Reviewer report

Anonymous Referee

December 30, 2019

1 Summary

My apologies for the delayed review and thanks for your patience. I hope the quality of my review warranted this delay.

Overall, the manuscript has much improved and in my opinion, there remain only minor points to be necessarily addressed.

- Page 14, below Eq. (62): How can Q be both a first integral for (61) (line 14) and not be conserved exactly (line 16–19)? This is particularly confusing in light of Sect. 6. I am guessing this is only a case of confusing wording?
- In the first submission, I had troubles understanding how the long-time conservation in the Hamiltonian was achieved given the impossibility remarks in the beginning of this section. I now understand that this property only holds in expectation, but not path wise. I am of the opinion that this needs to be established in the numerical experiments section as well.

  My personal proposal would be to replicate Fig. 9 only displaying step sizes $h = 0.2$ and $h = 0.05$ and plotting the Hamiltonian over time of a handful of indidivual sample paths. But maybe the authors find an even more suitable graphical representation.
- Sect. 9.3, mean-square convergence of MC estimators: I am of the strong opinion that there should be at least one experiment showing the variability of the estimator with respect to sample size M. Maybe this could be easiest achieved by plotting multiple curves in Fig. 6 with different sample sizes. Also include error bars.

However, I still find some parts of the manuscript confusing, in particular Sects. 6 and 7 which require more specialized background reading. I have collected my editorial suggestions below and invite both the authors and the editor to take these into considerations at their will and jugdement.

2 Optional Suggestions and Minor Points

- Page 2, paragraph "Chaotic differential equations": halfway through the paragraph, system variable is changed from y to x. Line 31: consider writing σ^2 instead of σ. Which are the three different scales? Fig. 1 requires y-labels.
- Page 5, Remark 1: I misremembered Conrad et al. to be only valid for explicit integrators. Given the same underlying integrator, I'd argue that the cost of the two methods is actually either the same (the cost of the RNG is probably neglible given the cost of the integrator) or incomparable (it is unclear whether the additive noise makes the implicit system harder or easier to solve). Maybe the editor agrees that this remark could either be removed or restated simply as "have the some computational complexity".
- Page 6, Assumption 2.(ii): I am still uncertain about this point. Consider Hairer, Norsett and Wanner, Sect. II.2, Theorem 2.11. I take this result as "if the vector field is sufficiently often differentiable, the integrator is sufficiently often differentiable". Wouldn't this result suffice to guarantee 2.(ii), given a sufficiently often differentiable vector field? On the other hand: f needs to be sufficiently often differentiable anyway in order to obtain high order convergence, correct? Or have I forgotten any other corner case?
- Page 6, Def. 1: This is the first instance of a final time $T = Nh$. Consider stating explicitly that you are considering a finite integration domain $[t_0, T]$.
- Page 7, Lemma 1: You mention the assumptions 1, 3 and 3 in this order. Why? Consider reordering the assumptions to ensure numerical ordering.
- Page 8, Assumption 4: you have changed notation from y to u. Why? Consider using y for the variable name. Same for Lemma 2.
- Page 9, Remark 6: consider creating a simple sketch for Remark 6, maybe showing 3 steps to make this point graphically obvious for the reader. This could look similar to Fig. 1.
- Page 10, line 21: Consider putting the comment about the proof outline of Lemma 2 to the previous section, i.e., to Lemma 2.
- Page 11, line 39: "number of number of"—remove one instance of "number of"
- Page 13, Proof of Corollary 1: this is not really a consequence, but analogous, no?

- Page 14, Eqs. (63) through (65): Is the symplecticity of an integrator a property of the integrator alone or are there integrators that are only symplectic for certain Hamiltonian systems? Is there a main argument, why (64) would be the new condition? In light of the first question: couldn't there be a case, where (64) is satisfied for a certain integrator and vector field, but not (65)?

- Page 15, lines 16–38: this is introduced with little background for verification and not followed-up upon later. For me, this has generated more distraction than intuition. Consider removing these sections, in particular Eq. (66). Consider simply introducing the modified Hamiltonian (69) and refer to Hairer, Lubich and Wanner for details.

- Lemma 7: may I propose to use other variable names for a_{jk}, b_j. These might get easily confused with the parameters of a RK method. Why not use directly δ, η?

- Theorem 6: is there any example in the literature where the constants can be simply stated? It would be nice to have an example where the minimum in Eq. (83) could be resolved to a concrete value.

- Page 18, Eq. (87), first line: superfluous point . at the end of the equation line.

- Page 18, line 33: referencing the Eqs. in the appendix breaks the flow of the proof. Consider putting them as particular results in Lemma 8 in the main text.

- Experimental section: Sects. 9.1, 9.2 and 9.3 correspond to Theorems 2, 1 and 3 resp. Consider reordering them to match the presentation order earlier in the manuscript.

- Page 27, line 52, "integrated": should probably be present tense.

- Page 29, last line: this sentence is hard to parse. You seem to be aware of this given that you explained the sentence in more detail as a comment to Reviewer 1. Consider rephrasing this sentence. Potentially, introduce variable names, e.g., "... initial conditions $y_{0,i}, i = 1, \ldots$ with corresponding energy levels $Q_{0,i} = Q(y_{0,i}) \neq Q(y_0)$, the true initial energy level, might be mapped by Heun's method to points $y_{1,i}$ with energy levels $Q_{1,i}$ much closer to $Q(y_0)$ thus leading to a posterior distribution ..."