

Reviewer 1

Thanks for your careful evaluation and your comments and suggestions. Please find below answers to your remarks.

- 1) *page 3, equation (7). The authors should provide the definition of the matrix \tilde{A}^ℓ .*
The definition is now provided by the unnumbered equation between equation (5) and (6).
- 2) *page 3, second column. Based on the definitions of \mathbf{X} and \mathbf{Y} given in the first column as well as equations (4) and (5), it is unclear how equation (6) could allow to forecast the signal at time $\frac{N-1+\ell}{f_s}$. Please give more details.*
The paragraph has been rephrased, in order to make the forecasting procedure clearer. In particular, we detail how we obtain the estimation of $\mathbf{x}_{k+\ell}$ by recursively applying the linear relation (4). The estimation of the signal at time $\frac{N-1+\ell}{f_s}$ is then directly deduced from the last element of this vector.
- 3) *page 7, second column. The authors wrote that the choice of the M parameter is especially crucial when the deterministic component of the signal is no longer stationary, like in the AHM model. However, when they present an application of the proposed method to a signal satisfying the AHM model, neither they discuss how they select the M value, nor they provide the analysis of the dependency of the performance of the proposed algorithm on the M value. Please include in section IV-C2 the evolution of the performance of the proposed method as a function of M and an explanation on how to tune its value in applications.*
We have added the new paragraph IV-C2a where we discuss the influence of the parameter M on the forecasting performance of an adaptive harmonic signal. In particular we mention that choosing a too small value for M does not provide enough information to forecast the signal satisfactorily, while a value of M that is too large makes the algorithm insensitive to local nonstationarities produced by frequency variations. The user must then make a compromise on the choice of M in order to optimize the performance of the algorithm. A choice of M such that the subsignals contain about a dozen oscillations is generally appropriate.
- 4) *page 8, Table I. The authors present the performance of the extension method when applied to a toy AHM model. But they do not provide an insight of the impact, in this specific case, on the associated "boundary-free" TFR. Please include in this table the "performance index D " estimation proposed in section "Evaluation Metrics".*
Table III now includes a new column containing the performance index D of the boundary-free STFT associated with each of the extension method that we tested. For the sake of clarity, the performance of the boundary-free versions of the other TF representations (SST, RS and ConceFT) are not provided in this section. Their study is referred to the analysis of the real physiological signals (section IV-D).
- 5) *page 10, Table III. Please provide the performance results associated with the first metric you proposed to use, the MSE.*
This results have been added to Table III.
- 6) *All real life signals proposed in the work are nonstationary signals which contain one single nonstationary component plus noise. In the toy example of a AHM signal, section IV-C2, you presented the case of a signal containing two nonstationary components. Please add the example of a real life signal containing two or more simple nonstationary components plus noise.*

Idea?

Reviewer 2

Thanks for your careful evaluation of the paper and your comments, which we tried to take into account as much as we could. You will find below answers to your remarks, suggestions and questions.

- 1) *The approach is similar to the shadowing approach suggested by C. Gregori in the Nonlinear Dynamics community, (see his Physics Review Letters in 1990 Shadowing of physical trajectories in chaotic dynamics) and subsequent papers on shadowing in multiple domains e.g. by Vargas et al in IEEE Transactions on Audio, Speech and Language in 2011.*

We have mentioned this approach in the introduction of our article. We precise that dynamic mode predictors have been proposed in speech processing by Vargas and McLaughlin, to forecast signals falling into the so-called source-filter model. This gradient-based technique rely on the shadowing approach proposed in Grebogi et al. Nevertheless, since the source-filter is not adapted to model physiological signals on which we focus, we do not implement this method in our article.

- 2) *The signal is assumed to be deterministic with additive noise, in my experience oscillatory signals of interest in many applications are a combination of deterministic and stochastic components. How robust is the scheme?*

A new paragraph has been added to precise the limitations of the theoretical justifications on the performance of **SigExt**. Indeed, the models of sections III-A and III-C consider the meaningful part of the signal to be a deterministic component. These models are purposely adapted to signals showing local line spectra. The physiological signals we are interested in, such as respiratory or cardiac signals, have this characteristic (see section IV-D). Signals with wider local spectra, such as electroencephalography signals, do not fall into this category, and are more faithfully modeled as random signals. Thus, the theoretical justifications proposed above are no longer applicable to guarantee the forecasting quality of **SigExt**. A more robust approach to the study of signals following this model is envisaged in our future work.

- 3) *The reference to Takens is slightly confusing could this be explained more clearly. You are not referring to nonlinear dynamic signals so I am not sure it is relevant.*

Hau-Tieng?

- 4) *The paper discusses asymptotic behaviour of the forecasting algorithm but the time series lengths in practice are quite short so what impact does that have on performance.*

In practice, the asymptotic behavior of **SigExt**, illustrated on a synthetic signal in the paper, can be generalized to the study of real-life signals, as long as the duration of the recorded signal largely exceeds that of the desired extension, which allows the user to set $K \gg L$. In other words, to ensure the relevance of the extension produced by **SigExt**, the user should be able to use the previously recorded signal for at least three times the duration of the desired extension, approximately. For example, in section IV-D2, the analysis of the PPG signal requires 7-second-long extensions to effectively reduce boundary effects in the TF domain. About twenty seconds of recording are thus necessary to ensure the fidelity of our algorithm.