USING LINEAR PREDICTION TO MITIGATE END EFFECTS IN EMPIRICAL MODE DECOMPOSITION

Steven Sandoval, Matthew Bredin, and Phillip L. De Leon

New Mexico State University
Klipsch School of Electrical and Computer Engineering
Las Cruces, New Mexico, U.S.A.
{spsandov,mbredin,pdeleon}@nmsu.edu

ABSTRACT

It is well known that empirical mode decomposition can suffer from computational instabilities at the signal boundaries. These "end effects" cause two problems: 1) sifting termination issues, i.e. convergence and 2) estimation error, i.e. accuracy. In this paper, we propose to use linear prediction in conjunction with a previous method to address end effects, to further mitigate these problems. We compare the proposed mitigation to the existing method and provide simulations which demonstrate that the new approach improves intrinsic mode function estimation accuracy while significantly improving convergence.

***Index Terms*—** Signal analysis, Empirical mode decomposition

1. INTRODUCTION

In [1], Huang proposed the original empirical mode decomposition (EMD) and sifting algorithms to sequentially determine a set of intrinsic mode functions (IMFs). Since the original publication, many improvements to EMD have been proposed to address various computational issues and other issues related to the signal decomposition. For example, the ensemble EMD (EEMD) [2] introduced ensemble averaging in order to address the mode mixing problem via an additive noise and an averaging of IMF estimates. The complete EEMD (CEEMD) was proposed to address some of the undesirable features of EEMD by averaging at the IMF level as each IMF is estimated rather than averaging at the conclusion of EEMD [3]. The improved CEEMD (ICEEMD) [4] was proposed to reduce the noise present in each IMF estimate and to reduce the occurrence of spurious IMFs as was observed with CEEMD. More recently, we proposed [5] additional improvements to CEEMD which include 1) a modification to the ensemble averaging which guarantees that the average IMF is a true IMF [2] and 2) a change from the additive noise used in ensemble averaging to a complimentary pair of narrowband tones which we term "tone masking."

In addition to the aforementioned improvements to EMD, improvements to the sifting algorithm have also been proposed. In particular, it is well known that the sifting algorithm may suffer from computational instabilities at the signal boundaries. These "end effects" cause two problems: 1) sifting termination issues, i.e. convergence and 2) estimation error, i.e. accuracy. One previous method to mitigate these problems was proposed by Rato [6,7]. In that method, extrema were extrapolated in order to improve performance of the extrapolation in sifting. In this paper, we consider the use of linear prediction (LP) in conjunction with Rato's method in order to mitigate the problem of end effects.

The remainder of this paper is organized as follows. In Section 2, we briefly review EMD and sifting algorithms. We also motivate the problems related to end effects and review Rato's method for mitigation. In Section 3, we describe our proposed approach which uses LP in conjunction with the previous method. In Section 4, we provide simulations and results which illustrate the efficacy of the proposed method. Finally, in Section 5 we provide conclusions.

2. EMPIRICAL MODE DECOMPOSITION

The original EMD and sifting algorithms were proposed by Huang [1]. The EMD algorithm acts as a "wrapper" and repeatedly calls the sifting algorithm, given in Algorithm 1. The purpose of the sifting algorithm is to iteratively identify and remove the trend from the real-valued signal, acting as an adaptive high pass filter. This process repeats to remove additional IMFs from the signal if they exist. The resulting decomposition is complete and sparse [1, 8, 9].

The sifting algorithm may be viewed as an iterative way of removing the asymmetry between the upper and lower envelopes in order to transform the input $r(t)$ into an IMF [7]. By doing so, low frequency content is discarded at every sifting iteration, effectively making the sifting algorithm behave as a high frequency filter or high frequency component tracker.

Algorithm 1 Sifting Algorithm

procedure $\varphi(t) = \mathrm{SIFT}(\ r(t)\ )$
    while $\int |e(t)|^2 \,\mathrm{d}t \neq 0$ **do**
        find all local maxima: $u_p = r(t_p),\ p = 1, 2, \ldots$
        find all local minima: $l_q = r(t_q),\ q = 1, 2, \ldots$
        interpolate: $u(t) =$ CublicSpline($\{t_p, u_p\}$)
        interpolate: $l(t) =$ CublicSpline($\{t_q, l_q\}$)
        $e(t) = [u(t) + l(t)]/2.$
        $r(t) \leftarrow r(t) - e(t).$
    end while
    $\varphi(t) = r(t)$
end procedure

2.1. End Effects

The sifting algorithm relies on the cubic spline interpolator to determine the envelopes from the extrema. At the signal boundaries, an interpolation is not possible because the extrema do not extend up to or beyond the signal boundaries. Thus the interpolation practically speaking, becomes extrapolation. This causes erratic behavior, termed end effects, in the envelope estimates near the boundaries and consequently errors in IMF estimation and convergence issues.

In order to understand the end effects, we refer the reader to Fig. 1(a) which illustrates a simple example. The figure shows the

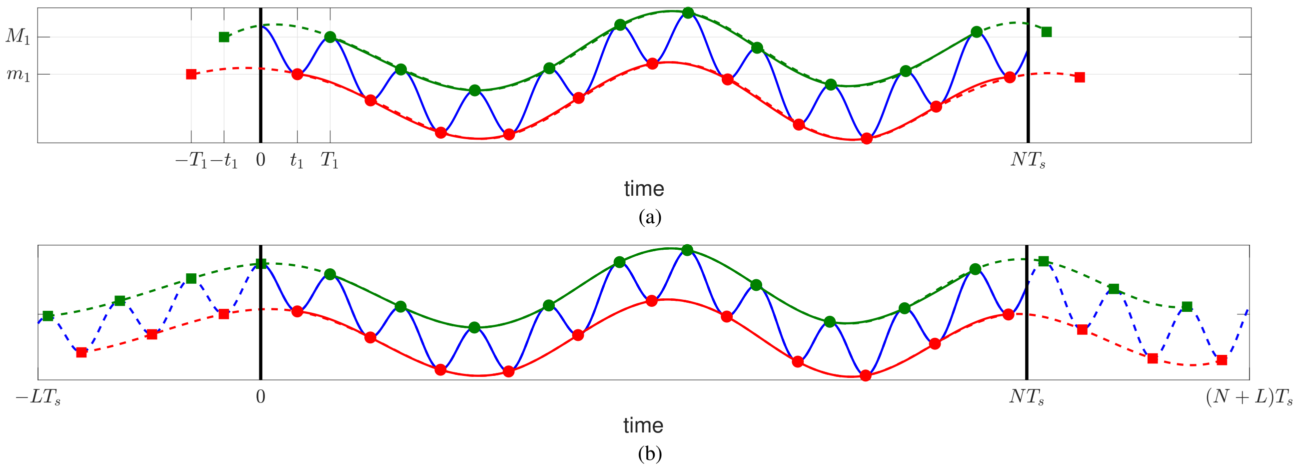

Fig. 1. For a signal segment, the residue signal $r(t)$ (—), maxima $\{t_p, u_p\}$ (●) and minima $\{t_q, l_q\}$ (●), upper $u(t)$ (—) and lower $l(t)$ (—) envelopes, and signal boundaries (❙) at $t = 0$ (left) and $t = NT_s$ (right). Beyond the signal boundaries, we also illustrate (a) artificially-inserted maxima (■) and minima (■) obtained using Rato's mitigation method, as well as the subsequently estimated upper envelope $\breve{u}(t)$ (- - -) and lower envelope $\breve{l}(t)$ (- - -), and (b) the extension of the residue signal (- - -) using LP, as well as the maxima $\{\tilde{t}_p, \tilde{u}_p\}$ (■), minima $\{\tilde{t}_q, \tilde{l}_q\}$ (■), upper envelope $\tilde{u}(t)$ (- - -), and lower envelope $\tilde{l}(t)$ (- - -) that are obtained from the extended residue $\tilde{r}(t)$.

residue signal $r(t)$ (—), maxima $\{t_p, u_p\}$ (●) and minima $\{t_q, l_q\}$ (●) extrema points[1], upper $u(t)$ (—) and lower $l(t)$ (—) envelopes, and signal boundaries (❙) at $t = 0$ (left) and $t = NT_s$ (right). The envelopes do not extend to the signal boundaries because by definition, an interpolator estimates a value within the domain of a set of observations. Thus, the envelopes cannot be *interpolated* outside the outermost extrema points. In practice however, the envelopes are often *extrapolated* from the outermost extrema points to the signal boundaries using the cubic spline interpolator. We term this "solution" as *no mitigation*. However, this is known to lead to sporadic behavior such as divergence toward $\pm\infty$, i.e. end effects because a good interpolator is not necessarily a good extrapolator.

2.2. Prior Methods for End Effect Mitigation

The goal in mitigating the end effects is to accurately estimate the upper and lower envelopes right up to the signal boundaries in a robust way. Several strategies for end effect mitigation have been proposed. For example, [7, 10, 11] considered symmetrical extrema extensions and [12] considered using both support vector machines to predict one maximum and one minimum on each side of the signal and also considered a mirroring procedure. As Rato [7] noted "several attempts [of mitigating the end effects] were made such as repetition or reflection of the signal. The results were not encouraging, because we were really extrapolating the signal. Instead we decided to extrapolate the maxima and the minima." Thus Rato proposed [7] to *insert artificial extrema* beyond the signal boundaries in order to constrain the behavior of the interpolator in the sifting algorithm so that the cubic spline may be used as an interpolator up to the signal boundaries. Referring to Fig. 1(a) at the left boundary, suppose the first maxima is at (T_1, M_1) and the first minima is at (t_1, m_1). We then extrapolate an artificial maxima at $(-t_1, M_1)$ (■) and an artificial minima at $(-T_1, m_1)$ (■). An analogous procedure may be performed at the right boundary. By including the artificial extrema in the sets used for interpolation, the upper envelope $\breve{u}(t)$ (- - -) and lower envelope $\breve{l}(t)$ (- - -) can now be "interpolated" up to the signal boundaries. Close inspection of Fig. 1(a) shows a slight deviation between the envelopes estimated with the artificial extrema [$\breve{u}(t)$ and $\breve{l}(t)$] and without the artificial extrema [$u(t)$ and $l(t)$]. Moreover, even within the outermost extrema the envelopes estimated using cubic spline interpolation, are affected by artificially-added extrema. This deviation is a result of changing boundary conditions imposed by the cubic spline interpolator.

[1] On the boundaries, the signal may be viewed as an extrema or not, because the value of the signal is unknown beyond the boundary. In our implementations, we choose the latter view.

3. LINEAR PREDICTION TO MITIGATE END EFFECTS

Another approach to mitigating the end effects is to *artificially extend the residue signal* $r(t)$ beyond the boundaries (as an initialization step in sifting) then trim the IMF estimate at the boundaries upon termination of the sifting algorithm. We propose to use a LP model, popular for signal modelling in speech analysis and other fields, to extend the residue signal [13]. The auto-regressive (AR) model used in LP is then utilized as both a forward and backward predictor in order to extend the residue beyond the signal boundaries resulting in an extended residue signal

$$\tilde{r}(nT_s) = \begin{cases} \sum_{p=1}^{P} a^*_{P-p} r(nT_s + pT_s), -L \leq n < 0 \\ r(nT_s), 0 \leq n \leq N \\ \sum_{p=1}^{P} a_p r(nT_s - pT_s), N < n \leq N + L \end{cases} \quad (1)$$

where P is the order of the AR model, a_p are the forward linear prediction coefficients (LPCs) obtained using Burg's algorithm, and L is the length of the extension in samples for each side. We note that P and L are design parameters and must be appropriately chosen.

As an illustration of using LP, consider Fig. 1(b) which shows the residue signal $r(t)$ (—) as well as the maxima $\{t_p, u_p\}$ (●) and minima $\{t_q, l_q\}$, upper envelope $u(t)$ (—), and lower envelope $l(t)$ (—) that may be obtained from the residue signal $r(t)$. Also shown are the extension of the residue signal (- - -) as well as the maxima $\{\tilde{t}_p, \tilde{u}_p\}$ (■), minima $\{\tilde{t}_q, \tilde{l}_q\}$ (■), upper envelope $\tilde{u}(t)$ (- - -), and lower envelope $\tilde{l}(t)$ (---) that are obtained from the extended residue $\tilde{r}(t)$. Again, close inspection of Fig. 1(b) shows a slight deviation between the envelope estimates even within the signal due to the boundary conditions in the cubic spline interpolator.

Mitigation using LP requires an appropriate choice for the extrapolation length parameter L. If the parameter L is chosen to be too small, the extended resuide may not display any additional extrema. *We note that Rato's method and the above LP are complimentary and can be used in conjunction.* Thus, as precaution for the event that L is chosen too small, we propose the use of LP in conjuction with Rato's method as follows. Prior to the while loop in Algorithm 1, compute the LPCs a_p using $r(t)$, compute the left and right residue extensions using the forward and backward predictors as in (1), and replace $r(t)$ with $\tilde{r}(t)$. After obtaining extrema from $\tilde{r}(t)$, insert artificial extrema *at the extended boundaries* of $\tilde{r}(t)$ using Rato's mitigation and compute the cubic spline between the outermost extrema, i.e. $-LT_s \leq t \leq (N+L)T_s$. Finally, prior to termination of the sifting algorithm the IMF estimate is trimmed off outside the interval $0 \leq t \leq NT_s$. The pseudocode for the sifting algorithm with the proposed end effect mitigation is given in Algorithm 2 where we also replace the original termination condition[2] with the average power over the interval $0 \leq t \leq NT_s$ below a threshold ε.

Algorithm 2 Sifting Algorithm with Proposed Mitigation

```
procedure φ(t) = SIFT( r(t), L, P )
    compute r̃(t) via Equation (1)
    while (1/(NT_s)) ∫_0^{NT_s} |ẽ(t)|^2 dt ≥ ε do
        find all local maxima: ũ_p = r̃(t̃_p), p = 1, 2, ...
        find all local minima: l̃_q = r̃(t̃_q), q = 1, 2, ...
        insert artificial extrema (per Rato)
        interpolate: ũ(t) = CublicSpline({t̃_p, ũ_p})
        interpolate: l̃(t) = CublicSpline({t̃_q, l̃_q})
        ẽ(t) = [ũ(t) + l̃(t)]/2
        r̃(t) ← r̃(t) − ẽ(t).
    end while
    φ(t) = r̃(t), 0 ≤ t ≤ NT_s
end procedure
```

4. SIMULATIONS AND RESULTS

As described earlier, the presence of end effects results in two problems: 1) IMF estimation error, i.e. accuracy and 2) sifting termination issues, i.e. convergence. We illustrate the efficacy of the proposed method for end effect mitigation with two simulations.

4.1. Simulation 1: Accuracy

In the first simulation, following [16] we constructed a signal consisting of two components $x(t) = s_0(t) + s_1(t)$ with $s_0(t) = \cos(2\pi t + \phi_0)$ and $s_1(t) = a\cos(2\pi f t + \phi_1)$ where $\phi_0, \phi_1 \in \mathcal{U}(-\pi, \pi]$ and a and f are the amplitude and frequency ratios. Again following [16], we define an error surface as a function of log amplitude ratio $-2 \leq \log_{10}(a) \leq 2$ and of frequency ratio $0 \leq f \leq 1$ as

$$J(a,f) = \begin{cases} \dfrac{|\varphi_0(t) - s_0(t)|}{|s_1(t)|}, & af^2 \leq 1 \\ \dfrac{|\varphi_0(t) - x(t)|}{|s_0(t)|}, & af^2 > 1 \end{cases} \tag{2}$$

[2]Various stopping criteria have been reported in the literature [7, 14, 15].

with a minimum value of 0 and a maximum value of 1, and the mean error as

$$\bar{J} = \underset{a,f}{\text{mean}}\,[J(a,f)] \tag{3}$$

where $\underset{a,f}{\text{mean}}[\cdot]$ denotes the average over a and f.

In the first simulation, we use sampling frequency $f_s = 500$, a signal duration of 10 seconds, extrapolation length $L = 2f_s$, order $P = 200$, number of trials $I = 100$. We report performance in terms of IMF estimation accuracy by defining the expected mean error surface $\mathbb{E}[J(a,f)]$ and the expected mean error $\mathbb{E}[\bar{J}]$ across the I trials.

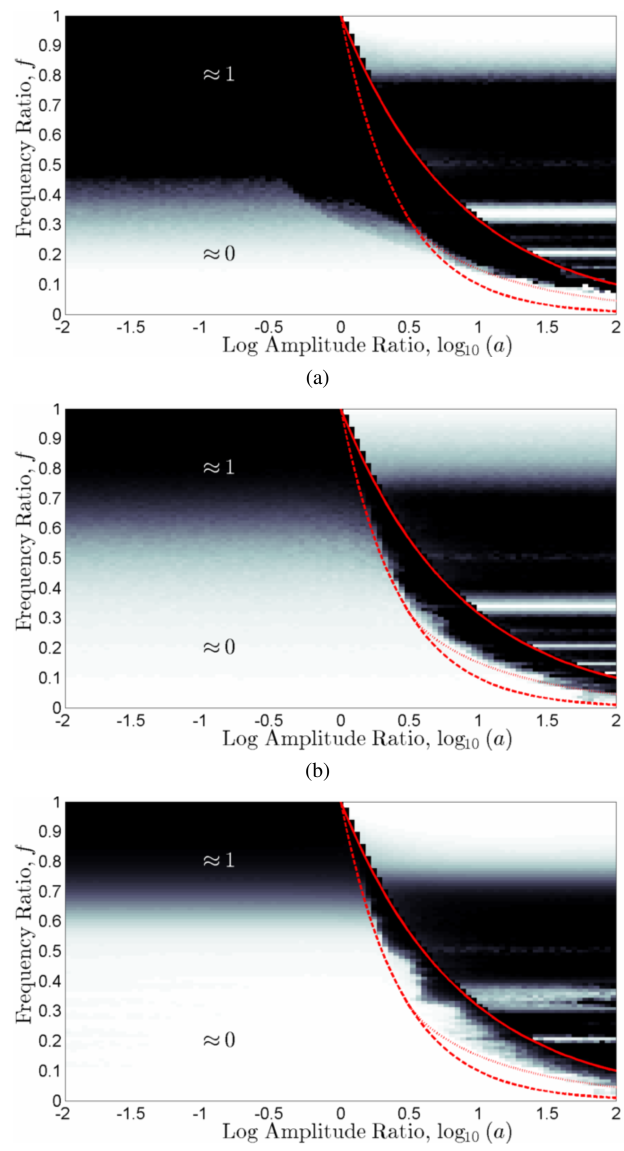

Fig. 2. The expected error surface $\mathbb{E}[J(a,f)]$ when using (a) no mitigation, (b) Rato's mitigation, and (c) the proposed mitigation (LP+Rato's). Additionally, the theoretical bifurcation curves $af^2 = 1$ (—), $af = 1$ (---), and $af\sin(3\pi f/2) = 1$ (·····) derived in [16] are overlaid.

Fig. 2(a) shows the expected error surface $\mathbb{E}[J(a, f)]$ with no mitigation of end effects. We note that the figure does not exactly match Figure 3 in the published reference [16] because the authors removed the end effects from the analysis, i.e. the first and last quarters of the signal were removed prior to error computation. Fig. 2(b) shows the expected mean error surface $\mathbb{E}[J(a, f)]$ when using Rato's method to mitigate the end effects which shifts the transition occurring at approximately $f = 0.37$ to $f = 0.67$ (back in agreement with [16]) and removes the area of high error near $\log_{10}(a) = 0$ and $f = 0.3$. We also note that the transition from low to high error aligns closer to the theoretical curves $af = 1$ (---) when $f > 1/3$ and $af \sin(3\pi f/2) = 1$ (····) when $f < 1/3$. Fig. 2(c) shows a smoother expected mean error surface $\mathbb{E}[J(a, f)]$ when using the proposed method (e.g. near $f = 0.67$) which implies a more consistent performance. Additionally, the proposed method lowers the error between the theoretical bounds so that the transition from low to high error aligns closer to the theoretical curve $af^2 = 1$ (—). Finally, Table 1 gives the expected mean errors $\mathbb{E}[\bar{J}]$ for the various methods including no mitigation as a function of the number of sifting iterations. The proposed method has a lower expected mean error regardless of the number of sifting iterations.

	Number of Sifting Iterations		
	1	10	100
No Mitigation	0.64	0.62	0.62
Rato's Method	0.52	0.49	0.51
Proposed Method	**0.46**	**0.40**	**0.37**

Table 1. The expected mean error $\mathbb{E}[\bar{J}]$ for Simulation 2 over 100 trials for various number of sifting iterations. The proposed method has the lowest expected mean error regardless of the number of sifting iterations. Rato's method shows evidence of instability by increasing error as the sifting iterations increaces from 10 to 100.

4.2. Simulation 2: Convergence

In the second simulation, we construct a signal consisting of two components $x(t) = s_0(t) + s_1(t)$. The first component $s_0(t)$ is a Gaussian AM, sinusoidal FM signal and the second component $s_1(t)$ is a sinusoidal AM signal. Mathematically these components are both described as

$$s_k(t) = \mathrm{Re}\left\{a_k(t)\mathrm{e}^{\mathrm{j}\left[2\pi f_k + \int_{-\infty}^{t} m_k(\tau)\,\mathrm{d}\tau + \phi_k\right]}\right\} \quad (4)$$

with first component parameters $a_0(t) = 0.25 + \exp\left[-\frac{1}{2}\left(\frac{t-0.25}{4000}\right)^2\right]$, $f_0 = 40$, $m_0(t) = 25\sin(4\pi t + \phi_m)$, and ϕ_0, ϕ_m drawn from $\mathcal{U}(-\pi, \pi]$ and second component parameters $a_1(t) = \frac{1}{5}\sin(20t) + 0.7$, $f_1 = 250$, $m_1(t) = 0$, and ϕ_1 drawn from $\mathcal{U}(-\pi, \pi]$.

Our convergence metric is the average power in the mean of the envelopes over the interval $0 < t < NT_s$

$$C = \frac{1}{NT_s}\int_0^{NT_s} |\tilde{e}(t)|^2 dt \quad (5)$$

in decibels (dB) as a function of the number of sifting iterations. Fig. 3 summarizes the results of 5000 trials with a sampling rate $f_s = \frac{1}{T_s} = 16000$ where a one half second signal segment was considered. For the proposed method, we choose extrapolation length $L = 200$ and order $P = 200$. First, the — shows the trial with the worst convergence and demonstrates a case where instability occurred. We note that with no mitigation, although the instability may be severe, additional iterations allow for convergence whereas with only Rato's method, additional iterations may not be beneficial. Finally, with the proposed method, the instability is small and additional iterations allow for convergence. Second, the ▬ shows the range of the convergence metric across the trials as a function of the number of iterations. We note that with no mitigation the range is about 100 dB, whereas with only Rato's method it is about 75 dB, and with the proposed method it is only about 50 dB. A smaller range of C implies more consistent convergence performance. Third, the ▬ shows the mean dB value of C across the trials as a function of the number of iterations. We note that the no mitigation case does only slightly worse then the proposed method, while Rato's method on average performs worse than the no mitigation case, in terms of convergence. Interpretation of the results from Experiment 1 implies, that use of Rato's method for mitigation on average takes more computation time (requires more iterations) than either no mitigation or the proposed mitigation, when the termination condition is based on average power of the mean of the envelope estimates.

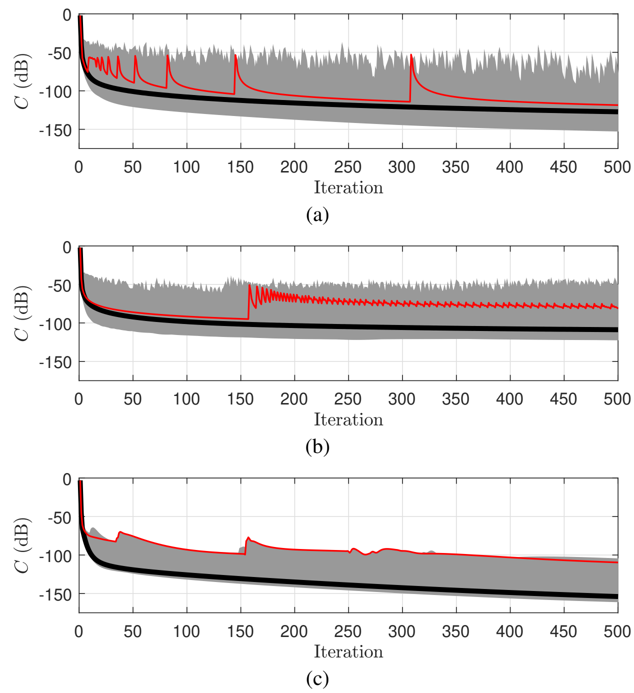

Fig. 3. For 5000 trials of the convergence metric C in dB as a function of iteration, the mean value (▬) and the range (▬) for (a) no mitigation, (b) Rato's mitigation only, (c) the proposed mitigation (LP+Rato's). The trial with the wost convergence (—) is also shown, note convergence instability.

5. CONCLUSION

In this paper, we proposed to use linear prediction in conjunction with a previously proposed extrema extrapolation to mitigate end effect problems in EMD. Using two simulations, we illustrate the efficacy of the proposed method. In the first simulation, we show that the expected mean error is reduced and in the second, both the convergence mean and worst case trial are smaller and convergence is more robust. The proposed method is anticipated to have the most impact in cases where the area of interest within the signal extends up to the signal boundary.

6. REFERENCES

[1] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London Ser. A*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[2] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 01, pp. 1–41, 2009.

[3] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2011, pp. 4144–4147.

[4] G. Colominas, Marcelo A. and Schlotthauera and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal. Process. Control*, vol. 14, pp. 19–29, Nov. 2014.

[5] S. Sandoval and P. L. De Leon, "Advances in empirical mode decomposition for computing instantaneous amplitudes and instantaneous frequencies," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2017, pp. 4311–4315.

[6] R. Rato and M. Ortigueira, "A modified emd algorithm for application in biomedical signal processing," in *Int. Conf. Comp. Intelligence Med. Healthcare, (CIMED)*, Jul. 2005.

[7] R. Rato, M. Ortigueira, and A. Batista, "On the HHT, its problems and some solutions," *Mechanical Syst. Signal Process.*, vol. 22, no. 6, pp. 1374–1394, 2008.

[8] T. Y. Hou and Z. Shi, "Adaptive data analysis via sparse time-frequency representation," *Adv. Adapt. Data Anal.*, vol. 3, no. 01n02, pp. 1–28, 2011.

[9] ——, "Data-driven time–frequency analysis," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 2, pp. 284–308, 2013.

[10] S. Zhongping and Y. Zhichun, "A better method for effectively suppressing end effect of empirical mode decomposition (EMD)," *J. North Polytechnical University*, vol. 24, no. 5, p. 639, 2006.

[11] F. Wu and L. Qu, "An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery," *J. Sound and Vibration*, vol. 314, no. 3-5, pp. 586–602, 2008.

[12] Y. Su, Z. Liu, K. Li, and B. Huo, "A new method for end effect of EMD and its application to harmonic analysis," *Advanced Technology of Electrical Engineering and Energy*, vol. 27, no. 2, p. 33, 2008.

[13] S. S. Haykin, *Adaptive Filter Theory*. Pearson, 2008.

[14] N. E. Huang, M.-L. C. Wu, S. R. Long, S. S. Shen, W. Qu, P. Gloersen, and K. L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *Proc. R. Soc. London Ser. A*, vol. 459, no. 2037, pp. 2317–2345, 2003.

[15] G. Rilling, P. Flandrin, and P. Goncalves, "On empirical mode decomposition and its algorithms," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Process. NSIP*, vol. 3, 2003, pp. 8–11.

[16] G. Rilling and P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, 2008.