QML Blind Deconvolution: Asymptotic Analysis

Alexander M. Bronstein, Michael M. Bronstein, Michael Zibulevsky, and Yehoshua Y. Zeevi *

Technion - Israel Institute of Technology, Department of Electrical Engineering, 32000 Haifa, Israel
{alexbron, bronstein}@ieee.org, {mzib, zeevi}@ee.technion.ac.il

Abstract. Blind deconvolution is considered as a problem of quasi maximum likelihood (QML) estimation of the restoration kernel. Simple closed-form expressions for the asymptotic estimation error are derived. The asymptotic performance bounds coincide with the Cramér-Rao bounds, when the true ML estimator is used. Conditions for asymptotic stability of the QML estimator are derived. Special cases when the estimator is super-efficient are discussed.

1 Introduction

Blind deconvolution arises in various applications related to acoustics, optics, medical imaging, geophysics, communications, control, etc. In the noiseless setup of single-channel blind deconvolution, the observed sensor signal x is created from the *source signal* s passing through a convolutive system with impulse response a, $x = a * s$. The setup is termed *blind* if only x is accessible, whereas no knowledge on w and s is available. Blind deconvolution attempts to find such a deconvolution (restoration) kernel w, that produces a possibly delayed waveform-preserving source estimate $\hat{s}_n = (w * x)_n \approx c \cdot s_{n-\Delta}$, where c is a scaling factor and Δ is an integer shift. Equivalently, the *global system response* $g = a * w$ should be approximately a Kroenecker delta, up to scale factor and shift. A commonly used assumption is that s is non-Gaussian.

Asymptotic performance of maximum-likelihood parameter estimation in blind system identification and deconvolution problems was addressed in many previous studies (see, for example, [1–4]). 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 [5, 6], asymptotic performance analysis of QML estimators for blind source separation was presented.

In this study, we derive asymptotic performance bounds for a QML estimator of the restoration kernel in the single-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.

* This research has been supported by the HASSIP Research Network Program HPRN-CT-2002-00285, sponsored by the European Commission, and by the Ollendorff Minerva Center.

2 QML blind deconvolution

Under the assumption that the restoration kernel w 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 in the noise-free case is [7]

$$\ell(x;w) = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|W(e^{i\theta})\right| d\theta + \frac{1}{T}\sum_{n=0}^{T-1}\varphi\left(y_n\right), \tag{1}$$

where $W(e^{i\theta})$ stands for the discrete Fourier transform of w, $y = x * w$ is a source estimate, $\varphi(s) = -\log p\,(s)$ and $p(s)$ is the probability density function (PDF) of the source s_n. We will henceforth assume that the restoration kernel w_n has a finite impulse response, supported on $n = -N, ..., N$. We also assume without loss of generality that $\mathbb{E}s_n = 0$.

Consistent estimator can be obtained by minimizing $\ell(x;w)$ even when $\varphi(s)$ is not exactly equal to $-\log p\,(s)$. Such QML estimation has been shown to be practical in instantaneous blind source separation [5, 8] and blind deconvolution [9, 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 $\varphi(s)$ [8]. 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 [8]. $\varphi_\lambda(s) \to |s|$ as $\lambda \to 0^+$.

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 $\varphi(s)$ [9, 10].

2.1 Equivariance

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

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

Proof. Observe that for an invertible h,

$$\begin{aligned}\ell(h*x;h^{-1}*w) &= -\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left|\frac{W(e^{i\theta})}{H(e^{i\theta})}\right| d\theta + \frac{1}{T}\sum_{n=0}^{T-1}\varphi\left((x*w)_n\right)\\ &= \ell(x;w) + \frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left|H(e^{i\theta})\right| d\theta.\end{aligned}$$

Let $w = \operatorname{argmin} \ell(x; w)$. Then $\ell(h * x; h^{-1} * w) = \ell(x; w) + \text{const}$, hence w is a minimizer of $\ell(h * x; h^{-1} * w)$ as well. Consequently, $\hat{w}(h * x) = h^{-1} * \hat{w}(x)$. □

Equivariance implies that the parameters to be estimated (in our case, the coefficients, w_n, 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 $\ell(w * x; \delta_n)$ instead of $\ell(x; w)$.

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

The gradient and the Hessian of $\ell(x; w)$ in (1) are given by

$$\begin{aligned} \frac{\partial \ell(x;w)}{\partial w_k} &= \frac{\partial}{\partial w_k}\left(-\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|W(e^{i\theta})\right| d\theta + \frac{1}{T}\sum_{n=0}^{T-1}\varphi\left(y_n\right)\right) = \\ &= -\frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{e^{-i\theta k}}{W(e^{i\theta})} + \left(\frac{e^{-i\theta k}}{W(e^{i\theta})}\right)^*\right) d\theta + \frac{1}{T}\sum_{n=0}^{T-1}\varphi'\left(y_n\right)\frac{\partial y_n}{\partial w_k} \\ &= -w_{-k}^{-1} + \frac{1}{T}\sum_{n=0}^{T-1}\varphi'\left((x*w)_n\right) x_{n-k}, \end{aligned} \tag{4}$$

and

$$\begin{aligned} \frac{\partial^2 \ell(x;w)}{\partial w_k \partial w_l} &= \frac{\partial}{\partial w_l}\left(-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{-i\theta k}}{W(e^{i\theta})} d\theta + \frac{1}{T}\sum_{n=0}^{T-1}\varphi'\left((x*w)_n\right) x_{n-k}\right) \\ &= w_{-(k+l)}^{-2} + \frac{1}{T}\sum_{n=0}^{T-1}\varphi''\left((x*w)_n\right) x_{n-k} x_{n-l}, \end{aligned} \tag{5}$$

where w^{-1} denotes the impulse response of the inverse of w, and $w^{-2} = w^{-1} * w^{-1}$. At the solution point, where $w = ca^{-1}$, it holds that $x * w = cs$. Consequently the Hessian of $\ell(cs; \delta_n)$ is

$$\left(\nabla^2 \ell\right)_{kl} = \delta_{k+l} + \frac{c^2}{T}\sum_{n=0}^{T-1}\varphi''\left(cs_n\right) s_{n-k} s_{n-l}.$$

For a large sample size T, the average $\frac{1}{T}\sum_{n=0}^{T-1}\varphi''\left(cs_n\right) s_{n-k} s_{n-l}$ approaches the expected value $\mathbb{E}\varphi''\left(cs_n\right) s_{n-k} s_{n-l}$. Since s_n is assumed to be zero-mean i.i.d., the following structure of the Hessian at the solution point is obtained asymptotically:

$$\nabla^2 \ell(cs; \delta_n) \approx \begin{pmatrix} \ddots & & & & \iddots \\ & \gamma\sigma'^2 & & 1 & \\ & & \alpha c^2 + 1 & & \\ & 1 & & \gamma\sigma'^2 & \\ \iddots & & & & \ddots \end{pmatrix}, \tag{6}$$

where $\sigma^2 = \mathbb{E}s^2$, $\sigma'^2 = (c\sigma)^2$, $\alpha = \mathbb{E}\varphi''(cs)s^2$, and $\gamma = \mathbb{E}\varphi''(cs)$.

3 Asymptotic error covariance matrix

Let the restoration kernel, w, be estimated by minimizing the minus log likelihood function $\ell(x;w)$ defined in (1), where the true $-\log p(s)$ of the source is replaced by some other function $\varphi(s)$. We assume that w has sufficient degrees of freedom to accurately approximate the inverse of a. For analytic tractability, we assume that $\mathbb{E}\varphi''(cs)$, $\mathbb{E}s^2$, $\mathbb{E}\varphi''(cs)s^2$, $\mathbb{E}\varphi'^2(cs)$, $\mathbb{E}\varphi'(cs)s$ and $\mathbb{E}\varphi'^2(cs)s^2$ exist and are bounded. Note that the expected values are computed with respect to the true PDF of s.

Let $w^* = ca^{-1}$ be the exact restoration kernel (up to a scaling factor). It can be shown that w^* satisfies [11] $w^* = \operatorname{argmin}_w \ \mathbb{E}_x \ell(x;w)$. Let $\hat{w}$ be the estimate of the exact restoration kernel w^*, based on the finite realization of the data x, $\hat{w} = \operatorname{argmin}_w \ \ell(x;w)$. Note that $\nabla\ell(x;\hat{w}) = 0$, whereas $\nabla\ell(x;w^*) \neq 0$; yet $\mathbb{E}\nabla\ell(x;w^*) = 0$. Denote the estimation error as $\Delta w = w^* - \hat{w}$. Then, assuming $\|\Delta w\|$ is small, second-order Taylor expansion yields

$$\nabla\ell(x;w^*) \approx \nabla^2\ell(x;w^*)\cdot(w^* - \hat{w}) \ = \ \nabla^2\ell(x;w)\big|_{w=a^{-1}} \cdot \Delta w.$$

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

$$\nabla\ell(w^* * x;\delta_n) \approx \nabla^2\ell(w^* * x;\delta_n)\cdot\Delta w.$$

Since $w^* = ca^{-1}$, we can substitute $w^* * x = cs$, and obtain $\nabla\ell(cs;\delta_n) \ \approx \ \nabla^2\ell(cs;\delta_n)\cdot \Delta w$, or, alternatively, $\Delta w \approx \nabla^2\ell(cs;\delta_n)^{-1}\cdot\nabla\ell(cs;\delta_n)$. For convenience, we will denote $\nabla\ell(cs;\delta_n)$ and $\nabla^2\ell(cs;\delta_n)$ as g and $\nabla^2\ell$, respectively. The covariance matrix of Δw is therefore given by

$$\Sigma_{\Delta w} = \mathbb{E}\Delta w \Delta w^T \ \approx \ \left(\nabla^2\ell\right)^{-1}\cdot\mathbb{E}\nabla\ell\nabla\ell^T\cdot\left(\nabla^2\ell\right)^{-T} \ = \ \left(\nabla^2\ell\right)^{-1}\cdot\Sigma_{\nabla\ell}\cdot\left(\nabla^2\ell\right)^{-1}.$$

For a large sample size, the asymptotic Hessian structure (6) can be used, allowing to split the asymptotic covariance matrix, $\Sigma_{\Delta w}$, into a set of 2×2 symmetric matrices of the form

$$\Sigma^{(k)}_{\Delta w} = \begin{pmatrix} \mathbb{E}(\Delta w_{-k})^2 & \mathbb{E}\Delta w_k \Delta w_{-k} \\ \mathbb{E}\Delta w_k \Delta w_{-k} & \mathbb{E}(\Delta w_k)^2 \end{pmatrix} \ \approx \ \begin{pmatrix} \gamma\sigma'^2 & 1 \\ 1 & \gamma\sigma'^2 \end{pmatrix}^{-1} \Sigma^{(k)}_{\nabla\ell} \begin{pmatrix} \gamma\sigma'^2 & 1 \\ 1 & \gamma\sigma'^2 \end{pmatrix}^{-1} \quad (7)$$

for $k \neq 0$, where $\Sigma^{(k)}_{\nabla\ell}$ is the covariance matric of g_{-k}, g_k, and an additional 1×1 element

$$\Sigma^{(0)}_{\Delta w} = \frac{\mathbb{E}g_0^2}{\left(\alpha c^2+1\right)^2}. \quad (8)$$

That is, the asymptotic error covariance matrix has a digaonal-anti-diagonal form. This implies that $\operatorname{cov}\Delta w_k \Delta w_{k'}$, for $k \neq k'$, $k \neq -k'$, decreases in the order of $1/T^2$ as $T \to \infty$. Taking the expectation of the gradient g_k, one obtains $\mathbb{E}g_k = -\delta_k + \mathbb{E}\varphi'(cs_n)cs_{n-k}$. Demanding $\mathbb{E}g_k = 0$, we obtain the following condition:

$$\mathbb{E}\varphi'(cs)cs = 1, \quad (9)$$

from where the scaling factor c can be found. Let us now evaluate the 2×2 gradient covariance matrix, $\Sigma_{\nabla\ell}^{(k)}$, for $k \neq 0$. Substituting $w = \delta_n$, $x = cs$ into (4) yields

$$g_k = \frac{\partial \ell(cs;\delta_n)}{\partial w_k} = -\delta_k + \frac{c}{T}\sum_n \varphi'(cs_n)s_{n-k}, \tag{10}$$

which for $k \neq 0$ reduces to $g_k = \frac{1}{T}\sum_n \varphi'(cs_n)cs_{n-k}$. Taking the expectation w.r.t. s, and neglecting second-order terms, we obtain

$$\mathbb{E}g_k^2 = \frac{c^2}{T^2}\sum_{n,n'} \mathbb{E}\left\{\varphi'(cs_n)\varphi'(cs_{n'})\ s_{n-k}s_{n'-k}\right\} \ \approx\ \frac{c^2}{T}\mathbb{E}\varphi'^2(cs)\mathbb{E}s^2 = \frac{1}{T}\beta\sigma'^2$$

$$\mathbb{E}g_{-k}\ g_k = \frac{c^2}{T^2}\sum_{n,n'} \mathbb{E}\left\{\varphi'(cs_n)\varphi'(cs_{n'})\ s_{n+k}s_{n'-k}\right\} \ \approx\ \frac{1}{T}\mathbb{E}^2\varphi'(cs)cs = \frac{1}{T},$$

that is,

$$\Sigma_{\nabla\ell}^{(k)} \approx \frac{1}{T}\cdot\begin{pmatrix}\beta\sigma'^2 & 1\\ 1 & \beta\sigma'^2\end{pmatrix},$$

where $\beta = \mathbb{E}\varphi'^2(cs)$. Substituting the former result to (7) yields after some algebraic manipulations

$$\mathrm{var}\Delta w_k \approx \frac{\beta\sigma'^2\left(\gamma^2\sigma'^4+1\right) - 2\gamma\sigma'^2}{T\left(\gamma^2\sigma'^4-1\right)^2} \tag{11}$$

$$\mathrm{cov}\Delta w_{-k}\Delta w_k \approx \frac{\gamma\sigma'^2\left(\gamma\sigma'^2 - 2\beta\sigma'^2\right)+1}{T\left(\gamma^2\sigma'^4-1\right)^2} \tag{12}$$

for $k \neq 0$. Note that the asymptotic variance depends on the sample size T and on parameters β, γ, c and σ'^2, which depend on the source distribution and on $\varphi(s)$ only.

Let us now address the case of $k = 0$. Neglecting second-order terms, the second moment of g_0 is given by

$$\mathbb{E}g_0^2 \approx -1 - 2\mathbb{E}\varphi'(cs)cs + \mathbb{E}^2\varphi'(cs)cs + \frac{1}{T}\left(\mathbb{E}\varphi'^2(cs)(cs)^2 - \mathbb{E}^2\varphi'(cs)cs\right) = \frac{c^2\vartheta - 1}{T},$$

where $\vartheta = \mathbb{E}\varphi'^2(cs)s^2$. Hence, $\Sigma_{\nabla\ell}^{(0)} \approx (c^2\vartheta - 1)/T$. Substituting $\Sigma_{\nabla\ell}^{(0)}$ into (8) yields

$$\mathrm{var}\Delta w_0 \approx \frac{c^2\vartheta - 1}{T(\alpha c^2+1)^2}. \tag{13}$$

Using $\mathrm{var}\Delta w_k$, an asymptotical estimate of restoration quality in terms of signal-to-interference ratio (SIR) can be expressed as

$$\mathrm{SIR} = \frac{\mathbb{E}\|cs\|_2^2}{\mathbb{E}\|w*x - cs\|_2^2} = \frac{|w_0^*|^2}{\mathbb{E}\|\Delta w\|_2^2} \ \approx\ \frac{T\ \left(\gamma^2\sigma'^4-1\right)^2}{2N\left(\beta\sigma'^2\left(\gamma^2\sigma'^4+1\right) - 2\gamma\sigma'^2\right)}. \tag{14}$$

3.1 Cramér-Rao Lower Bounds

We now show that the asymptotic variance of the estimation error in (11), (13) matches the CRLB on the asymptotic variance of $\hat{w}_k$, when the true MLE procedure is used, i.e., when $\varphi(s) = -\log p(s)$. In this case, $c = 1$, $\sigma'^2 = \sigma^2$, and under the assumption that $\lim_{s\to\pm\infty} p(s) = 0$, it can be shown [12] that $\gamma = \beta$. Substituting c, σ'^2, γ into (11), we obtain for $k \neq 0$

$$\mathrm{var}\Delta w_k \approx \frac{\beta\sigma^2}{T\left(\beta^2\sigma^4 - 1\right)} = \frac{1}{T} \cdot \frac{\mathcal{L}}{\mathcal{L}^2 - 1},$$

where $\mathcal{L} = \sigma^2 \cdot \mathbb{E}\varphi'^2(s)$ is known as Fisher's information for location parameter [4]. This result coincides with the CRLB on w_k developed in [4]. Similarly, under the assumption that $\lim_{s\to\pm\infty} p(s)s = 0$, it can be shown that $\theta = \alpha + 2$ [12]. Substituting $c = 1$ and the latter result into (13) yields

$$\mathrm{var}\Delta w_0 \approx \frac{1}{T} \cdot \frac{\vartheta - 1}{(\alpha + 1)^2} = \frac{1}{T} \cdot \frac{1}{\alpha + 1} \;=\; \frac{1}{T\mathcal{S}},$$

where $\mathcal{S} = \mathrm{cum}\left\{\varphi'(s), \varphi'(s), s, s\right\} + \mathcal{L} + 1$ is the Fisher information for the scale parameter [4]. This result coincides with the CRLB on w_0 in [4]. Substituting the obtained β and γ into (14), yields

$$\mathrm{SIR} \approx \frac{T\left(\mathcal{L}^2 - 1\right)}{2N \cdot \mathcal{L}} \leq \frac{T\mathcal{L}}{2N}.$$

This result coincides with the asymptotic performance bound derived in [4].

3.2 Super-efficiency

Let us now consider the particular case of *sparse* sources, such sources that take the value of zero with some non-zero probability $\rho > 0$. An example of such distribution is the Gauss-Bernoully (sparse normal) distribution [12]. 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,

$$\gamma = \mathbb{E}\varphi''(cs) \approx \frac{1}{\lambda} \int_{-\lambda/c}^{+\lambda/c} p(s)\, ds \approx \frac{\rho}{\lambda},$$

whereas β and c are bounded. Consequently, for $k \neq 0$

$$\underset{T\to\infty}{\mathrm{plim}}\; T \cdot \mathrm{var}\Delta w_k \leq \frac{\beta}{\gamma^2\sigma'^2} \leq \mathrm{const} \cdot \lambda^2, \tag{15}$$

where plim denotes the probability limit. Observe that this probability limit vanishes for $\lambda \to 0^+$, which means that the estimator $\hat{w}_k$ of w_k is *super-efficient*. Similarly, the sub-Gaussian QML estimator with $\varphi_\mu(s)$ defined in (3) is super-efficient for sources with compactly supported PDF.

4 Asymptotic stability

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

Proposition 2. *Let $\hat{w}(x)$ be the QML estimator of w. $\hat{w}(x)$ is asymptotically stable if the following conditions hold:*

$$\gamma > 0, \tag{16}$$
$$\gamma^2\sigma'^4 > 1, \tag{17}$$
$$\alpha c^2 > -1. \tag{18}$$

Proof. The QML estimator is asymptotically stable if in the limit $T \to \infty$, $w = w^*$ is a local minimizer of $\ell(x;w)$, or due to equivariance, $w = \delta_n$ is a local minimizer of $\ell(cs;w)$. The first- and the second-order Karush-Kuhn-Tucker conditions

$$\operatorname*{plim}_{T\to\infty} \nabla\ell(cs;\delta_n) = 0 \tag{19}$$
$$\operatorname*{plim}_{T\to\infty} \nabla^2\ell(cs;\delta_n) \succ 0 \tag{20}$$

are the necessary and the sufficient conditions, respectively, for existence of the local minimum. The necessary condition (19) requires that $\nabla\ell = 0$ as the sample size approaches infinity. For $k \neq 0$ we obtain from (10) that $\operatorname{plim}_{T\to\infty} g_k = \mathbb{E}\varphi'(cs) \cdot \mathbb{E}\, cs = 0$, and for $k = 0$, by choice of c, $\operatorname{plim}_{T\to\infty} g_0 = \mathbb{E}\varphi'(cs)cs - 1 = 0$. The sufficient condition (20) requires that $\nabla^2\ell \succ 0$ as the sample size approaches infinity. Using the asymptotic Hessian given in (6), this condition can be rewritten as

$$\begin{pmatrix} \gamma\sigma'^2 & 1 \\ 1 & \gamma\sigma'^2 \end{pmatrix} \succ 0, \qquad \alpha c^2 + 1 \succ 0.$$

The latter holds if and only if $\gamma > 0$, $\gamma^2\sigma'^4 > 1$ and $\alpha c^2 > -1$. □

It is observed that when $\varphi(s)$ is chosen to be proportional to $-\log p(s)$, $\hat{w}(x)$ is never asymptotically unstable. When $\varphi(s)$ is chosen according to (3), it can be shown that $c = (\mu \cdot \mathbb{E}|s|^\mu)^{-1/\mu}$, $\alpha = \mu(\mu-1)c^\mu \cdot \mathbb{E}|s|^\mu$, and $\gamma = \mu(\mu-1)c^{\mu-2} \cdot \mathbb{E}|s|^{\mu-2}$. For $\mu > 2$, it can be easily checked that conditions (16), (18) hold, hence, the asymptotic stability condition is $\mathbb{E}|s|^\mu < (\mu-1)\mathbb{E}s^2\ \mathbb{E}|s|^{\mu-2}$. In the particular case when $\mu = 4$, the latter condition becomes $\kappa < 0$, where κ is the *kurtosis excess*, meaning that the estimator is asymptotically stable for sub-Gaussian sources.

When $\varphi(s)$ is chosen according to (2), there exists no analytic expression for the asymptotic stability conditions, except the case when $\lambda \to 0^+$. In the latter case, $\varphi' = \operatorname{sign}(s)$ and $\varphi''(s) = 2\delta(s)$, from where $c = 1/\mathbb{E}|s|$, $\alpha = 2\mathbb{E}\delta(s)(cs)^2 = 0$, and $\gamma = 2\mathbb{E}\delta(s) = 2p(0)$. Observe that conditions (16), (18) hold again, hence the estimator is asymptotically stable if $\mathbb{E}|s| < 2p(0)\sigma^2$.

5 Conclusion

In order to be in a position to utilize the QML estimator of the restoration kernel in blind deconvolution, and to gain insight into the effect of the source distribution and the choice of $\varphi(s)$, 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 matrix has a diagonal-anti-diagonal 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 $\varphi(s)$ essentially through the ratio $\mathbb{E}\varphi'^2(s)/\mathbb{E}^2\varphi''(s)$ of non-linear moments of the source. We demonstrated that for the true ML estimator, our asymptotic performance bounds coincide with the CRLB. Asymptotic stability conditions for the QML estimator have been shown as well. Extension to the MIMO case is presented in [13]. Particular cases wherein the families of functions $\varphi_\lambda(s)$ and $\varphi_\mu(s)$ yield super-efficient estimators were highlighted. More delicate analysis is required to determine whether zero variance can be achieved on a *finite* sample, and what is its minimum size. Such a result is important from both theoretical and practical viewpoints.

References

1. Bellini, S., Rocca, F.: Near optimal blind deconvolution. In: Proc. of IEEE Conf. Acoust., Speech, Sig. Proc. (1988)
2. Cardoso, J.F., Laheld, B.: Equivariant adaptive source separation. IEEE Trans. Sig. Proc. **44** (1996) 3017–3030
3. Shalvi, O., Weinstein, E.: Maximum likelihood and lower bounds in system identification with non-Gaussian inputs. IEEE Trans. Information Theory **40** (1994) 328–339
4. Yellin, D., Friedlander, B.: Multichannel system identification and deconvolution: performance bounds. IEEE Trans. Sig. Proc. **47** (1999) 1410–1414
5. Pham, D., Garrat, P.: Blind separation of a mixture of independent sources through a quasi-maximum likelihood approach. IEEE Trans. Sig. Proc. **45** (1997) 1712–1725
6. Cardoso, J.F.: Blind signal separation statistical principles. Proc. IEEE. Special issue on blind source separation **9** (1998) 2009–2025
7. Amari, S.I., Cichocki, A., Yang, H.H.: Novel online adaptive learning algorithms for blind deconvolution using the natural gradient approach. In: Proc. SYSID. (1997) 1057–1062
8. Zibulevsky, M.: Sparse source separation with relative Newton method. In: Proc. ICA2003. (2003) 897–902
9. Amari, S.I., Douglas, S.C., Cichocki, A., Yang, H.H.: Multichannel blind deconvolution and equalization using the natural gradient. In: Proc. SPAWC. (1997) 101–104
10. Bronstein, A.M., Bronstein, M.M., Zibulevsky, M.: Blind deconvolution with relative Newton method. Technical Report 444, Technion, Israel (2003)
11. Kisilev, P., Zibulevsky, M., Zeevi, Y.: Multiscale framework for blind source separation. JMLR (2003) In press.
12. Bronstein, A.M., Bronstein, M.M., Zibulevsky, M., Zeevi, Y.Y.: Quasi maximum likelihood blind deconvolution: Asymptotic performance analysis. IEEE Info. Theory (2004) Submitted. [Online] http://visl.technion.ac.il/bron/alex.
13. Bronstein, A.M., Bronstein, M.M., Zibulevsky, M., Zeevi, Y.Y.: Asymptotic performance analysis of MIMO blind deconvolution. Technical report, Technion, Israel (2004)