

ON THE INFLUENCE OF SAMPLING ON THE EMPIRICAL MODE DECOMPOSITION

Gabriel Rilling^(*) and Patrick Flandrin^(*)

(*) Laboratoire de Physique (UMR CNRS 5672), École Normale Supérieure de Lyon
46, allée d'Italie 69364 Lyon Cedex 07, France
gabriel.rilling@ens-lyon.fr, flandrin@ens-lyon.fr

ABSTRACT

The rationale underlying the nonlinear Empirical Mode Decomposition method is intrinsically a continuous-time approach. The method can however only be applied in practice to discrete-time signals. EMD is obtained through iterating a basic nonlinear operator for which we derive an upper bound for the effects of sampling. Finally the effects of sampling for a complete EMD are assessed using the knowledge on the basic operator.

1. INTRODUCTION

Empirical Mode Decomposition (EMD) has been recently pioneered by Huang *et al.* [1, 2] for adaptively decomposing nonstationary and/or nonlinear time series. Its unique properties already awarded the technique a certain popularity in various domains, especially in complex systems analysis. However the method still lacks a theoretical basis and in particular there is no precise guarantees on the confidence that is to be granted to EMD. Extending previous work in the case of sinusoidal signals [3, 4], the purpose of this paper is to investigate a possible confidence level for EMD with respect to sampling.

2. EMD: A CONTINUOUS PRINCIPLE, A DISCRETE IMPLEMENTATION

2.1. EMD algorithm

Basically, Empirical Mode Decomposition (EMD) [1] considers a signal at the scale of its local oscillations. The main idea of EMD is then to formalize the idea that, locally: “signal = fast oscillations superimposed to slow oscillations”. Looking at one local oscillation (defined, e.g., as the signal between two consecutive local minima), EMD is designed to define a local “low frequency” component as the *local trend* $m_1[x](t)$ that supports a local “high frequency” component as a zero-mean oscillation or *local detail* $d_1[x](t)$ so that we can express $x(t)$ as:

$$x(t) = m_1[x](t) + d_1[x](t). \quad (1)$$

By construction, $d_1[x](t)$ is an oscillatory signal and, if it is furthermore required to be locally zero-mean everywhere, it corresponds to what is referred to as an *Intrinsic Mode Function* (IMF) [1]. Practically, this mostly implies that all its maxima are positive and all its minima negative. On the other hand, all we know about $m_1[x](t)$ is that it locally oscillates slower than $d_1[x](t)$. We can then apply the same decomposition to it, leading to: $m_1[x](t) =$

$m_2[x](t) + d_2[x](t)$ and, recursively applying this on the $m_k[x](t)$, we get a representation of $x(t)$ of the form:

$$x(t) = m_K[x](t) + \sum_{k=1}^K d_k[x](t). \quad (2)$$

At each step of the decomposition, we can state that $d_{k+1}[x](t)$ contains approximatively as many oscillations as $m_k[x](t)$. Correspondingly, $m_{k+1}[x](t)$ (and so $d_{k+2}[x](t)$) contains less, and locally slower, oscillations than $m_k[x](t)$ (and so $d_{k+1}[x](t)$). The decomposition usually ends when the signal $m_K[x](t)$ does not contain enough oscillations to define a meaningful local trend. EMD performs thus a multi-scale decomposition that is fully data-driven and that can be applied to all oscillatory time series, including nonstationary ones and/or those generated by a nonlinear system.

Practically the discrimination between “fast” and “slow” oscillations is obtained through an algorithm referred to as the *sifting process* [1] which iterates a nonlinear elementary operator \mathcal{S} on the signal (and then on the $m_k[x](t)$) until some stopping criterion is met. Given a signal $x(t)$, the operator \mathcal{S} is defined by the following procedure:

1. identify all extrema of $x(t)$
2. interpolate (cubic spline) between minima (resp. maxima), ending up with some “envelope” $e_{min}(t)$ (resp. $e_{max}(t)$)
3. compute the mean $m(t) = \frac{(e_{min}(t) + e_{max}(t))}{2}$
4. subtract to the signal to obtain $\mathcal{S}[x](t) = x(t) - m(t)$

If the convergence criterion is met after n iterations, the *local detail* and the *local trend* are defined as: $d_1[x](t) = \mathcal{S}^n[x](t)$ and $m_1[x](t) = x(t) - d_1[x](t)$.

2.2. The sampling issue

In the previous description of the EMD algorithm, the signals have implicitly been considered as continuous-time ones whereas in practice signals are usually known in discrete-time only. When dealing with linear operators, applying a continuous-time operator on a signal and applying its discrete-time version on the discretized signal are two ways to perform the same task as there is a direct link between the two outputs. As the sifting operator \mathcal{S} is nonlinear, there is no obvious link between the theoretic continuous-time EMD and its practical discrete-time version. Therefore it is generally recommended to apply EMD to signals with a large amount of oversampling which allows to consider with a limited prejudice that it behaves like a continuous-time operator. Given a continuous-time signal, our purpose here is to define an upper

bound for the difference between a continuous-time EMD and its discrete time version as a function of the sampling frequency.

To address properly this sampling issue we first have to define some minimum requirements on the sampling parameters for a continuous-time signal to be processed by EMD. As extrema play a major role in the sifting operator, a natural requirement would be for the discrete-time signal to keep as many extrema as in the continuous-time version. In fact, losing an extremum during the sampling process usually means losing a pair of maximum/minimum, which means to EMD losing a local oscillation. To ensure there is no loss of extrema the minimum requirement is for the sampling period to be at most one half of the minimum distance between extrema in the signal. In all the following we will implicitly consider that this requirement is met. It is worth noticing the fact that this requirement is completely independent from Shannon's criterion for band-limited signals. Indeed a band-limited signal can have arbitrary close extrema whatever the band.

3. A BOUND OF THE EFFECT OF SAMPLING ON THE EMD ELEMENTARY OPERATOR

Considering a continuous-time signal $x(t)$ with a minimum distance Δ between its extrema, let us define its discretized versions $x_{f_s, \varphi}[n] = x(n/f_s + \varphi)$. Given these our goal is to characterize the behaviour of the EMD of $x_{f_s, \varphi}$ as a function of the sampling frequency $f_s > 2/\Delta$. More particularly we mainly interest ourselves into the deviation between the sampled IMFs and the theoretic continuous ones. As EMD is obtained by iterating the elementary operator \mathcal{S} , it seems natural to start off with studying the effects of sampling on this elementary operator. In this section we investigate the case of a simplified elementary operator for which we can derive an upper bound for the sampling effects. The obtained results are then assessed by simulations using the original operator.

3.1. Model

We start off with analyzing the effects of sampling on the extrema. If t_0 is the position of a local maximum in $x(t)$, the condition $f_s > 2/\Delta$ ensures that there is also a maximum in $x_{f_s, \varphi}[n]$ for n such that $|n/f_s + \varphi - t_0| < 1/f_s$, i.e. either the closest sampling point on the right or on the left of the continuous-time maximum is a maximum for the sampled signal. Without any additional information concerning the analyzed signal, we can only state that the discrete maximum must belong to the curve $(t, x(t))$ for $t \in [t_0 - 1/f_s, t_0 + 1/f_s]$. If we separate uncertainties in abscissa and ordinate, this results in an $\Delta_{abs.} = 1/f_s$ uncertainty in abscissa and $\Delta_{ord.} = \max\{x(t_0) - x(t_0 - 1/f_s), x(t_0) - x(t_0 + 1/f_s)\}$ in ordinate, the latter depending on the behavior of the signal around the considered maximum.

Taking these uncertainties into account, the next step in our analysis is to evaluate the impact of the latters on the envelopes e_{min} and e_{max} . Usually these envelopes are computed using cubic spline interpolation [1, 3]. In our model however, we will use *piecewise linear interpolation*. This is the only simplification we consider for the elementary operator \mathcal{S} . The reason for this is simply that the value of a cubic spline interpolation between two knots depends not only on a few knots around but also on all the knots defining the interpolation. In this context, calculating an uncertainty on the value of the interpolation at a specific position given uncertainties on all the knots is rather complicated whereas it is

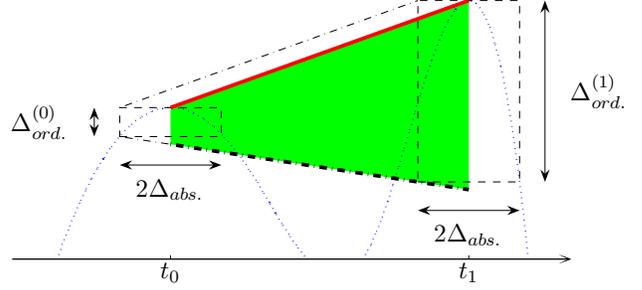


Fig. 1. Evaluating the uncertainty for the envelopes. For both maxima, the dashed boxes delimit the areas where the corresponding maxima in the sampled signal can be located. The thick plain line stands for the interpolation (piecewise linear in the model) based on the extrema in the continuous signal. Then the dash-dot lines delimit the area where the interpolation based on the extrema in the sampled signal can be located. Finally, the thick dash-dot line stands for the case that leads to the largest error.

much simpler with piecewise linear interpolation while guaranteeing that the obtained uncertainties generally match the observations.

If we consider two maxima $(t_0, x(t_0))$ and $(t_1, x(t_1))$ in the continuous-time signal, the corresponding maxima in the sampled signal are located in the boxes $[t_i - \Delta_{abs.}, t_i + \Delta_{abs.}] \times [x(t_i) - \Delta_{ord.}^{(i)}, x(t_i) + \Delta_{ord.}^{(i)}]$ ($i = 0, 1$) (see figure 1). In this context it is clear that the largest error for the envelope is obtained for the thick dash-dot line case in figure 1. Integrating this error over the range $[t_0, t_1]$ (corresponding to the shaded surface in figure 1) we get a bound of the sampling error for the upper envelope $e_{max}(t)$ (referred to as $\Delta_{e_{max}}(t)$) over the latter range:

$$\int_{t_0}^{t_1} |\Delta_{e_{max}}(t)| dt \leq \frac{(t_1 - t_0)(\Delta_{ord.}^{(0)} + \Delta_{ord.}^{(1)})}{2} + |x(t_1) - x(t_0) - \Delta_{ord.}^{(1)} + \Delta_{ord.}^{(0)}| \Delta_{abs.}, \quad (3)$$

which leads in turn to the looser bound:

$$\int_{t_0}^{t_1} |\Delta_{e_{max}}(t)| dt \leq \frac{(t_1 - t_0)(\Delta_{ord.}^{(0)} + \Delta_{ord.}^{(1)})}{2} + |x(t_1) - x(t_0)| \Delta_{abs.} + |\Delta_{ord.}^{(0)} - \Delta_{ord.}^{(1)}| \Delta_{abs.}, \quad (4)$$

In the last formula the first term can be read as the average uncertainty caused by the uncertainties in ordinate only. Its amplitude depends only on the behaviour of the signal around its extrema. On the contrary, the second term depending on the uncertainties in abscissa does not only depend on the behaviour of the signal around its extrema but on the relative amplitudes of successive extrema too. If the analyzed signal is e.g., an amplitude modulated sinusoidal signal, then the amplitude of the first term will only depend on the sampling frequency and the frequency of the AM signal while the second one will depend on the frequency and amplitude of the modulation too. Finally the last term, depending on both uncertainties, is generally much smaller than the two others and can therefore be neglected in most cases.

If we compute the bound (4) on each range delimited by two consecutive maxima, we obtain a bound of the L_1 norm of the sampling related effects on the upper envelope. Repeating the same

operation for the lower envelope then results in a bound of the error on the mean of the envelopes, which is simply half the sum of bounds on each envelope. Finally this bound is also valid for the elementary operator defined as the difference between the signal and the mean of the envelopes.

3.2. Estimating the parameters from the signal

Given a continuous-time signal $x(t)$, computing the upper bound requires only a limited set of parameters: the position $(t_i, x(t_i))$ and the uncertainty $\Delta_{ord}^{(i)}$ for each extremum. These are rather easy to measure from the continuous-time signal. If we only have a discrete version of the signal, some of these parameters can not be evaluated unless we have more knowledge on the signal. Indeed the $(t_i, x(t_i))$ can be estimated by the values corresponding to the same extrema in the discrete signal $\hat{t}_i = n_i/f_s + \varphi$ and $\hat{x}_i = \widehat{x(t_i)} = x_{f_s, \varphi}[n_i]$, where n_i is the index of the i^{th} extremum in the discrete signal. Concerning the $\Delta_{ord}^{(i)}$ however, we need information on the behaviour of the signal around its extrema. Therefore it seems natural to use some information on the regularity of the signal. If we e.g. require the signal to be two times continuously differentiable, then we can use second order Taylor expansions to estimate the behaviour of the signal around its extrema. Thus, we can locally use the approximation $x(t) \approx 0.5(t - t_i)^2 x''(t_i)$ which leads to $\Delta_{ord}^{(i)} \approx 0.5|x''(t_i)|/f_s^2$. Finally the second derivatives $x''(t_i)$ can be estimated using finite difference operators at the extrema of the sampled signal, leading to \hat{x}_i'' . Putting these approximations back into (4) results in an estimate of the maximum sampling error over the range $[t_i, t_{i+1}]$ for the upper envelope:

$$\int_{t_i}^{t_{i+1}} |\Delta_{ema}(t)| dt \leq \frac{(\hat{t}_{i+1} - \hat{t}_i)|\hat{x}_i'' + \hat{x}_{i+1}''|}{4f_s^2} + \frac{|\hat{x}_{i+1} - \hat{x}_i|}{f_s} + \frac{|\hat{x}_{i+1}'' - \hat{x}_i''|}{2f_s^3}, \quad (5)$$

we are thus led to the central result of this paper, i.e. we get an upper bound for the L_1 norm of the sampling error for the elementary operator ΔS of the form:

$$\int |\Delta S(t)| dt \leq \frac{\lambda}{f_s} + \frac{\mu}{f_s^2} + \frac{\nu}{f_s^3}, \quad (6)$$

with

$$\lambda = \frac{1}{2} \left(\sum_i |\hat{x}_{i+1}^m - \hat{x}_i^m| + \sum_i |\hat{x}_{i+1}^M - \hat{x}_i^M| \right), \quad (7)$$

$$\mu = \frac{1}{8} \left(\sum_i (\hat{t}_{i+1}^m - \hat{t}_i^m) |\hat{x}_i^{m''} + \hat{x}_{i+1}^{m''}| + \sum_i (\hat{t}_{i+1}^M - \hat{t}_i^M) |\hat{x}_i^{M''} + \hat{x}_{i+1}^{M''}| \right), \quad (8)$$

$$\nu = \frac{1}{4} \sum_i |\hat{x}_{i+1}^{m''} - \hat{x}_i^{m''}| + |\hat{x}_{i+1}^{M''} - \hat{x}_i^{M''}|, \quad (9)$$

where superscripts m and M respectively refer to minima and maxima.

Remark: Our previous studies showed that in the specific case of sinusoidal signals the sampling error is upper bounded by a function proportional to f_s^{-2} [4]. The bound (6) obtained here generalizes this result as parameters λ and ν are simply zero in this case.

Nevertheless, the μ coefficient computed with (8) for sinusoidal signals leads to a bound that is looser than the one we obtained in [4]. There are two reasons for this. First a sinusoidal signal is symmetric with respect to each extremum. Therefore, the uncertainty in abscissa Δ_{abs} can be reduced to $1/(2f_s)$. This leads to the reduced parameter values: $\lambda' = \lambda/2, \mu' = \mu/4$ and $\nu' = \nu/8$. Next, there generally is a partial compensation between the error associated to the f_s^{-2} terms coming from the upper and the lower envelope. This can be easily controlled for the sinusoidal case, allowing a bound reduced by half which is consistent with our former results.

3.3. Validation

To assess the performance of the upper bound model (6), we performed a set of simulations using synthetic signals. The methodology can be summarized by the following procedure:

1. synthesize a $2N\Delta$ points discrete-time signal $x[n]$ with minimum distance between extrema Δ .
2. compute the downsampled signals: $x_{k,l}[n] \equiv x[kn + l]$ for $1 \leq k \leq \Delta$ and $0 \leq l \leq k - 1$.
3. apply the elementary sifting operator to each downsampled signal, leading to signals $\mathcal{S}_{k,l}[n]$.
4. for each downsampled signal compute the sampling error measure:

$$e(f_s = 1/k, \varphi = l) \equiv \frac{k}{N} \sum_n |\mathcal{S}_{k,l}[n] - \mathcal{S}_{1,0}[kn + l]|. \quad (10)$$

Ideally, the error measure should be obtained by comparing the discrete time elementary operator output $\mathcal{S}_{k,l}[n]$ with a theoretic continuous time one $\mathcal{S}(t)$. This is almost what is done in (10) if we assume that $\mathcal{S}_{1,0}[n]$ is very close to $\mathcal{S}(t)$. Besides we chose the l_1 norm but other norms lead to similar results.

The test signals we used are piecewise cubic polynomials obtained through interpolation of a random set of extrema. The main property of the underlying model is that it ensures that the minimum distance between extrema is set by a parameter Δ . The details are of lesser importance as the results seem to be weakly dependent on the model.

Simulation results are plotted figure 2 for two representative examples. As it can be seen on the figure, the behaviour of the bound as a function of the sampling frequency can generally be divided in two areas: for lower sampling frequencies, the bound usually behaves like f_s^{-2} whereas for higher frequencies it behaves like f_s^{-1} (cf. fig. 2 (b)). There are specific cases however where the f_s^{-1} area does not exist as coefficient λ (7) is typically zero when all the maxima/minima have the same amplitude (see fig. 2 (a)). This two-areas behaviour also applies to the measured sampling error but either the f_s^{-1} or f_s^{-2} area can be missing. Moreover when both are present the critical sampling frequency delimiting these areas is generally not the same as the corresponding critical frequency for the bound. The evolution of the bound however usually gives a good outline for the evolution of the sampling error. Besides, the f_s^{-3} behaviour that we could expect from (6) has never been observed distinctly. Quantitatively, the sampling error measured in our simulations is always rather far from the bound. The latter usually is around 12dB above in the f_s^{-1} area and 25dB in the f_s^{-2} area. This discrepancy comes mainly from the fact that the bound is obtained by considering a worst case error everywhere whereas this worst case is rather improbable. Another reason is

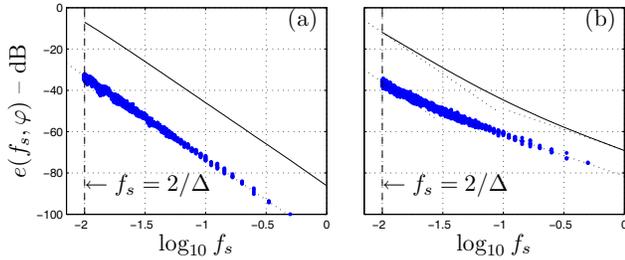


Fig. 2. Sampling error and bound estimate as a function of the sampling frequency. On both graphs, each dot stands for a measure of sampling error according to (10). For each sampling frequency and phase, a bound is also estimated according to (6) from the downsampled signal: the mean of these with respect to the phase is plotted against sampling frequency as a full line. Cases (a) and (b) correspond to rather close signals, the difference being that (a) has constant maxima/minima amplitudes while (b) has random maxima/minima values centered around 1/-1 with variance 0.1. Asymptotic f_s^{-1} or f_s^{-2} behaviours are evidenced by dotted lines.

that the sampling error on the upper envelope generally partially compensates the sampling error on the lower envelope while the two corresponding bounds are summed in the model.

4. IMPACT ON THE DECOMPOSITION

In the previous sections we analyzed the effects of sampling on the elementary operator underlying EMD. Here we intend to evaluate how the properties of this operator can explain the effects of sampling on a full decomposition, where the operator is iterated several times to compute each IMF. Therefore, some simulations have been carried out based on the same procedure as for the elementary operator in section 3.3. The only difference with the latter procedure is that the sampling error measure (10) no longer compares outputs of the elementary operator but IMFs. A representative example is proposed figure 3.

Analyzing the simulation results reveals a few interesting features. First, as in the case of the figure, there usually is a set of couples (f_s, φ) for which the sampling error for the IMFs (at least the first few) behaves almost as in the case of the elementary operator. To understand this result we have to remind that an IMF is obtained by iterating the elementary operator until some stopping criterion is met. When varying the sampling parameters the stopping criterion also varies and this results in a fluctuating number of iterations for each IMF. Unfortunately, the precise number of iterations performed to extract an IMF is generally unpredictable. Thus the behaviour close to the one of the elementary operator reported above occurs when the number of iterations for the observed IMF and the previous ones exactly match the corresponding numbers of iterations for the reference signal assumed to be continuous-time. Conversely, different numbers of iterations lead to larger discrepancies (see the “branches” for $\log_{10} f_s \lesssim -1$). Provided the numbers match however, it seems that iterating the elementary operator does not significantly increase the sampling error for the first IMF, as if the error mostly came from the first iteration. Moreover there is no significant increase when considering IMFs of higher indices. The reason for this is that IMFs of higher indices are extracted from the “slow” component of the signal which is therefore

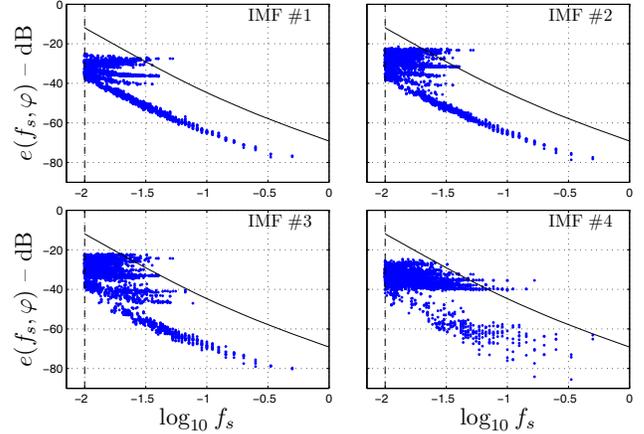


Fig. 3. Sampling error for a complete EMD. Example of sampling error for the 4 first IMF. For each IMF, dots stand for sampling error measures as a function of the sampling frequency for various phases. The bound computed from the downsampled signal according to (6) is also plotted as a full line. The signal used for this example is the same as the one used for figure 2 (b).

less sensitive to sampling. As this component is the difference between the signal and the first IMF however, the sampling error for the latter propagates in IMFs of higher indices.

5. CONCLUSION

Although based on a continuous-time rationale, EMD can only be applied in practice in a discrete-time context because of its non-analytic definition. When the effects of sampling can be well controlled for the basic operator that is iterated to obtain EMD, the impact on the final decomposition is difficult to assess because the number of iterations performed to obtain an IMF depends on the input signal and therefore on sampling. If there may be some ways to improve the robustness of EMD with respect to sampling, basic resampling is not recommended because it often creates low amplitude extrema pairs that might be not relevant for EMD.

6. REFERENCES

- [1] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, “The Empirical Mode Decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. Roy. Soc. London A*, vol. 454, pp. 903–995, 1998.
- [2] N.E. Huang and S.S.P. Shen, Eds., *Hilbert–Huang Transform and Its Applications*, World Scientific, 2005.
- [3] G. Rilling, P. Flandrin, and P. Gonçalvès, “On empirical mode decomposition and its algorithms,” in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03*, Grado (I) 2003.
- [4] G. Rilling and P. Flandrin, “Sur la décomposition modale empirique des signaux échantillonnés,” in *Proc. 20^e Coll. GRETSI sur le Traitement du Signal et des Images*, Louvain-la-Neuve (B) 2005, pp. 755–758.