

Revision of Manuscript Number: TPWRS-01391-2021.R1

August 12, 2022

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:

Authors are requested to revise and resubmit the manuscript after addressing all reviewer comments that persist.

Response: We.

Response to Reviewer #1:

R1.1. Comment by reviewer:

No further comments.

Response: Thank you for your time and efforts to help improve the manuscript.

Response to Reviewer #2:

R2.1. Comments by reviewer:

This reviewer appreciates the authors' effort in revising the manuscript based on the feedback in the previous round while clarifying and addressing the remarks with due diligence. The revised version reads well and adequately takes into account several relevant facets. The authors are kindly requested to proofread the article prior to final submission.

Response: Thank you, we have thoroughly proofread the article. We appreciate your comments from previous rounds to help us improve the manuscript.

Response to Reviewer #3:

R3.1. Comments by reviewer:

First and very importantly, the authors are suggested to readdress my previous two questions.

(1) Regarding my previous question about the classical model, I need to reemphasize here that the model is too simplistic that does not sufficiently reflect the responses from the exciter and the governor. Therefore, for a full paper on IEEE transactions on power systems, this model is incomplete, and cannot reflect reality. Also, for such an over-simplistic model, I still doubt the scalability of the method. Unfortunately, this problem is not addressed and improved after the revision.

(2) Regarding my previous questions about the comparison studies with the state-of-the-art method, indeed, I am very disappointed to find that the revised version still has added no comparison study with the state-of-the-art method at all. The ONLY comparison is still with Monte Carlo.

Response: [The](#)

R3.2. Comments by reviewer:

Second and most importantly, I disagree with the theory and the claim behind the SDE model used in this paper.1. Here, I am very familiar with the Ito SDE model in power systems that you cited here in your refs [16], [17], [45].

Although the authors claim the model you used here in equation (7) is the same as the ones used in ref[16] and ref[17] and extended from ref[45], I would like to tell the authors: NO, you are not using the same one. [16][17][45] are all right, but the one used in this article is incorrect. Please pay very special attention to [16][17], they only add uncertainties directly on differential equations, namely \dot{P} and \dot{Q} here, but not from P and Q . This is because transferring the power system model from DAE to Ito-SDE is still an unsolved problem where P and Q is contained in algebraic equations while Ito-SDE only considers differential equations. Also, [45] is one machine system that does not even have a network. That is different from your problem. Here, the authors might argue that you are using the Kron reductions to remove algebraic equations. However, I have to tell the authors the Kron reductions to apply to the SDE model here. It is because the Kron reduction is relying on constant P and Q to derive a constant admittance matrix. However, the P and Q in this manuscript are also assumed to be uncertain, and also claimed in the response letter to be time-varying stochastic processes. Then, its assumptions do not hold. Then, its Ito SDE form is inconsistent. I highly suggest to the authors study the following article carefully and learn more about the basics of Kron reduction.

Time-domain generalization of Kron Reduction, MK Singh, S Dhople, F Dorfler, GB Giannakis.

Response: [The.](#)

R3.3. Comments by reviewer:

The simulation results are weird. Just a simple check at fig 8. How a power system dynamic simulation with line 13 tripped can has no oscillations at all within the first 1 second? If so, it has no practical meaning at all.

Response: [The.](#)

R3.4. Comments by reviewer:

Third, my engineering motivation for me is not clear. The authors reclaim multiple times about the joint pdf for the output system state. I understand that we consider joint pdf in the model inputs and understand its practical meaning. However, among the current literature, to my best knowledge, I saw no previous efforts in obtaining the “joint” pdf of outputs. They only seed for marginal pdf for output states. So an important question is what motivates us to get joint pdf, instead of individual pdf, for model output? For

which exactly applications, you must know the joint pdf of the output states? What is its practical meaning and why it is important?

Response: [Thanks for this comment.](#) The

R3.5. Comments by reviewer:

Also, the authors should put a very clear position about some of the references that you are making use to argue with the reviewer. For example, [45] is a 10-year ago paper. Why do you claim it to the state-of-the-art? How can this argument convince the reviewer?

Response: [The](#)

R3.5. Comments by reviewer:

Regarding Ref [23]F. Milano, and R. Zarate-Minano, “A systematic method to model power systems as stochastic differential-algebraic equations”, IEEE transactions on Power Systems, Vol. 28, No. 4, pp. 4537–4544, 2013. I mentioned this article in the first-round review, not for the purpose of a simple cite, but trying to help you to improve the model and make your work realistic. However, a suggestion is ignored. This is indeed a pity. Besides, please do not argue again that this is a different work because this ref is well-known in modeling realistic stochastic power system dynamics, which should be the foundation of your model.

Response: [The](#)

R3.6. Comments by reviewer:

In general, I need to claim here that math is fine in itself, but only publishable when connected to or interpreted correctly in a convincing engineering application.

Response: [The](#)

Response to Reviewer #4:

R4.1. Comments by reviewer:

The authors have successfully addressed most of the reviewer's previous concerns. However, the reviewer still has some questions, mostly regarding the numerical simulations.

Response: Thank you. Please find our itemized responses to the questions related to numerical simulations.

R4.2. Comments by reviewer:

It is stated that "The Fig 9 shows that the computational time needed to perform the proximal updates, for all t_k , are indeed orders-of-magnitude faster than h ." This statement doesn't seem to hold, since $h = 10^{-3}$ and the computational time lies within the range of around 10^{-3} to 10^{-2} . How does the computational time needed to perform the proximal updates relate to time step h ? If they are not strongly related, does that mean the step size h is constrained to be greater than that computational time?

Response: The reviewer is right, we should have been more careful in phrasing that sentence. We meant to say that the computational time needed to perform a proximal update is statistically comparable (in the big O sense) to h . The reason why the precise computational time is probabilistic is twofold: random sampling, and randomness in computational latency, power and memory footprint. Fig. 9 shows that $h = 10^{-3}$ and the computational time needed to perform a single proximal update is $O(10^{-3})$ with high probability in the sense most time-steps incur $O(10^{-3})$ time. In Fig. 9, while instances when computational times are $O(10^{-2})$ do occur for Case I, they are rare. In the present revision, we have now rephrased the corresponding text (highlighted in blue) toward the end of Sec. V-A.

The proximal updates involve certain block coordinate ascent recursions in the dual co-ordinates \mathbf{y}, \mathbf{z} ; see lines 13-21 in Algorithm 1. This recursion is guaranteed to be contractive (hence existence, uniqueness and convergence are guaranteed) in the product orthant cone: $(\mathbf{y}, \mathbf{z}) \in \mathbb{R}_{>0}^N \times \mathbb{R}_{>0}^N$, see ref. [41, Sec. III-B, III-C] and especially Corollary 4 in that reference. We have cited that reference in this manuscript's Sec. IV. While h appears as parameter in this recursion (lines 5 and 14 in Algorithm 1), we do not have analytical relation between the computational time and h . We did not constrain or adapt h in Algorithm 1. We simply set a small value $h = 10^{-3}$ to promote statistical consistency (see the unnumbered equation after (27)). Depending on the random sampling, dynamics and computational platform, it is possible for the single-step proximal update computational time to fall below h (see e.g., ref. [41, Fig. 9] which used the same $h = 10^{-3}$). The standard desktop computer/processor we used for our simulation is reported in Sec. V, first paragraph, last sentence.

R4.3. Comments by reviewer:

In Fig. 10, could the authors please explain why the computational time of Case I is less than that of Case II? The computational cost for the 5-generator system is roughly 4 times that of Case II. Could the authors comment if such a scale makes sense intuitively considering the state dimension of the two systems?

Response: Thanks for this comment. The state dimensions for both Cases I and II in Fig. 10 are the same, i.e., $5 \times 2 = 10$ dimensions. Other than the difference in dynamics due to line failure in Case II, we do not think there is any specific physical insight to be gleaned on why the computational times for Case II are below Case I; they both use identical set of random initial samples although the Wiener process noise realizations are different.

Regarding the scaling comment, perhaps the reviewer meant to say that the total computational times for the 50 generator (100 state) system are roughly 4 times that of Case II (5 generator, 10 state). The possible explanation of why the scaling is not so bad is because the number of samples for the 50

generator case is 2000, which is twice the number of samples used in Cases I and II. This is made possible by our usage of scattered point clouds with explicit proximal computation of the time-varying probability weights, as opposed to gridding the state space and then approximating the joint PDF via histogram (which would have caused exponential in dimension scaling in the computational cost). Being able to directly evaluate the joint PDF values via proximal computation allows us to bypass construction of function approximation oracle over the high dimensional state space (which in general scales poorly with dimensions), and then to evaluate it. A direct comparison between the two cases is less meaningful due to difference in parameter values, initial joint PDF etc. However, the general trend that the proposed scattered point cloud-type proximal framework, similar to the one proposed here, can incur computational times which scale sublinearly with dimensions while resolving the joint PDF well, were reported in our earlier works in different application contexts: ref. [41, Sec. V-B and Fig. 14], ref. [42, Fig. 5].

R4.4. Comments by reviewer:

Both Fig. 12 and Fig. 13 demonstrate a decreasing trend as k increases. Could the authors comment on this and link this to the property of the proposed proximal algorithm? Similarly, why is the trend in Fig. 14? Comparing Figs. 12, 13 and Fig. 14, did the authors omit the initial time instant in Figs. 12, 13?

Response: Thanks for ...