FREQUENCY-DOMAIN CONTRAST FUNCTIONS FOR SEPARATION OF CONVOLUTIVE MIXTURES

Jean-Christophe Pesquet

Laboratoire Systèmes de Communication
Université de Marne-la-Vallée
5, Boulevard Descartes
77454 Marne la Vallée Cédex 2, France
e-mail:pesquet@univ-mlv.fr

Binning Chen and Athina P. Petropulu

ECE Department, Drexel University
Philadelphia, PA 19104, USA
e-mail: {chenbn,athina}@cbis.ece.drexel.edu

ABSTRACT

This paper addresses the problem of blind separation of convolutive mixtures via contrast maximization. New frequency domain contrast functions are constructed based on second and higher-order spectra of the observations. They allow to separate mixtures of sources which are spatially independent, and temporally possibly non i.i.d. linear or non-linear processes. The proposed criteria provide a framework for extending to the convolutive case contrasts that have been proposed in the context of instantaneous mixtures.

1. INTRODUCTION

Blind separation of convolutive mixture is a challenging signal processing problem with a wide variety of applications, such as multiuser multiaccess communications, bioengineering and seismology. Among the possible approaches [3, 4, 5, 8, 9], frequency-domain methods appear particularly appealing [1]. In the Fourier domain, the original problem is transformed into a set of instantaneous-like ones (although a frequency-dependent permutation ambiguity remains to be resolved).

We here propose a set of new contrast functions, which are designed based on second and fourth-order spectra of the observations. Extensions to any higher-order spectra can also be easily derived. The mathematical formulation allows for the separation of the sources independently of their temporal dependence structure. For example, the proposed functions could be helpful in separating convolutive mixtures of non-i.i.d or non-linear processes. A scenario where such non-linear sequences could arise, is the output of an error correcting device in a communication system. The non-linearity there is due to the redundancy introduced by the correction codes to the data flow. To the best of our knowledge, few methods exist in literature [7] allowing to separate convolutive mixtures of nonlinear processes.

The paper is organized as follows. In Section 2 we state the problem and introduce the necessary notations. In Section 3 we relate our frequency domain approach to a joint diagonalization problem for a polyspectrum tensor. Section 4 contains our main results and describes how frequency domain contrasts can be built. Some particular cases and extensions are also discussed in this section.

This work was supported by National Science Foundation under grant MIP-9553227

2. PROBLEM STATEMENT

We consider a convolutive mixture of $N \in \mathbf{N}^*$ unknown complex valued source signals $(s_1(n))_{n\in\mathbf{Z}}, \ldots, (s_N(n))_{n\in\mathbf{Z}}$. The signal output of the mixture is assumed to be an N-dimensional vector

$$\begin{aligned}\mathbf{x}(n) &= (x_1(n), \ldots, x_N(n))^T \\ &\triangleq \sum_{k=-\infty}^{\infty} \mathbf{h}(n-k)\mathbf{s}(k) + \mathbf{b}(n) \qquad (1)\end{aligned}$$

where $\mathbf{s}(n) = (s_1(n), \ldots, s_N(n))^T$, $\mathbf{b}(n)$ denotes some noise vector and $(\mathbf{h}(n))_{n\in\mathbf{Z}}$ is the unknown MIMO impulse response of the mixing system. For all $n \in \mathbf{Z}$, $\mathbf{h}(n)$ is an $N \times N$ matrix with elements $[h_{ij}(n)]_{1\leq i\leq N, 1\leq j\leq N}$.

In this work, we will make the following assumptions:

A1. the sources $s_i(n)$, $i \in \{1, \ldots, N\}$ are mutually independent random sequences which are uncorrelated, and have unit variance. The trispectrum of source $s_i(n)$ (i.e. the Fourier transform with respect to $(\tau_1, \tau_2, \tau_3) \in \mathbf{Z}^3$ of $\mathrm{Cum}[s_i(n), s_i^*(n+\tau_1), s_i(n+\tau_2), s_i^*(n+\tau_3)]$) is assumed to be defined and will be denoted by $\Gamma_i^4(\omega_1, \omega_2, \omega_3)$.

A2. For all $(i,j) \in \{1, \ldots, N\}^2$, the filter with impulse response $(h_{ij}(n))_{n\in\mathbf{Z}}$ is stable, which guarantees the existence of a bounded frequency response matrix

$$\mathbf{H}(\omega) \triangleq \sum_{n=-\infty}^{\infty} \mathbf{h}(n) e^{-\imath n\omega}, \qquad \omega \in [-\pi, \pi). \qquad (2)$$

A3. For all $\omega \in [-\pi, \pi)$, $\mathbf{H}(\omega)$ is invertible.

A4. The noise $(\mathbf{b}(n))_{n\in\mathbf{Z}}$ is Gaussian, zero-mean and stationary with known spectrum density matrix $\mathbf{S_b}(\omega)$.

In this context, our objective will be to estimate the MIMO frequency response, or, more precisely, its inverse so as to perform source separation, i.e. recover the sources from the observations. The considered problem possesses some inherent indeterminacies as the best one can expect is to find a solution $\mathbf{G}(\omega)=\widehat{\mathbf{H}}^{-1}(\omega)$ such that

$$\mathbf{G}(\omega)\mathbf{H}(\omega)=\mathbf{P}\,e^{\imath(\Theta+\mathbf{D}\omega)} \tag{3}$$

where $\mathbf{P}$ is a permutation matrix, Θ is a real diagonal matrix and $\mathbf{D}$ is an integer diagonal matrix. When such a property holds, we will say that a type-I separation is achieved. Sometimes however, we are only able to guarantee that

$$\mathbf{G}(\omega)\mathbf{H}(\omega)=\mathbf{P}\,e^{\imath\Phi(\omega)} \tag{4}$$

where $\Phi(\omega)$ is a real diagonal matrix, and we say that a type-II solution is obtained. This means that, if $\mathbf{g}(n)=[g_{ij}(n)]_{1\le i\le N,1\le j\le N}$ denotes the impulse response of the separation system,

$$y_i(n)\triangleq\sum_{j=1}^{N}\sum_{k=-\infty}^{\infty}g_{ij}(n-k)x_j(k),\quad i\in\{1,\dots,N\} \tag{5}$$

then corresponds to an all-pass filtered sequence of one of the sources. In order to estimate the sources up to a phase/delay ambiguity, a blind monodimensional deconvolution method must be applied to each of the signals $(y_i(n))_{n\in\mathbf{Z}}$. When the sources are not i.i.d., some additional assumptions (e.g. prior knowledge on the source statistics or the structure of the mixing system) are required to realize this operation. In the i.i.d. case, there exist many blind single-input single-output deconvolution algorithms that could be used to recover the sources. For these reasons, in this work we will mainly focus on type-II solutions.

3. A JOINT DIAGONALIZATION CRITERION

As $\mathbf{S_b}(\omega)$ is assumed to be known, it is possible to prewhiten the mixed signals. Let $\mathbf{V}(\omega)$ be the frequency response of a prewhitening filter and let

$$\mathbf{W}(\omega)\triangleq\mathbf{V}(\omega)\mathbf{H}(\omega)=(W_{i,j}(\omega))_{1\le i\le N,1\le j\le N} \tag{6}$$

be the remaining paraunitary frequency response to be identified. The 4th order cross-spectrum of the prewhitened data, observed at locations i,j,l_1,l_2, equals :

$$C^4_{ijl_1l_2}(\omega_1,\omega_2,\omega_3)=\sum_{p=1}^{N}\Gamma^4_p(\omega_1,\omega_2,\omega_3)$$
$$W_{ip}(-\omega_1-\omega_2-\omega_3)W^*_{jp}(-\omega_1)W_{l_1p}(\omega_2)W^*_{l_2p}(-\omega_3). \tag{7}$$

Let $\mathbf{C}^4_{l_1l_2}(\omega_1,\omega_2,\omega_3)$ denote the matrix whose (i,j)-th element is equal to $C^4_{ijl_1l_2}(\omega_1,\omega_2,\omega_3)$. We get:

$$\mathbf{C}^4_{l_1l_2}(\omega_1,\omega_2,\omega_3)=\mathbf{W}(-\omega_1-\omega_2-\omega_3)$$
$$\mathbf{D}_{l_1l_2}(\omega_1,\omega_2,\omega_3)\mathbf{W}(-\omega_1)^H \tag{8}$$

where $\mathbf{D}_{l_1l_2}(\omega_1,\omega_2,\omega_3)$ is a diagonal matrix.

Thus, the source separation problem can be addressed in the following way: For a fixed pair of frequencies, (ω_2,ω_3), find $\widehat{\mathbf{W}}(-\omega_1-\omega_2-\omega_3)$ and $\widehat{\mathbf{W}}(-\omega_1)$ maximizing

$$\mathcal{I}(\omega_1,\omega_2,\omega_3)\triangleq\sum_{l_1,l_2}\mathrm{on}(\widehat{\mathbf{W}}(-\omega_1-\omega_2-\omega_3)^H$$
$$\mathbf{C}^4_{l_1l_2}(\omega_1,\omega_2,\omega_3)\widehat{\mathbf{W}}(-\omega_1)) \tag{9}$$

where $\mathrm{on}(\mathbf{M})$ denotes the sum of the squared moduli of the on-diagonal terms of $\mathbf{M}$. This maximization amounts to a problem of joint diagonalization (or more exactly, joint singular value decomposition) of a set of matrices.

This constitutes the basic idea of a separation algorithm which was shown to be effective in the case of i.i.d. sources [1].

It is important to note that, $\mathcal{I}(\omega_1,\omega_2,\omega_3)$ is a function of $\widehat{\mathbf{W}}$. For notation concision, we will not make this dependence explicit for $\mathcal{I}(\omega_1,\omega_2,\omega_3)$ or any other criteria which will be derived from $\mathcal{I}(\omega_1,\omega_2,\omega_3)$.

We now study more carefully the criterion which is optimized in the approach we mentioned. In particular we have the following invariance property whose proof is omitted due to the lack of space:

Lemma 1 *Let us consider the global MIMO filters corresponding to the true one cascaded with the inverse of the estimated mixing system:*

$$\widetilde{\mathbf{W}}(\omega)\triangleq\widehat{\mathbf{W}}(\omega)^H\mathbf{W}(\omega). \tag{10}$$

Then,

$$\mathcal{I}(\omega_1,\omega_2,\omega_3)=\sum_{i,j}|\Gamma^4_j(\omega_1,\omega_2,\omega_3)|^2$$
$$|\widetilde{W}_{ij}(-\omega_1-\omega_2-\omega_3)|^2|\widetilde{W}_{ij}(-\omega_1)|^2$$
$$=\sum_{il_1l_2}|\widetilde{C}^4_{iil_1l_2}(\omega_1,\omega_2,\omega_3)|^2 \tag{11}$$

where $(\widetilde{C}^4_{iil_1l_2}(\omega_1,\omega_2,\omega_3))_{i,l_1,l_2}$ correspond to the cross-trispectra of the outputs of the global system.

This result shows in a simple way that, in general, we cannot expect to estimate the phase of the system from the only maximization of $\mathcal{I}(\omega_1,\omega_2,\omega_3)$ as the value of this criterion is only depending on the moduli of $\widetilde{W}_{ij}(-\omega_1-\omega_2-\omega_3)$ and $\widetilde{W}_{ij}(-\omega_1)$. Besides, if the quantities in (11) are not depending on frequency, which is equivalent to considering instantaneous mixtures of i.i.d. sources, the considered criterion reduces to the one used in the JADE source separation algorithm. This means that our approach extends the work in [2] to the convolutive case.

4. CONNECTIONS WITH CONTRASTS

In this section, the following assumption will be made:

A5. For at least $N-1$ sources and for almost all $(\omega,\nu)\in[0,2\pi)^2$, there exists $\alpha_{\omega,\nu}\in[-2\pi,2\pi)$ such that

$$\Gamma^4_j\Big(\omega,\frac{\nu+\alpha_{\omega,\nu}}{2},\frac{\nu-\alpha_{\omega,\nu}}{2}\Big)\neq 0. \tag{12}$$

The assumption we adopt here is fairly weak as it allows us to consider non i.i.d. sources. For this assumption to be satisfied, at most one of the sources can be Gaussian. Possible choices for $\alpha_{\omega,\nu}$ are $\alpha_{\omega,\nu} = \nu$, $\alpha_{\omega,\nu} = -\nu$ and $\alpha_{\omega,\nu} = 0$. For these choices, only simple 2D slices of the involved trispectra have to be determined.

4.1. Main results

We will next prove that

$$\bar{\mathcal{I}} \triangleq \int_0^{2\pi}\int_0^{2\pi} \mathcal{I}\Big(\omega, \frac{\nu+\alpha_{\omega,\nu}}{2}, \frac{\nu-\alpha_{\omega,\nu}}{2}\Big)\, d\omega\, d\nu \tag{13}$$

is a type-II contrast, in the sense that its maximization allows us to separate the sources up to a scalar all-pass filtering of each of them.

Proposition 1 *Under Assumption* A5, *we have*

$$\bar{\mathcal{I}} \le \bar{\mathcal{I}}_{\max} \triangleq \sum_j \int_0^{2\pi}\int_0^{2\pi} \Big|\Gamma_j^4\Big(\omega, \frac{\nu+\alpha_{\omega,\nu}}{2}, \frac{\nu-\alpha_{\omega,\nu}}{2}\Big)\Big|^2 d\omega\, d\nu \tag{14}$$

and the upper bound is attained iff a type-II separation property holds, i.e., for all ω, $\widetilde{\mathbf{W}}(\omega) = \mathbf{P}\, e^{\imath\Phi(\omega)}$ where $\mathbf{P}$ is a permutation matrix and $\Phi(\omega)$ a real diagonal matrix.

Proof. As a result of the paraunitarity of $\widetilde{\mathbf{W}}(\omega)$, we have

$$\sum_i |\widetilde{W}_{ij}(-\omega_1-\omega_2-\omega_3)|^2 |\widetilde{W}_{ij}(-\omega_1)|^2 \le 1. \tag{15}$$

Combining this inequality with (11) yields

$$\mathcal{I}(\omega_1,\omega_2,\omega_3) \le \sum_j |\Gamma_j^4(\omega_1,\omega_2,\omega_3)|^2. \tag{16}$$

This allows to deduce that

$$\int_0^{2\pi}\int_0^{2\pi} \mathcal{I}\Big(\omega, \frac{\nu+\alpha_{\omega,\nu}}{2}, \frac{\nu-\alpha_{\omega,\nu}}{2}\Big)\, d\omega\, d\nu \le \int_0^{2\pi}\int_0^{2\pi} \sum_j \Big|\Gamma_j^4\Big(\omega, \frac{\nu+\alpha_{\omega,\nu}}{2}, \frac{\nu-\alpha_{\omega,\nu}}{2}\Big)\Big|^2 d\omega\, d\nu \tag{17}$$

and (14) is proved.

It is clear that the upper bound is attained iff, for all j and almost all $(\omega,\nu) \in [0,2\pi)^2$,

$$\Big|\Gamma_j^4\Big(\omega, \frac{\nu+\alpha_{\omega,\nu}}{2}, \frac{\nu-\alpha_{\omega,\nu}}{2}\Big)\Big|^2 \Big(1 - \sum_i |\widetilde{W}_{ij}(-\omega-\nu)|^2 |\widetilde{W}_{ij}(-\omega)|^2\Big) = 0. \tag{18}$$

Now, let j be an arbitrary index such that (12) is satisfied. We have then, for almost all $(\omega,\nu) \in [0,2\pi)^2$,

$$\sum_i |\widetilde{W}_{ij}(-\omega-\nu)|^2 |\widetilde{W}_{ij}(-\omega)|^2 = 1. \tag{19}$$

which, by integrating over $[0,2\pi]^2$ and using Parseval's equality, leads to

$$\sum_i \|\widetilde{w}_{ij}\|^4 = \sum_i \big(\sum_k |\widetilde{w}_{ij}(k)|^2\big)^2 = 1. \tag{20}$$

Furthermore, the unitary assumption implies that

$$\forall j, \qquad \sum_i \|\widetilde{w}_{ij}\|^2 = 1 \tag{21}$$

where, obviously, $\|\widetilde{w}_{ij}\| \le 1$. This shows that equality can only arise in (14) if there exists an index i_j such that $\|\widetilde{w}_{i_j j}\| = 1$. Due to (21), this implies that, for all $i \neq i_j$, $\|\widetilde{w}_{ij}\| = 0$, i.e. $\widetilde{W}_{ij}(\omega) = 0$, for almost all $\omega \in [0,2\pi)$. Proceeding in the same way, for all the (at least $N-1$) values of j such that (12) is satisfied, we establish that, on the corresponding j-th column of $\widetilde{\mathbf{W}}(\omega)$, there is only one nonzero element at row i_j. As the unitary condition also introduces the constraint,

$$\forall i, \qquad \sum_j \|\widetilde{w}_{ij}\|^2 = 1 \tag{22}$$

two of those nonzero elements of $\widetilde{\mathbf{W}}(\omega)$ cannot be located on the same row. In other words, we have proved that $\widetilde{\mathbf{W}}(\omega) = \mathbf{P}\,\mathbf{\Lambda}(\omega)$ where $\mathbf{P}$ is a permutation and $\mathbf{\Lambda}(\omega)$ a diagonal matrix. As $\widetilde{\mathbf{W}}(\omega)$ is unitary, the diagonal elements of $\mathbf{\Lambda}(\omega)$ are necessarily with unit modulus. □

Under a slightly more restrictive condition, another form of contrast can be derived:

Proposition 2 *Let us assume that*

A5'. *for at least $N-1$ sources and for almost all $(\omega,\nu) \in [0,2\pi)^2$, there exists a set $\mathcal{E}(\omega,\nu) \subset [-2\pi,2\pi)$ such that*

$$\int_{\mathcal{E}(\omega,\nu)} \Big|\Gamma_j^4\big(\omega, \frac{\nu+\alpha}{2}, \frac{\nu-\alpha}{2}\big)\Big|^2 d\alpha \neq 0. \tag{23}$$

Then,

$$\mathcal{I}_{\mathcal{E}} \triangleq \int_0^{2\pi}\int_0^{2\pi} \Big(\int_{\mathcal{E}(\omega,\nu)} \mathcal{I}\big(\omega, \frac{\nu+\alpha}{2}, \frac{\nu-\alpha}{2}\big)\, d\alpha\Big)\, d\omega\, d\nu$$

is a type-II contrast.

It is worth noting that, provided some appropriate modifications of Assumptions A5 and A5', results similar to Propositions 1 and 2 can be proved for discretized versions (involving discrete frequencies) of Criteria $\bar{\mathcal{I}}$ and $\mathcal{I}_{\mathcal{E}}$.

4.2. Specific cases

4.2.1. Integration over the whole frequency space

A form of criterion which could appear more intuitive is given by:

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathcal{I}(\omega_1,\omega_2,\omega_3)\, d\omega_1\, d\omega_2\, d\omega_3.$$

After some calculations, this expression can be seen to be equal to $\mathcal{I}_{\mathcal{E}}/2$ when $\mathcal{E}(\omega,\nu) = [-2\pi, 2\pi)$. Furthermore, it can be observed that

$$\frac{1}{16\pi^3}\mathcal{I}_{[-2\pi,2\pi)} = \sum_{il_1 l_2}\sum_{\tau_1,\tau_2,\tau_3} |\mathrm{Cum}[y_i(n), y_i^*(n+\tau_1), y_{l_1}(n+\tau_2), y_{l_2}^*(n+\tau_3)]|^2.$$

where $y_i(n)$ denotes the i-th output of the global system. In the considered case, it is easy to see that Assumption A5' reduces to the fact that for at least $N-1$ sources, there exists time-delays $(\tau_{1j}, \tau_{2j}, \tau_{3j})$ such that

$$\mathrm{Cum}[s_j(n), s_j^*(n+\tau_{1j}), s_j(n+\tau_{2j}), s_j^*(n+\tau_{3j})] \neq 0.$$

4.2.2. i.i.d. sources

When the sources are i.i.d., Assumption A5 reduces to the more common assumption: at most one of the fourth-order cumulants of the sources is zero. Furthermore, according to (11), $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ becomes a function of the two variables ω_1 and $\omega_2 + \omega_3$ in this case. Consequently, the contrasts $\bar{\mathcal{I}}$ are no more depending on the choice of $\alpha_{\omega,\nu}$, which can be made arbitrary. Moreover, we have $\bar{\mathcal{I}} = \frac{1}{2\pi}\mathcal{I}_{[-2\pi,2\pi)}$ which is the criterion studied in the previous section.

4.2.3. Other contrasts

Using the same approach as in [6] other contrasts can be derived from the previous ones. Unlike [6], the i.i.d. assumption is not necessary for these criteria to be valid. In particular, extensions to the convolutive case of contrasts introduced by L. Delathauwer and E. Moreau for instantaneous mixtures can be obtained in this way.

5. SIMULATION RESULTS

We consider two non i.i.d. source signals of length 8192 obtained by linear filtering *and* subsampling of i.i.d. sequences with Laplacian distributions. These sources are mixed by a paraunitary MIMO system of order 2. The filters have been parametrized using a lossless lattice representation. The maximization of a discretized version of Contrast $\bar{\mathcal{I}}$ with $\alpha_{\omega,\nu} = 0$ has been carried out by a Jacobi-like algorithm. The trispectrum of the mixed data has been empirically estimated for each realization using an averaged 4th order periodogram method. A Monte Carlo study of the proposed method has been realized with SNR=30dB. The 3 rotation angles characterizing the lattice representation of the mixing system have been drawn following a uniform distribution and, for each run, the source signals have been generated randomly. Since $\mathbf{W}(\omega)$ is a 2x2 FIR paraunitary system, it holds that $|W_{11}(\omega)| = |W_{22}(\omega)|$ and $|W_{12}(\omega)| = |W_{21}(\omega)|$. Figure 1 illustrates the estimation result for $|W_{11}(\omega)|$ and $|W_{12}(\omega)|$ for $0 \leq \omega \leq \pi$ (mean$\pm$ standard deviation from 20 Monte Carlo simulations, on top of true magnitudes).

6. REFERENCES

[1] B. Chen and A.P. Petropulu, "Multiple-Input-Multiple-Output Blind System Identification Based on Cross-Polyspectra," *ICASSP'2000*, pp. 584-587.

[2] J.F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signals", *IEE Proceedings-F*, Vol. 40, pp. 362-370, 1993.

[3] P. Comon, "Contrasts for multichannel blind deconvolution", *IEEE Signal Processing Letters*, Vol. 3, No. 7, pp. 209-211, July 1996.

[4] Y. Inouye and K. Hirano, "Cumulant-Based Blind Identification of Linear MIMO Systems Driven by Colored Inputs", *IEEE Trans. on SP*, pp. 1543, 1997.

[5] E. Moreau and J.-C. Pesquet, "Generalized contrasts for multichannel blind deconvolution of linear systems," *IEEE Signal Processing Letters*, June 1997.

[6] N. Thirion-Moreau and E. Moreau,"New Criteria for Blind Signal Separation", SSAP'2000, Pennsylvania, USA, pp 344-348, August 2000.

[7] C. Simon, Ph. Loubaton, C. Vignat, C. Jutten, G. d'Urso, "Separation of a class of convolutive mixtures: a contrast function approach", *ICASSP'2000.*

[8] A. Swami, G.B. Giannakis and S. Shamsunder, "Multichannel ARMA Processes," *IEEE Trans. on Signal Processing*, vol. 42, No. 4, pp. 898-913, 1994.

[9] J.K. Tugnait, "Identification and Deconvolution of Multichannel Linear Non-Gaussian Processes Using Higher Order Statistics and Inverse Filter Criteria," *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 658–672, March. 1997.

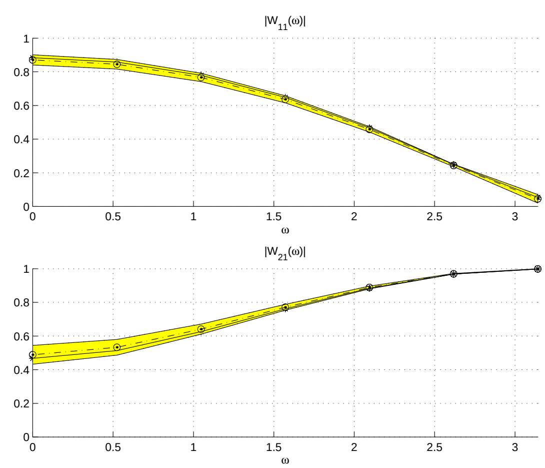

Figure 1: Estimation of $|W_{11}(\omega)|$ and $|W_{21}(\omega)|$.