

Review of
“Markov-switching state space models
for uncovering musical interpretation”

This paper develops a state space model for analyzing tempo changes in recorded classical music. The model is applied to 46 recordings of Chopin’s Mazurka in F (Op. 68, no. 3) using an available pre-processed data set containing local tempo determinations. The paper is clearly written, well-organized, and enjoyable to read. I believe that the material is an excellent fit to AOAS. I divide my comments between major and minor comments below.

Major comments:

1. The most serious concern has to do with viewing changes in tempo as additive rather than multiplicative. There is a tradition to view tempo relations as proportional rather than additive (see, for example, Mead, A. (2007). On Tempo Relations. *Perspectives of New Music*, 64-108). This would suggest that the authors’ state-space model might be more appropriately applied to the log of tempo.
2. I also had some concerns about the lack of checking the sensitivity to the prior distribution. The authors discuss in Section 3.6 the reasons for identifiability problems and the need for some of the prior distribution choices to address these problems, but it was difficult to understand how the particular distributional choices were likely to impact inferences. More is needed here. The decision to model variances as Gamma-distributed (as opposed to inverse-Gamma, or uniform as suggested by Gelman and others in similar settings) was a curious choice.
3. The material on clustering could be improved. For clustering according to Mahalanobis distances, I was not clear why the authors did not choose to standardize by the inverse of the posterior covariance matrix. Also, given the likely skewed distribution of the estimated variances across performers, it might have been more sensible to include the variance parameters in computing Mahalanobis distances on the log scale. It was also unclear how the specific number of clusters was determined. Methods exist (e.g., Tibshirani’s gap statistic) to make this choice.

Minor comments:

1. The abstract is clear, but does not specify that the focus of the paper is on modeling tempo changes within music. The abstract should be revised accordingly.

2. Page 5, para before section 2: "it's statistical estimation" should be "its statistical estimation"
3. Page 6: Just below equation (1), conventionally ϵ_i and η_i are not identically distributed.
4. Page 6: In equation (2), c and d are vectors, not matrices, as incorrectly stated below the equation. This reference to matrices occur at various points within the text.
5. Page 7: It would be helpful to explain up front that in the switching state-space model, the continuous hidden states are functions of both the current and previous switch states. Some rationale should be provided early on for this choice of a second-order Markov model, i.e., that the states depend on both velocity (tempo) and acceleration.
6. I think it would be helpful to discuss whether the switching model developed using four discrete states is more generally applicable beyond Chopin. The first three states seem reasonably generalizable, and perhaps single note stress is widely applicable in piano performance, but this is not entirely clear.
7. Small notational issue: Please use ℓ_i (\code{\ell}) instead of l_i to denote the duration of note i .
8. Page 13: The Fernhead and Clifford citation should appear as "Fernhead and Clifford (2003)". Several other similar instances of incorrectly displayed citations should be fixed.
9. What are the differences between the model in the current manuscript and the generative model in Gu and Raphael (2012)? This question arises in the context of section 2.5.