

On the FIR Inversion of an Acoustical Convolutive Mixing System: Properties and Limitations

Markus Hofbauer

Swiss Federal Institute of Technology, Zürich, Switzerland
hofbauer@isi.ee.ethz.ch

Abstract. In this paper we address the problem of Least-Squares (LS) optimal FIR inverse-filtering of an convolutive mixing system, given by a set of acoustic impulse responses (AIRs). The optimal filter is given by the LS-solution of a block-Toeplitz matrix equation, or equivalently by the time-domain Multi-Channel Wiener Filter. A condition for the minimum FIR filter length can be derived, depending on the number of sensors and sources and the AIR length, such that an exact FIR inverse exists, which perfectly separates and deconvolves all sources. In the general case, where an exact FIR solution does not exist, we discuss how SDR, SIR and SNR gains can be traded against each other. Results are shown for a set of AIRs, measured in an typical office room. Furthermore we present a method, which allows a time-domain shaping of the envelope of the global transfer function, reducing pre-echoes and reverberation.

1 Introduction

Blind Source-Separation methods aim at inverting a convolutive mixing system, using linear de-mixing filters. Examining the non-blind case, where the convolutive system of AIRs is perfectly known and determining the optimal FIR inverse (with respect to a quadratic cost function) demonstrates in principle the degree of achievable separation and deconvolution. The optimal FIR inverse is given by a block Toeplitz matrix equation. We derive a condition for the minimum FIR filter length, such that an exact FIR inverse exists, and also conditions for obtaining exact separation only, or exact deconvolution only. By appropriate weighting of the sources, the space of possible solutions, favoring the SIR, the SDR or SNR gain, can be sampled. Results are shown for a 4×4 set of AIRs, measured in an typical office room. In order to reduce reverberation, it is desirable to have an influence on the shape of the time-domain envelope of the global transfer function: we present a method using a weighting function.

1.1 Problem formulation

In a reverberant and noisy environment with M mutually uncorrelated point sources s_m the i th sensor of an array with K sensors receives the signal $x_i[k]$

$$x_i[k] = \sum_{m=1}^M (h_{im} * s_m) \Big|_k + v_i[k] \quad i = 1 \dots K, \quad (1)$$

where h_{im} is the acoustic impulse response (AIR) from source s_m , having n_h coefficients. The spatially incoherent sources and sensor noise appear as spatially uncorrelated components $v_i[k]$ at the i th sensor and are referred to as noise in the following.

We aim at finding a set of K FIR filters w_{mi} , $i = \{1, \dots, K\}$ with n_w coefficients, such that output $y_m[k]$ is an estimate $\hat{s}_m[k-d]$ of the delayed source s_m .

$$y_m[k] = \hat{s}_m[k-d] = \sum_{i=1}^K (w_{mi} * x_i) \Big|_k. \quad (2)$$

Each output y_m has its own set of filters w_{mi} , extracting one of the sources s_m , and thus can be addressed separately. In the following we consider only one of the outputs, extracting the target source $s_{\tilde{m}}$ ($m = \tilde{m}$). Perfect deconvolution of the target source $s_{\tilde{m}}$ and separation from the other jammer sources s_m is obtained, if

$$\sum_{i=1}^K (h_{im} * w_i) \Big|_k = t_m[k] \quad m = 1 \dots M, \quad k = 1 \dots n_h + n_w - 1 \quad (3)$$

where $t_m[k]$ is the total (global) response from the source s_m to output y and is chosen as $t_{\tilde{m}}[k] = \delta[k-d]$ for the target source $s_{\tilde{m}}$, and $t_m[k] = 0$ for the other sources s_m . The minimum possible total delay $d = d_{\min}$ corresponds to the propagation time from the source $s_{\tilde{m}}$ to the sensors.

Since we are looking for a signal independent inverse of the AIR set, which is not optimized on the individual source spectras, we assume all sources to be white with equal power. The source powers are however weighted properly according to the desired inversion task, e.g. according to their actual powers $\sigma_{s_m}^2$.

2 Optimal FIR inverse filters

2.1 Exact FIR inverses

The equation system (3), defining the optimal FIR inverse filters w_i is illustrated in Fig. 1. Each source/sensor adds a row/column of Toeplitz blocks to the block Toeplitz matrix $\bar{\mathbf{H}}$, which has dimension $(M(n_w + n_h - 1) \times Kn_w)$.

A plausible and common assumption [1],[2] for the AIR system is, that the filters h_{im} mutually do not share common zeros in the frequency domain and that the corresponding Matrix $\bar{\mathbf{H}}$ has full rank. Then there exists a set of FIR filters w_i , which fulfills (3) for the desired total transfer functions t_m , if Matrix $\bar{\mathbf{H}}$ is square or if $\dim(\bar{\mathbf{H}}) = (D_1 \times D_2)$, $D_1 \leq D_2$. This case applies if

$$M(n_w + n_h - 1) \leq Kn_w. \quad (4)$$

Solving (4) for n_w gives a condition for the minimum number of filter coefficients of w_i ,

$$n_w \geq \left\lceil \frac{(n_h - 1)M}{K - M} \right\rceil \quad K > M. \quad (5)$$

Thus, if there are less sources than sensors, there exist FIR filters w_i which achieve perfect separation and deconvolution of the sources. A proof for a similar condition for the case of only one source ($M = 1$) can be found in [3].

If for instance there is one more sensor than sources ($K = M + 1$), then choosing $n_w \geq M(n_h - 1)$ will allow perfect separation and deconvolution. If $K = aM$, $a \in \mathbb{Z}^+$, then $n_w \geq (n_h - 1)/(a - 1)$ coefficients are required, i.e. filters w_i may be even shorter than the AIRs in this case.

The diagram illustrates the equation system (3) defining the FIR inverse filters w_i of a $K \times M$ AIR system. It shows a grid of equations for M sources and K sensors. Each equation is scaled by a factor $\gamma(m, k)$. The equations are arranged in a grid where rows represent sources and columns represent sensors. The equations are: $\gamma(1,1)^\circ (h_{11}[1] w_1[1] + h_{11}[n_h] w_1[n_w]) = t_1[1]$; $\gamma(1, n_h)^\circ (h_{11}[1] w_1[1] + h_{11}[n_h] w_1[n_w]) = t_1[n_h + n_w - 1]$; $\gamma(2,1)^\circ (h_{12}[1] w_2[1] + h_{12}[n_h] w_2[n_w]) = t_2[1]$; $\gamma(2, n_h)^\circ (h_{12}[1] w_2[1] + h_{12}[n_h] w_2[n_w]) = t_2[n_h + n_w - 1]$; ... $\gamma(M,1)^\circ (h_{1M}[1] w_M[1] + h_{1M}[n_h] w_M[n_w]) = t_M[1]$; $\gamma(M, n_h)^\circ (h_{1M}[1] w_M[1] + h_{1M}[n_h] w_M[n_w]) = t_M[n_h + n_w - 1]$. The overall equation is $\tilde{\mathbf{H}} \tilde{\mathbf{w}} = \tilde{\mathbf{t}}$.

Fig. 1. Equation system (3) defining the FIR inverse filters w_i of an $K \times M$ AIR system (h_{im}) for the given desired total transfer functions t_m (\circledast denotes a scaling of an equation by $\gamma(m, k)$); n_w and n_h : filter lengths; length of t_m : $n_t = n_h + n_w - 1$

2.2 Least squares optimal FIR inverses

If the number of sources is equal or larger than the number of sensors then the optimal FIR inverse filters are given by the Least Squares solution of (3). This case ($K \leq M$) applies at any rate, if also the noise sources at the sensors are included in the LS-optimization since noise at the i th sensor can be regarded as a source ‘sitting’ at the sensor (with AIRs $h_{ii} = 1$ and $h_{ij} = 0$). Thus, when considering noise, K additional noise sources have to be placed in (3).

In the LS case, maximal separation (SIR), deconvolution (SDR) and a maximal SNR gain (see 2.4) can not be achieved simultaneously. However the SIR, SDR and SNR gains can be traded against each other. By weighting the sources with appropriate factors $\gamma(m, k)$

$$\gamma(m, k) = \gamma_m^\circ \gamma_m(k) \quad k = 1 \dots n_t = n_h + n_w - 1 \quad m = 1 \dots M, \quad (6)$$

the space of possible solutions, favoring the SIR, SDR or SNR gain, can be sampled. As shown in Fig. 1, both sides of equation (3) are scaled with $\gamma(m, k)$. While γ_m° is constant for each source (e.g. the source power $\sigma_{s_m}^2$), $\gamma_m(k)$ allows for a individual weighting of each equation. Choosing larger weights γ_m° for the noise sources will result in a solution favoring the SNR gain, larger weights for the target source will increase the SDR, and larger weights for the jammer sources will increase the SIR.

Maximal separation (best SIR) Maximal separation is achieved, if the jammer sources receive dominant weights γ_m° , and by choosing $t_m[k] = 0$ for all $m \neq \tilde{m}$. For the target source $s_{\tilde{m}}$ a total response $t_{\tilde{m}}[k]$ will result, which will introduce some additional distortion. A constraint, e.g. $t_{\tilde{m}}[d] = 1$ is required to avoid the all zero solution $w_i = 0$. In the *noiseless* case **perfect separation** (however *with distortion*) is possible, if

$$n_w \geq \left\lceil \frac{n_h(M-1) - M + 2}{K - M + 1} \right\rceil \quad K \geq M, \quad (7)$$

since $\dim(\tilde{\mathbf{H}}) = (D_1 \times D_2)$, $D_1 \leq D_2$ applies, if (7) holds. (Compared to (4) there are $n_w + n_h - 2$ fewer equations, since here $\gamma_{\tilde{m}}^\circ / \gamma_m^\circ \approx 0$ and $t_{\tilde{m}}[d] = 1$). If e.g. $K = aM$ (over-determined case), perfect separation is achieved with filters w_i shorter than the AIRs. In the square case ($K = M$) perfect separation is achievable, if $n_w \geq n_h(M-1) - M + 2$. With the definition of Matrix $\hat{\mathbf{H}}^{K \times K} = [h_{im}]$ and since $\text{adj}(\hat{\mathbf{H}})\hat{\mathbf{H}} = \det(\hat{\mathbf{H}})\mathbf{I}$ [4], it follows, that $\text{adj}(\hat{\mathbf{H}})$ is the perfect separating solution for the $K = M$ case. Filters w_{mi} are then given by $w_{mi} = \hat{\mathbf{H}}_{im}$, with $\hat{\mathbf{H}}_{im}$ being the cofactors of $\hat{\mathbf{H}}$. Thus w_{mi} will be a sum of different combinations of $K-1$ convolutions of h_{im} (e.g. $h_{12} * h_{23} * \dots$). The corresponding total target transfer function is causal if desired and is given by $t_{\tilde{m}} = \det(\hat{\mathbf{H}})$, which is a sum of combinations of K convolutions of h_{im} . Since each convolution with an AIR introduces additional reverberation, the resulting distortion can be significant.

Perfect separation is not possible in the under-determined case $K < M$. Also, when including the noise sources in (3), the LS-solution will reduce the best possible SIR gain, depending on the weighting of jammer and noise sources.

Maximal deconvolution (best SDR) If a maximal deconvolution (minimal distortion, high SDR) of the target source is desired, all sources, except for the target, have to receive small or even zero weighting γ_m° . From (4) follows, that when neglecting all unwanted sources ($M = 1$), **perfect deconvolution** is in principal achievable, if at least two sensors are available $K \geq 2$ (with $n_w \geq n_h - 1$). Complete deconvolution may however result in a large SNR and SIR loss, especially if the delay d is chosen small [5]. Multi-channel deconvolution is caused by two mechanisms: single-channel type *inversion* (in general IIR) and *elimination by addition* of different sensor observations (possible with FIR). The latter is mainly responsible for possible SNR and SIR losses.

A common criterion used by BSS source separation algorithms is the **Minimal Distortion Principle** [6], where separation of the sources is aimed for, with the constraint, that no additional distortion (on top of the distortion/reverberation caused by the AIRs) should be introduced: the target source at the output is ought to be identical or close to its observation in one of the sensors, i.e. $y_{s_{\tilde{m}}} \approx s_{\tilde{m}} * h_{i\tilde{m}}$. The corresponding solution is obtained by setting $t_{\tilde{m}} = h_{i\tilde{m}}$ as desired total response in (3). Note that the source is not completely deconvolved in this case.

The **Minimum Variance Distortionless Response** Beamformer (MVDR-BF) [7] aims at complete deconvolution of the target source ($t_{\tilde{m}}[k] = \delta[k-d]$), while minimizing the variance of all undesired sources. Generalized from a model

with simple propagation delays to general AIRs, the MVDR-BF filters w_i are calculated in the frequency domain [5] and are typically IIR. The MVDR-BF solution is obtained from (3) (approximately) by using a large n_w and dominant weights γ_m° for the target source.

Maximal SNR gain (best SNR) The solution leading to a maximal SNR gain is obtained by setting the weights γ_m° of all sources, except for the noise sources to small values or zero. A constraint, e.g. $t_{\tilde{m}}[d] = 1$ is required to avoid the all zero solution $w_i = 0$. The maximal SNR achieving solution is given by $w_i[\cdot] = h_{i\tilde{m}}[-\cdot]$, which follows from the Wiener solution (9) and (13), setting $\mathbf{R}_{\mathbf{xx}} = \mathbf{I}$. If the AIRs are simple delays, $w_i[\cdot]$ will reduce to the well known Delay-and-Sum Beamformer, which is the BF producing the highest white noise SNR gain [7].

2.3 Time-Domain Multi-Channel Wiener Filter (MCWF)

If all sources are white and are given the weighting $\gamma(m, k) = \gamma_m^\circ$, the least squares solution of (3) coincides with the time-domain FIR Multi-Channel Wiener Filter MCWF, which is the MMSE solution. Defining the stacked data and filter vectors

$$\begin{aligned}\mathbf{x}_i[k] &= [x_i[k] x_i[k-1] \cdots x_i[k-n_w+1]]^T \\ \mathbf{x}[k] &= [\mathbf{x}_1^T \mathbf{x}_2^T \cdots \mathbf{x}_K^T]^T \\ \mathbf{w}_i &= [w_i[0] w_i[1] \cdots w_i[n_w-1]]^T \\ \mathbf{w} &= [\mathbf{w}_1^T \mathbf{w}_2^T \cdots \mathbf{w}_K^T]^T,\end{aligned}\quad (8)$$

the Multi-Channel Wiener Filter (with K inputs and one output) is given by

$$\mathbf{w}_{\text{MCWF}} = \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{r}_{\mathbf{x}s_{\tilde{m}}}, \quad (9)$$

where $\mathbf{R}_{\mathbf{xx}}^{(Kn_w \times Kn_w)}$ is the autocorrelation matrix of the sensor signals

$$\mathbf{R}_{\mathbf{xx}} = \mathbb{E}\{\mathbf{x}[k]\mathbf{x}[k]^T\} = \sum_{m=1}^M \gamma_m^{\circ 2} \mathbf{R}_{\mathbf{x}_{s_m} \mathbf{x}_{s_m}} + \gamma_v^{\circ 2} \mathbf{R}_{\mathbf{vv}}, \quad (10)$$

and $\mathbf{r}_{\mathbf{x}s_{\tilde{m}}}^{(Kn_w \times 1)}$ is the cross-correlation vector of \mathbf{x} and $s_{\tilde{m}}$

$$\mathbf{r}_{\mathbf{x}s_{\tilde{m}}} = \gamma_{\tilde{m}}^{\circ 2} \mathbb{E}\{\mathbf{x}[k] s_{\tilde{m}}[k-d]\}. \quad (11)$$

For a given set of AIRs h_{im} , element $(\mathbf{R}_{\mathbf{x}_i \mathbf{x}_j})_{a,b}$ amounts to

$$\begin{aligned}(\mathbf{R}_{\mathbf{x}_i \mathbf{x}_j})_{a,b} &= r_{x_i x_j}[a-b] = \\ &= \sum_{m=1}^M \gamma_m^{\circ 2} (h_{im}[-\cdot] * h_{jm}[\cdot] * r_{s_m}[\cdot]) \Big|_{k'=a-b} + \gamma_v^{\circ 2} (\mathbf{R}_{\mathbf{v}_i \mathbf{v}_j})_{a,b},\end{aligned}\quad (12)$$

where $h_{im}[-\cdot]$ denotes time reversion. Element $\mathbf{r}_{\mathbf{x}_i s_{\tilde{m}}}(a)$ is given by

$$\mathbf{r}_{\mathbf{x}_i s_{\tilde{m}}}(a) = r_{x_i s_{\tilde{m}}}[a-1-d] = \gamma_m^{\circ 2} (h_{i\tilde{m}}[-\cdot] * r_{s_{\tilde{m}}}[\cdot]) \Big|_{k'=a-1-d}. \quad (13)$$

With (12) and (13), the MCWF (9) can be determined. $\mathbf{R}_{\mathbf{xx}}^{(Kn_w \times Kn_w)}$ and $\bar{\mathbf{H}}$ are matrices of *Toeplitz* blocks of typically very large dimension (e.g. 40000×40000). An algorithm, utilizing the block-Toeplitz structure can be used to solve (9) or (3) efficiently (Schur alg.).

2.4 SDR gain, SIR gain and SNR gain

The realized degree of deconvolution, separation and noise reduction is measured by the Signal-to-Distortion Ratio SDR, the Signal-to-Interference Ratio SIR and the SNR, respectively. Since we assume sources and the noise to be white, the SDR of source $s_{\tilde{m}}$ at sensor 1 is calculated from $h_{1\tilde{m}}$ as

$$\text{SDR}_{x_1} := 10 \log_{10} \left(\frac{\max(|h_{1\tilde{m}}[k]|^2)}{\sum_k |h_{1\tilde{m}}[k]|^2 - \max(|h_{1\tilde{m}}[k]|^2)} \right), \quad (14)$$

i.e. the SDR is the ratio of the power of the main peak to the reverberation part in the AIR $h_{1\tilde{m}}$. The SDR at the output y we obtain from the total response $t_{\tilde{m}}[k]$:

$$\text{SDR}_y := 10 \log_{10} \left(\frac{\max(|t_{\tilde{m}}[k]|^2)}{\sum_k |t_{\tilde{m}}[k]|^2 - \max(|t_{\tilde{m}}[k]|^2)} \right). \quad (15)$$

The reduction of distortion or reverberation is then $\text{SDR}_{\text{gain}} = \text{SDR}_y - \text{SDR}_{x_1}$. Similarly the $\text{SIR}_{\text{gain}} = \text{SIR}_y - \text{SIR}_{x_1}$ of the power of the target source $s_{\tilde{m}}$ to all jammer sources s_m is given by:

$$\text{SIR}_{\text{gain}} := 10 \log_{10} \left(\frac{\sum_k |t_{\tilde{m}}[k]|^2 \cdot \sum_{k,m \neq \tilde{m}} |h_{1m}[k]|^2}{\sum_{k,m \neq \tilde{m}} |t_m[k]|^2 \cdot \sum_k |h_{1\tilde{m}}[k]|^2} \right). \quad (16)$$

Finally, the white noise gain $\text{SNR}_{\text{gain}} = \text{SNR}_y - \text{SNR}_{x_1}$ is obtained by:

$$\text{SNR}_{\text{gain}} := 10 \log_{10} \left(\frac{\sum_k |t_{\tilde{m}}[k]|^2}{\sum_{i,k} |w_i[k]|^2 \cdot \sum_k |h_{1\tilde{m}}[k]|^2} \right). \quad (17)$$

3 Shaping the envelope of the total response

The LS-solution of (3) will result in some signal distortion, appearing in $t_{\tilde{m}}$. AIRs and filters w_i typically have several thousands of coefficients and thus will produce a long, noncausal global response $t_{\tilde{m}}$ with its main power concentrated at $k = d$ and slow decaying tails. These tails cause audible undesired artifacts: pre-echoes and (late) reverberation. It is therefore desirable to shape the envelope of $t_{\tilde{m}}$, such that the tails decay faster. We propose a method to shape the envelope of the total response $t_{\tilde{m}}$ by incorporating an appropriate weighting function $\gamma_{\tilde{m}}(k)$ in (3), scaling the equations associated with the target source:

$$\gamma_{\tilde{m}}(k) = (1 + \epsilon) - \exp(-\tau[k - d]^2) + \beta\delta[k - d] \quad \epsilon, \tau \approx 0, \quad \beta \approx 1. \quad (18)$$

This will cause the LS-solution to drive the tails of $t_{\tilde{m}}[\cdot]$ to zero, while permitting a degree of freedom for ‘ k around d ’, which is favorable for the LS-optimization. The effect of the total response shaping is shown in Fig. 2 f).

4 SDR, SIR and SNR gains for a typical office size room

We have measured a $K \times M = 4 \times 4$ set of AIRs ($n_h = 3600$) in an office room (5m \times 3.5m \times 2.5m, $T_{60} \approx 400$ ms, $f_s = 8$ kHz), with a speaker microphone distance range of [1m-3m] and a sensor array spacing of [4cm 14cm 4cm]. Results of the inversion are shown in Fig. 2. The total transfer function $t_{\tilde{m}}$ of the target

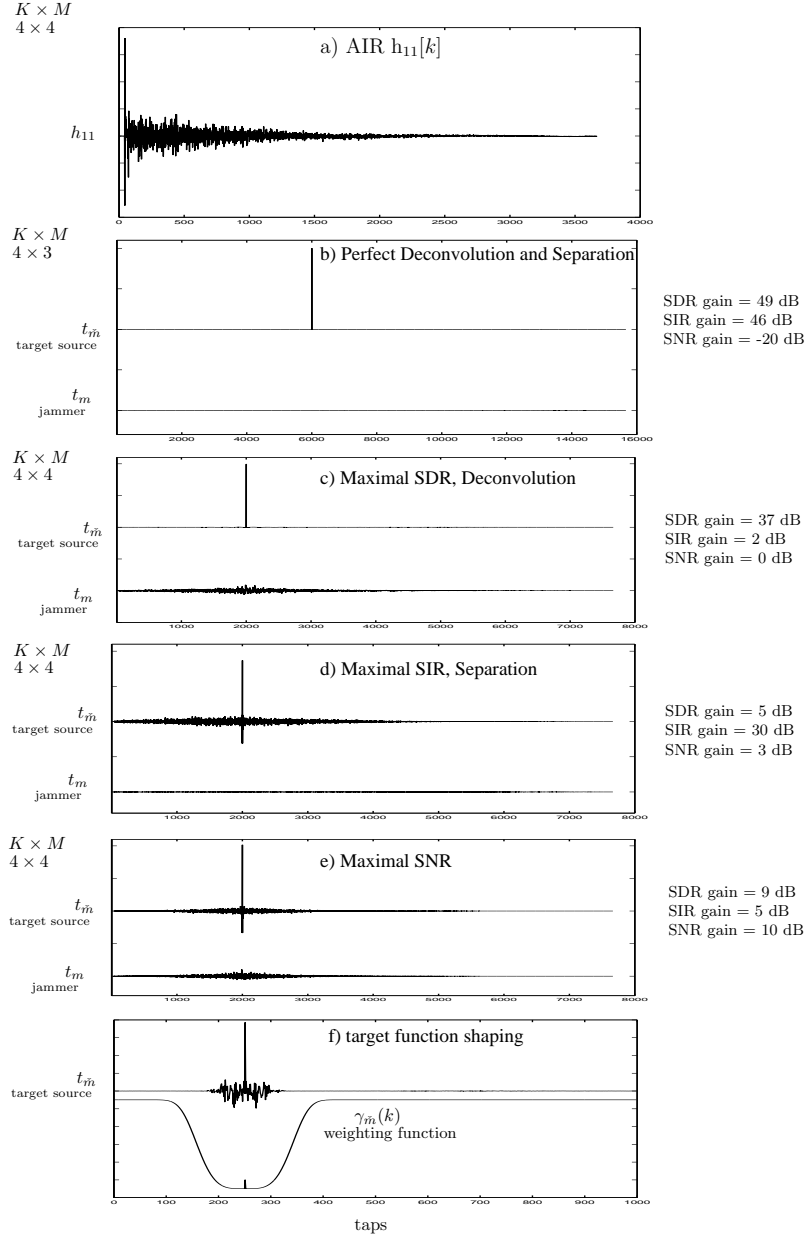


Fig. 2. FIR LS-inverse of an $K \times M = 4 \times 4$ AIR set ($n_h = 3600$, $T_{60} \approx 400\text{ms}$) of an office room: total transfer functions of target $t_{\bar{m}}$, and t_m of one of the jammers; **a)** AIR h_{11} ; **b)** perfect separation and deconvolution for the 4×3 case; **c)-e)** solutions favoring the SDR, SIR or SNR gain, obtained by different weighting γ_m^o of the sources; **f)** effect of introducing the weighting function $\gamma_{\bar{m}}(k)$: tails of $t_{\bar{m}}$ are suppressed, reducing late reverberation (a simulated AIR set with $n_h = 500$ was used for case f))

source, and exemplarily t_m of one of the jammers – the other jammers are comparable – are depicted. By applying a different weights γ_m° to the sources, the LS-solution of (3) favors the SDR, SIR or SNR gain. The achievable gains are indicated. **Case b)**: $K \times M = 4 \times 3$: in accordance with (5), which here demands $n_w \geq 10797$, (nearly) perfect deconvolution and separation is accomplished, since $n_w = 12000$ was chosen. The sensor noise sources were neglected here. **Case c)-e)**: $K \times M = 4 \times 4$, and noise being also considered (i.e. M is in fact increased by $K = 4$ noise sources to a total of $M=8$): now the Least Squares optimization applies, and the SDR, SIR, and SNR can be traded against each other, with maximal values of the SDR/SIR/SNR $\text{gain}_{\max} = 37/30/10\text{dB}$. Plot **f)** shows the effect of shaping $t_{\bar{m}}$ by incorporating the weighting function $\gamma_{\bar{m}}(k)$ given by (18). It can be observed, that the tails of $t_{\bar{m}}[k]$ vanish, at the cost of increased values ‘around $k = d$ ’. The weighting function $\gamma_{\bar{m}}(k)$ drives the tails of $t_{\bar{m}}[k]$ to zero, reducing late reverberation and also the artifacts due to the non-causal side tail.

5 Conclusion

AIRs are typically non-minimum phase, having non-causal and very long-tailed IIR inverses. Nevertheless, under certain conditions, there exist causal FIR inverses which perfectly invert the AIR mixing system. If $K > M$, perfect separation and deconvolution is achievable, while separation only requires $K \geq M$. Deconvolution only, demands $K \geq 2$. We derived conditions for the corresponding minimum FIR filter lengths n_w . For the most common case ($K \leq M$), the LS-optimization will give a tradeoff between the SDR, SIR and SNR gains, which can be controlled by appropriate weighting of the sources. Confirmative results are shown for a 4×4 set of AIRs, measured in a typical office room. Finding the LS-optimal filters for the non-blind case demonstrates what one can expect from a BSS-algorithm in a similar acoustical setup. In order to additionally reduce reverberation and artifacts we proposed a method using a weighting function $\gamma_m(k)$, which allows a time-domain shaping of the envelope of the global transfer function t_m , driving the undesired long tails to zero.

References

1. M. Miyoshi, Y. Kaneda: “Inverse filtering of room acoustics”. IEEE T. Acoust. Speech and Signal Processing. (1988)
2. K. Rahbar, J.P. Reilly, J. H. Manton: “Blind identification of MIMO FIR systems driven by quasi-stationary sources using second order statistics”: A frequency domain approach. IEEE T. Signal Processing. (2002)
3. G. Harikumar, Y. Bresler: “FIR Perfect Signal reconstruction from multiple convolutions: Minimum Deconvolver Orders”. IEEE T. Signal Processing. (1998)
4. R. A. Horn, Ch. R. Johnson: “Matrix Analysis”. Cambridge University Press. (1999)
5. M. Hofbauer, H.-A. Loeliger: “Limitations of FIR Multi-Microphone Speech Derivation in the Low-Delay Case”. IWAENC. Kyoto. (2003)
6. K. Matsuoka: “Independent Component Analysis and Its Application to Sound Signal Separation”. IWAENC. Kyoto. (2003)
7. M. Brandstein, D. Ward: “Microphone Arrays”. Springer. (2001)