CCIT Report #476 March 2004

Asymptotic Performance Analysis of QML Multichannel Blind Deconvolution

Alexander Bronstein[1], Michael Bronstein[2], Michael Zibulevsky[1], and Yehoshua Y. Zeevi[1]

[1]Technion - Israel Institute of Technology, Department of Electrical Engineering, 32000 Haifa, Israel.
[2]Technion - Israel Institute of Technology, Department of Computer Science, 32000 Haifa, Israel.

March 18, 2004

Abstract

We derive simple closed-form expressions for the asymptotic estimation error in multichannel quasi maximum likelihood (QML) blind deconvolution. The asymptotic performance bounds coincide with the Cramér-Rao bounds, when the true ML estimator is used. Special cases of single-channel blind deconvolution and instantaneous blind source separation are highlighted. Conditions for asymptotic stability of the QML estimator are derived.

Notation

A tensor notation is adopted in this work. Signals are represented as tensors and the Einstein summation convention is used, according to which repeated indices, one lower and one upper, are implicitly summed over:

$$a_i^{\ j} b_{jk} = \sum_j a_{ij} b_{jk}.$$

Multi-channel time signals are denoted as second-rank tensors, x_{it}, where i stands for the channel number and t is a discrete time index. Cross-talk kernels are expressed as third-rank tensors, $h_{ij\tau}$, denoting cross-talk from channel j to channel i at time τ.

Definition 1. *The discrete Fourier transform of $h_{ij\tau}$ is defined by*

$$\bar{h}_{ij}(\theta) = \mathcal{F}\left\{h_{ij\tau}\right\}(\theta) = h_{ij}^{\ \ \tau} e_\tau(\theta),$$

where

$$e_\tau(\theta) \quad = \quad \exp\{-\imath\tau\theta\}\,.$$

The Fourier transform is always calculated w.r.t. the time index.

Definition 2. *Convolutive mixture between two cross-talk kernels $a_{ij\tau}$ and $b_{ij\tau}$ is defined as*

$$c_{ij\tau} \quad = \quad a_{ij\tau} * b_{ij\tau} = a_i^{\;k\tau'} b_{kj,\tau-\tau'},$$

or, in the Fourier transform domain,

$$\bar{c}_{ij}(\theta) \quad = \quad \bar{a}_i^{\;k}(\theta)\cdot\bar{b}_{kj}(\theta).$$

Definition 3. *Standard (non-mixing) convolution between two kernels $a_{ij\tau}$ and $b_{ij\tau}$ is defined as*

$$c_{ij\tau} \quad = \quad a_{ij\tau} \star b_{ij\tau} = a_{ik}^{\;\;\tau'} b_{kj,\tau-\tau'},$$

or, in the Fourier transform domain,

$$\bar{c}_{ij}(\theta) \quad = \quad \bar{a}_{ik}(\theta)\cdot\bar{b}_{kj}(\theta).$$

Definition 4. *Let $h_{ij\tau}$ be a kernel and let $h^{-1}_{ij\tau}$ be such a kernel that*

$$h^{-1}_{ij\tau} * h_{ij\tau} \quad = \quad \delta_{ij}\delta_\tau.$$

Then, $h^{-1}_{ij\tau}$ is termed the inverse kernel of $h_{ij\tau}$, and $h_{ij\tau}$ is termed invertible.

In the Fourier domain,

$$\bar{h}^{-1}_{ik}(\theta)\cdot\bar{h}^{k}_{\;j}(\theta) \quad = \quad \delta_{ij},$$

i.e., the inverse kernel $\bar{h}^{-1}_{ik}(\theta)$ is the matrix inverse of $\bar{h}_{ik}(\theta)$, for every θ.

1 Introduction

Multichannel blind deconvolution arises in various applications related to acoustics, optics, medical imaging, geophysics, communications, control, etc. In comunications, the term *blind channel equalization* is more common, as the main interest lies in retrieving the data transmitted over a dispersive communication channel [20, 5, 14, 24]. In control, blind deconvolution is usually known as *blind identification*, since the main interest lies in obtaining a model of the system [23, 25, 18], whereas in acoustics, optics and geophysics the term *blind deconvolution* is more adequate, since the goal is to ”undo” the influence of a convolution system by finding its stable inverse.

In the noiseless setup of multichannel blind deconvolution, the observed sensor signals x_{it} ($i = 1, ..., N$) are created from the *source signals* s_{it} passing through a convolutive mixing system, described by a matrix of input responses, $w_{ij\tau}$, sometimes termed as the *crosstalk matrix*:

$$x_{it} = w_{ijt} * s_{it}.$$

The diagonal elements of $w_{ij\tau}$ are responsible for convolution of each channel, whereas the off-diagonal elements result in cross-talk between the channels. The setup is termed *blind* if only x are accessible, whereas no knowledge on w and s is available.

Blind deconvolution attempts to find such a deconvolution (or restoration) kernel $h_{ij\tau}$, that produces a an estimate of s, up to a possible delay and permutation:

$$\hat{s}_{it} = h_{ijt} * x_{it} \approx c_i \cdot s_{\pi_i, t-\Delta_i},$$

where c_i are scaling factors, Δ_i is an integer shift and π_i is a permutation of $\{1, ..., N\}$. Equivalently, the *global system response* should be approximately an identity, up to scale factor, shift and permutation:

$$h_{ij\tau} * w_{ij\tau} \approx c_i \pi_{ij} \delta_{\tau-\Delta_i},$$

where π_{ij} is a perturbed identity. A commonly used assumption is that s are non-Gaussian.

Asymptotic performance of ML parameter estimation in blind system identification and deconvolution problems was addressed in many previous studies (see, for example, [4, 21, 12, 22, 27, 15, 26]). In all these studies, the Cramér-Rao lower bound (CRLB) for the system parameters are found, and lower bounds on signal reconstruction quality are derived. However, sometimes the true source distribution is either unknown, or not suitable for optimization, which makes the use of ML estimation impractical. In these cases, a common solution is to replace the true source PDF by some other function, leading to a *quasi ML* estimator. Such an estimator generally does not achieve the CRLB and a more delicate performance analysis is required. In [19, 11], asymptotic performance analysis of QML estimators for blind source separation was presented. Single-channel QML blind deconvolution was studied in [10, 9].

In this study, we derive asymptotic performance bounds for a QML estimator of the restoration kernel in the multi-channel blind deconvolution problem, and state the asymptotic stability conditions. We show that in the particular case when the true ML procedure is used, our bounds coincide with the CRLB, previously reported in literature.

2 QML blind deconvolution

Let $h_{ij\tau}$ be the restoration kernel. Under the assumption that $\det \bar{h}_{ij}(\theta)$ has no zeros on the unit circle, and the source signal is real and i.i.d., the normalized minus-log-likelihood function of the observed signal x_{it} in the noise-free case is [1, 2, 17]

$$\ell(x;h) = -\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left|\det \bar{h}_{ij}(\theta)\right|\,d\theta + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\varphi_i\left(y_{it}\right), \tag{1}$$

where

$$y_{it} \;=\; h_i^{j\tau}x_{j,t-\tau}$$

is a source estimate, $\varphi_i(s) = -\log p_i\left(s\right)$ and $p_i(s)$ is the probability density function (PDF) of the i-th source. We will henceforth assume that the restoration tensor $h_{ij\tau}$ has a finite impulse response, supported on $\tau = -M, ..., M$. We also assume without loss of generality that the sources are zero-mean. Cost function (1) can be also derived using negative joint entropy [1, 2] and information maximization [3] considerations.

Consistent estimator can be obtained by minimizing $\ell(x;h)$ even when $\varphi_i(s_i)$ is not exactly equal to $-\log p_i\left(s_i\right)$. Such quasi-ML estimation has been shown to be practical in instantaneous blind source separation [19, 30, 29, 16] and blind deconvolution [2, 7, 10] when the source PDF is unknown or not well-suited for optimization. For example, when the source is super-Gaussian (e.g. it is sparse or sparsely representable), a smooth approximation of the absolute value function is a good choice for $\varphi(s)$ [13, 30]. It is convenient to use a family of convex smooth functions, e.g.

$$\varphi_\lambda(s) \;=\; |s| - \lambda\log\left(1 + \frac{|s|}{\lambda}\right) \tag{2}$$

with λ being a positive smoothing parameter, to approximate the absolute value [28]. $\varphi_\lambda(s) \to |s|$ as $\lambda \to 0^+$. In the limit $\lambda \to 0^+$, this choice yields a super-efficient QML estimator of sources with $P(s=0) > 0$ [9].

In case of sub-Gaussian sources, the family of functions

$$\varphi_\mu(s) = |s|^\mu \tag{3}$$

with the parameter $\mu > 2$ is usually a good choice for $\varphi(s)$ [2, 7]. For example, when sources are uniformly distributed, an increase in the value of μ refines the approximation and in the limit $\mu \to \infty$, $\varphi_\mu(s)$ approaches $-\log p(s)$. In the limit $\mu \to \infty$, this choice yields a super-efficient QML estimator of sources, whose PDF vanishes outside an interval [9].

2.1 Equivariance

A remarkable property of the QML estimator $\hat{h}(x)$ of a restoration kernel $h_{ij\tau}$ given the observation x, obtained by minimization of $\ell(x;h)$ in (1), is its *equivariance*, stated in the following proposition:

Proposition 1. *The estimator $\hat{h}(x)$ obtained by minimization of $\ell(x;h)$ is equivariant, i.e., for every invertible a, $\hat{h}(a*x) = a^{-1} * \hat{h}(x)$.*

Proof. Observe that for an invertible a,

$$\begin{aligned}
&\ell(a*x; a^{-1}*h) = \\
&\quad -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|\det\left(\bar{a}^{-1}_{ik}(\theta)\bar{h}^{k}_{j}(\theta)\right)\right| d\theta + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\varphi_i\left(a^{-1}*a*h*x_{it}\right) \\
&= -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|\det \bar{h}_{ij}(\theta)\right| d\theta + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\varphi_i\left(y_{it}\right) \\
&\quad -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|\det \bar{a}^{-1}_{ij}(\theta)\right| d\theta \\
&= \ell(x;h) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|\det \bar{a}^{-1}_{ij}(\theta)\right| d\theta.
\end{aligned}$$

Let $h = \operatorname{argmin} \ell(x;h)$. Then

$$\ell(a*x; a^{-1}*h) = \ell(x;h) + \text{const},$$

that is, h is a minimizer of $\ell(a*x; a^{-1}*h)$ as well. Consequently, $\hat{h}(a*x) = a^{-1} * \hat{h}(x)$. □

Equivariance implies that the parameters to be estimated (in our case, the coefficients, $h_{ij\tau}$, specifying the restoration kernel) form a group. This is indeed the case for invertible kernels with the convolution operation. In view of equivariance, we may analyze the properties of $\ell(h*x; \delta_{ij}\delta_\tau)$ instead of $\ell(x;h)$.

2.2 The gradient and the Hessian of $\ell(x;h)$

The gradient of the first term of $\ell(x;h)$ in (1), containing the integral of log spectrum, is given by

$$\begin{aligned}
&\frac{\partial}{\partial h_{ij\tau}}\left(-\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left|\det \bar{h}_{mn}(\theta)\right| d\theta\right) = \\
&\quad -\frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(\frac{d\log\left|\det \bar{h}_{mk}(\theta)\right|}{d\bar{h}_{mk}(\theta)}\frac{\partial \bar{h}_{n}{}^{k}(\theta)}{\partial h_{ij\tau}}\right) d\theta \\
&= -\frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left(e^{-\imath\tau\theta}\cdot \bar{h}^{-1}_{km}(\theta)\delta_{n-i}\delta^{\,k-j}\right) d\theta \\
&= -\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-\imath\tau\theta}\cdot \bar{h}^{-1}_{ji}(\theta) d\theta \\
&= -h^{-1}_{ji,-\tau}. \qquad (4)
\end{aligned}$$

The gradient of the second term of $\ell(x;h)$, containing the sum of φ's, is given by

$$\frac{\partial}{\partial h_{ij\tau}}\left(\frac{1}{T}\sum_{t=0}^{T-1}\sum_{k=1}^{N}\varphi_i\left(y_{kt}\right)\right) \quad = \quad \frac{1}{T}\sum_{t=0}^{T-1}\sum_{k=1}^{N}\varphi_k'\left(y_{kt}\right)\frac{\partial y_{kt}}{\partial h_{ij\tau}}, \tag{5}$$

where

$$\frac{\partial y_{kt}}{\partial h_{ij\tau}} \quad = \quad \frac{\partial}{\partial h_{ij\tau}}\left(h_k{}^{j'\tau'}x_{j',t-\tau'}\right) = x_{j,t-\tau}\delta_{ik}.$$

Combining (4) and (5), yields

$$\begin{aligned}\frac{\partial \ell}{\partial h_{ij\tau}} \quad &= \quad -h^{-1}_{ji,-\tau} + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{k=1}^{N}\varphi_k'\left(y_{kt}\right)x_{j,t-\tau}\delta_{ik} \\ &= \quad -h^{-1}_{ji,-\tau} + \frac{1}{T}\sum_{t=0}^{T-1}\varphi_i'\left(y_{it}\right)x_{j,t-\tau} \end{aligned} \tag{6}$$

Differentiating $h^{-1}_{ji,-\tau}$ w.r.t. $h_{i'j'\tau'}$ yields

$$\begin{aligned}\frac{\partial h^{-1}_{ji,-\tau}}{\partial h_{i'j'\tau'}} \quad &= \quad \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-\imath\tau\theta}\cdot\frac{\partial\ \mathrm{tr}\left(h^{-1}_{mk}\delta^{k-i}\delta_{n-j}\right)}{\partial h_{i'j'\tau'}}d\theta \\ &= \quad \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-\imath\tau\theta}\cdot\mathrm{tr}\left(\frac{d\ \mathrm{tr}\left(h^{-1}_{mk}\delta^{k-i}\delta_{l-j}\right)}{dh^{-1}_{mk}}\frac{\partial(h^{-1})^l{}_n}{\partial h_{i'j'\tau'}}\right)d\theta \\ &= \quad \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-\imath\tau\theta}\cdot\mathrm{tr}\left(\delta_{m-i}\delta_{k-j}(h^{-1})^k{}_l\frac{\partial h^l{}_r}{\partial h_{i'j'\tau'}}(h^{-1})^r{}_n\right)d\theta \\ &= \quad \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-\imath(\tau+\tau')\theta}\cdot\mathrm{tr}\left(\delta_{m-i}\delta_{k-j}(h^{-1})^k{}_l\delta^{l-i'}\delta_{r-j'}(h^{-1})^r{}_n\right)d\theta \\ &= \quad \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-\imath(\tau+\tau')\theta}\cdot h^{-1}_{ji'}(\theta)h^{-1}_{j'i}(\theta)d\theta \\ &= \quad h^{-1}_{ji',-(\tau+\tau')} * h^{-1}_{j'i,-(\tau+\tau')}. \end{aligned} \tag{7}$$

Differentiating the second term of (6), containing the sum, yields

$$\frac{\partial}{\partial h_{i'j'\tau'}}\left(\frac{1}{T}\sum_{t=0}^{T-1}\varphi_i'\left(y_{it}\right)x_{j,t-\tau}\right) \quad = \quad \frac{1}{T}\sum_{t=0}^{T-1}\varphi_i''\left(y_{it}\right)x_{j,t-\tau}x_{j',t-\tau'}\cdot\delta_{ii'}. \tag{8}$$

Combining (7) and (8), yields

$$\frac{\partial^2 \ell}{\partial h_{ij\tau}\partial h_{i'j'\tau'}} = h^{-1}_{ji',-(\tau+\tau')} * h^{-1}_{j'i,-(\tau+\tau')} + \frac{1}{T}\sum_{t=0}^{T-1}\varphi_i''\left(y_{it}\right)x_{j,t-\tau}x_{j',t-\tau'}\cdot\delta_{ii'}. \tag{9}$$

2.3 Asymptotic Hessian

At the solution point, where $h_{ij\tau} = c_i w^{-1}_{ij\tau}$, it holds that $h_{ijt} * x_{it} = c_i s_{it}$. Consequently the Hessian of $\ell(c_i s_{it}; \delta_{ij}\delta_\tau)$ is

$$\left(\nabla^2 \ell\right)_{ij\tau i'j'\tau'} = \delta_{ij'}\delta_{ji'}\delta_{k,-k'} + \frac{\delta_{ii'}}{T}\sum_{t=0}^{T-1}\varphi''_i\left(c_i s_{it}\right) c_j s_{j,t-\tau} c_{j'} s_{j',t-\tau'}.$$

For a large sample size T, the average

$$\frac{1}{T}\sum_{t=0}^{T-1}\varphi''_i\left(c_i s_{it}\right) c_j s_{j,t-\tau} c_{j'} s_{j',t-\tau'}$$

approaches the expected value

$$\mathbb{E}\varphi''\left(c_i s_{it}\right) c_j s_{j,t-\tau} c_{j'} s_{j',t-\tau'}.$$

Since each source s_{it} is assumed to be zero-mean i.i.d., and two distinct sources are independent, the following structure of the Hessian at the solution point is obtained asymptotically:

$$\left(\nabla^2 \ell\right)_{ij\tau i'j'\tau'} = \begin{cases} \alpha_i + 1 & : \quad i = i' = j = j', \tau = \tau' = 0, \\ \gamma_i \sigma'^2_j & : \quad i = i', j = j', \tau = \tau', \\ 1 & : \quad i = j', i' = j, \tau = -\tau', \\ 0 & : \quad otherwise. \end{cases} \tag{10}$$

where

$$\begin{aligned} \sigma_i^2 &= \mathbb{E}s_i^2 \\ \sigma'^2_i &= (c_i\sigma_i)^2 \\ \alpha_i &= \mathbb{E}\varphi''_i(c_i s_i)(c_i s_i)^2 \\ \gamma_i &= \mathbb{E}\varphi''_i(c_i s_i). \end{aligned}$$

Such Hessian structure suggests that the relative Newton algorithm [28, 7, 10], which uses the asymptotic Hessian approximation at the solution point, can be performed in a block-coordinate manner, as it has been previously done in the context of blind source separation [8]. Observe that the Newton system decouples into N independent linear equations, additional $\frac{1}{2}N(N-1)+MN$ systems of size 2×2, and $\frac{1}{2}MN(N-1)$ systems of size 4×4 .

2.4 The scaling factors c_i

In order to find the scaling factors c_i, observe that the gradient of $\ell(c_i s_{it}; \delta_{ij}\delta_\tau)$ is given by

$$\left(\nabla \ell\right)_{ij\tau} = -\delta_{ij}\delta_\tau + \frac{1}{T}\sum_{t=0}^{T-1}\varphi'_i\left(c_i s_{it}\right) c_j s_{j,t-\tau},$$

hence,

$$\begin{aligned}\mathbb{E}\left(\nabla\ell\right)_{ij\tau} &= -\delta_{ij}\delta_\tau + \mathbb{E}\varphi_i'\left(c_i s_{it}\right)\left(c_j s_{j,t-\tau}\right)\\ &= -\delta_{ij}\delta_\tau + \delta_{ij}\delta_\tau\zeta_i,\end{aligned}$$

where

$$\zeta_i = \mathbb{E}\varphi_i'\left(c_i s_i\right)\left(c_i s_i\right)$$

Demanding $\mathbb{E}\left(\nabla\ell\right)_{ij\tau} = 0$, one obtains the set of equations

$$\zeta_i = 1, \tag{11}$$

from where the scaling factors c_i can be found.

3 Asymptotic error covariance tensor

We present now an asymptotic analysis of the QML estimation error. Let the restoration kernel, h, be estimated by minimizing the minus log likelihood function $\ell(x;h)$ defined in (1), where the true $-\log p_i(s_i)$ of the sources are replaced by some other functions $\varphi_i(s_i)$. We assume that h has sufficient degrees of freedom to accurately approximate the inverse of w. For analytic tractability, we assume that γ_i, σ_i^2, $\sigma_i'^2$, α_i, $\beta_i = \mathbb{E}\varphi_i'^2(c_i s_i)$, $\zeta_i = \mathbb{E}\varphi_i'(c_i s_i)(c_i s_i)$ and $\theta_i = \mathbb{E}\varphi_i'^2(c_i s_i)(c_i s_i)^2$ exist and are bounded. Note that the expected values are computed with respect to the true PDFs of s_i.

Let $h^*_{ij\tau} = c_i w^{-1}_{ij\tau}$ be the exact restoration kernel (up to scaling). It can be shown that h^* satisfies [16]

$$h^* = \underset{h}{\operatorname{argmin}}\ \mathbb{E}_x \ell(x;h).$$

Let $\hat{h}$ be the estimate of the exact restoration kernel h^*, based on the finite realization of the data x,

$$\hat{h} = \underset{h}{\operatorname{argmin}}\ \ell(x;h).$$

Note that $\nabla\ell(x;\hat{h}) = 0$, whereas $\nabla\ell(x;h^*) \neq 0$; yet $\mathbb{E}\nabla\ell(x;h^*) = 0$. Denote the estimation error as $\Delta h = h^* - \hat{h}$. Then, assuming $\|\Delta h\|$ is small, second-order Taylor expansion yields

$$\begin{aligned}\left(\nabla\ell(x;h^*)\right)_{ij\tau} &\approx \left(\nabla^2\ell(x;h^*)\right)_{ij\tau}^{\ \ i'j'\tau'}\left(h^*_{i'j'\tau'} - \hat{h}_{i'j'\tau'}\right)\\ &= \left(\nabla^2\ell(x;w^{-1})\right)_{ij\tau}^{\ \ i'j'\tau'}\Delta h_{i'j'\tau'}.\end{aligned}$$

Due to the equivariance property, the former relation can be rewritten as

$$\left(\nabla\ell(h^* * x;\delta_{ij}\delta_\tau)\right)_{ij\tau} \approx \left(\nabla^2\ell(h^* * x;\delta_n)\right)_{ij\tau}^{\ \ i'j'\tau'}\Delta h_{i'j'\tau'}.$$

Since $h^*_{ij\tau} = c_i w^{-1}_{ij\tau}$, we can substitute $h^*_{ijt} * x_{it} = c_i s_{it}$, and obtain

$$(\nabla \ell(c_i s_{it}; \delta_{ij}\delta_\tau))_{ij\tau} \approx \left(\nabla^2 \ell(c_i s_{it}; \delta_{ij}\delta_\tau)\right)_{ij\tau}^{\quad i'j'\tau'} \Delta h_{i'j'\tau'}.$$

For convenience, we will denote the gradient and the Hessian at the solution point as

$$\begin{aligned} g_{ij\tau} &= (\nabla \ell(cs; \delta_n))_{ij\tau} \\ H_{ij\tau i'j'\tau'} &= \left(\nabla^2 \ell(cs; \delta_n)\right)_{ij\tau i'j'\tau'}, \end{aligned}$$

and rewrite the former equation as

$$g_{ij\tau} \approx H_{ij\tau}^{\quad i'j'\tau'} \Delta h_{i'j'\tau'}. \tag{12}$$

For a large sample size, the asymptotic Hessian structure (10) can be used, allowing to split (12) into the following three sets of problems:

1. For $i = j, \tau = 0$:

$$g_{ii0} = (a_i + 1)\Delta h_{ii0} \tag{13}$$

2. For $i = j, \tau \neq 0$:

$$\begin{aligned} g_{ii\tau} &= \gamma_i \sigma_i'^2 \Delta h_{ii\tau} + \Delta h_{ii,-\tau} \\ g_{ii,-\tau} &= \gamma_i \sigma_i'^2 \Delta h_{ii,-\tau} + \Delta h_{ii,\tau} \end{aligned} \tag{14}$$

3. For $i \neq j, \tau \neq 0$:

$$\begin{aligned} g_{ij\tau} &= \gamma_i \sigma_j'^2 \Delta h_{ij\tau} + \Delta h_{ji,-\tau} \\ g_{ij,-\tau} &= \gamma_i \sigma_j'^2 \Delta h_{ij,-\tau} + \Delta h_{ji\tau} \\ g_{ji\tau} &= \gamma_j \sigma_i'^2 \Delta h_{ji\tau} + \Delta h_{ij,-\tau} \\ g_{ji,-\tau} &= \gamma_j \sigma_i'^2 \Delta h_{ji,-\tau} + \Delta h_{ij\tau}, \end{aligned} \tag{15}$$

or, alternatively, in terms of Δh,

1. For $i = j, \tau = 0$:

$$\Delta h_{ii0} = \frac{1}{a_i + 1} g_{ii0} \tag{16}$$

2. For $i = j, \tau \neq 0$:

$$\begin{pmatrix} \Delta h_{ii\tau} \\ \Delta h_{ii,-\tau} \end{pmatrix} = \frac{1}{\gamma_i^2 \sigma_i'^4 - 1} \begin{pmatrix} \gamma_i \sigma_i'^2 & -1 \\ -1 & \gamma_i \sigma_i'^2 \end{pmatrix} \begin{pmatrix} g_{ii\tau} \\ g_{ii,-\tau} \end{pmatrix} \tag{17}$$

3. For $i \neq j, \tau \neq 0$:

$$\begin{pmatrix} \Delta h_{ij\tau} \\ \Delta h_{ij,-\tau} \\ \Delta h_{ji\tau} \\ \Delta h_{ji,-\tau} \end{pmatrix} =$$
$$\frac{1}{\gamma_i \gamma_j \sigma_i'^2 \sigma_j'^2 - 1} \begin{pmatrix} \gamma_j \sigma_i'^2 & & & -1 \\ & \gamma_j \sigma_i'^2 & -1 & \\ & -1 & \gamma_i \sigma_j'^2 & \\ -1 & & & \gamma_i \sigma_j'^2 \end{pmatrix} \begin{pmatrix} g_{ij\tau} \\ g_{ij,-\tau} \\ g_{ji\tau} \\ g_{ji,-\tau} \end{pmatrix} \qquad (18)$$

Consequently, the asymptotic covariance tensor of Δh has the following separable structure

$$\begin{aligned} (\Sigma_{\Delta h})_{ij\tau i'j'\tau'} &= \mathbb{E}\left(\Delta h_{ij\tau} - \mathbb{E}\Delta h_{ij\tau}\right)\left(\Delta h_{i'j'\tau'} - \mathbb{E}\Delta h_{i'j'\tau'}\right) \\ &= \mathbb{E}\Delta h_{ij\tau}\Delta h_{i'j'\tau'} \\ &= \begin{cases} \mathbb{E}\Delta h_{ij\tau}\Delta h_{ij\tau'} & : \ i = i', j = j', \tau = \pm\tau', \\ \mathbb{E}\Delta h_{ij\tau}\Delta h_{ji\tau'} & : \ i = j', i' = j, \tau = \pm\tau', \\ 0 & : \ otherwise. \end{cases} \end{aligned}$$

This implies that the covariance terms, for which $(\Sigma_{\Delta h})_{ij\tau i'j'\tau'} = 0$, decrease in the order of $1/T^2$ as $T \to \infty$.

Let us now evaluate the covariance tensor $(\Sigma_{\Delta h})_{ij\tau i'j'\tau'}$. For $i = j, \tau = 0$, we obtain:

$$\mathbb{E}\Delta h_{ii0}^2 = \frac{1}{(\alpha_i + 1)^2}\, \mathbb{E}g_{ii0}^2. \qquad (19)$$

Neglecting second-order terms, the second moment of g_{ii0} is

$$\begin{aligned} \mathbb{E}g_{ii0}^2 &\approx -1 - 2\mathbb{E}\varphi_i'(c_i s_i)(c_i s_i) + \mathbb{E}^2\varphi_i'(c_i s_i)(c_i s_i) \\ &\quad + \frac{1}{T} \cdot \left(\mathbb{E}\varphi_i'^2(c_i s_i)(c_i s_i)^2 - \mathbb{E}^2\varphi'(c_i s_i)(c_i s_i)\right) \\ &= (\zeta_i - 1)^2 + \frac{\vartheta_i - \zeta_i^2}{T} = \frac{\vartheta_i - 1}{T}, \end{aligned}$$

where

$$\begin{aligned} \zeta_i &= \mathbb{E}\varphi_i'(c_i s_i)(c_i s_i) = 1 \\ \vartheta_i &= \mathbb{E}\varphi_i'^2(c_i s_i)(c_i s_i)^2. \end{aligned}$$

Hence,

$$(\Sigma_{\Delta h})_{ii0\ ii0} \approx \frac{\vartheta_i - 1}{T(\alpha_i + 1)^2}.$$

For $i = j, \tau \neq 0$, we obtain:

$$\begin{pmatrix} \mathbb{E}\Delta h_{ii\tau}\Delta h_{ii\tau} & \mathbb{E}\Delta h_{ii\tau}\Delta h_{ii,-\tau} \\ \mathbb{E}\Delta h_{ii,-\tau}\Delta h_{ii\tau} & \mathbb{E}\Delta h_{ii,-\tau}\Delta h_{ii,-\tau} \end{pmatrix} = \frac{1}{\left(\gamma_i^2\sigma_i'^4 - 1\right)^2} \cdot \begin{pmatrix} \gamma_i\sigma_i'^2 & -1 \\ -1 & \gamma_i\sigma_i'^2 \end{pmatrix} \begin{pmatrix} \mathbb{E}\Delta g_{ii\tau}\Delta g_{ii\tau} & \mathbb{E}\Delta g_{ii\tau}\Delta g_{ii,-\tau} \\ \mathbb{E}\Delta g_{ii,-\tau}\Delta g_{ii\tau} & \mathbb{E}\Delta g_{ii,-\tau}\Delta g_{ii,-\tau} \end{pmatrix} \cdot \begin{pmatrix} \gamma_i\sigma_i'^2 & -1 \\ -1 & \gamma_i\sigma_i'^2 \end{pmatrix} \tag{20}$$

Neglecting second-order terms,

$$\begin{aligned} \mathbb{E}g_{ii\tau}^2 &\approx \frac{1}{T}\mathbb{E}\varphi_i'^2(c_is_i)\mathbb{E}(c_is_i)^2 = \frac{\beta_i\sigma_i'^2}{T}, \\ \mathbb{E}g_{ii\tau}g_{ii,-\tau} &\approx \frac{1}{T}\mathbb{E}^2\varphi_i'^2(c_is_i)(c_is_i) = \frac{\zeta_i^2}{T} = \frac{1}{T}, \end{aligned}$$

where

$$\beta_i = \mathbb{E}\varphi_i'^2(c_is_i).$$

Substituting the covariance of $g_{ii\tau}$ into the covariance of $\Delta h_{ii\tau}$, after some algebraic manipulations, one obtains

$$\begin{aligned} (\Sigma_{\Delta h})_{ii\tau ii\tau} &\approx \frac{\beta_i\sigma_i'^2\left(\gamma_i^2\sigma_i'^4 + 1\right) - 2\gamma_i\sigma_i'^2}{T(\gamma_i^2\sigma_i'^4 - 1)^2} \\ (\Sigma_{\Delta h})_{ii\tau ii,-\tau} &\approx \frac{\gamma_i\sigma_i'^2\left(\gamma_i\sigma_i'^2 - 2\beta_i\sigma_i'^2\right) + 1}{T(\gamma_i^2\sigma_i'^4 - 1)^2}. \end{aligned}$$

For $i \neq j, \tau \neq 0$, we obtain:

$$\begin{pmatrix} \mathbb{E}\Delta h_{ij\tau}\Delta h_{ij\tau} & \mathbb{E}\Delta h_{ij\tau}\Delta h_{ij,-\tau} & \mathbb{E}\Delta h_{ij\tau}\Delta h_{ji\tau} & \mathbb{E}\Delta h_{ij\tau}\Delta h_{ji,-\tau} \\ \mathbb{E}\Delta h_{ij,-\tau}\Delta h_{ij\tau} & \mathbb{E}\Delta h_{ij,-\tau}\Delta h_{ij,-\tau} & \mathbb{E}\Delta h_{ij,-\tau}\Delta h_{ji\tau} & \mathbb{E}\Delta h_{ij,-\tau}\Delta h_{ji,-\tau} \\ \mathbb{E}\Delta h_{ji\tau}\Delta h_{ij\tau} & \mathbb{E}\Delta h_{ji\tau}\Delta h_{ij,-\tau} & \mathbb{E}\Delta h_{ji\tau}\Delta h_{ji\tau} & \mathbb{E}\Delta h_{ji\tau}\Delta h_{ji,-\tau} \\ \mathbb{E}\Delta h_{ji,-\tau}\Delta h_{ij\tau} & \mathbb{E}\Delta h_{ji,-\tau}\Delta h_{ij,-\tau} & \mathbb{E}\Delta h_{ji,-\tau}\Delta h_{ji\tau} & \mathbb{E}\Delta h_{ji,-\tau}\Delta h_{ji,-\tau} \end{pmatrix} = \frac{1}{\left(\gamma_i\gamma_j\sigma_i'^2\sigma_j'^2 - 1\right)^2} \begin{pmatrix} \gamma_j\sigma_i'^2 & & & -1 \\ & \gamma_j\sigma_i'^2 & -1 & \\ & -1 & \gamma_i\sigma_j'^2 & \\ -1 & & & \gamma_i\sigma_j'^2 \end{pmatrix} \cdot \begin{pmatrix} \mathbb{E}\Delta g_{ij\tau}\Delta g_{ij\tau} & \mathbb{E}\Delta g_{ij\tau}\Delta g_{ij,-\tau} & \mathbb{E}\Delta g_{ij\tau}\Delta g_{ji\tau} & \mathbb{E}\Delta g_{ij\tau}\Delta g_{ji,-\tau} \\ \mathbb{E}\Delta g_{ij,-\tau}\Delta g_{ij\tau} & \mathbb{E}\Delta g_{ij,-\tau}\Delta g_{ij,-\tau} & \mathbb{E}\Delta g_{ij,-\tau}\Delta g_{ji\tau} & \mathbb{E}\Delta g_{ij,-\tau}\Delta g_{ji,-\tau} \\ \mathbb{E}\Delta g_{ji\tau}\Delta g_{ij\tau} & \mathbb{E}\Delta g_{ji\tau}\Delta g_{ij,-\tau} & \mathbb{E}\Delta g_{ji\tau}\Delta g_{ji\tau} & \mathbb{E}\Delta g_{ji\tau}\Delta g_{ji,-\tau} \\ \mathbb{E}\Delta g_{ji,-\tau}\Delta g_{ij\tau} & \mathbb{E}\Delta g_{ji,-\tau}\Delta g_{ij,-\tau} & \mathbb{E}\Delta g_{ji,-\tau}\Delta g_{ji\tau} & \mathbb{E}\Delta g_{ji,-\tau}\Delta g_{ji,-\tau} \end{pmatrix} \cdot \begin{pmatrix} \gamma_j\sigma_i'^2 & & & -1 \\ & \gamma_j\sigma_i'^2 & -1 & \\ & -1 & \gamma_i\sigma_j'^2 & \\ -1 & & & \gamma_i\sigma_j'^2 \end{pmatrix} \tag{21}$$

Neglecting second-order terms,

$$\begin{aligned}
\mathbb{E}g^2_{ij\tau} &\approx \frac{1}{T}\mathbb{E}\varphi'^2_i(c_is_i)\mathbb{E}(c_js_j)^2 = \frac{\beta_i\sigma'^2_j}{T}, \\
\mathbb{E}g_{ij\tau}g_{ji,-\tau} &\approx \frac{1}{T}\mathbb{E}\varphi'^2_i(c_is_i)(c_is_i)\mathbb{E}\varphi'^2_j(c_js_j)(c_js_j) = \frac{\zeta_i\zeta_j}{T} = \frac{1}{T}, \\
\mathbb{E}g_{ij\tau}g_{ij,-\tau} &\approx 0.
\end{aligned}$$

Substituting the covariance of $g_{ij\tau}$ into the covariance of $\Delta h_{ij\tau}$, after some algebraic manipulations, one obtains

$$\begin{aligned}
(\Sigma_{\Delta h})_{ij,-\tau,ij,-\tau} = (\Sigma_{\Delta h})_{ij\tau ij\tau} &\approx \frac{\sigma'^2_i\left(\gamma_j^2\beta_i\sigma'^2_i\sigma'^2_j - 2\gamma_j + \beta_j\right)}{T\left(\gamma_i\gamma_j\sigma'^2_i\sigma'^2_j - 1\right)^2} \\
(\Sigma_{\Delta h})_{ij\tau ji,-\tau} = (\Sigma_{\Delta h})_{ij,-\tau,ji\tau} &\approx \frac{\gamma_i\gamma_j\sigma'^2_i\sigma'^2_j - \sigma'^2_i\sigma'^2_j\left(\gamma_i\beta_j + \gamma_j\beta_i\right) - 1}{T\left(\gamma_i\gamma_j\sigma'^2_i\sigma'^2_j - 1\right)^2}.
\end{aligned}$$

Combining the previous results, the following asymptotic covariance tensor of the estimation error is obtained:

$$(\Sigma_{\Delta h})_{ij\tau i'j'\tau'} \approx \frac{1}{T}\cdot\begin{cases}
\frac{\vartheta_i - 1}{(\alpha_i+1)^2} & : \; i = i' = j = j', \tau = \tau' = 0, \\
\frac{\beta_i\sigma'^2_i\left(\gamma_i^2\sigma'^4_i+1\right) - 2\gamma_i\sigma'^2_i}{(\gamma_i^2\sigma'^4_i - 1)^2} & : \; i = j' = i' = j, \tau = \tau' \neq 0, \\
\frac{\gamma_i\sigma'^2_i\left(\gamma_i\sigma'^2_i - 2\beta_i\sigma'^2_i\right)+1}{(\gamma_i^2\sigma'^4_i - 1)^2} & : \; i = j' = i' = j, \tau = -\tau' \neq 0, \\
\frac{\sigma'^2_i\left(\gamma_j^2\beta_i\sigma'^2_i\sigma'^2_j - 2\gamma_j + \beta_j\right)}{\left(\gamma_i\gamma_j\sigma'^2_i\sigma'^2_j - 1\right)^2} & : \; i = i' \neq j = j', \tau = \tau', \\
\frac{\gamma_i\gamma_j\sigma'^2_i\sigma'^2_j - \sigma'^2_i\sigma'^2_j(\gamma_i\beta_j + \gamma_j\beta_i) - 1}{\left(\gamma_i\gamma_j\sigma'^2_i\sigma'^2_j - 1\right)^2} & : \; i = j' \neq j = i', \tau = -\tau', \\
0 & : \; otherwise.
\end{cases} \tag{22}$$

Observe that for every τ, τ',

$$\begin{aligned}
(\Sigma_{\Delta h})_{ii\tau jk\tau'} = 0 &\quad : \quad i, j \neq k \\
(\Sigma_{\Delta h})_{ii\tau jj\tau'} = 0 &\quad : \quad i \neq j,
\end{aligned}$$

which implies lack of correlation between the estimates of the diagonal and off-diagonal elements of $h_{ij\tau}$ for every τ. Therefore, the asymptotic error covariance tensor decouples into the source separation terms ($h_{ij\tau}$ for $i \neq j$), determining the quality of cross-talk elimination, and the N independent single-channel deconvolution terms ($h_{ii\tau}$). Observe also that for every τ, τ', and $i \neq j$, $(\Sigma_{\Delta h})_{ij\tau i'j'\tau'} \neq 0$ only for $i' = j, j' = i$. Therefore, the source separation terms are further decoupled into independent 2×2 terms ($h_{ij\tau}$, $h_{ji\tau}$). This decoupling has been previously observed in the special case of true ML multichannel blind deconvolution [26]. Since for every i, $(\Sigma_{\Delta h})_{ii\tau ii\tau'} \neq 0$ only for $\tau = \pm\tau'$, each of the single-channel blind deconvolution terms in the covariance tensor also decouples into N independent 2×2 terms ($h_{ii\tau}$, $h_{ii,-\tau}$) and one 1×1 term (h_{ii0}). The latter property is rather surprising, since combined with asymptotic normality of the estimation error, it implies that $\Delta h_{ii\tau}$ are statistically independent for all $\tau \geq 0$.

3.1 Particular cases

In several particular cases, the asymptotic error covariance tensor (22) can be significantly simplified. We discuss here the cases of single-channel blind deconvolution, the instantaneous blind source separation, and the true ML deconvolution. We also highlight two special cases of Gaussian sources and sources allowing superefficient deconvolution.

3.1.1 Single-channel blind deconvolution

In the case of *single-channel* blind deconvolution ($N = 1$), the asymptotic covariance reduces to

$$(\Sigma_{\Delta h})_{\tau\tau'} \approx \frac{1}{T} \cdot \begin{cases} \frac{\vartheta-1}{(\alpha+1)^2} & : \quad \tau = \tau' = 0, \\ \frac{\beta_i \sigma'^2\left(\gamma^2\sigma'^4+1\right)-2\gamma\sigma'^2}{(\gamma^2\sigma'^4-1)^2} & : \quad \tau = \tau' \neq 0, \\ \frac{\gamma\sigma'^2\left(\gamma\sigma'^2-2\beta\sigma'^2\right)+1}{(\gamma^2\sigma'^4-1)^2} & : \quad \tau = -\tau' \neq 0, \\ 0 & : \quad otherwise. \end{cases}$$

This result coincides with the asymptotic covariance derived in [10, 9] for the single-channel case.

3.1.2 Instantaneous blind source separation

In the case of *instantaneous* blind source separation ($K = 0$), the asymptotic covariance reduces to

$$(\Sigma_{\Delta h})_{iji'j'} \approx \frac{1}{T} \cdot \begin{cases} \frac{\vartheta_i-1}{(\alpha_i+1)^2} & : \quad i = i' = j = j', \\ \frac{\sigma_i'^2\left(\gamma_j^2\beta_i\sigma_i'^2\sigma_j'^2-2\gamma_j+\beta_j\right)}{\left(\gamma_i\gamma_j\sigma_i'^2\sigma_j'^2-1\right)^2} & : \quad i = i' \neq j = j', \\ \frac{\gamma_i\gamma_j\sigma_i'^2\sigma_j'^2-\sigma_i'^2\sigma_j'^2(\gamma_i\beta_j+\gamma_j\beta_i)-1}{\left(\gamma_i\gamma_j\sigma_i'^2\sigma_j'^2-1\right)^2} & : \quad i = j' \neq j = i', \\ 0 & : \quad otherwise. \end{cases}$$

This result is equivalent to the asymptotic performance estimate for QML blind source separation derived in [19].

3.1.3 True ML restoration

True ML restoration is obtained for $\varphi_i(s) = -\log p_i(s)$. In this case, $c_i = 1$, $\sigma_i'^2 = \sigma_i^2$, and under the assumption that $\lim_{s_i \to \pm\infty} p_i'(s)s = 0$,

$$\begin{aligned}
\gamma_i &= \mathbb{E}\varphi_i''(s) = -\int_{-\infty}^{\infty} \frac{p_i'(s)p_i(s) - p_i'^2(s)}{p_i^2(s)} p_i(s)\, ds \\
&= -p_i'(s)\big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{p_i'^2(s)}{p_i^2(s)} p_i(s)\, ds = \mathbb{E}\varphi_i'^2(s) = \beta_i.
\end{aligned}$$

Observe that differentiating the equation

$$\int_{-\infty}^{\infty} p_i(s)s^2 ds \;=\; \sigma_i^2$$

w.r.t. s yields

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty} p_i'(s)s^2 ds + 2\int_{-\infty}^{\infty} p_i(s) s ds \\
&= \int_{-\infty}^{\infty} \frac{p_i'(s)}{p_i(s)} s^2 p_i(s) ds + 2\mathbb{E}s_i \\
&= -\int_{-\infty}^{\infty} \varphi_i'(s) s^2 p_i(s) ds.
\end{aligned}$$

Differentiating again w.r.t. s, we obtain, using integration by parts,

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty} \left(\varphi_i''(s)s^2 + 2\varphi_i'(s)s\right) p_i(s)ds + \int_{-\infty}^{\infty} \varphi_i'(s)s^2 p_i'(s)ds \\
&= \mathbb{E}\varphi_i''(s)s^2 - \int_{-\infty}^{\infty} \left(2p_i'(s)s + \left(\frac{p_i'(s)}{p_i(s)}\right)^2 s^2 p_i(s)\right) ds \\
&= \mathbb{E}\varphi_i''(s)s^2 - \mathbb{E}\varphi_i'^2(s)s^2 - 2p_i(s)s\big|_{-\infty}^{\infty} + 2\int_{-\infty}^{\infty} p_i(s)ds \\
&= \mathbb{E}\varphi_i''(s)s^2 - \mathbb{E}\varphi_i'^2(s)s^2 + 2 = \alpha_i - \theta_i + 2.
\end{aligned}$$

Hence, $\theta_i = \alpha_i + 2$. Let us also define

$$\begin{aligned}
\mathcal{L}_i &= \sigma_i^2 \cdot \mathbb{E}\varphi_i'^2(s) = \beta_i \sigma_i^2 \\
\mathcal{S}_i &= \mathrm{cum}\left\{\varphi_i'(s_i), \varphi_i'(s_i), s_i, s_i\right\} + \mathcal{L}_i + 1 = \theta_i - 1.
\end{aligned}$$

$\mathcal{L}_i$ and $\mathcal{S}_i$ are known as Fisher's information for location and scale parameters, respectively. Substituting $\sigma_i'^2 = \sigma_i^2$, $\gamma_i = \beta_i$, $\theta_i = \alpha_i + 2$, $\theta_i - 1 = \mathcal{S}_i$ and

$\beta_i \sigma_i^2 = \mathcal{L}_i$ into (22), we obtain

$$(\Sigma_{\Delta h})_{ij\tau i'j'\tau'} \approx \frac{1}{T} \cdot \begin{cases} \frac{1}{\mathcal{S}_i} & : \; i = i' = j = j', \tau = \tau' = 0, \\ \frac{\mathcal{L}_i}{\mathcal{L}_i^2 - 1} & : \; i = j' = i' = j, \tau = \tau' \neq 0, \\ \frac{1}{(1-\mathcal{L}_i)(1+\mathcal{L}_i)^2} & : \; i = j' = i' = j, \tau = -\tau' \neq 0, \\ \frac{\sigma_i^2}{\sigma_j^2} \cdot \frac{\mathcal{L}_j}{\mathcal{L}_i \mathcal{L}_j - 1} & : \; i = i' \neq j = j', \tau = \tau', \\ -\frac{1}{\mathcal{L}_i \mathcal{L}_j - 1} & : \; i = j' \neq j = i', \tau = -\tau', \\ 0 & : \; otherwise. \end{cases}$$

This result coincides with the CRLB on $h_{ij\tau}$ developed in [26].

3.1.4 Gaussian sources

Let us assume that the source is Gaussian, i.e.,

$$p(s) \quad = \quad \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \cdot \exp\left\{-\frac{s^2}{2\sigma^2}\right\}.$$

Under mild assumption on $\varphi(s)$, using integration by parts one obtains

$$\begin{aligned} \gamma \quad &= \quad \int_{-\infty}^{\infty} \varphi''(cs)p(s)ds = \frac{\varphi'(cs)}{c}\, p(s)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\varphi'(cs)}{c}\, p'(s)ds \\ &= \quad \frac{1}{c\sigma^2}\int_{-\infty}^{\infty} \varphi'(cs)\, p(s)sds = \frac{1}{(c\sigma)^2}\int_{-\infty}^{\infty} \varphi'(cs)cs\, p(s)ds, \end{aligned}$$

from where by the condition on the scaling factor $\zeta = 1$, it follows that $\gamma\sigma'^2 = 1$. Hence, for a Gaussian source, the denominator of the covariance tensor in (22) vanishes, resulting in infinite variance of the estimation error, which implies the well-known fact that deconvolution of Gaussian sources is impossible.

3.1.5 Super-efficient deconvolution

Let us now consider the particular case of *sparse* sources, such sources that take the value of zero with some non-zero probability $\rho = P(s_n = 0) > 0$. An example of such distribution is the Gauss-Bernoully (sparse normal) distribution [10]. When $\varphi(s)$ is chosen according to (2), $\varphi'_\lambda(s) \to \mathrm{sign}(s)$ and $\varphi''_\lambda(s) \to 2\delta(s)$ as $\lambda \to 0^+$. Hence, for a sufficiently small λ,

$$\gamma \quad \approx \quad \frac{1}{c\lambda}\int_{-\lambda}^{+\lambda} p\left(s\right)\, ds \approx \frac{\rho}{c\lambda},$$

whereas β and c are bounded. Consequently,

$$\begin{aligned}
\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ii\tau ii\tau} &\leq \frac{\beta_i}{\gamma_i^2\sigma_i'^2} \leq \text{const}\cdot\lambda^2 \qquad \tau\neq 0,\\
\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ij\tau ij\tau} &\leq \frac{\beta_i}{\gamma_i^2\sigma_j'^2} \leq \text{const}\cdot\lambda^2, \qquad i\neq j.
\end{aligned}$$

Observe that

$$\begin{aligned}
\lim_{\lambda\to 0^+}\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ii\tau ii\tau} &= 0 \qquad \tau\neq 0,\\
\lim_{\lambda\to 0^+}\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ij\tau ij\tau} &= 0, \qquad i\neq j,
\end{aligned}$$

that is, the estimator $\hat{h}_{ij\tau}$ of $h_{ij\tau}$ is *super-efficient* in the limit $\lambda\to 0^+$.

Similar result is obtained for the sub-Gaussian quasi ML estimator $\varphi_\mu(s)$ defined in (3), with the uniform distribution. Let us assume that $s\sim\mathcal{U}\,[-a,a]$, for $a>0$. By the condition on the scaling factor, we obtain

$$\begin{aligned}
1 &= \zeta = \mathbb{E}\varphi'(cs)(cs) = c\int_{-\infty}^{\infty}\mu|cs|^{\mu-1}\text{sign}(s)s\,p(s)ds\\
&= \frac{\mu c^\mu}{2a}\int_{-a}^{a}|s|^\mu ds = \frac{\mu c^\mu\cdot a^\mu}{\mu+1},
\end{aligned}$$

from where

$$c = \left(\frac{\mu+1}{a^\mu\mu}\right)^{1/\mu}.$$

Consequently,

$$\begin{aligned}
\gamma &= \mathbb{E}\varphi''(cs) = \int_{-\infty}^{\infty}\mu(\mu-1)|cs|^{\mu-2}p(s)ds\\
&= a^2(\mu-1)\left(\frac{\mu}{\mu+1}\right)^{2/\mu}.
\end{aligned}$$

For a sufficiently large μ, $\gamma\approx a^2\mu$, hence,

$$\begin{aligned}
\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ii\tau ii\tau} &\leq \frac{\beta_i}{\gamma_i^2\sigma_i'^2} \leq \text{const}\cdot\frac{1}{\mu^2} \qquad \tau\neq 0,\\
\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ij\tau ij\tau} &\leq \frac{\beta_i}{\gamma_i^2\sigma_j'^2} \leq \text{const}\cdot\frac{1}{\mu^2}, \qquad i\neq j.
\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
\lim_{\mu\to\infty}\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ii\tau ii\tau} &= 0 \qquad \tau\neq 0,\\
\lim_{\mu\to\infty}\operatorname*{plim}_{T\to\infty} T(\Sigma_{\Delta h})_{ij\tau ij\tau} &= 0, \qquad i\neq j,
\end{aligned}$$

that is, the estimator is super-efficient in the limit $\mu\to\infty$. The latter result can be simply generalized for distributions, whose PDF vanishes outside some interval.

4 Asymptotic restoration quality

The quality of the restored signal is degraded by crosstalk due to imperfect channel separation, and from distortions caused by imperfect deconvolution of the channel itself. A common measure of the signal reconstruction quality is the interference-to-signal (ISR) power ratio. Consider the i-th output, for which s_{it} is the desired output, and s_{jt}, $j \neq i$, are the interfering signals. Denoting the global system response by $g_{ij\tau} = h_{ij\tau} * w_{ij\tau}$, the average power of the desired signal is

$$P_{ii} \quad = \quad \mathbb{E} \left\| g_{iit} \star s_{it} \right\|_2^2 \approx c_i \sigma_i^2 = \sigma_i'^2,$$

whereas the average power of the j-th interferer is

$$\begin{aligned} P_{ij} &= \mathbb{E} \| g_{ijt} \star s_{jt} \|_2^2 = \mathbb{E} \| \Delta h_{ijt} \star s_{jt} \|_2^2 \approx \sigma_j^2 \cdot \sum_\tau \left(\Sigma_{\Delta h} \right)_{ij\tau ij\tau} \\ &\approx \frac{(2M+1)\, \sigma_j^2 \sigma_i'^2 \left(\gamma_j^2 \beta_i \sigma_i'^2 \sigma_j'^2 - 2\gamma_j \zeta_i \zeta_j + \beta_j \right)}{T \left(\gamma_i \gamma_j \sigma_i'^2 \sigma_j'^2 - 1 \right)^2}. \end{aligned}$$

Since the sources are assumed to be independent,

$$\mathrm{ISR}_i \quad = \quad \sum_{j \neq i} \frac{P_{ij}}{P_{ii}} \approx \frac{2M+1}{T} \cdot \sum_{j \neq i} \frac{\sigma_j^2 \left(\gamma_j^2 \beta_i \sigma_i'^2 \sigma_j'^2 - 2\gamma_j \zeta_i \zeta_j + \beta_j \right)}{\left(\gamma_i \gamma_j \sigma_i'^2 \sigma_j'^2 - 1 \right)^2}.$$

Observe that for the true ML deconvolution, the latter expression reduces to

$$\mathrm{ISR}_i \quad = \quad \frac{2M+1}{T} \cdot \sum_{j \neq i} \frac{\mathcal{L}_j}{\mathcal{L}_i \mathcal{L}_j - 1},$$

which corresponds to the bound reported in [26].

Another commonly used measure for restoration quality is the *mean square error* (MSE), given by

$$\begin{aligned} \mathrm{MSE}_i &= \mathbb{E} \left(y_{it} - c_i s_{it} \right)^2 = \sum_j \mathbb{E} \left\| g_{ijt} \star s_{jt} - c_i \delta_t \right\|_2^2 = \sum_j \sigma_j^2 \cdot \mathbb{E} \left\| \Delta h_{ijt} \right\|_2^2 \\ &\approx \sum_j \sigma_j^2 \cdot \sum_{\tau=-M}^{M} \left(\Sigma_{\Delta h} \right)_{ij\tau ij\tau} \\ &\approx (2M+1) \cdot \sum_{j \neq i} \frac{\sigma_j^2 \sigma_i'^2 \left(\gamma_j^2 \beta_i \sigma_i'^2 \sigma_j'^2 - 2\gamma_j \zeta_i \zeta_j + \beta_j \right)}{T \left(\gamma_i \gamma_j \sigma_i'^2 \sigma_j'^2 - 1 \right)^2} \\ &\quad + 2M \cdot \frac{\beta_i \sigma_i^2 \sigma_i'^2 \left(\gamma_i^2 \sigma_i'^4 + 1 \right) - 2\gamma_i \sigma_i'^2 \zeta_i^2}{T (\gamma_i^2 \sigma_i'^4 - 1)^2} + \frac{\sigma_i^2 \left(\vartheta_i - \zeta_i^2 \right)}{T (\alpha_i + 1)^2} \end{aligned}$$

5 Asymptotic stability

The estimator $\hat{h}(x)$ of h^*, obtained by minimization of $\ell(x;h)$, is *asymptotically stable* if $h = h^*$ is a local minimizer of $\ell(x;h)$ for infinitely large sample size. Any consistent estimator is asymptotically stable. Asymptotic error analysis, presented in Section 3, is valid only when the QML estimator is asymptotically stable.

Proposition 2. *Let $\hat{h}(x)$ be the QML estimator of h. $\hat{h}(x)$ is asymptotically stable if for all i, j, the following conditions hold:*

$$\begin{aligned} \gamma_i &> 0, \\ \gamma_i\gamma_j\sigma_i'^2\sigma_j'^2 &> 1, \\ \alpha_i + 1 &> 0. \end{aligned} \tag{23}$$

$\hat{h}(x)$ is asymptotically unstable if there exist i, j, for which one of the following conditions hold:

$$\begin{aligned} \gamma_i &< 0, \\ \gamma_i\gamma_j\sigma_i'^2\sigma_j'^2 &< 1, \\ \alpha_i + 1 &< 0. \end{aligned} \tag{24}$$

Proof. The QML estimator is asymptotically stable if in the limit $T \to \infty$, $h = h^*$ is a local minimizer of $\ell(x;h)$, or due to equivariance, $h_{ij\tau} = \delta_{ij}\delta_\tau$ is a local minimizer of $\ell(c_i s_i; h)$. The first- and the second-order Karush-Kuhn-Tucker conditions [6]

$$\operatorname*{plim}_{T\to\infty} \nabla\ell(c_i s_i; \delta_{ij}\delta_\tau) = 0 \tag{25}$$

$$\operatorname*{plim}_{T\to\infty} \nabla^2\ell(c_i s_i; \delta_{ij}\delta_\tau) \succ 0 \tag{26}$$

are the necessary and the sufficient conditions, respectively, for existence of the local minimum.

The necessary condition (25) requires that $\nabla\ell = 0$ as the sample size approaches infinity. By the law of large numbers,

$$\operatorname*{plim}_{T\to\infty} g_{ij\tau} = \mathbb{E} g_{ij\tau} = 0,$$

by choice of c_i. The sufficient condition (26) requires that $\nabla^2\ell \succ 0$ as the sample size approaches infinity, that is, for all $x_{ij\tau} \neq 0$,

$$\operatorname*{plim}_{T\to\infty} x^{ij\tau} H_{ij\tau}{}^{i'j'\tau'} x_{i'j'\tau'} > 0.$$

Using the asymptotic Hessian given in (10), this condition can be rewritten as: $\forall x_{ij\tau} \neq 0$,

$$
\begin{aligned}
& x^{ij\tau} H_{ij\tau}{}^{i'j'\tau'} x_{i'j'\tau'} = \\
& \quad \sum_i (\alpha_i + 1) x_{ii0}^2 + \sum_{i,\tau\neq 0} \left(\gamma_i^2 \sigma_i'^2 x_{ii\tau}^2 + x_{ii\tau} x_{ii,-\tau} \right) + \sum_{i\neq j,\tau} \left(\gamma_i^2 \sigma_j'^2 x_{ij\tau}^2 + x_{ij\tau} x_{ji,-\tau} \right) \\
& \quad = \sum_i (\alpha_i + 1) x_{ii0}^2 + \sum_{i,\tau>0} \begin{pmatrix} x_{ii\tau} \\ x_{ii,-\tau} \end{pmatrix}^T \begin{pmatrix} \gamma_i \sigma_i'^2 & 1 \\ 1 & \gamma_i \sigma_i'^2 \end{pmatrix} \begin{pmatrix} x_{ii\tau} \\ x_{ii,-\tau} \end{pmatrix} \\
& \quad + \sum_{j\neq i} \begin{pmatrix} x_{ij0} \\ x_{ji0} \end{pmatrix}^T \begin{pmatrix} \gamma_i \sigma_j'^2 & 1 \\ 1 & \gamma_i \sigma_j'^2 \end{pmatrix} \begin{pmatrix} x_{ij0} \\ x_{ji0} \end{pmatrix} \\
& \quad + \sum_{j\neq i,\tau>0} \begin{pmatrix} x_{ij\tau} \\ x_{ij,-\tau} \\ x_{ji\tau} \\ x_{ji,-\tau} \end{pmatrix}^T \begin{pmatrix} \gamma_i \sigma_j'^2 & & & 1 \\ & \gamma_i \sigma_j'^2 & 1 & \\ & 1 & \gamma_j \sigma_i'^2 & \\ 1 & & & \gamma_j \sigma_i'^2 \end{pmatrix} \begin{pmatrix} x_{ij\tau} \\ x_{ij,-\tau} \\ x_{ji\tau} \\ x_{ji,-\tau} \end{pmatrix} > 0.
\end{aligned}
$$

The latter holds if and only if for all i, j

$$
\begin{aligned}
a_i + 1 &> 0 \\
\begin{pmatrix} \gamma_i \sigma_j'^2 & 1 \\ 1 & \gamma_i \sigma_j'^2 \end{pmatrix} &\succ 0 \\
\begin{pmatrix} \gamma_i \sigma_j'^2 & & & 1 \\ & \gamma_i \sigma_j'^2 & 1 & \\ & 1 & \gamma_j \sigma_i'^2 & \\ 1 & & & \gamma_j \sigma_i'^2 \end{pmatrix} &\succ 0,
\end{aligned}
$$

which holds if and only if conditions (23) are satisfied. The necessary second-order condition for existence of local minimum of ℓ is

$$
\operatorname*{plim}_{T\to\infty} \nabla^2 \ell(c_i s_i; \delta_{ij}\delta_\tau) \succeq 0.
$$

Hence, if one of conditions (24) hold, the estimator is asymptotically unstable. □

6 Discussion and conclusion

In order to be in a position to utilize the QML estimator of the restoration kernel in multichannel blind deconvolution, and to gain insight into the effect of the source distribution and the choice of $\varphi_i(s)$, it is important to quantify the asymptotic performance and establish stability conditions. For this purpose we derived simple closed-form expressions for the asymptotic estimation error, and showed that its covariance tensor has a simple separable form. An asymptotic estimate of the restoration quality in terms of SIR was also presented. The main conclusion from the performance analysis is that the asymptotic performance depends on the choice of $\varphi_i(s)$ essentially through the ratios $\frac{\mathbb{E}\varphi_i'^2(s_i)}{\mathbb{E}^2\varphi_i''(s_i)}$ of non-linear moments of each of the sources. This provides insight of how to choose a "good" function $\varphi_i(s)$, and can be also used to predict the performance degradation (compared to the CRLB) due to choice of a sub-optimal $\varphi_i(s)$. We demonstrated that in the special cases of single-channel blind deconvolution and instantaneous blind source separation, our asymptotic performance analysis coincide with the results previously reported in literature. We also demonstrated that for the true ML estimator, our asymptotic performance bounds coincide with the CRLB.

The explicit stability conditions on the QML estimator show that asymptotic stability is fully determined by the signs of

$$\mathbb{E}\varphi_i''(c_i s_i),$$
$$\mathbb{E}\varphi_i''(c_i s_i)\mathbb{E}(c_i s_i)^2 \cdot \mathbb{E}\varphi_j''(c_i s_j)\mathbb{E}(c_j s_j)^2 - 1,$$

and

$$\mathbb{E}\varphi_i''(c_i s_i)(c_i s_i)^2 + 1,$$

where c_i are scaling factors, obtained from the solution of $\mathbb{E}\varphi_i'(c_i s_i)(c_i s_i) = 1$. Extension of these results to the case of multi-dimensional sources is straighforward.

References

[1] S.-I. Amari, A. Cichocki, and H. H. Yang. Novel online adaptive learning algorithms for blind deconvolution using the natural gradient approach. In *Proc. SYSID*, pages 1057–1062, July 1997.

[2] S.-I. Amari, S. C. Douglas, A. Cichocki, and H. H. Yang. Multichannel blind deconvolution and equalization using the natural gradient. In *Proc. SPAWC*, pages 101–104, April 1997.

[3] A.J. Bell and T.J. Sejnowski. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[4] S. Bellini and F. Rocca. Near optimal blind deconvolution. In *Proc. of IEEE Conf. Acoust., Speech, Sig. Proc.*, 1988.

[5] A. Benveniste, M. Goursat, and G. Ruget. Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications. *IEEE Trans. Automat. Contr.*, 25(3):385–399, 1980.

[6] D. P. Bertsekas. *Nonlinear Programming (2nd edition)*. Athena Scientific, 1999.

[7] A. M. Bronstein, M. M. Bronstein, and M. Zibulevsky. Blind deconvolution with relative Newton method. Technical Report 444, Technion, Israel, October 2003.

[8] A. M. Bronstein, M. M. Bronstein, and M. Zibulevsky. Block-coordinate relative Newton method for blind source separation. Technical Report 445, Technion, Israel, October 2003.

[9] A. M. Bronstein, M. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi. Quasi maximum likelihood blind deconvolution: asymptotic performance analysis. *IEEE Trans. Information Theory*. submitted.

[10] A. M. Bronstein, M. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi. Quasi maximum likelihood blind deconvolution of images using optimal sparse representations. Technical Report 455, Technion, Israel, December 2003.

[11] J.-F. Cardoso. Blind signal separation statistical principles. *Proc. IEEE. Special issue on blind source separation*, 9(10):2009–2025, 1998.

[12] J.-F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Trans. Sig. Proc.*, 44(12):3017–3030, 1996.

[13] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20(1):33–61, 1998.

[14] D. N. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Trans. Commun.*, 28(11):1867–1875, 1980.

[15] A. Gorokhov and J.-F. Cardoso. Equivariant blind deconvolution of MIMO-FIR channels. In *Proc. of SPAWC*, 2003.

[16] P. Kisilev, M. Zibulevsky, and Y.Y. Zeevi. Multiscale framework for blind source separation. *JMLR*, 2003. In press.

[17] E. Moulines, J.-F. Cardoso, and E. Gassiat. Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models. 1997.

[18] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel fir filters. *IEEE Trans. Sig. Proc.*, 43:516–525, 1995.

[19] D. Pham and P. Garrat. Blind separation of a mixture of independent sources through a quasi-maximum likelihood approach. *IEEE Trans. Sig. Proc.*, 45:1712–1725, 1997.

[20] Y. Sato. A method of self-recovering equalization for multilevel amplitude-modulation systems. *IEEE Trans. Computers*, pages 679–682, 1975.

[21] D. Sengupta and S. Kay. Efficient estimation of parameters for non-Gaussian autoregressive processes. *IEEE Trans. Acoust., Speech, and Sig. Proc.*, 37(6):785–794, 1989.

[22] O. Shalvi and E. Weinstein. Maximum likelihood and lower bounds in system identification with non-Gaussian inputs. *IEEE Trans. Information Theory*, 40(2):328–339, 1994.

[23] L. Tong, G. Xu, and T. Kailath. A new approach to blind identification and equalization of multipath channel. In *Proc. 25th Asilomar Conf. Signal, Syst., Comput.*, 1991.

[24] J. R. Treichler and B. G. Agee. A new approach to the multi-path correction of constant modulus signals. *IEEE Trans. Acoust. Speech Sig. Proc.*, 31(2):331–344, 1983.

[25] G. Xu, H. Liu, L. Tong, and T. Kailath. Least squares approach to blind channel identification. *IEEE Trans. Sig. Proc.*, 43(12):2982–2993, 1995.

[26] D. Yellin and B. Friedlander. Multichannel system identification and deconvolution: performance bounds. *IEEE Trans. Sig. Proc.*, 47(5):1410–1414, 1999.

[27] D. Yellin and E. Weinstein. Multichannel signal separation: methods and analysis. *IEEE Trans. Sig. Proc.*, 44(1):106–118, 1996.

[28] M. Zibulevsky. Sparse source separation with relative Newton method. In *Proc. ICA2003*, pages 897–902, April 2003.

[29] M. Zibulevsky, P. Kisilev, Y. Y. Zeevi, and B. A. Pearlmutter. Blind source separation via multinode sparse representation. In *Proc. NIPS*. MIT Press, 2002.

[30] M. Zibulevsky, B. A. Pearlmutter, P. Bofill, and P. Kisilev. Blind source separation by sparse decomposition. In S. J. Roberts and R. M. Everson, editors, *Independent Components Analysis: Principles and Practice*. Cambridge University Press, 2001.