Blind, Non-stationary Source Separation Using Variational Mode Decomposition with Mode Culling

Seema Sud *seema.sud@aero.org*
Communication Technologies and Engineering Division
The Aerospace Corporation
Chantilly, VA 20151, USA

Abstract

The Fourier Transform (FT) is the single best-known technique for viewing and reconstructing signals. It has many uses in all realms of signal processing, communications, image processing, radar, optics, etc. The premise of the FT is to decompose a signal into its frequency components, where a coefficient is determined to represent the amplitude of each frequency component. It is rarely ever emphasized, however, that this coefficient is a constant. The implication of that fact is that Fourier Analysis (FA) is limited in its accuracy at representing signals that are time-varying, e.g. non-stationary. Another novel technique called empirical mode decomposition (EMD) was introduced in the late 1990s to overcome the limits of FA, but the EMD was shown to have stability issues in reconstructing non-stationary signals in the presence of noise or sampling errors. More recently, a technique called variational mode decomposition (VMD) was introduced that overcomes the limitations of both aforementioned methods. This is a powerful technique that can reconstruct non-stationary signals blindly. It is only limited in the choice of the number of modes, K, in the decomposition. In this paper, we discuss how K may be determined a priori, using several examples. We also present a new approach that applies VMD to the problem of blind source separation (BSS) of two signals, estimating the strong powered signal, termed the interferer, first and then extracting the weaker one, termed the signal-of-interest (SOI). The baseline approach is to use all the predetermined K modes to reconstruct the interferer and then subtract its estimate from the received signal to estimate the SOI. We then devise an approach to choose a subset of the K modes to better estimate the interferer, termed culling, based on a very rough a priori frequency estimate of the weak SOI. We show that the VMD method with culling results in improvement in the mean-square error (MSE) of the estimates over the baseline approach by nearly an order of magnitude.

Keywords: Blind Source Separation (BSS), Culling, Non-stationary, Variational Mode Decomposition (VMD).

1. INTRODUCTION

The problem of blind source separation (BSS) using a single receiver is a challenging one. The problem is compounded by real world artifacts of signals that are non-stationary, pass through non-linear and/or non-stationary channels, and perhaps are distorted by passing through non-linear amplifiers. The non-stationarity is by far a bigger limiting factor than non-linearity [1]. Two novel techniques have been developed to perform signal extraction in non-stationary conditions, far surpassing the performance of conventional methods including Fourier analysis, wavelet processing, principal components analysis (PCA), and singular value decomposition (SVD). The first method, called empirical mode decomposition (EMD), constructs a signal using a series of intrinsic mode decompositions, by isolating frequency components from high to low. This is performed by an averaging process of the envelopes created by local minima and maxima of the signals [2]. The second method, called variational mode decomposition (VMD) is a novel technique that overcomes some of the limitations associated with EMD, such as noise and

Approved for public release. OTR 2020-00547.

sampling rate sensitivities [1]. VMD operates by constructing a signal from a set of modes, where each mode is formed using a Wiener filter constructed around a particular center frequency. Both the modes and frequencies are estimated. The modes themselves are selected to be narrowband about the center frequency.

In [1], the authors demonstrated superior performance of VMD in reconstructing a signal in the presence of noise. However, there has been limited application to studying the algorithm's usage to separate two signals, whereby both a signal-of-interest (SOI) and the interferer are to be estimated. This is the classical problem of BSS. In [3], VMD was applied to separation of a sinusoidal signal from speech, using PCA also, but the application was limited to this single use case. In addition, the issue of selection of the number of modes, K, is yet to be addressed.

In this paper, we describe applying VMD using several unknown, non-stationary SOI and interferer combinations, where we assume without loss of generality that the interferer is stronger than the SOI. Through this, we observe a natural selection of K based on an estimate of the number of narrowband components of signals. We generalize this to selecting K based on the type of interferer present. We then use the selected value of K to estimate the higher powered interferer, and then subtract this estimate from the received signal to extract the weaker powered signal-of-interest (SOI); this is our baseline VMD BSS algorithm. Next, we propose a modification to this VMD baseline approach, wherein certain modes of the decomposition are removed in estimating the interferer, based on an assumption of the SOI's amplitude or frequency, so as to not include spurious modes that belong to the SOI. This is shown to provide an improvement in MSE of the interferer and hence also of the SOI. Significant improvement over the baseline algorithm is seen over a range of carrier-to-interference ratios (CIRs), which is the power ratio between the SOI and interferer, as well as signal-to-noise-ratios (SNRs).

An outline of the paper is as follows: Section 2 first discusses the VMD algorithm, introduced in [1]. Section 3 presents the signal model and problem. Section 4 shows a performance analysis with several signal types, including tones, chirps, speech, and a digitally modulated quaternary phase shift keyed (QPSK) signal to demonstrate how the number of modes may be selected. Section 5 describes the two proposed algorithms: the baseline VMD BSS algorithm and the improved culling method and presents their performance with some examples. Finally, concluding remarks are given in Section 6.

2. BACKGROUND: VARIATIONAL MODE DECOMPOSITION (VMD)

As discussed in [3] and originally presented in [1], a signal x(t) may be viewed as the sum of a set of modes, constructed from the VMD using a series of transformations. First, we use a Hilbert Transform (HT) to compute an analytic (complex) version of the real mode, denoted u_k, to be estimated; its associated center frequency will be denoted ω_k. Next, the mode is translated via an exponential shift to baseband, using ω_k as the shifting frequency. Third, the center frequency and bandwidth of the mode are estimated, by assuming the baseband signal takes on a narrow band about the zero frequency. The bandwidth is estimated by computing the magnitude squared, often referred to as the L^2 norm, of the gradient of this translated signal. Hence, the problem formulation is based on minimizing this bandwidth. The steps are repeated to compute subsequent modes, until K, the predetermined number of modes, has been reached. The sum of all the modes approximates the original signal, x(t), resulting in a constrained problem written as:

$$\underset{u_k,\omega_k}{arg\ min} ||\partial_t[(\delta(t) + \frac{j}{\pi t}) * u_k(t)]e^{-j\omega_k t}||^2$$
$$s.t. \sum_{k=1}^{K} u_k(t) = x(t), \quad (1)$$

where k = 1, 2, …, K. The term in parenthesis is applied to compute the Hilbert Transform of $u_k(t)$, * denotes convolution, the exponential term provides the translation of component ω_k to baseband, and the partial derivative ∂t is the gradient of the result. Finally, the norm of the entire term is computed using $||\cdot||^2$. This problem is solved by augmenting with additional constraints: The first constraint is a parameter, α, used to adjust the weight of the first term in the Eq. (1) depending on the noise. The more the noise, the smaller the α. The second constraint is a Lagrange multiplier, λ, used to adhere strictly to the constraint dictated by the second term in Eq. (1). These two constraints result in the updated problem formulation.

$$\underset{u_k,\omega_k}{arg\ min}\ \ \alpha||\partial_t[(\delta(t)+\frac{j}{\pi t})*u_k(t)]e^{-j\omega_k t}||^2 + ||x(t)-\sum_{k=1}^{K}u_k(t)+\frac{\lambda(t)}{2}||^2. \tag{2}$$

After some mathematical manipulation (details provided in [1]), the solution is obtained by initializing: n = 0, k = 1, $u_1^0 = \lambda^0 = \omega_1^0 = 0$, and computing the modes and frequencies as:

$$u_k^{n+1}(\omega) = \frac{\hat{X}(\omega) - \sum_{j<k}\hat{u}_j^{n+1}(\omega) - \sum_{j>k}\hat{u}_j^{n}(\omega) + \lambda^n(\omega)/2}{1+2\alpha(\omega-\omega_k^n)^2}, \tag{3}$$

and

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega|\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega}, \tag{4}$$

where

$$\lambda^{n+1}(t) = \lambda^n(t) + \tau[x(t) - \sum_{k=1}^{K}u_k^{n+1}(t)], \tag{5}$$

and τ is a time step constant. We iterate from k = 1, 2, …, K and also from n = 0, 1, …, N-1 until the modes and frequencies converge; for our simulations, we let N = 500. Note that one limit with the VMD algorithm is in selection of the number of modes, similar to selection of the number of signals when using PCA. If underselected, the interferer is not fully captured but if overestimated, then partial SOI cancellation occurs. In both cases, errors in estimating the interferer and SOI occur. In the next section, we discuss how K can be selected based on the expected types of signal.

3. SIGNAL MODEL

Let us assume we collect a signal that takes the form:

$$y(t) = A_x x(t) + I(t) + \sigma_n n(t), \tag{6}$$

where x(t) is a SOI, I(t) is an interfering signal, and n(t) is noise, modeled as additive white Gaussian noise (AWGN). The goal is to blindly separate and estimate both the non-stationary SOI and non-stationary interferer in the presence of the noise, hence this is a BSS problem. The amplitude of the interferer is set to unity, without loss of generality. The amplitude of the SOI, A_x, is set to achieve a certain SOI to interferer ratio, denoted as a carrier-to-interference (CIR), in dB, computed by:

$$A_x = \sqrt{10^{CIR/10}}. \tag{7}$$

Furthermore, the noise standard deviation, σ_n, is set to achieve a desired carrier-to-noise ratio (C/N), also in dB, given by:

$$\sigma_n = \sqrt{10^{-(C/N)/10} \cdot 0.5}, \tag{8}$$

Without loss of generality, we assume: (1) that the amplitude of I(t) is unity, so that we vary the amplitude of the SOI alone; and (2) $A_x < 1$, meaning the interferer is stronger than the SOI; if $A_x > 1$ we would treat the SOI as the interferer. An important consequence of the second assumption is that the VMD will estimate the interfering signal first; we limit our attention therefore to the case where -10 < CIR < 0 dB, resulting in $0.1 < A_x < 1$. We also limit our attention to 0 < C/N < 15 dB.

For our first proposed, baseline, approach, we input y(t) to the VMD algorithm, and the output modes are used to compute the estimate of I(t), denoted $\hat{I}(t)$ as:

$$\hat{I}(t) = \sum_{k=1}^{K} u_k(t). \tag{9}$$

We again emphasize that this equation is based on the assumption that the interferer power is higher, hence it will be estimated first by VMD. We then estimate the SOI as:

$$\hat{x}(t) = y(t) - \hat{I}(t). \tag{10}$$

Finally, performance is determined by computing the mean-square error (MSE) between the true signals and their estimates, using:

$$MSE_{I(t)} = \overline{(I(t) - \hat{I}(t))^2}, \tag{11}$$

and

$$MSE_{x(t)} = \overline{(x(t) - \hat{x}(t))^2}, \tag{12}$$

For the purpose of analysis, we assume four types of SOI and interferer types. These are as follows:

- Tone: x(t)/I(t) = $e^{2\pi f_t t}$, where f_t is the tone frequency;
- Chirp: x(t)/I(t) = $e^{j2\pi f_d t} e^{j\pi\beta t^2}$, where f_d is the initial frequency of the chirp, and β is the rate of change in the frequency in Hz/sec;
- QPSK: x(t)/I(t) = A(t) + jB(t), where A, B take on values of {-1, +1} and are randomly generated;
- Speech: x(t)/I(t) = $\sum_{i=1}^{N_s} s_i(t)$, where

$$s_i(t) = a_i(t) cos(2\pi[f_{c_i} + \int_0^t f_i(\tau) d\tau] + \theta); \tag{13}$$

and where $a_i(t)$ is the amplitude of component i, which will be time-varying, i.e. non-stationary, f_{ci} is its center frequency, $f_i(t)$ is the frequency modulation, i = 1, 2, …, N_s, and N_s is the number of signals combined to form the speech signal [4].

For simulation purposes, we let $1 < f_t < 5$ kHz, $f_d = 40$ kHz, $\beta = 10$ kHz/s, and $1 < f_c < 1; 000$ Hz, and $N_s = 6$.

In the next section, we study the performance of VMD as a BSS algorithm, using Eqs. (11) and (12) as the performance metrics, with combinations of the above four signal types representing the interferer and SOI pair. After that, we discuss how to cull some modes in Eq. (9) to improve the MSEs of both signals.

4. PREDICTION OF THE NUMBER OF MODES FOR BLIND SOURCE SEPARATION (BSS)

Fig. 1 shows several examples of interferer and SOI pairs, where the number of modes K is plotted vs. the MSE between the tru/e interferer (or SOI) and its estimate, using CN = 15 dB and CIR = -10 dB. The first two cases show one and three tone interferers, respectively, with a speech SOI. The optimum number of modes K when the interferer is a tone is equal to the number of tones. If its underestimated, then a tone may not be cancelled, resulting in errors to both tone and SOI estimates. Overestimation also results in slight increase in errors; this is expected because now a portion of the SOI is being cancelled, but the increase is small because the speech signal contains many spectral components.

The third plot shows a QPSK interferer. Here, the number of modes is roughly equal to the number of lobes in the QPSK signal spectrum. A larger hit to the MSE of the SOI is seen if the number of modes is underestimated, because this does not result in adequate cancellation of an in-band lobe. If the QPSK signal is wideband, a few modes are still needed to capture the wider main lobe. Hence, a good value of K is about 6 to 10. In the fourth plot, a speech signal is the interferer. This is the most unstructured of all the interferers, with several narrowband spectral components, and hence the most modes are required. Still, the number of modes needed to estimate the signal can be assumed to be fairly small, i.e. $K < 10$.

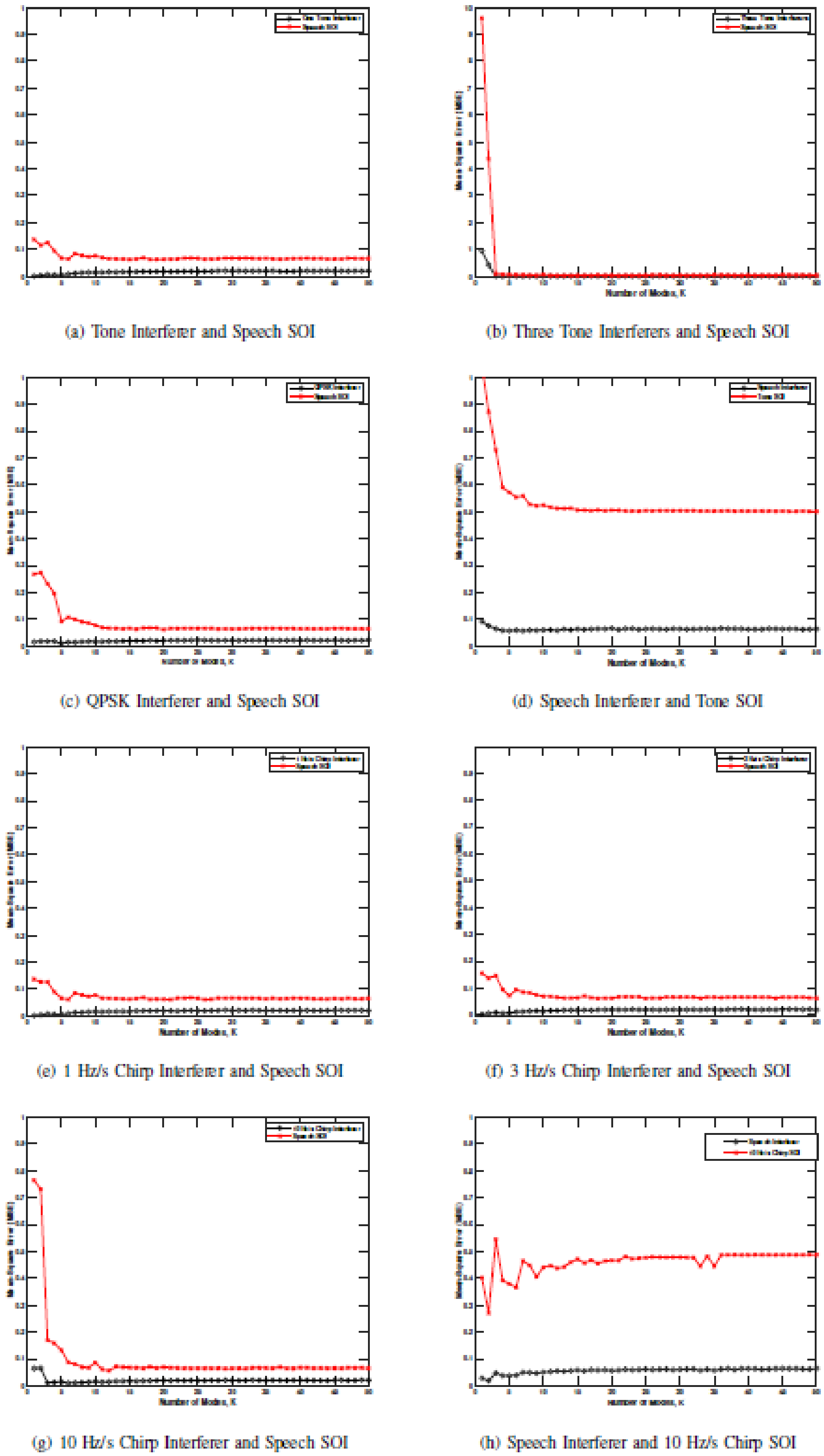
When the interferer is a chirp, as in the fifth, sixth, and seventh plots, the number of modes required depends on the chirp bandwidth, with more modes for wider chirps. However, even in this case, only about K = 6 to 10 modes is needed. Note that if the number of modes is estimated to be too high, then slight errors occur in the estimation of both the interferer and speech signals. If it's too low, larger errors can occur for estimation of the SOI, as the interferer is not being adequately cancelled. Furthermore, for the wideband chirp case, underestimating the modes results in significant errors for both signals, so overestimation is better.

In the final plot where the interferer is a speech signal and the SOI is a chirp, we still require K = 6 to 10. This is the most stressing of all the cases with significant spectral overlap of the signals, so we see a degradation in the MSE performance. However, the algorithm can still separate the signals and does not require any increase in K. Hence, we see that VMD in general operates with small values of K, smaller for tones/narrowband chirps, larger for wideband chirps/QPSK, and largest for speech. A summary of the number of modes for the signal pairs is in Table I.

The next section describes our second proposed approach to reduce the MSE by frequency- or amplitude-based mode culling.

5. IMPROVED ALGORITHM USING FREQUENCY-BASED MODE CULLING

Revisiting Eqs. (6) and (9), we explore ways to improve the MSEs of interferer and SOI. The selected method involves an initial estimate of the frequency bands at which a strong SOI component may exist. Then, the modes u_k associated with those frequencies, ω_k, are removed from Eq. (9) to obtain an improved estimate of the interferer. We write the summation as before, but subtract every mode j, where ω_j is estimated to be a frequency of $X(\omega)$, and where $X(\omega) = FFT\{x(t)\}$, i.e.



(a) Tone Interferer and Speech SOI

(b) Three Tone Interferers and Speech SOI

(c) QPSK Interferer and Speech SOI

(d) Speech Interferer and Tone SOI

(e) 1 Hz/s Chirp Interferer and Speech SOI

(f) 3 Hz/s Chirp Interferer and Speech SOI

(g) 10 Hz/s Chirp Interferer and Speech SOI

(h) Speech Interferer and 10 Hz/s Chirp SOI

FIGURE 1: Number of Modes, K, vs. Mean-Square Error (MSE); C/N = 15 dB, CIR = -10 dB.

Interferer	Signal-of-Interest (SOI)	Number of Modes, K
Tone	Speech	1
Three tones	Speech	3
QPSK	Speech	6 to 10
Speech	Tone	10
Chirp	Speech	6 to 10
Speech	Chirp	6 to 10

TABLE 1: Interferer, SOI, and Number of Modes.

$$\hat{I}(t)_{culled} = \sum_{\substack{k=1 \\ k \neq j, \omega_j \epsilon X(\omega)}}^{K} u_k(t). \tag{14}$$

Since certain modes have been removed in estimating the higher powered interferer, we term this new approach as VMD with mode culling. Note that the ωj's have to be estimated or guessed at, since we do not know X(ω). This could be done by knowing what bands a particular SOI lies in or knowing the type of signal under search, since most signals occupy specific frequency bands designated by the International Telecommunications Union (ITU) internationally; in the United States the Federal Communications Commission (FCC) determines frequency bands for specific services. The SOI is still obtained from the new interferer estimate using Eq. (10), and MSEs are computed as before from Eqs. (11) and (12).

Another way to do the mode culling is to try to estimate the amplitude of a SOI, which of course is easiest to estimate with a narrowband or tone SOI, and then remove the $u_k(t)$'s whose amplitude is closest, i.e. the k^{th} component where $A_x \approx \max(u_k(t))$. We can write:

$$\hat{I}(t)_{culled} = \sum_{\substack{k=1 \\ k \neq j, u_j(t) \epsilon A_x \approx max(u_j(t))}}^{K} u_k(t). \tag{15}$$

From the analysis conducted, the selection should be based on which parameter is easier to estimate, i.e. more stationary over time. Removal of estimated SOI modes by culling has the objective of better estimating the interferer, which in turn results in an improved estimate of the SOI.

Fig. 2 illustrates the culling concept for a speech interferer overpowering a narrow tone SOI. We take a guess as to the frequency of the SOI, choosing the ω_k that most closely reflects that frequency. We then remove the associated mode, u_2 in this case, from the interferer estimation. We compute the MSE of the resulting signals. The figure illustrates the MSEs of both SOI and interferer, before and after culling. We also show the ratio of MSEs to determine how much improvement the culling provides. An improvement of one to two orders of magnitude in both the interferer and SOI are seen. This is explained by noting that the tone has a power that is higher than some modes of the interferer, hence it will be included in the mode decomposition of the interferer. By using its estimated frequency or amplitude to remove (cull) it, we can improve the estimates of both signals. The improvement is seen across the whole range of C/N and CIR, where the interferer MSE is substantially reduced because estimation error due to the presence of the SOI is removed. Also, note that when the SOI is culled, the MSE performance of the interferer is more constant across the range of CIR, which is expected because the SOI is removed in obtaining the interferer estimate, so its power is not important. The resultant culling also improves the SOI MSE for all C/N's and CIR's.

The next example is a more stressing case where there is a speech interferer riding on a chirp SOI. There is overlap between the two in the spectral domain, and the amplitude of the SOI is high enough so as to cause errors in estimating the interferer, and hence the SOI itself. Estimating that the chirp is present over the first three frequencies, and culling these, we obtain the result in Fig. 3. Both sets of plots again show the results over a range of C/N and CIR for completeness. For the baseline algorithm, when CIR is very small, the MSE of the interferer is low. This is because the interferer is much stronger than the SOI, so it can be easily estimated; hence, the MSE of the SOI is smallest in this range, too. MSE of both interferer and SOI degrade gracefully as C/N reduces. With culling, we see a similar improvement in the interferer's MSE as in the previous example; it is improved across the whole range of C/N and CIR, and is overall reduced. Culling also significantly reduces the MSE of the SOI overall, except at very low CIR. At low CIR, i.e. CIR = -10 dB, the culling approach does not work as well, since the modes estimate the interferer, not the SOI, so culling introduces cancellation errors. The most significant improvement in SOI MSE is seen at higher CIR. However, performance improvement at low CIR is difficult. Hence, future effort can be directed at joint algorithms to improve performance at higher CIR. If we look at the ratio of culled MSE to the baseline MSE, which is also plotted in Figs. 2 and 3, both for the interferer and SOI, we see that up to one or two orders magnitude improvement is with the culling approach. This occurs with both tone and chirp SOIs.

These results demonstrate that the culling approach offers improvement when compared to the conventional VMD approach presented in [1]; EMD and other conventional methods such as Fourier analysis, wavelet analysis, and PCA fail to perform here due to the presence of the interfering signal. Thus, the VMD culling approach offers promise in blindly separating non-stationary signals, surpassing the best currently available approach of the conventional VMD.

6. CONCLUSION

This paper describes the application of the variational mode decomposition (VMD) algorithm to the problem of blind source separation (BSS). A higher powered interferer is estimated using VMD first, and then subtracted from the received mixture to extract a lower powered signal-of-interest (SOI). Typically, as we show with numerous examples and signal types, the number of VMD modes can be less than K = 10. If the interferer can be effectively broken down into N tones, then we should choose K = N for optimum performance. For wideband interferers, we typically choose K = 6 to 10. We further describe an improvement to this blind estimation problem by using a culling approach, wherein high powered frequency components of the SOI that may contribute to the modes used to estimate the interferer are culled, by excluding those modes k whose frequencies ω_k correspond to frequency bands of the SOI. The culling approach improves estimation of the interferer and SOI, typically producing mean-square error (MSE) reduction by an order of magnitude or more. This is tested over a range of power levels and noise levels. Future work includes the development of a joint blind source separation algorithm using the VMD method and testing the algorithm with real data, such as speech signals collected in the presence of other non-stationary interferers.

7. ACKNOWLEDGMENTS

The author thanks The Aerospace Corporation for funding this work and Alan Foonberg for reviewing the paper.

8. REFERENCES

[1] K. Dragomiretskiy, and D. Zosso, "Variational Mode Decomposition," IEEE Trans. on Signal Proc., Vol. 62, No. 3, pp. 531-544, 2014.

[2] 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. Royal Soc. A: Math., Phys. Eng. Sci., Vol. 454, No. 1971, pp. 903-995, Mar. 1998.

[3] P. Dey, U. Satija, and B. Ramkumar, "Single Channel Blind Source Separation Based on Variational Mode Decomposition and PCA," Proc. Annual IEEE India Conference (INDICON), New Delhi, India, pp. 1-5, 2015.

[4] E. Lai, "Practical Digital Signal Processing", Elsevier Science and Technology, Oxford, UK, Jan. 2004.

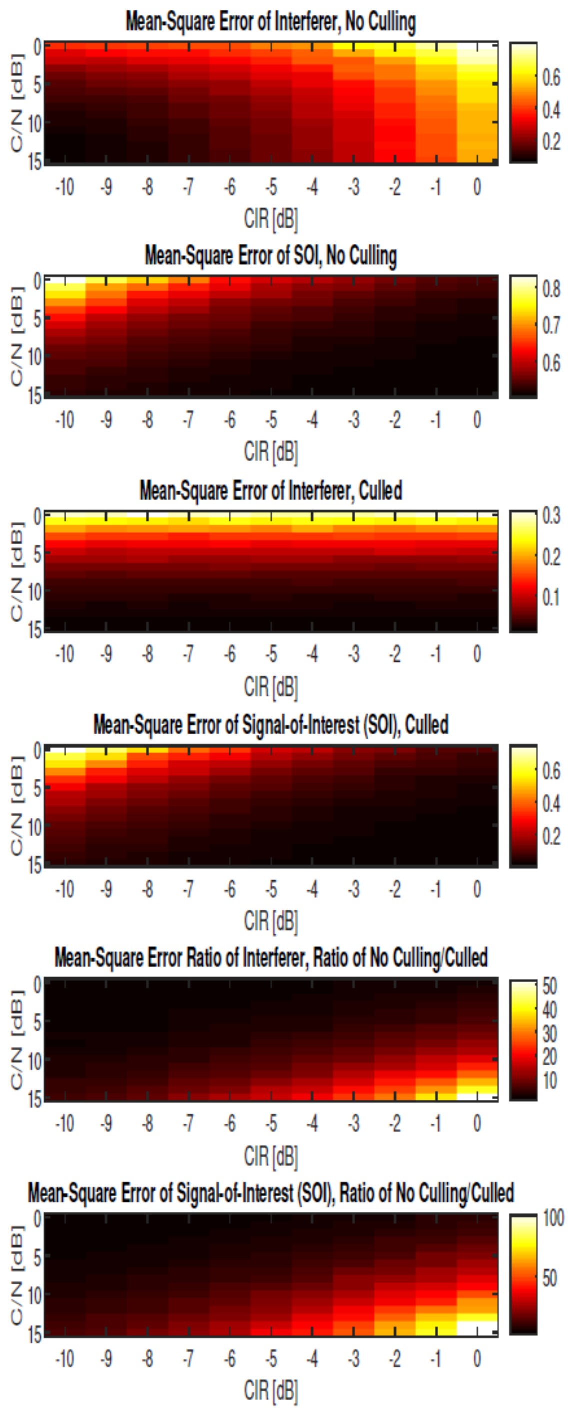


FIGURE 2: Mean-Square Errors (MSEs); Speech Interferer and Tone SOI; K = 10.

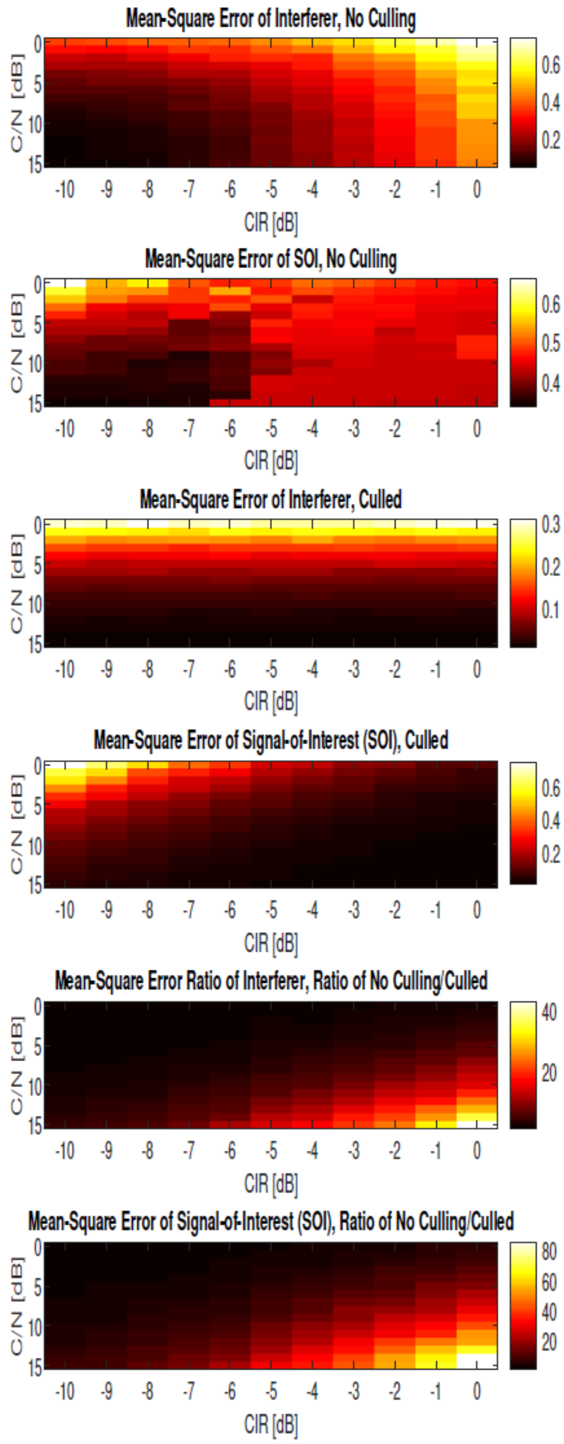


FIGURE 3: Mean-Square Errors (MSEs); Speech Interferer and 10 Hz/s Chirp SOI; K = 10.