

Response to Reviewers

PONE-D-22-22196 Compressed computations using wavelets for hidden Markov models with continuous observations

In the following I will address the reviewer's comments designated as mandatory by the academic editor in detail. It seems that we failed to communicate sufficiently clearly that our method is the *very first* and *only* method to allow compressive computations for frequentist HMM algorithms with continuous emissions (cf. below). We added exposition to the manuscript to clarify this point.

For the authors,
Alexander Schliep

1. There should be a separate section in the manuscript highlighting the novelty of the proposed method.

The novelty of the method has previously been highlighted in the final paragraph of the introduction. We added a section header to indicate it.

2. The proposed method should be compared with other state-of-the-art techniques to highlight its effectiveness.

Ours is the *first* method to perform compressed computations of Likelihood, Viterbi paths, and Maximum-likelihood estimation with the Baum-Welch algorithm for HMMs with *continuous* emissions. Prior approaches, including ours for Bayesian HMMs with discrete and continuous emissions, tackle different problems or different types of emissions. As such, our method is the only method for this problem, making an evaluation benchmark impossible as no competing methods exist.

We added some text to abstract, introduction and conclusion to make this more obvious.

3. At the end of the introduction, it is difficult to understand your research motivation due to the lack of necessary summary and innovation of the method.

To summarize, the research motivation is as follows. Compressed computations for Hidden Markov Models have been used with great success for three out of the four cases resulting from two types of emissions (discrete vs. continuous) and two types of computations (frequentist—Viterbi, likelihood and Baum-Welch—vs. Bayesian MCMC) by us and one other group. The submission focuses on the fourth and last open case (frequentist and continuous). This has been motivated in the abstract “Here we extend the compressive computation approach for the first time to the classical frequentist HMM algorithms on continuous-valued observations”. In the introduction we give a complete summary of the subfield “Mozes et al. ... for discrete observations achieve considerable speed-ups”, “Also for discrete observations, Mahmud et al. substantially accelerate Forward-Backward Gibbs

(FBG) sampling ...” and “substantial improvement [18] in the running times of the FBG sampler for continuous-valued observation”, to “In the following, we introduce compressive computations based on wavelets for Rabiner’s three original problems [1]: likelihood computation, computing Viterbi paths, and Maximum-likelihood estimation with the Baum-Welch algorithm.” We also added corresponding language to the discussion.

The last paragraph of the introduction is a succinct, complete and exhaustive summary of what we accomplish in the manuscript, now highlighted by a separate section header.

4. There is a lack of concatenation among the chapters of the article. It is suggested that the motivation between the chapters be supplemented.

We added some additional exposition to the manuscript.

5. Formula (3) lacks the introduction of variable definition and it is suggested to be supplemented.

It is not clear to us which variable(s) the reviewer might be referring to. Every variable used in equations (3a)–(3b) has been defined and explained in lines 86–97, resp. lines 72–84 (conditional probability), and line 102 (the likelihood function). We additionally made the definition of y_s^t explicit by repeating and adapting the definition of q_s^t .

6. It is suggested that the next research direction should be supplemented in the conclusion, and the reasonable prospect of this study should be carried out in order to inspire the follow-up scholars.

As a matter of fact we already do discuss potential follow-up research in the conclusion (note, we follow PLOS ONE suggested article structure), in particular suggesting an extension to heteroscedastic noise and a potential use of the symbolic indexing by Keogh et al..

7. What are the meanings of these variables in formula (2)? Please add.

It is not clear to us which variable(s) the reviewer might be referring to. Every variable used in equation (2) has been defined and explained in lines 86–97, resp. lines 72–84 (conditional probability), and line 102 (the likelihood function). We additionally made the definition of y_s^t explicit by repeating and adapting the definition of q_s^t .

8. Why the method proposed by the author has achieved the best effect, please make a specific analysis.

We don’t make the claim that our method has the “best effect”, as we do not perform a

benchmark evaluation against other methods. The reason why we don't perform a benchmark evaluation is that ours is the *first* method to perform compressed computations of likelihood, Viterbi paths, and Maximum-likelihood estimation with the Baum-Welch algorithm for HMMs with continuous emissions. Prior approaches, including ours for Bayesian HMMs with discrete and continuous emissions, tackle different problems or different types of emissions. As such, our method is the only and thus *the* state-of-the-art method for this problem, making an evaluation benchmark impossible as no competing method exists.

What we did benchmark is a comparison of uncompressed (prior art) and compressed (our contribution) versions; to reiterate there are no other compression method for the continuous case and the particular HMM algorithm we consider.