Blind Estimation of the Ocean Acoustic Channel by Time-Frequency Processing

Nelson E. Martins and Sérgio M. Jesus[1]

Abstract

A blind estimator of the ocean acoustic channel impulse response envelope is presented. The signal model is characterized by a deterministic multipath channel excited by a highly non-stationary deterministic source signal. The time-frequency representation of the received signal allows for the separation between the channel and the source signal. The proposed estimator proceeds in two steps: first, the unstable initial arrivals allow for the estimation of the source signal instantaneous frequency, by maximization of the radially Gaussian kernel distribution; then, the Wigner-Ville distribution is sequentially windowed and integrated, where the window is defined by the previously estimated instantaneous frequency. The integral gives the channel impulse response envelope, which turns to be an approximation to the blind conventional matched-filter. The blind channel estimator is applicable upon the following conditions: that the multipath channel contains at least one dominant arrival well separated from the others, and that the instantaneous frequency of the source signal is an one-to-one function. Results obtained on real data from the INTIMATE '96 underwater experiment, where the acoustic channel was driven by an LFM signal, show that the channel's envelope detailed structure could be accurately and consistently recovered, with the correlation of the estimates ranging from 0.796 to 0.973, as compared to the matched-filter result.

Index Terms

Blind channel estimation, multipath, time-frequency, LFM.

EDICS Category: 2-TIFR, 3-CEQU

Work supported by FCT –MCES, scholarship SFRH/BD/9032/2002–, Portugal.

[1]N.E. Martins and S.M. Jesus are with SiPLAB - FCT – Universidade do Algarve, Campus de Gambelas, PT-8000 Faro, Portugal. E-mails: {nmartins,sjesus}@ualg.pt.

I. INTRODUCTION

Blind channel estimation is a topic of intense research in the underwater acoustics community[1], [2], as well as in signal processing-related fields such as wireless communications[4], geophysics[5], etc.. Being essentially transparent to sound propagation, the ocean is a favorable medium for transmitting information by acoustic signals, either for biological or human purposes. This information is contained on the signal emitted by the acoustic source, and is sensed by an acoustic receiver, which can be positioned from meters to hundreds of kilometers from the source. The received signal obviously contrasts with the emitted signal, due to the severe distortion introduced by the medium boundaries, the space- and time-variable sound-speed, bottom properties, source and receiver coordinates, among other factors. Both emitted and received signals, and the ocean distortion, can be modeled by time- (or space-) dependent functions. Assuming the ocean acoustic channel as linear time-invariant, the received signal can be modeled as a noisy convolution of the emitted signal with the channel impulse response. At the receiver end, the purpose is usually to interpret the information contained in the source signal, for example in digital communications[3], or to estimate the physical or geometric (*e.g.* water depth, source coordinates) parameters responsible for signal distortion, which is the purpose of ocean acoustic tomography[25], [26]. This purpose can be achieved without major difficulty, if the source signal is known at the receiver. Otherwise, in ocean monitoring scenarios where the discretion of the receiver is important, as in biological studies or military applications, the problem of channel estimation transforms itself onto a blind estimation problem, in that the only information available to the receiver is the received signal, which encapsulates both the emitted signal, the channel distortion and the inevitable noise. Consequently, there is an increasing need for the development of blind receivers capable of estimating the ocean physical parameters, or the emitted signal, in presence of this unknown excitation source signal. Often, these signal is a non-stationary transient that propagates from the source to the receiver, after severe channel distortion and noise corruption. Recovering the ocean distortion amounts to solving a blind signal deconvolution problem, by modelling the ocean channel as linear and time-invariant. Blind deconvolution is itself a research area, with specific principles and methods.

Classical deconvolution is a well-established inversion method, which allows for the estimation of either the channel or the source signal. When the source signal is known, a correlator

receiver can be used for channel estimation, which constitutes an application of the classical matched-filter. As is well known, the performance of this estimator is dependent on the source signal's bandwidth, which has a direct impact on the estimate quality. For low signal-to-noise ratio environments, the correlator can be improved with a prior denoising step. This has been done in [6], by using a chirp signal as training sequence to estimate the impulse response of the channel. The proposed method makes use of the high concentration of the chirp signal when compared to the noise spreading in the joint time-frequency plane. In essence, a time-variant filter is applied to the received signal, resorting to the Gabor transform, prior to conventional channel estimation. If the channel has a multipath structure, which is a reliable approximation in the ocean channel case, then channel estimation can be transformed into a time-delay estimation problem[7], [8], [9]. This can also be addressed in the spectral domain, as proposed by Kirsteins[10]. This method is based on fitting weighted complex exponentials to the spectrum of the received signal. In this case, the drawbacks of ill-conditioning and bias, and the requirement of considering contiguous frequency samples, led to the development of other algorithms, which attempt to find least-squares unbiased estimates of the time-delays[1]. In this last reference, the techniques were extended to the blind deconvolution of a gated sinusoid, where the sinusoid and channel parameters are searched simultaneously. A common problem in signal deconvolution is ill-poseness. In underwater acoustics, this problem was dealt with by observing the source signal through a multichannel array[11], [12]. The method presented in [11] is based on *maximum a posteriori* likelihood estimation, assuming a stochastic model for the source signal. The approach proposed by Smith *et al.*[12] solves an iterative minimization problem, using simulated annealing. In this last reference, the authors mention that the proposed method can be extended to the case of the simultaneous estimation of the source signal and the multichannel impulse response.

Blind deconvolution methods that deviate from classical deconvolution generally involve higher--order statistics of the source signal or the channel. One important requirement is that the involved signals satisfy ergodicity and stationarity. The techniques are based on the principle that the higher-order moments of the input signal are decreased by convolution with a linear and time-invariant channel. The problem is solved by maximization of the higher-order moments (equivalently, minimization of the entropy) of the deconvolved quantities. Finding obvious applications, for example, in wireless communications, blind deconvolution has, in this field, exploited also the cyclostationary properties of the received signal. As an example, Gardner[13]

has proposed a method for channel identification based on the second-order cyclic autocorrelation function. This method involves the use of a training period, during which, the unobserved channel input is transmitted at a slow rate, making the inter-symbol interference negligible. The need for a training period can however fall off (increasing the available bandwidth), for certain channel models, as shown by Tong *et al.*[14], using extensions of Gardner's original method. A comprehensive study on the use of blind deconvolution on communication systems was presented in [4]. Blind deconvolution methods dealing with higher-order statistics have been applied also in geophysics[5], to estimate the reflection coefficients of a layered Earth model. The same principles were recently applied, in underwater acoustics, to the deconvolution of signals with a moderate degree of temporal non-stationarity[2]. However, the deconvolution result is influenced by a free parameter –the filter length–, and there remains an open question on whether such principles could be applied to signals with a high degree of non-stationarity.

This paper presents a blind estimator of the ocean acoustic channel, excited by a highly non-stationary source (probe) signal. Assuming deterministic models for the channel and source signal, the method is divided into two steps. The first step consists of the estimation of the source signal instantaneous frequency. In the second step, the Wigner-Ville distribution of the received signal is integrated along delayed versions of the source signal instantaneous frequency estimate, giving as output the channel impulse response envelope estimate.

The paper extends the work presented in [15], [16], and is structured as follows. Sec. II defines the problem and briefly sketches its solution, by time-frequency methods; the time-frequency estimator is presented and characterized in Sec. III. Experimental results with underwater acoustic data are presented in Sec. IV, and some conclusions and perspectives are drawn on Sec. V.

II. Problem Statement and Time-Frequency Approach

Consider both an acoustic source and receiver positioned at given locations in the ocean. Assuming that their positions are stationary along the propagation time (here, $\approx$4.5 s), and that the environmental conditions do not change significantly along this same period, the acoustic channel linking the source to the receiver can be characterized by a linear and time-invariant impulse response. Note that for different source and/or receiver positions, and/or for environmental temporal variations on time scales that are larger as compared to the observation time (in the considered experimental setup, this time is $\approx$1.5 min, see section IV), as *e.g.* tidal-induced

water depth variation, the source and received signals will be linked by time-varying impulse responses. Thus, during the observation time, the received signal $r(t)$ can be modeled by a noisy and channel distorted version of the deterministic source signal:

$$\begin{aligned} r(t) &= x(t) + n(t) \\ &= s(t) * h(t) + n(t) \end{aligned} \tag{1}$$

where $h(t)$ is the impulse response of the ocean channel, $n(t)$ is zero-mean stationary additive noise, $*$ stands for convolution, and $s(t)$ is the complex representation (analytic signal) of the real source signal:

$$s(t) = a_i(t)\mathrm{e}^{\mathrm{j}\varphi_i(t)}, \tag{2}$$

where $a_i(t)$ and $\varphi_i(t)$ are the instantaneous amplitude (IA) and phase, respectively. The instantaneous frequency (IF) of $s(t)$ is readily given by

$$f_i(t) = \frac{1}{2\pi}\frac{\mathrm{d}\varphi_i(t)}{\mathrm{d}t}. \tag{3}$$

One key point, in considering an impulse response model for the ocean acoustic channel, is its structure, which, when given by a series of delayed impulses, generally constitutes a realistic model. Further, the initial impulses are typically very close in time, and at least one of them has a high amplitude as compared to the remaining well-separated impulses. The initial high-amplitude impulse(s) is(are) due to direct eigenrays or eigenrays with a small number of reflections, while the well-separated impulses are due to several surface and/or sediment reflections.

The problem at hand is, knowing $r(t)$, to recover the impulse response envelope $|h(t)|$, assuming no knowledge about $s(t)$, which constitutes a classical blind channel (envelope) estimation problem. The present time-frequency approach does not allow to estimate the phase of the impulse response, since the procedure is based on time-frequency integration, which gives a measure of signal power. Nevertheless, if this impulse response estimate is to be used *e.g.* in travel-time acoustic tomography, where only the time-delays of the impulse response are important, the envelope estimate is sufficient for post-processing. Hence, the parameter of interest is designated by

$$\theta(t) = |h(t)|\,. \tag{4}$$

To solve this problem by a time-frequency approach, it is important to evidence the following key ideas, which emerge as conditions of application:

- The acoustic channel is described by a set of arrivals, with at least one having a high amplitude relatively to the other arrivals;
- The acoustic source emits a deterministic and non-stationary signal, whose instantaneous frequency is an injective (one-to-one) mathematical function.

The channel estimation procedure is schematized in figure 1. As illustrated in this figure, the first step relies on the representation of the received signal by a signal-dependent time-frequency distribution. The maximization of this distribution gives the source signal instantaneous frequency estimate. The second step consists of coherently measuring the Wigner-Ville distribution energy of the received signal, on time-frequency supports defined by the source signal instantaneous frequency estimate. The obtained estimator is a sub-optimal blind channel estimator, as compared to the matched-filter estimator.

III. Channel Estimation

Taking into account that non-stationary signals are common in practical applications, TF processing appears in such cases as a natural tool. TF representations indeed provide a means for detailed analysis of the local TF structure of the received signal. As a consequence, if the channel impulse response and the source signal have 'distinct' TF patterns, they can be separated and recovered in the TF domain. In this work, the Wigner-Ville and the radially Gaussian kernel[17] TF distributions (TFDs) are used.

A. Correlator Channel Estimator

Let us take, as a reference channel estimator, the correlator receiver. This estimator makes use of the principle behind the matched-filter (MF), and can be used only when both received and source signals are available. The correlator receiver output is given by

$$\hat{\theta}_{MF}(t) \quad = \quad \left| \int_{-\infty}^{\infty} r(\tau) s^*(\tau - t) \mathrm{d}\tau \right| \tag{5}$$

and is used as a term of comparison to evaluate the performance of the TF blind channel estimator, concerning the real data results presented in Sec. IV.

B. Time-Frequency Blind Channel Estimator

This section describes the TF blind channel estimator. A non-blind directly related estimator is first presented, and its similarity to the matched-filter is discussed. Then, it is seen that the multipath nature of the channel allows for the derivation of an estimator of the IF of $s(t)$, and the consecutive definition of the blind channel estimator.

1) Non-Blind Estimator: Let us state the definition of a TF non-blind channel estimator, by first considering the formulation of the MF in the TF domain. This can be accomplished by means of the classical Moyal formula, which relates the overlap of two signals in the time domain with the overlap of their respective bilinear unitary TFDs[18], [19]:

$$\begin{aligned}&\iint_{-\infty}^{\infty} C_{x_1x_2}(\tau,f)C^*_{x_3x_4}(\tau,f)\ \mathrm{d}\tau\ \mathrm{d}f\\ &\qquad=\left[\int_{-\infty}^{\infty} x_1(\tau)x_3^*(\tau)\mathrm{d}\tau\right]\left[\int_{-\infty}^{\infty} x_2(\tau)x_4^*(\tau)\mathrm{d}\tau\right]^*\end{aligned} \tag{6}$$

where $C_{yz}(\tau,f)$ designates any unitary cross-distribution of the generic signals $y(\tau)$ and $z(\tau)$. For the particular cases of $x_1(\tau)=x_2(\tau)=r(\tau)$ and $x_3(\tau)=x_4(\tau)=s(\tau-t)$, and when the unitary distribution $C_{yz}(\tau,f)$ is the Wigner-Ville distribution $WV_{yz}(\tau,f)$, equation (6) transforms into

$$\begin{aligned}&\iint_{-\infty}^{\infty} WV_{rr}(\tau,f)WV^*_{ss}(\tau-t,f)\ \mathrm{d}\tau\mathrm{d}f\\ &\qquad=\left|\int_{-\infty}^{\infty} r(\tau)s^*(\tau-t)\mathrm{d}\tau\right|^2,\end{aligned} \tag{7}$$

where the second term is simply a squared version of the MF channel estimate (5). This expresses the equivalence between the MF and a correlation in the TF domain with respect to the temporal variable. This equivalence has already been used in the context of detection, in [20]. Thus, if the source signal was known at the receiver, an MF estimator could be implemented in the time domain, by (5), or equivalently in the TF domain, by the TF-based channel estimator

$$\hat{\theta}_{MF}(t)\equiv\sqrt{\iint_{-\infty}^{\infty} WV_{rr}(\tau,f)WV^*_{ss}(\tau-t,f)\mathrm{d}\tau\mathrm{d}f}. \tag{8}$$

In the case that the IA of the source signal does not change significantly with time, the spread in frequency of $WV_{ss}(\tau, f)$ can be approximated by[21]:

$$\sigma^2_{f|\tau} = \frac{1}{2}\left[\left(\frac{a_i'(\tau)}{a_i(\tau)}\right)^2\right], \tag{9}$$

where $a_i'(\tau)$ designates the derivative of $a_i(\tau)$. The value of $\sigma^2_{f|\tau}$ will be small, for a slowly--varying function $a_i(\tau)$, meaning that the Wigner-Ville distribution (WV) is highly concentrated along the IF. This will be the case when considering the experimental data, in section IV. Upon this assumption, one can obtain an approximation to (8), by approximating $WV_{ss}(\tau, f)$ as:

$$WV_{ss}(\tau, f) \approx \delta\left[f - f_i(\tau)\right]. \tag{10}$$

Upon insertion of (10) into (8), a simply structured TF non-blind channel estimator is obtained:

$$\hat{\theta}_{TFNB}(t) = \sqrt{\int_{-\infty}^{\infty} WV_{rr}\left[\tau, f_i(\tau - t)\right] \mathrm{d}\tau}. \tag{11}$$

Note that this estimator is sub-optimal with respect to the MF, in terms of output signal-to-noise ratio, due to the approximation (10). Let us define the conditions upon which, (11) can be applied as a channel estimator. Taking into account (1) and the fact that the WV satisfies the convolution property, (11) can be developed as

$$\hat{\theta}_{TFNB}(t) = \sqrt{\mathcal{S}(t) + \mathcal{N}(t) + \mathcal{SN}(t)}, \tag{12}$$

where $\mathcal{S}(t)$ refers to the signal component,

$$\mathcal{S}(t) = \iint_{-\infty}^{\infty} WV_{hh}(u, f)\Psi(u, f, t)\mathrm{d}f\,\mathrm{d}u \tag{13}$$

with

$$\Psi(u, f, t) = \int_{-\infty}^{\infty} \delta\left[f - f_i(\tau - t)\right] WV_{ss}(\tau - u, f)\mathrm{d}\tau, \tag{14}$$

and $\mathcal{N}(t)$ and $\mathcal{SN}(t)$ refer to the noise and crossed signal-noise components,

$$\begin{aligned}\mathcal{N}(t) &= \iint_{-\infty}^{\infty} \delta\left[f - f_i(\tau - t)\right] WV_{nn}(\tau, f)\mathrm{d}\tau\,\mathrm{d}f \\ &= \int_{-\infty}^{\infty} WV_{nn}\left[\tau, f_i(\tau - t)\right] \mathrm{d}\tau\end{aligned} \tag{15}$$

and

$$\mathcal{SN}(t) = 2\iint_{-\infty}^{\infty} \delta\left[f - f_i(\tau - t)\right] \mathrm{Re}\left\{WV_{xn}(\tau, f)\right\} \mathrm{d}\tau \mathrm{d}f$$
$$= 2\int_{-\infty}^{\infty} \mathrm{Re}\left\{WV_{xn}\left[\tau, f_i(\tau - t)\right]\right\} \mathrm{d}\tau, \tag{16}$$

respectively. Let us analyze only the signal component $\mathcal{S}(t)$. Note that if, at every frequency f,

$$\int_{-\infty}^{\infty} \delta\left[f - f_i(t)\right] WV_{ss}(t - \tau, f)\mathrm{d}t \quad \approx \quad \delta(\tau) \tag{17}$$

then (13) can be approximated by:

$$\mathcal{S}(t) \quad \approx \quad \iint_{-\infty}^{\infty} WV_{hh}(u, f)\delta(u - t)\mathrm{d}f\,\mathrm{d}u, \tag{18}$$

which, for non-noisy data, leads to the approximate channel estimate

$$\hat{\theta}_{TFNB}(t) \quad \approx \quad |h(t)|\,. \tag{19}$$

Approximation (17) holds, as long as:

(i) at every frequency f, the distribution $WV_{ss}(t, f)$ has at maximum one peak;

(ii) the distribution $WV_{ss}(t, f)$ is infinitely concentrated along $f_i(t)$.

For condition (i) to hold, it is necessary that $f_i(t)$ be an injective (or one-to-one) mathematical function –a function which maps distinct values of t to distinct values of $f_i(t)$. For condition (ii) to hold, it is necessary that the WV of the signal $s(t)$ be defined by a bi-dimensional Dirac distribution centered on $f_i(t)$:

$$WV_{ss}(t, f) \quad = \quad |c|^2\,\delta\left[f - f_i(t)\right] \tag{20}$$

where c is a complex scalar. It is well established that the class of deterministic source signals for which, both conditions are satisfied, entails the linear frequency modulation signal (LFM)

$$s(t) \quad = \quad c\,\mathrm{e}^{\mathrm{j}2\pi\left(\frac{\alpha}{2}t^2 + f_0 t + \varphi_0\right)} \tag{21}$$

where α, f_0 and φ_0 designate the modulation rate, initial frequency and initial phase, respectively, and the Dirac distribution

$$s(t) \quad = \quad c\,\delta(t). \tag{22}$$

Note that the requirements for optimality of the TF channel estimator are very similar to the requirement of an impulsive autocorrelation of the emitted signal, for the MF estimator. In the former case, the emitted signal is required to have a concentrated signature in the TF plane, while in the later, it is required to have a concentrated autocorrelation signature. In practice, none of these conditions can be fully verified, which leads to sub-optimal estimators. In Sec. IV, the results obtained with a finite-duration smooth-amplitude LFM signal in a real underwater environment are shown to provide reliable and consistent channel estimates.

2) Blind Estimator: To construct a blind channel estimator, it is reasonable to modify (10) according to an estimate $\hat{f}_i(t)$ of the IF:

$$WV_{ss}(\tau, f) \approx \delta\left[f - \hat{f}_i(\tau)\right]. \tag{23}$$

Using this approximation in (8), the final TF blind channel estimator is given by:

$$\hat{\theta}_{TFB}(t) = \sqrt{\int_{-\infty}^{\infty} WV_{rr}\left[\tau, \hat{f}_i(\tau - t)\right] \mathrm{d}\tau}. \tag{24}$$

Thus, contrarily to the non-blind channel estimator, the blind estimator requires prior estimation of the source signal instantaneous frequency $f_i(t)$.

3) Instantaneous Frequency Estimator: In several real underwater scenarios, the ocean is reliably modeled by a multipath channel, with impulse response

$$h(t) = \sum_{m=1}^{M} a_m \delta(t - \tau_m), \tag{25}$$

where a_m and τ_m designate each of the M channel amplitudes and time-delays, respectively[1]. For this case, an estimator of $f_i(t)$ has been derived, taking advantage of a reduced cross-terms TFD. As is well known, TFDs of this type are usually derived as a tradeoff between linearity and high resolution. First, let us consider an ideal (linear and infinitely concentrated) TFD $I_s(t, f)$ of $s(t)$ (assumed a monocomponent signal), defined as

$$I_s(t, f) = |c|\, \delta\left[f - f_i(t)\right]. \tag{26}$$

[1]Here, frequency dispersion effects are neglected, because the ratios water depth-maximum wavelength and source-receiver distance-maximum wavelength, are of orders 10 and 10^3, respectively

The expected value of the ideal distribution of the received signal, assuming that this expected distribution would represent the noise term as a constant I_n, would be given by

$$\mathrm{E}[I_r(t,f)] \;=\; \sum_{m=1}^{M} a_m I_s(t-\tau_m, f) + I_n. \tag{27}$$

Consider now that the channel highest amplitude impulse is well separated and significantly stronger than the closely spaced impulses. In this case, at least one of the replicas of $s(t)$ is represented in $\mathrm{E}[I_r(t,f)]$ by a large amplitude along the delayed IF of $s(t)$. Thus, maximization of $\mathrm{E}[I_r(t,f)]$ with respect to t, within a given band of interest B, would select the strongest arrival, giving an unbiased estimate[2] of $f_i(t)$. Obviously, within the available non-linear TFDs, signal analysis is constrained by the particular characteristics of the kernel, and finite data lengths. Nonetheless, it seems reasonable to apply the maximization with respect to t, to a TFD with reduced cross-terms of the received signal, which will give an accurate estimate[3] of $f_i(t)$, if this distribution is a good approximation to $I_r(t,f)$, *i.e.*, it attains a significant cross-terms rejection, without causing a significant broadening of the signal components. In this work, the distribution used for IF estimation was the signal-dependent radially Gaussian kernel distribution (RGK) of the received signal, $RGK_{rr}(t,f)$. The signal-dependence of this distribution allows a reduced cross-terms TF representation of the multipath received signal, even when the IF of $s(t)$ is a non-linear function. This is due to the distribution capability of adapting the shape of a radially Gaussian low-pass filter to a broad class of signals[17], hence attenuating the TF spurious cross-terms. The volume of the distribution kernel was set to 1, since this is the volume of every spectrogram kernel. In the ideal case in which the calculated kernel of the RGK would equal the Doppler-reversed ambiguity function of $s^*(t)$, the RGK would coincide with the matched-filter spectrogram. This point is discussed below in more detail. Obviously, other distributions could be applied as well, due to their cross-terms rejection capabilities, such as the polynomial or modified Wigner-Ville distribution[22], or TFDs with adaptive window width, to mention only a few.

Considering the presence of N snapshots at reception –what is the case in the real data presented in the next section–, each one obeying to model (1), the estimation of the IF may

[2]It is expected that this estimate, seen as a function of frequency, is also a close estimate of the group delay of $s(t)$.

[3]The IF estimate is not rigorously a function of t, due to its definition as the 'inverse function' of a non-invertible function.

take into account the information contained in the set of snapshots, in order to construct a TF representation that is close to $\mathrm{E}[I_r(t,f)]$. This TF representation is defined as the average RGK over the snapshots:

$$\overline{RGK}_{rr}(t,f) \;=\; \sum_{n=1}^{N} RGK_{r_n r_n}(t,f) \tag{28}$$

where $r_n(t)$ designates the n^{th} snapshot of the received signal. The multi-snapshot estimator is defined by means of the maximization of $\overline{RGK}_{rr}(t,f)$:

$$\begin{aligned}&\left[t_i, \hat{f}_i(t_i)\right]\\ &\quad= \left\{(t,f) : t = \arg\left\{\max_t \overline{RGK}_{rr}(t,f)\right\}, f \in B\right\}.\end{aligned} \tag{29}$$

This signal-dependent maximization on the TF plane will be illustrated with experimental data in section IV (see figures 9 and 10, and discussion). It is expected that the average $\overline{RGK}_{rr}(t,f)$ preserves the replicas of $s(t)$, while the noise contributes as a simple additive term, in the TF plane. This is difficult to assert theoretically, since the RGK is a signal-dependent TFD. However, a simple analysis can be done by considering a TFD whose kernel is perfectly adapted to the multicomponent received signal. Such a distribution, considering the multipath nature of $r(t)$, is the matched-filter spectrogram (MFS)[17], whose analysis window is thus coincident with $s^*(t)$. The kernel of the MFS can be expressed as

$$\Phi_{MFS}(\nu,\tau) \;= \int_{-\infty}^{\infty} s^*\left(t+\frac{\tau}{2}\right) s\left(t-\frac{\tau}{2}\right) \mathrm{e}^{-\mathrm{j}2\pi\nu t}\mathrm{d}t. \tag{30}$$

The MFS represents the best attainable representation of $r(t)$ by the RGK, in terms of the tradeoff signal-terms information/cross-terms rejection. Finally, one can interpret the average RGK as an estimator of

$$
\begin{aligned}
&\mathrm{E}[MFS_{rr}(t,f)] \\
&\quad = \frac{1}{2\pi}\left[\iint_{-\infty}^{\infty} AF_{xx}(\nu,\tau)AF_{ss}(-\nu,-\tau)\mathrm{e}^{-\mathrm{j}2\pi(\nu t+f\tau)}\mathrm{d}\nu\mathrm{d}\tau\right. \\
&\qquad \left.+\sigma_n^2\right] \\
&\quad = \frac{1}{2\pi}\left[\sum_{m=1}^{M}\sum_{n=1}^{M} a_m a_n \iint_{-\infty}^{\infty} AF_{ss}(\nu,\tau+\tau_m-\tau_n)\right. \\
&\qquad \left.\times AF_{ss}^{*}(\nu,\tau)\mathrm{e}^{-\mathrm{j}2\pi[\nu(t-\tau_{mn})+f\tau]}\mathrm{d}\nu\mathrm{d}\tau+\sigma_n^2\right], \qquad (31)
\end{aligned}
$$

where

$$
\tau_{mn} \quad = \quad (\tau_m+\tau_n)/2. \qquad (32)
$$

For a monocomponent signal $s(t)$, it is expected that, for the terms in (31) for which, $\tau_m \neq \tau_n$, the different delays in the ambiguity functions will render the integrand approximately null. For the M terms in which $\tau_m = \tau_n$, the integrals transform into

$$
\iint_{-\infty}^{\infty} |AF_{ss}(\nu,\tau)|^2\,\mathrm{e}^{-\mathrm{j}2\pi[\nu(t-\tau_m)+f\tau]}\mathrm{d}\nu\mathrm{d}\tau, \qquad (33)
$$

what corresponds to slightly broadened signal terms in the TF plane. Thus, $AF_{ss}^{*}(\nu,\tau)$ is to be viewed as a low-pass filter, which will retain only the eigen-terms of $AF_{xx}(\nu,\tau)$, leading to an approximate linear TF distribution $\mathrm{E}\,[MFS_{rr}(t,f)]$. Ideally, the maximization of the expected MFS [requiring the knowledge of $s(t)$] would be meaningful as an IF estimator; in practice, the expected MFS is approximated by the average RGK.

A final concern about IF estimation relates to source signals with non-linear frequency modulation. In this case, each of the source signal replicas (in the received signal) contains its own cross-terms in the TF plane, which will be added to the implicit cross-terms due to the multipath channel. It is difficult to predict the relative amplitude between the signal terms and the cross-terms, and the maximization of the signal-dependent TFD can be trapped into the cross-terms, at given frequencies, instead of the signal term. To avoid this problem, it may be necessary to decrease the volume of the RGK distribution kernel, in order to filter out all the high-amplitude

cross-terms, and guarantee that the global maximization will pick only the signal term(s) due to the initial strong channel arrival(s).

Note that the required minimum time separation (≈20 ms, in the considered data) between the strong channel impulses(s) and the remaining impulses is much smaller than the source signal duration (2 s, see figure 3), since the considered impulse response structure corresponds to a shallow water scenario. This implies that each arrival overlaps several other arrivals, in the received signal, in the time domain. This is opposed to the TF domain, where the inter-arrival overlap is spread in the two dimensions of the plane. For extremely short-duration source signals, or in deep water scenarios, if there was a completely isolated arrival in the received signal, the channel impulse response could be estimated, for example, by simply cross-correlating that arrival with the remaining portion of the received signal.

IV. Experimental Results

The experimental results presented in this section concern data from the acoustic tomography experiment INTIMATE '96, whose details have appeared in [23]. The experiment was conducted in the continental platform near the town of Nazaré, off the west coast of Portugal, during June 1996. For the data considered here, the acoustic source and receiving hydrophone were respectively located at 92 m and 35 m depth, 5.6 km apart in a 135 m-depth approximately range independent shallow water area, as shown in figure 2. The signal-to-noise ratio estimate at the receiver was 10 dB, within the frequency band of interest. The sound source was emitting a 300–800 Hz, 2 s-duration LFM sweep $l(t)$, repeated every 8 s. At the receiver, $N = 10$ consecutive sweeps were processed to form one single channel estimate, under stationarity assumptions. This process was repeated every 5 min, during 19 h of transmissions, with the objective of analyzing the time variations of the acoustic channel due to tidal waves. The frequency response of the electro-acoustic transducer used as sound source, presented a main resonance at 650 Hz and a secondary resonance at 350 Hz. The convolution of one LFM sweep $l(t)$ with the transducer response is shown in figure 3. Straightforward lengthy calculations show that this convolution may be approximately expressed as a product, leading to

$$s(t) \approx l(t)v(t) \tag{34}$$

where $v(t)$ is a factor due to the transducer response, which effectively modulates the source signal in amplitude. As a matter of fact, figure 3 clearly shows the influence of the transducer

resonances on the amplitude of the emitted signal. Note that the actual signal amplitude, as modified by the transducer, is not taken into account when performing blind channel estimation, and is another contribution (together with the non-infinite concentration of the TFD) to the deviation from the optimal case. Moreover, signal amplitude discrepancy should be taken into account when comparing the (non-blind) MF channel estimate (which 'knows' the transducer response) to the blind TF estimate, which does not know the transducer response.

A. Channel Estimation Results

Taking into account that, for an observation time of the order of minutes, the acoustic multipath channel can be assumed non-stationary in the amplitudes, and stationary in the time-delays[24], and that an accurate estimate of the time-delays is sufficient for post-processing, *e.g.*, in ocean acoustic tomography[26], the presented results privilege time-delay estimation. Thus, due to the similarity between the non-blind TF and MF channel estimators, and proceeding as in [24], for the derivation of the MF, in the presence of a set of snapshots at the reception, the effectively used channel estimators were given by

$$\hat{\theta}_{CE}(t) \;=\; \frac{1}{N}\sum_{n=1}^{N}\hat{\theta}_{CE,n}(t) \tag{35}$$

where CE is to be replaced by MF or TFB, in order to obtain averaged versions of the estimators (5) and (24), respectively. This is expected to reduce the bias and variance of the final estimators, by comparison to the single snapshot estimators $\hat{\theta}_{CE,n}(t)$. The IF estimate to be inserted in (24) was obtained by maximization of the average RGK within the band $B = [300, 800]$ Hz, coincident with the LFM's band. Note that it is not necessary to have *a priori* knowledge of the emitted signal's band, provided that B is chosen large enough, since the energy of the received signal will lie essentially within the emitted signal's band, in the TF plane.

The results obtained along the 19 h-duration data set are shown in figures 4 and 5, for the MF and TF estimators, respectively. It can be seen that the blind channel estimates are, in general, similar to the homologous (non-blind) MF estimates. Of most concern here is the ability to (i) discriminate closely spaced arrivals and (ii) follow the arrivals waving through time. It can be noted that these concerns are both met: after the initial non-separated arrivals ($t \leq 25$ ms), four distinct arrivals can be seen on each of the following seven packets of arrivals; the long period waving due to tidal influence can clearly be seen on the later arrivals, on both figures.

B. Performance Analysis

It is now important to quantify the accuracy of the blind TF channel estimator, taking, for example, the (non-blind) MF estimator, as reference. For each TF channel estimate, the quality measure is defined as the maximum of the normalized correlation between the TF and MF channel estimates. In the same way, the quality χ of the IF estimate (29) is defined as

$$\chi \;=\; 1 - \frac{\sqrt{\int_B \left[t_{iN}(f) - \hat{t}_{iN}(f)\right]^2 \mathrm{d}f}}{2} \tag{36}$$

where $t_{iN}(f)$ and $\hat{t}_{iN}(f)$ are normalized versions of the instants corresponding to the true IF and its estimate, respectively. Both channel and IF estimates quality measures are shown in figure 6, for the whole 19 h data set. It can be seen in this figure that, in general, the IF and the channel estimates are directly related, *i.e.*, the channel estimate quality is highly sensitive to the IF estimate quality, as can be seen for *e.g.* the estimates obtained at times 17 h 07 min and 04 h 09 min. This fact can be explained by the high concentration of $s(t)$ in the TF plane, which leads to meaningless values for the integral (24), when the IF is not accurately estimated. In order to closely characterize the proposed IF and blind TF estimators, it is important to look at the worst and best cases over the whole data set.

The reference channel given by the MF, corresponding to the worst blind channel estimate, is shown in figure 7 (a), while the corresponding blind estimate, with quality 0.796, is depicted in figure 7 (b). In this case, the acoustic channel consists of a set of leading closely spaced arrivals with large amplitude, followed by a peaky pattern with smaller amplitude and well separated in time. It can be seen that the MF and TF estimates differ mainly on the amplitude envelope but not on the arrival times. However, the total number of arrivals seems to be higher in the TF estimator than in the MF estimator.

According to the measure of quality plot shown in figure 6, the best blind TF channel estimate is obtained at time 20 h 08 min, and the corresponding MF and TF estimates are shown in figure 8 (a) and (b), respectively. The most important difference when comparing figure 8 (a) to the MF estimate in the worst case of figure 7 (a) is the rapid progressive amplitude attenuation, leading to a greater amplitude ratio between the initial and the remaining arrivals. The best blind TF estimate, depicted in figure 8 (b), shows that both the envelope and arrivals positions are well estimated and the MF and TF estimates are indeed very similar. As in the worst case,

the number of arrivals is over-estimated on the TF estimator, by comparison to the MF estimator.

Figure 9 shows the IF estimates for the worst and best cases as discussed above. Clearly, the worst case IF estimate shows a piecewise line where the IF oscillates well away from the due straight line IF, and the best channel estimate corresponds to an accurate IF estimate close to the true straight line. These IF estimates were obtained by maximization of the average RGKs shown in figure 10, where one can see that, unlike for the best case in (b), the worst case in (a) corresponds to initial high amplitude overlapped replicas of $s(t)$ in the average RGK.

Considering the available data, one can deduce that the quality of the channel estimate is largely dependent on the quality of the IF estimate. A set of close unresolved strong arrivals with similar amplitude, as is the case corresponding to the worst estimate, corresponds to energy spreading in the TF plane, which, added to the low-pass filtering effect on the WV, in the calculus of the RGK, gives a biased estimate $\hat{f}_i(t)$. This difficulty is not observed for the case corresponding to the best estimate. An additional concern is about the apparent higher resolution of the TF estimator, leading to a larger number of small peaks along the arrival pattern, by comparison to the MF estimator. This can be explained by the use of the IA information of the emitted signal in the MF, while that information is not used in the blind TF. Using the IA means using a reduced bandwidth of the transducer frequency response when compared to the blind TF, where all frequencies are assumed to equally belong to the emitted signal spectrum: reduced bandwidth implies larger correlation functions and therefore a lower resolution.

V. Conclusion

In this paper, a blind sub-optimal time-frequency channel envelope estimator was proposed. This estimator was tested on a 19 h-duration real data set from the INTIMATE ’96 sea trial, where a severe multipath underwater channel was driven by a deterministic linear frequency modulated signal. Three main topics emerge as conclusions. The first concerns the restrictions imposed by the blind channel envelope estimator. This estimator requires the multipath channel structure to consist of at least one leading strong arrival well separated from the highly-attenuated remaining arrivals. Also, the emitted signal is required to have a concentrated time-frequency signature, and an injective (one-to-one) instantaneous frequency function. The second topic concerns the influence of the instantaneous frequency estimate on the channel envelope estimate. The quality of the time-frequency channel envelope estimate is largely sensitive to the source signal instan-

taneous frequency estimate. Indeed, a small inaccuracy in the instantaneous frequency estimate causes a large degradation in the channel envelope estimate, due to the high time-frequency concentration of the source signal. The last topic is that the time-frequency blind channel envelope estimator is reliable in that it gives similar results to the matched-filter, in terms of the arrivals structure and resolution. In addition, though the blind estimator does not achieve an optimal output signal-to-noise ratio, its higher resolution can be viewed as an advantage, when emphasis is given to the estimation of the channel time-delays.

As future trends, the time-frequency channel envelope estimator will possibly be extended to a broader class of source signals, for which, the instantaneous frequency estimator could be reformulated by using the product high-order ambiguity function. The integration in the time-frequency domain could be operated on a highly concentrated time-frequency distribution, *e.g.* the polynomial or modified Wigner-Ville distribution.

ACKNOWLEDGMENTS

Thanks are due to the INTIMATE ’96 team for the real data acquisition, and to André Quinquis and Cédric Gervaise for the kind valuable reception at ENSIETA, Brest (France).

REFERENCES

[1] R.J. Vaccaro, C.S. Ramalingam and D.W. Tufts, “Least-squares time-delay estimation for transient signals in a multipath environment”, *J. Acoust. Soc. Am.*, vol. 92, pp. 210–218, 1992.

[2] M.K. Broadhead and L.A. Pflug, “Performance of some sparseness criterion blind deconvolution methods in the presence of noise”, *J. Acoust. Soc. Am.*, vol. 107, pp. 885–893, 2000.

[3] A. Silva and S.M. Jesus, “Underwater communications using virtual time reversal in a variable geometry channel”, in Proc. IEEE/MTS Oceans’2002, Biloxi (USA), October.

[4] S. Haykin, *Blind Deconvolution*, PTR Prentice Hall, 1994.

[5] A.T. Walden, “Non-Gaussian reflectivity, entropy, and deconvolution”, *Geophysics*, vol. 50, no. 12, pp. 2862–2888, 1985.

[6] X.-G. Xia, “System identification using chirp signals and time-variant filters in the joint time-frequency domain”, *IEEE Trans. Signal Processing*, vol. 45, pp. 2072–2084, 1997.

[7] C.H. Knapp and G.C. Carter, “The generalized correlation method for estimation of time delay”, *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 24, pp. 320–327, 1976.

[8] J.P. Ianniello, “Time delay estimation via cross-correlation in the presence of large estimation errors”, *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, no. 6, pp. 998–1003, 1982.

[9] I.M.G. Lourtie and J.M.F. Moura, “Multisource delay estimation: nonstationary signals”, *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, pp. 1033–1048, 1991.

[10] I.P. Kirsteins, “High resolution time delay estimation”, in *IEEE Proceedings ICASSP*, pp. 451–454, 1987.

[11] P. Mignerey and S. Finette, "Multichannel deconvolution of an acoustic transient in an oceanic waveguide", *J. Acoust. Soc. Am.*, vol. 92, pp. 351–364, 1992.

[12] J.F. Smith, III and S. Finette, "Simulated annealing as a method of deconvolution for acoustic transients measured on a vertical array", *J. Acoust. Soc. Am.*, vol. 94, pp. 2315–2325, 1993.

[13] W.A. Gardner, "A new method of channel identification", *IEEE Trans. Commun.*, vol. 39, pp. 813–817, 1991.

[14] L. Tong, G. Xu and T. Kailath, "A new approach to blind identification and equalization of multipath channels", in *Proceedings of the Twenty-Fifth Asilomar Conference on Signals, Systems, and Computers*, 1991.

[15] N. Martins, S.M. Jesus, C. Gervaise and A. Quinquis, "Time-frequency approach to multipath underwater channel deconvolution", in *IEEE Proceedings ICASSP*, vol. II, SPTM-L06, pp. 1225–1228, 2002.

[16] N.E. Martins and S.M. Jesus, "Blind channel estimation with data from the INTIMATE'96 sea trial", in *Proceedings MED*, Lisboa, Portugal, p. 62, 2002.

[17] R.G. Baraniuk and D.L. Jones, "Signal-dependent time-frequency analysis using a radially Gaussian kernel", *Signal Processing*, vol. 32, no. 3, pp. 263–284, 1993.

[18] J.E. Moyal, "Quantum mechanics as a statistical theory", *Proc. Camb. Philos. Soc.*, vol. 45, pp. 99–124, 1949.

[19] F. Hlawatsch, "Regularity and Unitarity of Bilinear Time-Frequency Signal Representations", *IEEE Trans. Information Theory*, vol. 38, pp. 82–94, 1992.

[20] P. Flandrin, "A time-frequency formulation of optimum detection", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 9, pp. 1377–1384, 1988.

[21] L. Cohen, *Time-frequency analysis*, Prentice Hall PTR, 1995.

[22] C. Ioana and A. Quinquis, "On the signal interference structure generated by modified Wigner-Ville distribution", in *IEEE Proceedings ICASSP*, vol. II, pp. 1449–1452, 2002.

[23] X. Démoulin, Y. Stéphan, S.M. Jesus, E.F. Coelho and M.B. Porter, "Intimate'96: a shallow water tomography experiment devoted to the study of internal tides", in *Zhang R. and Zhou J., editors, Proceedings of the Shallow Water Acoustics Conference (SWAC'97), Beijing*, pp. 485–490, 1997.

[24] S.M. Jesus, M.B. Porter, Y. Stéphan, X. Démoulin, O.C. Rodríguez and E.F. Coelho, "Single hydrophone source localization", *IEEE Journal of Oceanic Engineering*, vol. 25, no. 3, pp. 337–346, 2000.

[25] M.D. Collins and W.A. Kupperman, "Focalization: environmental focusing and source localization", *J. Acoust. Soc. Am.*, vol. 90, pp. 1410–1420, 1991.

[26] O.C. Rodríguez, "Application of ocean acoustic tomography to the estimation of internal tides on the continental platform", PhD thesis, University of Algarve, 2000, URL: http://www.ualg.pt/fct/adeec/siplab/pubs/orlando_PHD.pdf.

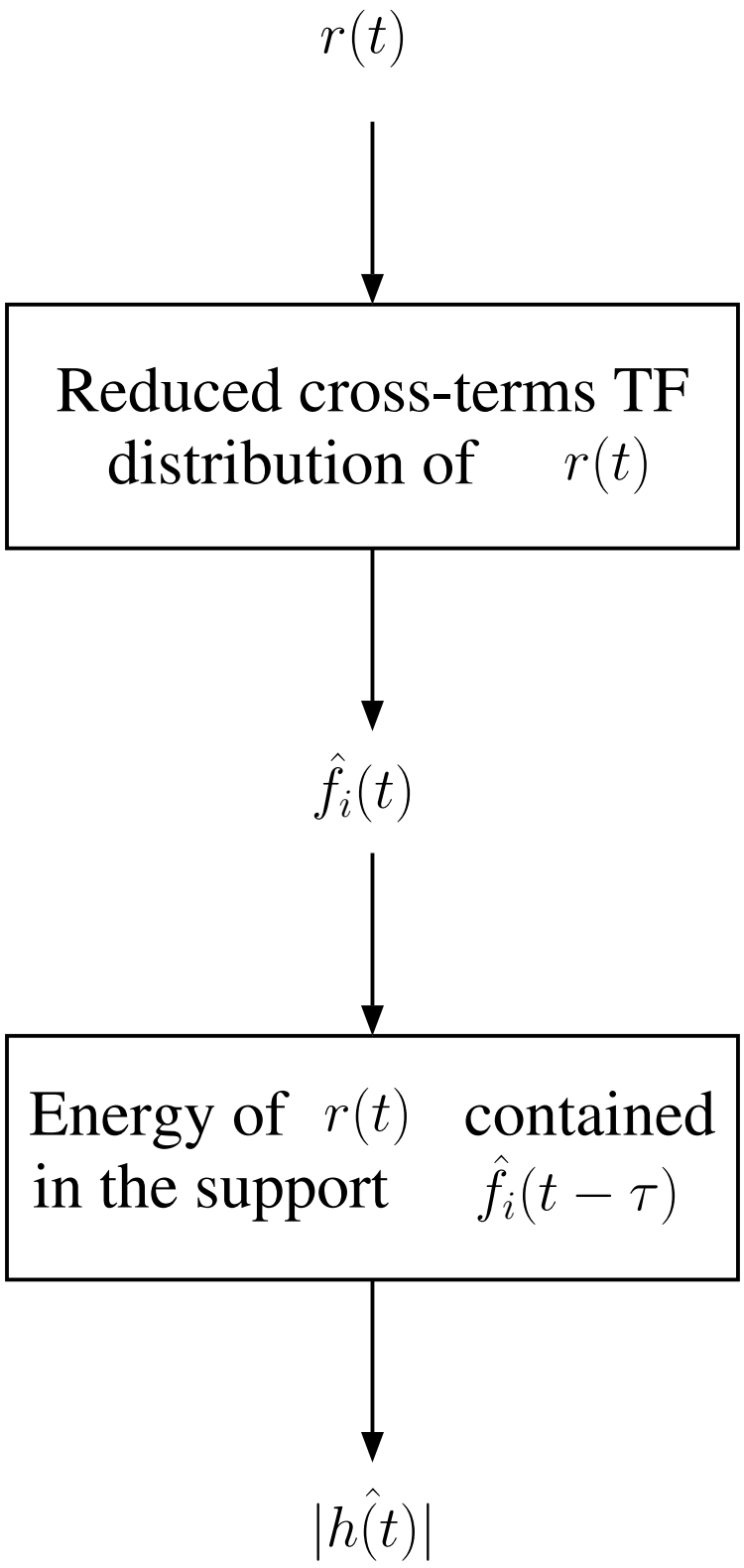


Fig. 1. Blind channel estimation procedure.

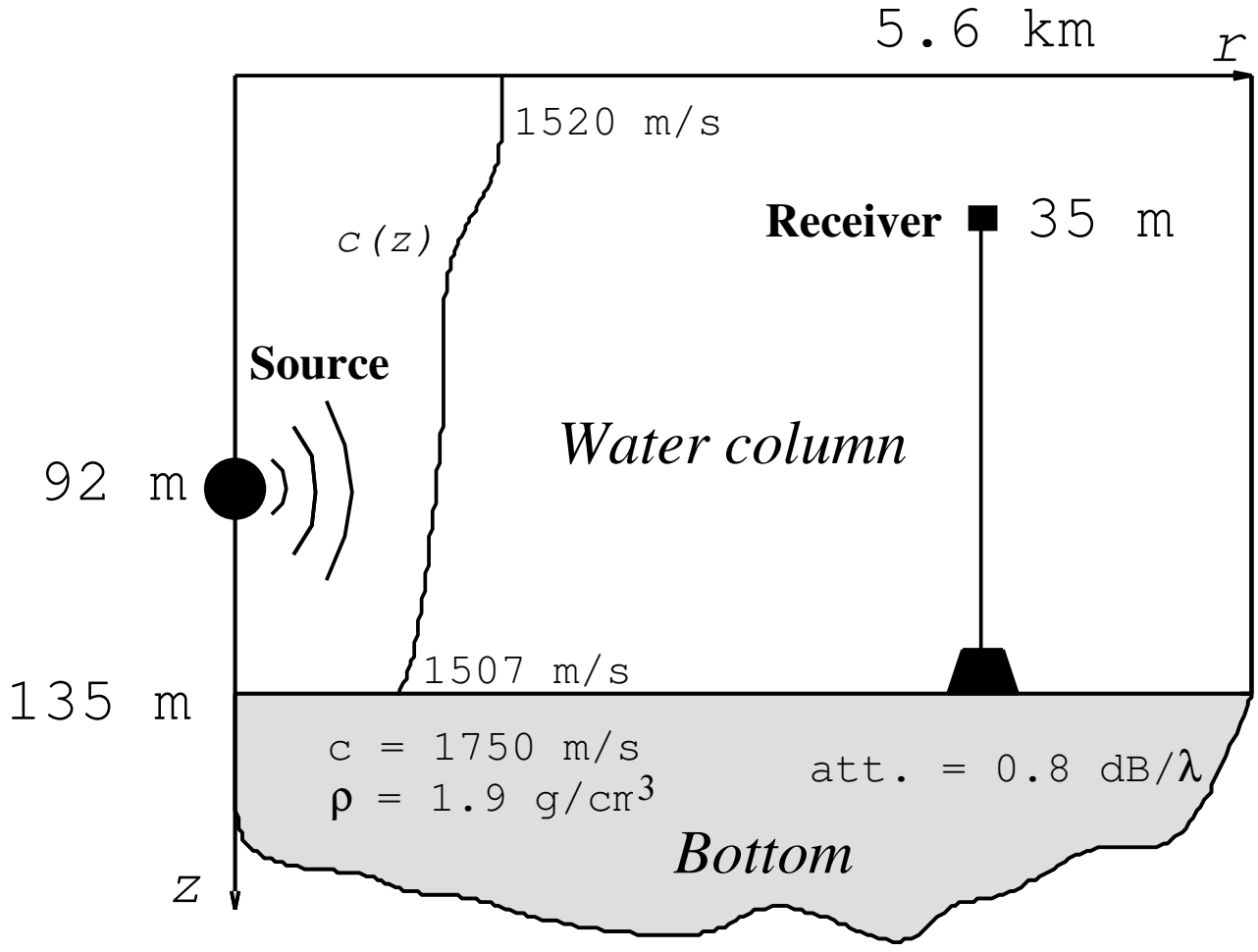


Fig. 2. INTIMATE ’96 real data environment scenario.

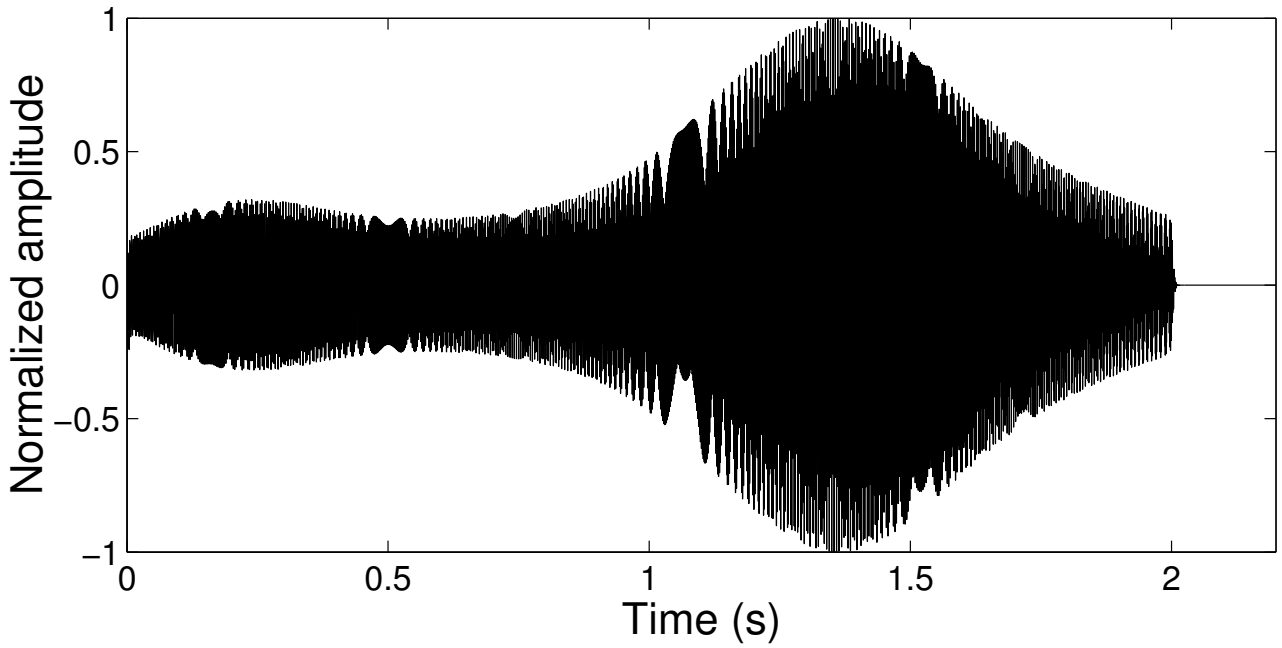


Fig. 3. Computer generated transducer response to the 300-800 Hz 2 s-duration LFM sweep [real part of $s(t)$] used during the INTIMATE '96 sea trial.

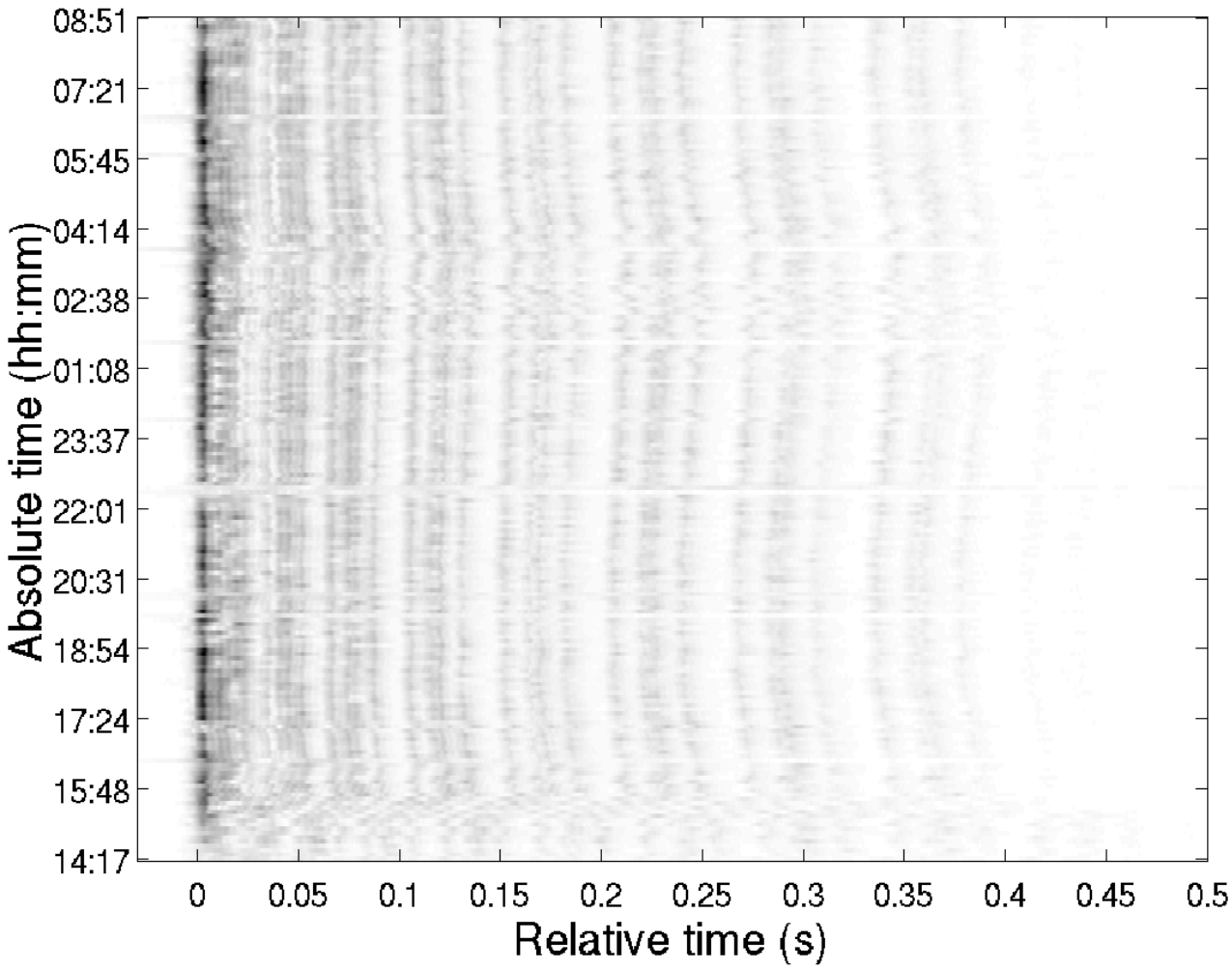


Fig. 4. INTIMATE '96 data set: (non-blind) channel estimates obtained with the matched-filter [eq. 5 arranged as in (35), with N=10].

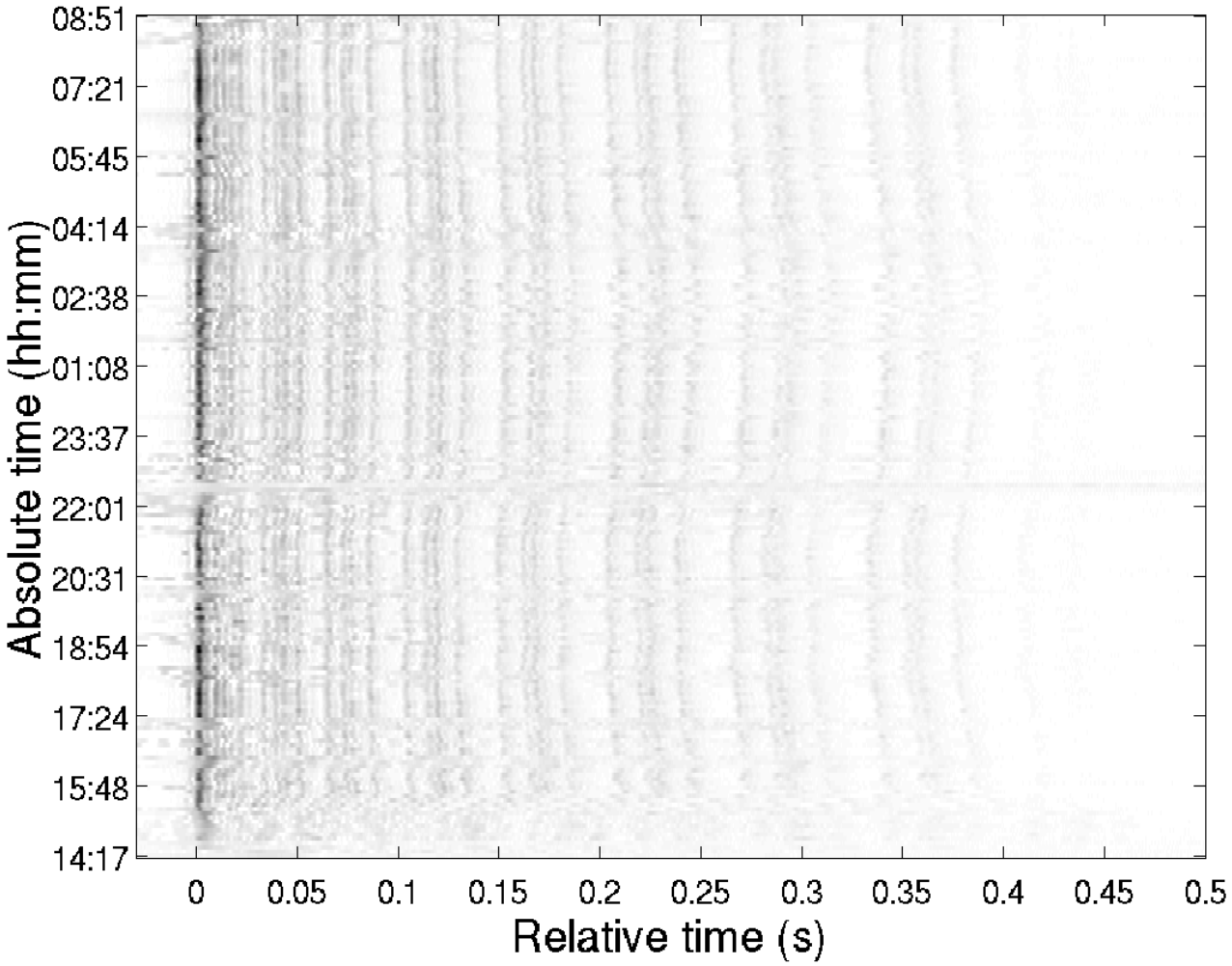


Fig. 5. INTIMATE '96 data set: blind channel estimates obtained by time-frequency processing [eq. 24 arranged as in (35), with N=10].

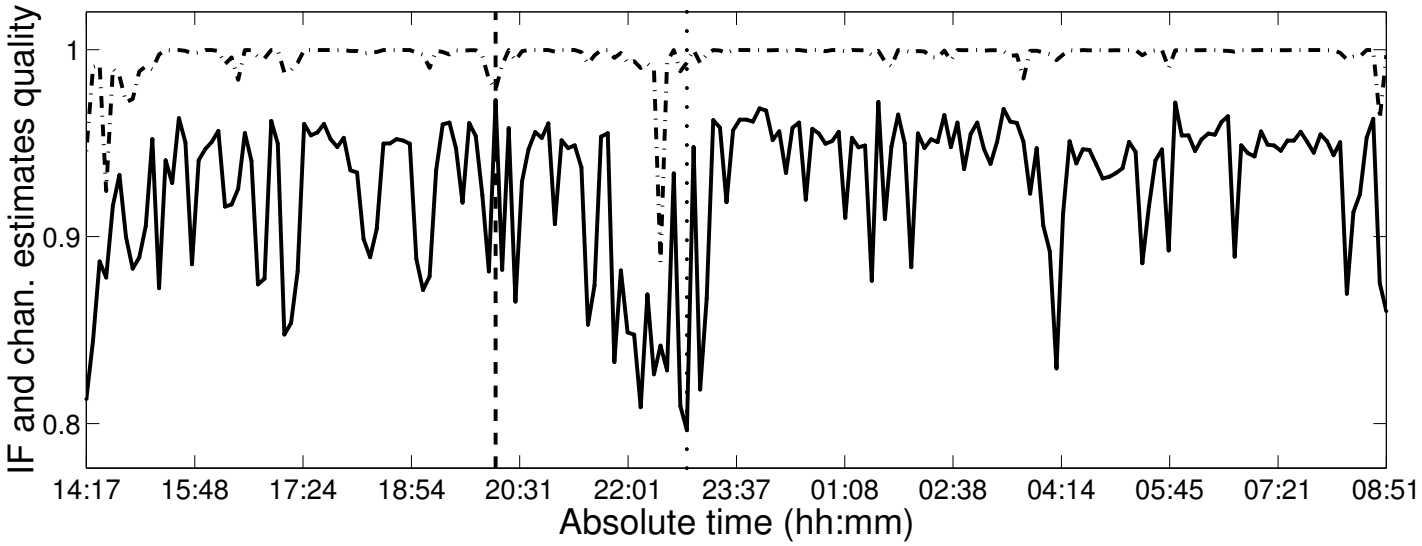
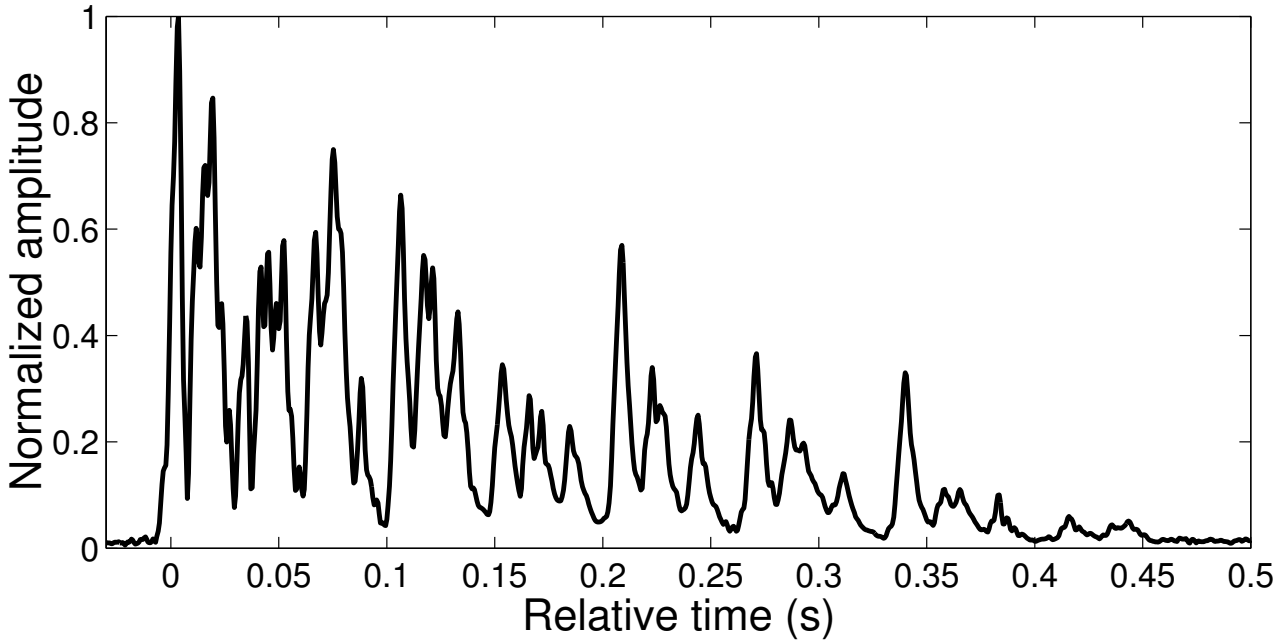


Fig. 6. Measure of quality for the instantaneous frequency estimate, using eq. 36 (dash dots), and for the blind time-frequency channel estimate, as the correlation between the blind time-frequency and the matched-filter estimates (solid), over the whole INTIMATE ’96 data set. The vertical dashed and dotted lines indicate 20 h 08 min and 22 h 52 min as the times corresponding to the best and worst blind channel estimates, respectively.

(a)



(b)

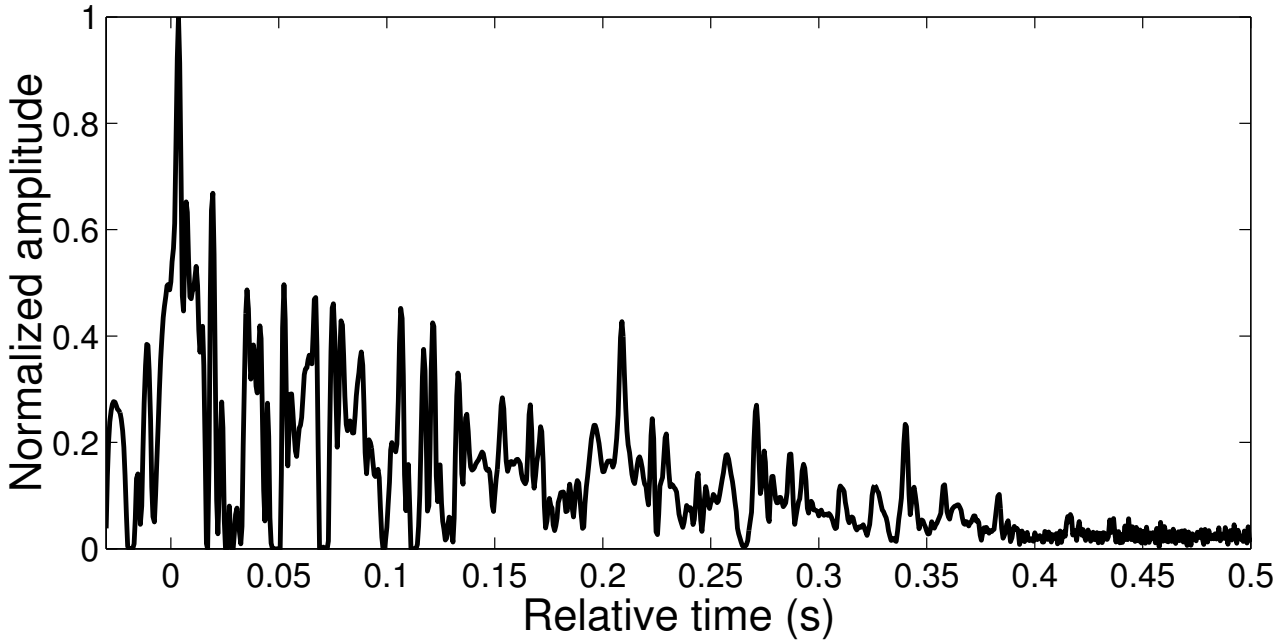
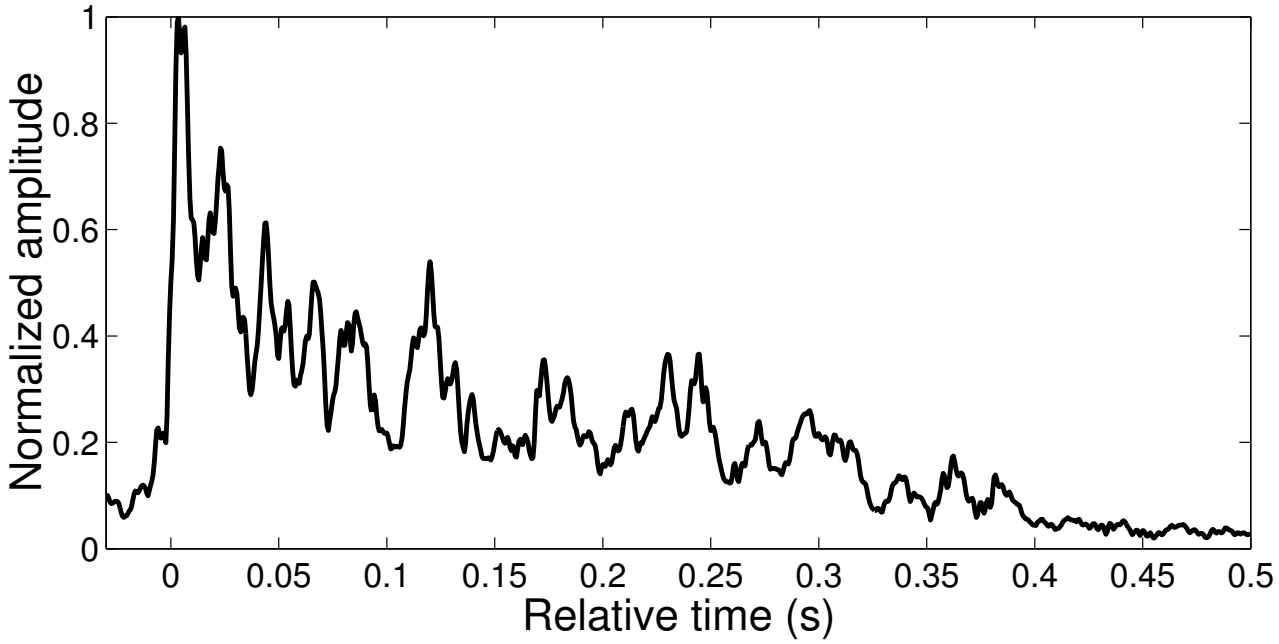


Fig. 7. INTIMATE '96: channel estimates given by the (a) (non-blind) matched-filter and (b) blind time-frequency channel estimators, corresponding to the worst blind channel estimate, at time 22 h 52 min of figure 6. The blind estimate quality [correlation coefficient between (b) and (a)] is 0.796.

(a)



(b)

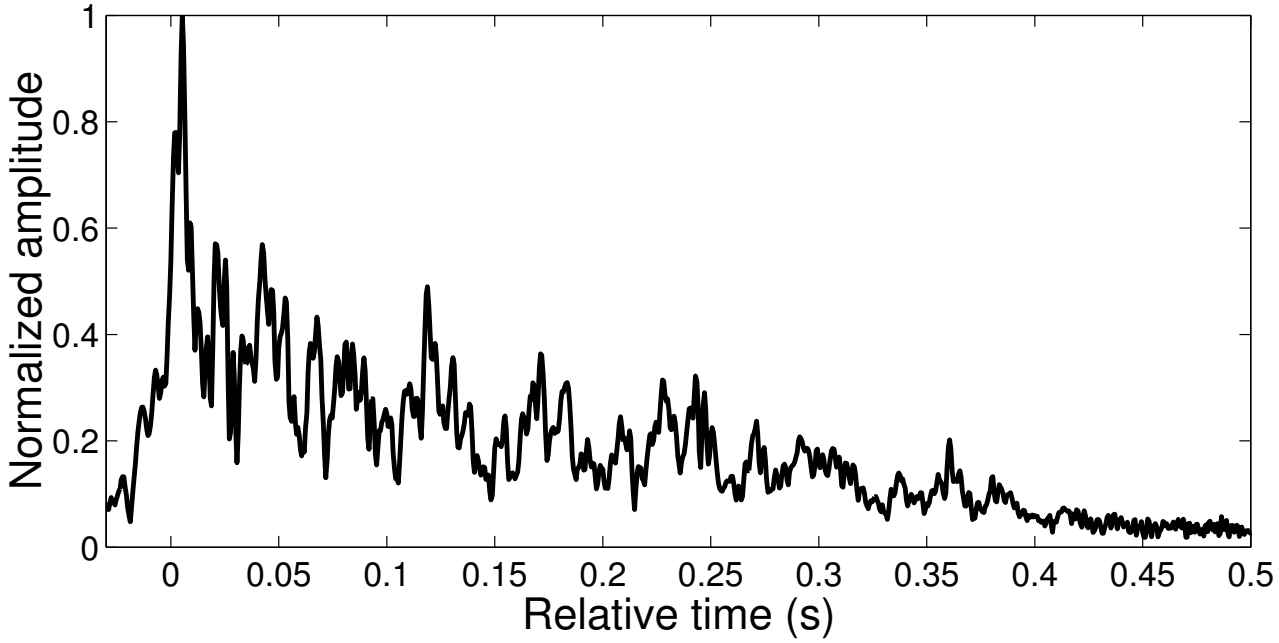


Fig. 8. INTIMATE ’96: channel estimates given by the (a) (non-blind) matched-filter and (b) blind time-frequency channel estimators, corresponding to the best blind channel estimate, at time 20 h 08 min of figure 6. The blind estimate quality [correlation coefficient between (b) and (a)] is 0.973.

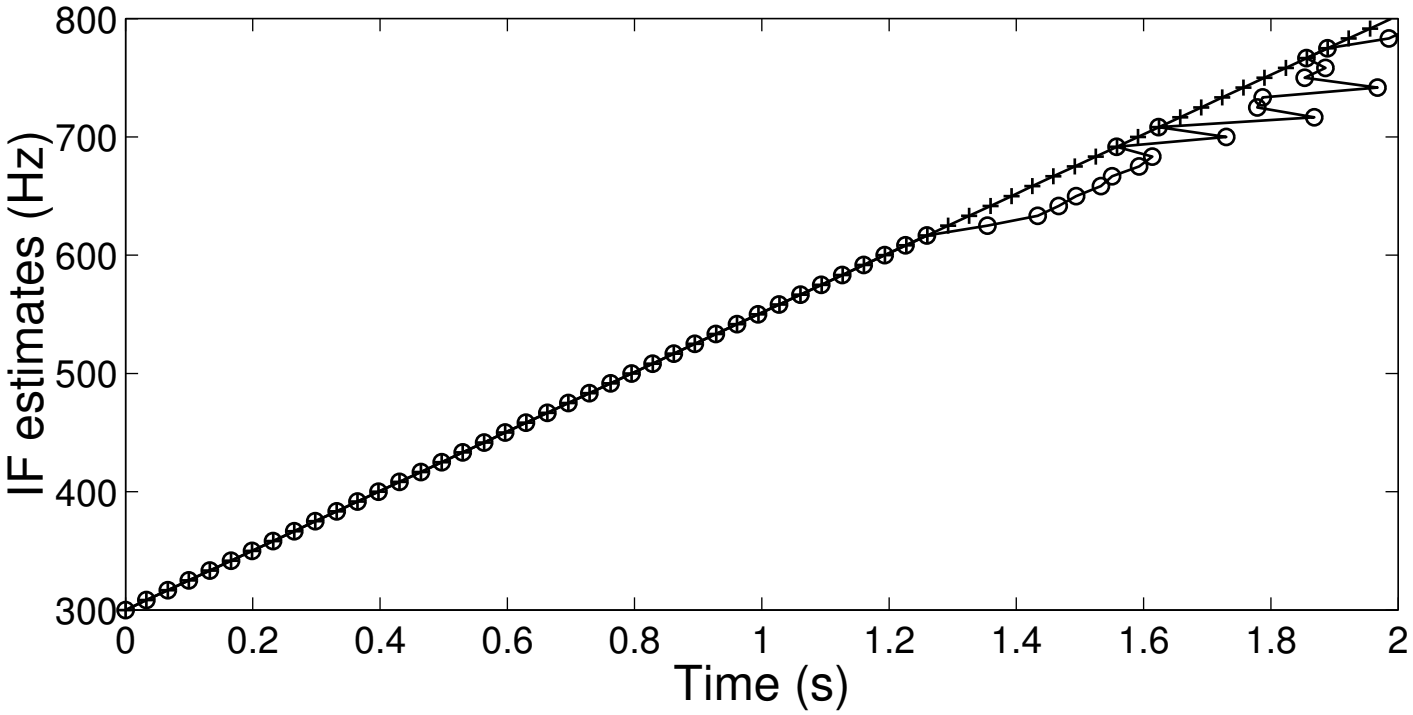
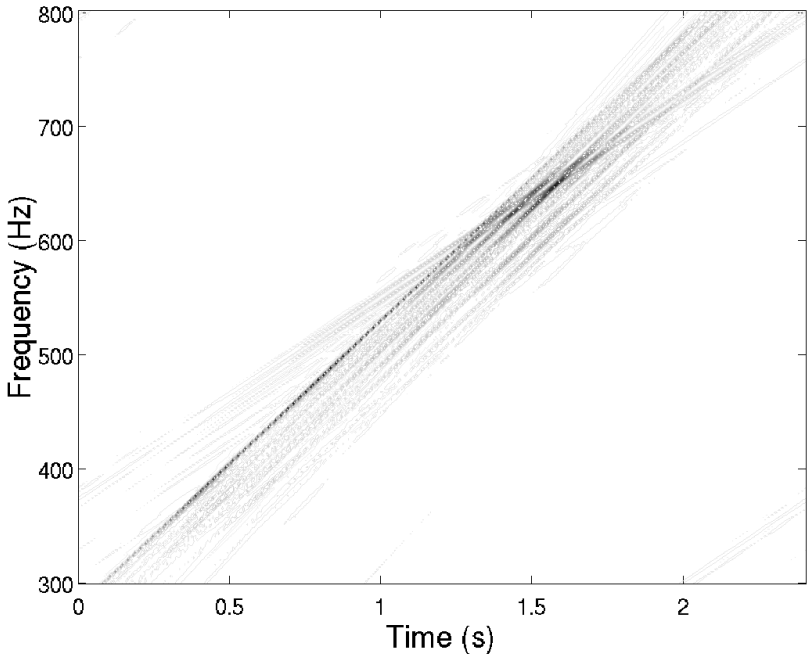


Fig. 9. INTIMATE '96: instantaneous frequency estimates (eq. 29) for the worst (O) and best (+) cases as shown in figures 7 and 8, respectively. The underlying straight line represents the true instantaneous frequency.

(a)



(b)

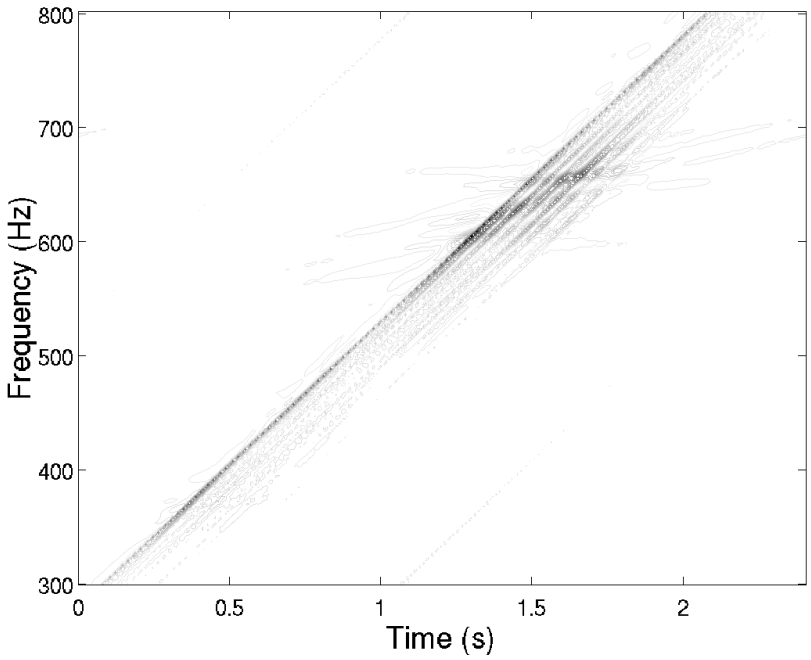


Fig. 10. INTIMATE ’96: contour plots of the average radially Gaussian kernel distributions from which, the instantaneous frequency estimates for the worst (a) and best (b) cases as shown in figures 7 and 8, respectively, were obtained.

LIST OF FIGURES