# THE GENERALIZED SIDELOBE DECORRELATOR


*Craig Fancourt and Lucas Parra*

Adaptive Image & Signal Processing Group
Sarnoff Corp.
201 Washington Rd
Princeton, NJ 08543
cfancourt@sarnoff.com lparra@sarnoff.com



### ABSTRACT

We introduce a new hybrid algorithm called the generalized sidelobe decorrelator (GSD) that combines elements of geometric beamforming and blind source separation. On the one hand, it is an extension of the generalized sidelobe canceller (GSC), also known as the Griffiths-Jim beamformer, from the standard criteria of power minimization to a decorrelation criteria. On the other hand, it can be seen as an extension of a particular blind source separation (BSS) algorithm for non-stationary signals to include prior information about the location of one of the sources. However, unlike GSC, performance doesn't degrade with leakage of the source outside the primary beam and, unlike BSS, it performs well independent of whether the sources are simultaneously active. This makes it ideal for noise reduction in a continuously running on-line operation. We demonstrate its superior performance in a real-room audio experiment.


## 1. INTRODUCTION

The fields of geometric beamforming and blind source separation (BSS) have progressed on largely independent tracks. And yet, even though the literature in these two areas is now quite mature (see [1]-[4]), their assumptions, goals, and methods share many commonalities that have not been fully explored. The fundamental starting point for both is an array of sensors, each of which measures a different mixture of several sources. In beamforming, it is often the case that only one source is of interest and the others are designated as noise. In BSS, all sources that carry information are usually of interest. However, these distinctions are largely superficial. *The goal of both fields is to filter and combine the sensor signals so as to best recover the source(s) of interest*. The only fundamental differences between the two fields are the prior information that is exploited and the criteria that is used for recovery.

When the only prior information available is that the sources are statistically independent, then this defines the blind source separation problem. Independence then becomes an adaptation criteria for the filtered sensor signals. Strict independence requires an infinite amount of data to measure, therefore in practice the criteria is relaxed to some subset of higher-order statistics (HOS).

When the prior information includes the array geometry and the angular position of a single source of interest, then this defines a beamforming problem. Under the assumption that sources coming from any other direction are noise, the sensor signals are then filtered to minimize power subject to the constraint that a delay-sum beam points in the direction of the known source. When this constraint is appended to the optimization criteria, there results the *linearly constrained minimum-variance* (LCMV) algorithm introduced by Frost [5]. When the constraint is explicitly embedded in the architecture, there results the *generalized sidelobe canceller* (GSC) introduced by Griffiths and Jim [6].

However, there are many problems where the distinction between BSS and beamforming is not so clear. For example, consider the case where the prior information is the following:

(a) there are multiple sources of interest;
(b) the sources are independent;
(c) the angular position of one of the sources is known;
(d) the array geometry is known.

By themselves, (a) and (b) define the BSS problem while (c) and (d) define a beamforming problem. Therefore, right away we can phrase this as one of two equivalent problems: (1) how can we modify BSS to include the case when the angular location of one of the sources and the array geometry are known? or (2) how can we modify adaptive beamforming to recover other sources in addition to the one that lies in the known direction?

We have previously considered these priors from the standpoint of the first problem formulation. That is, we began with a multiple-input/multiple-output feedforward system utilizing finite impulse response (FIR) filters adapted under an independence criteria on the outputs. We then appended a linear constraint that the filters associated with one of the outputs represent a beamformer with a constant unit gain response in the known source direction. The resulting algorithm is called *geometric source separation* (GSS) [7], and can be considered a HOS analog of the LCMV algorithm. From the point of view of BSS, the geometric constraint regularizes the solutions, resolving the ambiguities due to the frequency permutation problem [7] and the indeterminacies in the case of more sensors than sources. From the point of view of beamforming, GSS solves many problems associated with the so-called "leakage" problem caused by a power minimization criteria, which we will discuss in greater detail in a later section.

In this paper, we consider the priors from the standpoint of the second problem formulation. We propose an alternative and more efficient way of implementing the linear constraints following the approach proposed by Griffiths and Jim in their alternative implementation of the LCMV. In the same spirit, we call this new algorithm the *generalized sidelobe decorrelator* (GSD).

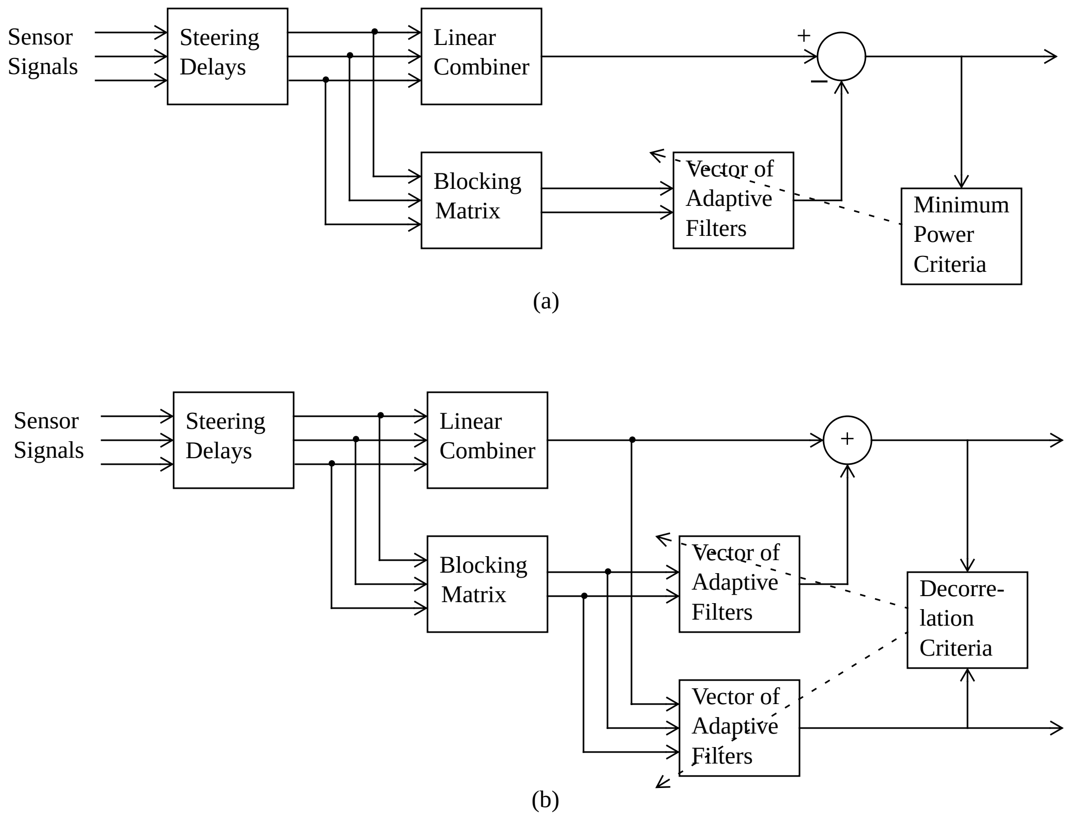


Figure 1. (a) *generalized sidelobe canceller*; (b) *generalized sidelobe decorrelator.*

## 2. PROBLEM STATEMENT

The problem we seek to solve is the following: $N$ unknown source signals are convolutively mixed and measured by $M$ sensors

$$\mathbf{x}(t) = \mathbf{A} * \mathbf{s}(t) \tag{1}$$

where $\mathbf{s}$ is an unknown ($N$x1) vector of source signals, $\mathbf{A}$ is an unknown ($M$x$N$) mixing matrix of channel impulse responses, and $\mathbf{x}$ is a measured ($M$x1) vector. The convolution operator * here implies both matrix multiplication and convolution. We then seek a matrix of filters operating on the sensor measurements

$$\mathbf{y}(t) = \mathbf{W} * \mathbf{x}(t) \tag{2}$$

such that the components of the ($N$x1) output $\mathbf{y}$ recover the source signals, where $\mathbf{W}$ is a ($N$x$M$) matrix of filter impulse responses.

If we stop here and define successful recovery of the signals to be when the components of $\mathbf{y}$ are independent, then this defines the blind source separation problem. If we add other prior information, such as the array geometry and the position of one or more of the sources, then this defines a geometric beamforming problem. We first review each of these as separate solutions and then show how they can be combined.

### 2.1 Generalized Sidelobe Canceller Solution

The fundamental principle of beamforming is that prior knowledge of the sensor and source geometry can be exploited to design filters that delay the sensor signals so that they add in-phase for a desired direction, greatly increasing the sensitivity of the array to a source in that direction. However, while the delay-sum beamformer is optimal for a point source in the presence of spatially diffuse noise, it is not optimal if the noise also originates from point sources.

As we mentioned previously, one possibility for suppressing these noise sources is the LCMV algorithm, which adaptively filters the sensor signals so as to minimize power, subject to a constraint that a delay-sum beam points in the direction of the source of interest.

An alternative but equivalent approach is the GSC, shown in Fig. 1a. It also implements a power minimization criteria on the filtered sensor signals. However, unlike the LCMV, the requirement that a beam point in the direction of interest is enforced in the architecture rather than the criteria. Specifically, the GSC utilizes a delay-sum beam through the use of steering delays followed by a linear combiner. The linear combiner is a window that can be designed to vary the trade-off between main lobe width and sidelobe energy. After the steering delays but prior to the linear combiner, the $M$ delayed sensor signals are all in phase. This is exploited to form $M$-1 secondary beams orthogonal to the primary beam through the use of a "blocking matrix". Each row of the blocking matrix is constrained to sum to zero to ensure that the resulting secondary beams will all have a null in the direction of the primary beam. During adaptive power minimization, the secondary beams are adapted out of the primary beam but are prevented by the blocking matrix from canceling any signal that *exclusively* resides in the primary beam.

The GSC approach has the advantage that the resulting optimization can be carried out using unconstrained power minimization, such as the least mean squares (LMS) algorithm. Unlike the LCMV, the constraint is always enforced and no extra steps have to be taken to ensure that the filter weights don't stray from the constraint over time due to finite precision effects.

### 2.2 Blind Source Separation Solution

In order to understand what an independence criteria can accomplish, it suffices to determine the set of all operations on **s** such that the resulting signals are still independent. Clearly a reordering of the components of **s** does not affect their independence. The components of **s** can also be separately filtered, either linearly or nonlinearly, without affecting their independence. Thus, **y** can only approximate **s** to within a permutation and filtering operation. The latter limitation means that blind source separation is distinct from the problem of blind deconvolution. That is, blind source separation by itself cannot recover the components of **s** from filtered versions of themselves.

We have previously studied this problem for the special case when the source signals are non-stationary processes [10]. BSS is primarily based on the assumption of statistical independence of the source signals. For stationary signals, second-order statistics (decorrelation) is not sufficient to identify and invert the mixing coefficients, and higher-order statistics have to be considered. However, for non-stationary signals, varying second-order statistics provides a sufficient constraint for separation.

In the time domain, independence must be tested not only at the same instant of time, but for all possible combinations of delays of the components of **y**. This problem can be ameliorated by performing the separation in the frequency domain. In the frequency domain, convolution becomes multiplication and (2) becomes

$$\mathbf{Y}(\omega, t) = \mathbf{W}(\omega, t) \cdot \mathbf{X}(\omega, t) \tag{3}$$

Note that because the signals are *assumed* non-stationary, we have written their frequency response as an *implicit* function of time. We have written the ($N$x$M$) matrix of filter frequency responses, $\mathbf{W}(\omega,t)$, as an implicit function of time with an eye towards online adaptation.

Equation (3) describes *any* linear system. Ultimately, we must implement it in a specific architecture. In this paper, we use finite impulse response (FIR) filters because this allows the actual filtering operation to be carried out in the frequency domain.

In [8], we adopt a decorrelation criteria that is the sum of the squares of the *coherence functions* between all $N$x($N$-1)/2 distinct pairs of outputs:

$$J = \sum_t \sum_{i,j} \frac{\left|S_{Y_iY_j}(\omega, t)\right|^2}{S_{Y_iY_i}(\omega, t) S_{Y_jY_j}(\omega, t)} \tag{4}$$

where $S_{Y_iY_j}(\omega, t)$ is the cross-power spectral density between outputs $i$ and $j$ at time $t$. We estimate $S_{Y_iY_j}$ directly in the frequency domain using a recursive estimator, which in matrix form can be written

$$\mathbf{S}_{\mathrm{YY}}(\omega, t) = \gamma \ \mathbf{S}_{\mathrm{YY}}(\omega, t-T) + (1-\gamma) \ \mathbf{Y}(\omega, t) \cdot \mathbf{Y}^H(\omega, t) \tag{5}$$

where $T$ is a block processing time. Stochastic gradient descent on (4) using (5) leads to the weight update equation in matrix form

$$\Delta\mathbf{W} = -\eta \Lambda_{\mathrm{YY}}^{-1} \cdot [\mathbf{S}_{\mathrm{YY}} - \Lambda_{\mathrm{YY}}] \cdot \Lambda_{\mathrm{YY}}^{-1} \cdot \mathbf{S}_{\mathrm{YX}} \tag{6}$$

where $\Lambda_{\mathrm{YY}} = \mathrm{diag}[\mathbf{S}_{\mathrm{YY}}]$ and $\mathbf{S}_{\mathrm{YX}}$ is the cross-power spectral density between the output and input, also estimated recursively:

$$\mathbf{S}_{\mathrm{YX}}(\omega, t) = \gamma \ \mathbf{S}_{\mathrm{YX}}(\omega, t-T) + (1-\gamma) \ \mathbf{Y}(\omega, t) \cdot \mathbf{X}^H(\omega, t) \tag{7}$$

## 3. GENERALIZED SIDELOBE DECORRELATOR

Both the GSC and BSS approaches each have their strengths and weaknesses. While the GSC exploits the available prior geometric information, it does not exploit the independence prior and is thus subject to the leakage problem associated with power minimization. That is, any leakage of the primary source into the secondary beams will result in cancellation of the primary source and a degradation of the SNR improvement. This leakage can be due to any of several factors, including: (1) array calibration errors; (2) primary source location error; (3) a main beam that is narrower than the primary source, caused by a large array aperture; (4) spatial aliasing lobes, caused by an insufficiently spaced sensor array; (5) reverberation, caused by reflections of the primary source coming from directions outside the primary beam.

BSS, on the other hand, can clearly separate in the presence of leakage but does not exploit all available prior information. Also, when there are more sensors than sources, the separation problem is highly underdetermined. Furthermore, when the separation is performed in the frequency domain, the permutation and scaling problem exists at every frequency band.

To overcome these deficiencies, we combine aspects of the generalized sidelobe canceller and blind source separation to create an algorithm we call the generalized sidelobe decorrelator (GSD), shown in Fig. 1b. Like the GSC, it consists of steering delays that place all the sensor signals in-phase, a linear combiner that forms the primary beam, and a blocking matrix that forms secondary orthogonal beams. However, unlike the GSC, instead of adopting a power minimization criteria that adapts the secondary beams out of the primary beam, we adopt a *cross-power* minimization criteria, as described in Section 2.2, that *decorrelates* the secondary beams from the primary beam. This allows for removing leakage of the primary source into the secondary beams, while the blocking matrix guarantees the integrity of the primary beam independent of whether the sources are active.

## 4. EXPERIMENT

We conducted an acoustic experiment designed to demonstrate the superior performance of the algorithm for noise reduction. A 2-D rectangular sensor array of dimension 10 cm x 7 cm was formed, corresponding to the dimensions of a personal digital assistant (PDA), using inexpensive omnidirectional lapel microphones (Audio-Technica ATR35S).

The array was located in a room of dimension 3.0 m x 3.6 m x 2.3 m. A loudspeaker was placed 0.5 m directly in front of the array, which was used to replay a quiet recording of a male speaking 300 short commands over a period of twenty minutes, with a pause of ~2-3 seconds between commands. The recording was automatically segmented into speech/non-speech for the purpose of measuring signal to noise ratio (SNR), and the speaker had previously trained an automatic speech recognition system for the purpose of measuring speaker-dependent command error rate (CER). The recognizer and all algorithms operated at 11.025 kHz.

Also in the room but in the corner and facing the wall ~2.5 m from the array, a loudspeaker played babble (the sounds of many voices). Outside the room, another loudspeaker played a recording of street noises. The nominal SNR at the microphones was 1.2 dB, which corresponded to a CER of 77.6%. We then applied four on-line adaptive algorithms to the array signals, the results of which are shown in Table 1.

Table 1. *Results of real-room experiment.*

| Algorithm | SNR | CER |
|---|---|---|
| none | 1.2 dB | 77.6% |
| fixed delay-sum beam | 1.3 dB | 19.4% |
| generalized sidelobe canceller | 3.0 dB | 73.9% |
| blind source separation | 3.6 dB | 100.0% |
| generalized sidelobe decorrelator | 4.6 dB | 5.4% |

Because the source was directly in front of the array, the fixed delay-sum beam could be obtained by a simple averaging of the four sensors. Although the fixed beam does not provide much SNR improvement, it does provide significant CER improvement. This is partly because it does not distort the speech at all.

Next, we implemented the GSC using a “Walsh” blocking matrix (see [6]) to form three secondary beams orthogonal to the primary delay-sum beam. The secondary beams were adapted out of the primary beam using the frequency domain LMS algorithm with filter sizes of 512 taps. Although there is improvement in the SNR, there is degradation in the CER relative to the delay-sum beam, most likely due to spectral distortion of the speech.

Next, we applied BSS on the 4 raw inputs signals, using the algorithm of Section 2.2, using 2 outputs and filter sizes of 512 taps. Although BSS provides a small SNR improvement over GSC, the algorithm completely destroys the recognition performance. Part of the problem is that BSS requires that the sources be simultaneously active, and thus the filters start to degrade during the silent periods between commands. In addition, the frequency permutation problem can distort the speech spectrum.

Finally, we applied our new hybrid GSD by performing BSS on the fixed delay-sum beam and blocking matrix outputs using filter sizes of 512 taps, and obtained very encouraging results. In addition to obtaining the largest SNR improvement of any of the algorithms, the CER was a very respectable 5.4%, approaching the single microphone CER of 2.0% in a quiet environment.

## 5. CONCLUSIONS

We introduced a new algorithm called the generalized sidelobe decorrelator that combines elements of geometric beamforming and blind source separation. It solves the leakage problem of the generalized sidelobe canceller and allows for recovering other sources in addition to the one that lies in the known direction. It also allows blind source separation to run continuously independent of whether the sources are simultaneously active.

Although the BSS algorithm we have used is based on decorrelation of nonstationary signals, we could equally have used higher-order statistics in the separation criteria. Finally, instead of relying on prior knowledge of the primary source location, we could use a direction of arrival (DOA) finding algorithm to locate the strongest source. We shall consider all these in future papers.

## 6. ACKNOWLEDGEMENTS

We would like to thank Uday Jain for training the recognizer.

## 7. REFERENCES

[1] S. Roberts and R. Everson (eds.), *Independent component analysis: principles and practice*, Cambridge Univ. Press, Cambridge, U.K., 2001.

[2] S. Haykin (ed.), *Unsupervised adaptive filtering, vol. 1: blind source separation*, Wiley & Sons, New York, 2000.

[3] S. Haykin (ed.), *Array signal processing*, Prentice Hall, Englewood Cliffs, NJ, 1985.

[4] D. Johnson and D. Dudgeon, *Array signal processing: concepts and techniques*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[5] O.L. Frost, An algorithm for linearly constrained adaptive array processing, *Proc. IEEE*, v. 60, no. 8, pp. 926-935, 1972.

[6] L.J. Griffiths and C.W. Jim, An alternative approach to linearly constrained adaptive beamforming, *IEEE Trans. Antennas Propagation*, vol. AP-30, no. 1, pp. 27-34, 1982.

[7] L. Parra and C. Alvino, Geometric source separation: merging convolutive source separation with geometric beamforming, submitted to *IEEE Trans. Speech Audio Processing*.

[8] C. Fancourt and L. Parra, The coherence function in blind separation of convolutive mixtures of nonstationary signals, accepted at *Neural Networks for Signal Processing*, 2001.

[9] L. Parra, C. Spence, and B. De Vries, Convolutive blind source separation based on multiple decorrelation, in *IEEE Workshop on Neural Networks for Signal Processing*, Cambridge, UK, 1998.

[10] L. Parra and C. Spence, Convolutive blind separation of nonstationary sources, *IEEE Trans. Speech Audio Processing*, vol. 8, no. 3, pp. 320-327, 2000.

[11] L. Parra and C. Spence, On-line convolutive blind source separation of non-stationary signals, *Journal of VLSI Signal Processing*, vol. 26, pp. 39-46, 2000.