

**Revision of Manuscript Number:** IEEE-TPWRS-01391-2021, version 1 December 25, 2021

**Journal:** IEEE Transactions on Power Systems

**Author:** Abhishek Halder, Kenneth F. Caluya, Pegah Ojaghi, Xinbo Geng

### **Response to decision:**

The perceptive input from all reviewers and the AE are much appreciated.

Please find below itemized responses to the individual comments addressing the concerns and detailing the revisions. All the edits in the revised manuscript are marked in blue.

Sincerely,  
The authors

### **Response to AE:**

#### **AE.1. Comments by AE:**

*The authors are requested to revise and resubmit their paper after addressing reviewer feedback. Of particular importance is that reviewers point out the model leveraged for analysis is overly simplistic.*

**Response:** Thanks for this comment.

In addition to the itemized responses to the reviewers' comments and the corresponding revisions marked in blue in the manuscript, we have made the following additional fixes and improvements (also marked in blue in the revised manuscript):

- (i) In equation (8), we corrected a typo in its right hand side: the second summand runs over pairs  $(i, j)$  such that  $i < j$ .
- (ii)

#### **AE.2. Comments by AE:**

*Authors are also encouraged to revisit their overview of prior art.*

**Response:** Please see

#### **AE.3. Comments by AE:**

*Presentation of technical points and results can also be improved.*

**Response:** Please see

## Response to Reviewer #1:

### R1.1. Comment by reviewer:

*The original contribution of the article is not clear. Indeed, what is original new from this article, not the summary from existing techniques?*

**Response:** This is.

### R1.2. Comments by reviewer:

*The structure of the article should be highly improved. A large portion of the contexts is simply repeating the well-known power system classical model, kroon reduction, and the FPF equations. Only from section IV, page 6, I can see some of the proposed weighted points methods.*

**Response:** We have.

### R1.3. Comments by reviewer:

*I have major doubts about the model used in the article for the following reasons.*

- (1) The paper claim the proposed method can propagate stochastic uncertainties in realistic power networks. However, only a classical model is used, which is not realistic at all.*
- (2) Why do active power and reactive power from load have no randomness at all. This is too unrealistic.*
- (3) The diffusion term of this article is only considered in the frequency. The process noise in the rotor angle is fully ignored. This is also unreasonable.*

*To improve your work, I highly suggested the authors follow [R1], which is missing in your literature but can help the authors better know Power system SDE. Also, it is suggested to introduce the OU process to model the randomness in the load, which cannot be ignored in practice.*

*R1. A systematic method to model power systems as stochastic differential-algebraic equations, F. Milano, R. Zarate-Minano, IEEE Transactions on Power Systems, 28(4), 4537-4544.*

**Response:** TBD.

### R1.4. Comments by reviewer:

*There is no comparison of the proposed method compared with other state-of-the-art methods. Therefore, it is hard for the readers to judge the importance of this article.*

**Response:** TBD.

### R1.5. Comments by reviewer:

*Contributions are over-claimed in this article:*

- (1) The scalability of the proposed method cannot convince the readers since you are using a too simple model and ignored many important random dimensions from the loads.*
- (2) The proposed method claims to be nonparametric, however, all the inputs are assumed to be either wiener process or uniform distribution, which are all parametric distributions.*
- (3) If the “nonparametric” is only referred to the output PDF, I would not think of it as a major contribution since there are a lot of existing techniques that can depict nonparametric PDF.*

**Response:** TBD.

### R1.6. Comments by reviewer:

*The simulation time is too short to convince the readers. Only 1 second is considered here. Note that the errors can be quickly accumulated in the iteration.*

**Response:** TBD.

**R1.7. Comments by reviewer:**

*It is suggested to count the total computing time, not on each step.*

**Response:** While the depiction of the computational times in Figs. 8 and 9 in the revised manuscript are somewhat unconventional, we explain why we choose to plot that way.

At each fixed physical time  $t_k = kh$ ,  $k \in \mathbb{N}$ , where  $h$  is the fixed step-size, we perform the proximal update  $q_{k-1} \mapsto q_k$  via Algorithm 1. This proximal update involves execution of a contractive fixed point recursion. The physical time  $t_k$  is “frozen” (zero-order-hold) during the fixed point recursions underlying this one-step proximal update. To claim that the proposed method is practical, we therefore, need to demonstrate that the computational timescale is faster than the physical timescale, i.e., computational times needed to complete the zero-order-hold fixed point recursions are orders-of-magnitude faster than  $h$ , the step-size of physical time. On the other hand, theoretical consistency (see the unnumbered equation after (26), before Sec. IV) requires that  $h$  itself should be small. Thus, even if one has a provably convergent algorithm that can perform the one step proximal update, it is not clear *a priori* if the computational time needed to do so can be much faster than  $h$ . If not, then the method would be statistically consistent but practically not useful. The Figs. 8 and 9 in the revised manuscript show that the computational time needed to perform the proximal updates, for all  $t_k$ , are indeed orders-of-magnitude faster than  $h$ .

**R1.8. Comments by reviewer:**

*If the proposed method depends on the Kronecker product, how can it be scaled to a higher dimension?*

**Response:** The proposed method does not entail computation of a Kronecker product. In the manuscript, the Kronecker product with the  $2 \times 2$  identity matrix appeared in the definition of the change-of-variable (17). Its only purpose is to define that change-of-variable, and thereby derive the PDE in the new variables. The proximal recursion (for the PDE thus derived) detailed in Algorithm 1 does not have any Kronecker product computation.

**R1.9. Comments by reviewer:**

*beta is undefined in equation (15a).*

**Response:** Thanks for the comment. The  $\beta$  in (15a) is the same  $\beta$  in (16), as in part of the same sentence. To avoid any confusion, in the revised manuscript, we have moved the phrase “for some  $\beta > 0$ ” in (15a).

## Response to Reviewer #2:

### R2.1. Comments by reviewer:

*Overall, the paper is well-written and theoretically sound.*

**Response:** Thank you for the kind and encouraging remarks.

### R2.2. Comments by reviewer:

*The paper's language and presentation make it seem like it is more oriented towards the applied mathematics and control communities. It is highly suggested that the authors revise the exposition of ideas and concepts, putting more emphasis on the power system so that it can be well-understood by the power system experts.*

**Response:** TBD.

### R2.3. Comments by reviewer:

*How will the proposed approach be used for the cases in which dependencies among uncertainty sources exist? An explanation and illustration would be helpful.*

**Response:** In our paper, we do not make any assumption on statistical independence of the uncertainties. In fact, all stochastic uncertainties are considered to be “joint”, i.e., statistically correlated. Please note that the initial joint PDFs (28) and (30) are taken to be in that form only for the sake of concrete example simulations. Their only purpose, in the simulations reported in Sec IV, is to generate joint samples, and to evaluate the initial joint PDF values at the joint samples thus generated. If this initial scattered point cloud data is supplied to us only as a numerical table, without specifying any functional form whatsoever, nothing changes in the proposed algorithm. Similarly, the parameters in Table I could be correlated random vectors, and all we need are the joint parametric weighted point cloud data. In fact, this is how we expect the proposed computational framework to be applied in practice.

### R2.4. Comments by reviewer:

*Incorporation of power system dynamics into the proposed uncertainty propagation framework is a bit vague. The proposed study seems not to have utilized any tool capable of simulating power system dynamics. Could the authors explain the part pertaining to the dynamic simulation in more detail?*

**Response:** TBD.

### R2.5. Comments by reviewer:

*With regard to the point above, this reviewer would expect to see uncertainty bounds for the system dynamic response such as the trajectory of the rotor angle. Without such information, the method's usability is difficult to gauge.*

**Response:** TBD.

### R2.6. Comments by reviewer:

*Instead of using a purely synthetic system, the authors are highly encouraged to choose available test cases such as the IEEE 118-bus system and a larger one, if possible.*

**Response:** TBD.

## Response to Reviewer #3:

### **R3.1. Comments by reviewer:**

*My main concern with this work is that all its derivations are deeply model-dependent and, unfortunately, the model that the author consider is not adequate to represent modern power systems. Moreover, second order model of synchronous machines are adequate only in a time scale of few seconds, so, even if the system were effectively including only conventional power plants, the model would be acceptable only for very short-term transients, i.e., transients for which the propagation of the noise is immaterial. In other words, the proposed model is useless.*

**Response:** TBD.

### **R3.2. Comments by reviewer:**

*The literature is incomplete, many recent works on SDAEs for power system applications are not mentioned.*

**Response:** TBD.

### **R3.3. Comments by reviewer:**

*There are several typos in the test (especially in the references).*

**Response:** TBD.

### **R3.4. Comments by reviewer:**

*Figures 5 and 6 are unreadable.*

**Response:** TBD.

## Response to Reviewer #4:

### R4.1. Comments by reviewer:

*The main contribution of this paper lies in applying results on the governing partial differential equation (PDE) for the joint PDF to avoid discretization of the state space, thereby circumventing computational challenges. The reviewer appreciates that this paper is rigorous in development and thinks it deals with an important problem.*

**Response:** We appreciate the positive remark.

### R4.2. Comments by reviewer:

*The key contribution of this paper is that the proposed method can reduce the computation burden when solving for the transient PDF. However, no comparison is provided with existing methods. The authors should provide comparisons to the exiting approaches, for example, based on finite element discretization or other numerical methods. The test example could be performed using the 5-generator system or larger systems. Besides, referring to Figs. 8 and 9, the total computation time should be reported for a longer horizon.*

**Response:** TBD.

### R4.3. Comments by reviewer:

*Regarding the presentation, the reviewer suggests making the paper more self-contained. Some discussion in this paper highly relies on the authors' previous work, which can be improved by either adding a summarization or trimming out the materials.*

*Following the same vein, there are some unnecessary information and technical points provided in this paper that may be distracting. For example, on page 3, why would the authors mention the Kuramoto oscillator model? On page 4, is the detailed discussion regarding Wasserstein-2 metric necessary? Similarly, on page 5, the discussion about the natural energy functional and how the gradient drift case does not apply also seems too detailed.*

**Response:** TBD FOR FIRST PARA.

We appreciate the reviewer's point that the sentence containing equation (6) and the mention of Kuramoto model can be removed without affecting the rest of the development. However, we prefer to keep it since it adds didactic value, and many engineers may feel comfortable by the explicit mention that (1) is simply a rewriting of (6). Furthermore, an astute reader, encountering that sentence will realize that the scope of the technical contribution of this paper is broad because the algorithmic framework developed here will be applicable to several other non-power system applications (mentioned in the two cited references in that sentence) where the second order Kuramoto model is used. Since this is just a matter of one sentence, we will prefer to keep it. We hope you approve.

On the other hand, we believe the detailed discussion regarding Wasserstein gradient flow and why that standard development cannot be applied in this case, is in fact, the heart of this paper. Please note that ours is not a recipe paper where a numerical algorithm is proposed backed by some theory. The key contribution of this paper, from the authors' viewpoint, is to help distill out the core technical difficulty that hinders simply importing standard techniques from the optimal transport literature. We believe this clarification of what really is the technical crux of the matter—versus what are peripheral issues—should be illuminating to not just power systems researchers, but also to applied mathematicians underscoring how the standard theory and algorithms for measure-valued gradient flows need to be adapted in engineering applications.

**R4.4. Comments by reviewer:**

*Along with [25], there are recent developments that over-approximates the reachable set for nonlinear systems under uncertain initial conditions, parameters, and external disturbances: “Reach-Set Estimation for DAE Systems under Uncertainty and Disturbances Using Trajectory Sensitivity and Logarithmic Norm”.*

**Response:** Thanks for providing that interesting recent reference. We have cited the same in the revised manuscript.

**R4.5. Comments by reviewer:**

*On page 7, the authors stated about “To account the pertinent geometry of  $\mathbb{T}^n \times \mathbb{R}^n$ ”. What exactly does this mean? Do the authors indicate that in (28) the two groups of state values are independent? Or do the authors mean the specific choices of the product von Mises PDF and the uniform distribution? If so, what are the connections?*

**Response:** We simply mean that the joint PDF should be supported on  $\mathbb{T}^n \times \mathbb{R}^n$ , or a submanifold thereof. For instance, one cannot assume the initial joint PDF to be a  $2n$  dimensional multivariate normal since that is supported over  $\mathbb{R}^{2n}$ , not over  $\mathbb{T}^n \times \mathbb{R}^n$ . It is, however, perfectly fine if some or all of the  $2n$  components of the initial state vector are statistically correlated. As long as the support is  $\mathbb{T}^n \times \mathbb{R}^n$ , there is no issue if the angular variables are correlated with the Euclidean (i.e., angular velocity) variables.

**R4.6. Comments by reviewer:**

*On page 5, is there a mistake in “random samples from (28), and evaluated at (28)”?*

**Response:** That was not a typo. Indeed, equation (28) was used in two ways: it was used to generate initial state samples, and then the initial joint PDF values at those samples were evaluated via the functional form of (28), i.e., as in function evaluation. This is needed in our case because unlike standard Monte Carlo, we co-evolve the state samples and joint PDF values at those state samples. To clarify this further, in the revised manuscript, we have slightly expanded that sentence.

**R4.7. Comments by reviewer:**

*As a validation, it would be good to also provide information of higher-order moments along with the mean given in Fig. 10.*

**Response:** TBD.

**R4.8. Comments by reviewer:**

*There are some typos in this paper:*

*Page 4: in therms of*

*Page 5: in the right-hand side of the arrow in the expression following “converges to the flow generated by (22)”, it should be plain  $\rho$*

**Response:** Corrected both. Thanks for pointing them out.