J. Wind Eng. Ind. Aerodyn.*, Vol. 90, Issue 3, pp. 231-246, 2002
- [16] E. Lorenz, J. Hurka, G. Karampela, D. Heinemann, H. Georg Beyer, M. Schneider, "Qualified Forecast of Ensemble Power Production by Spatially Dispersed Grid connected PV Systems", IEA Task 36, 2007
- [17] CAISO, "Integration of renewable resources – Technical appendices for California ISO Renewable Integration Studies", Oct 2010.
- [18] H. Holttinen et al., "Design and operation of power systems with large amounts of wind power", IEA Wind Task 25, Final report, Phase I, 2006-08, Helsinki, 2009.
- [19] J. M. Bogoya, A. Böttcher, and S. M. Grudsky, "Eigenvalues of Hermitian Toeplitz matrices with polynomially increasing entries", *J. Spectr. Theory*, Vol 2, pp. 267–292, 2012.