Time delay and amplitude estimation in underwater acoustics: a Gibbs Sampling approach

Zoi-Heleni Michalopoulou

Department of Mathematical Sciences
New Jersey Institute of Technology
Newark, NJ 07102

Michele Picarelli

Department of Mathematics
St. Peter's College
2641 Kennedy Blvd.
Jersey City, NJ 07306

Time delay and amplitude estimation in underwater acoustics: a Gibbs Sampling approach

Zoi-Heleni Michalopoulou
Department of Mathematical Sciences,
New Jersey Institute of Technology, Newark, NJ 07102
E-mail: michalop@njit.edu

Michele Picarelli
Department of Mathematics
St. Peter's College
2641 Kennedy Blvd,
Jersey City, NJ 07306
E-mail: mpicarelli@optonline.net

September 10, 2004

Abstract

Multipath arrivals of a signal at a receiving sensor are frequently encountered in many areas of signal processing, including sonar, radar, and communication problems. In underwater acoustics, numerous approaches to source localization, geoacoustic inversion, and tomography rely on accurate multipath arrival extraction. In this work, a novel method for estimation of time delays and amplitudes of arrivals with Maximum A Posteriori (MAP) estimation is presented. MAP estimation is optimal if appropriate statistical models are selected for

1

the received data; implementation, requiring maximization of a multi-dimensional function, is computationally demanding. Gibbs Sampling is proposed as an efficient means for estimating the necessary posterior probability distributions, bypassing analytical calculations. The Gibbs Sampler estimates posterior distributions through an iterative process and includes as unknowns noise variance and number of arrivals as well as time delays and amplitudes of multipaths. Through Monte Carlo simulations, the method is shown to have a performance very close to that of analytical MAP estimation. The method is shown to be superior to Expectation-Maximization, which is often applied to time delay estimation. The Gibbs Sampling - MAP approach is demonstrated to be more informative than other time delay and amplitude estimation methods, providing complete posterior probability distributions compared to just point estimates; the probability distributions illustrate the uncertainty in the problem, presenting likely values of the unknowns that are different from simple point estimates.

I Introduction

In underwater acoustics, Matched Field Processing approaches [1, 2] are frequently employed for the estimation of the source location of a sound emitting source. Such methods produce estimates by numerically calculating the full acoustic field and obtaining a measure of correlation between the computed field (replica) and received data; they are inherently dependent on assumptions necessary for the acoustic field computations.

A good match and consequently, accurate estimates, between full replica and true fields are difficult to attain, even when uncertainty on environmental factors is integrated in matched-field methods, as is often the case, when the propagation medium is complex and challenging to mathematically describe. Under such circumstances, simple approaches that do not rely on full field calculations can be implemented with excellent results. Such approaches depend on identification of individual arrivals-paths in the received field. Source location, bottom depth, sediment depths, and sound speed information can be extracted from the arrival times of these paths. The amplitudes of the arrivals provide information on geoacoustic properties of the sediments. In Refs. [3, 4, 5, 6, 7, 8, 9, 10, 11] it has been shown that arrival information can

be employed for efficient and accurate source and receiver localization and tracking and environmental parameter estimation. The quality of the estimates, however, is interwoven with the quality of time delay and amplitude estimation of the explored arrivals. It is, therefore, of great interest to develop methods for the extraction of accurate information on distinct arrivals. In a noisy received time series, however, arrival time differences between distinct paths and corresponding amplitudes can be difficult to identify.

A matched filter between source waveform and received time series is the simplest time delay and amplitude estimation approach, but suboptimal, especially for closely spaced arrivals [12]. Many other methods, several focusing on high resolution approaches, have been presented in the literature. (For a thorough presentation of different methods and applications, the reader is referred to [13, 14].)

In this work interest is on multipath propagation of deterministic signals. As shown in [12], an optimal approach for time delay and amplitude estimation is maximization of the posterior probability distribution function of delays and amplitudes. This maximization given the observed data leads to an analytical expression for amplitudes; using those estimates, time delays can be obtained by identifying the maximum of an M dimensional function, where M is the (known) number of paths at the receiver [12]. When M is large, these calculations can become a computationally cumbersome task. A simpler method, which also requires M-dimensional optimization and approximates Maximum A Posteriori (MAP) estimation, has been proposed in [15]. Approximate maximum likelihood approaches have been proposed in [16, 17]. Simulated annealing as a tool for optimization of the time delay estimation problem was suggested in [18].

To reduce the computational requirements of maximum a posteriori estimation of time delays and amplitudes, Expectation Maximization (EM) has been implemented [19, 20]. EM is an efficient way of maximizing log-likelihoods (equivalent under certain assumptions to posterior probability distributions), when the exact likelihood function is difficult to compute. A drawback is that EM is a hill climbing algorithm, generally converging to *local* maxima.

Here, we propose implementation of a MAP approach for time delay and amplitude estimation using Gibbs sampling for the efficient computation of

the full, joint posterior distributions [21]. The method was first introduced in [22] and is here, evaluated, and studied in terms of convergence. Results from this approach are compared to estimates obtained with the analytical maximum a posteriori approach, which is feasible to implement when M is small and the exhaustive search over time delays for distribution maximization is manageable.

Section II discusses the derivation of the joint posterior probability distribution of time delays and amplitudes and what is entailed in its maximization. Section III introduces Gibbs Sampling and derives conditional marginal distributions of the unknown parameters necessary for the operation of the Gibbs Sampler for time delay and amplitude estimation. Section IV presents a performance evaluation of the Gibbs Sampling approach comparing it to the analytical MAP method. Section V evaluates the novel approach through a comparison to EM. Section VI discusses the case of an unknown number of arrivals. Section VII addresses convergence issues for the Gibbs Sampler. Conclusions are presented in Section VIII.

II The analytical Maximum A Posteriori estimator

Estimates of unknown parameters of a statistical model can be obtained through the maximization of the posterior probability distribution of these parameters given the observed data and quantitatively described prior knowledge. Assuming a received signal $r(n)$, consisting of M multipaths and noise, we can write:

$$r(n) = \sum_{i=1}^{M} a_i s(n - n_i) + w(n), \tag{1}$$

where $n = 1, \ldots, N$ (N is the duration of the received signal), a_i is the amplitude of the ith path, and n_i is the arrival time of the ith path. Quantity $w(n)$ is additive, white normally distributed noise with zero mean and variance σ^2. It is assumed that the number of arrivals is known. Initially, it is also assumed that σ^2 is known as well. The amplitudes are real numbers (positive or negative, the sign indicating polarity of the arrivals).

One might assume that prior information is available for the arrival am-

plitudes. For example, in multipath propagation, the paths that undergo multiple reflections off the boundaries are attenuated; an exponential decay model might be a suitable representation for the amplitudes. However, any such model does not always describe sufficiently the problem at hand. When source and receiver are at the same depth, for example, some paths arrive simultaneously. The two simultaneous arrivals appear as a single arrival with a doubled amplitude, which is larger than that of the preceding arrival; a model of decay would then be unsuitable. Here, to avoid erroneous assumptions, we consider no prior information on the amplitudes. We assign to them uniform, improper prior distributions [23]:

$$p(a_i) = 1, \quad -\infty < a_i < \infty, \ i = 1, \ldots, M. \tag{2}$$

We set uniform priors for the delays:

$$p(n_i) = \frac{1}{N}, \quad 1 \le n_i \le N, \ i = 1, \ldots, M. \tag{3}$$

We can write the posterior probability distribution function of all amplitudes and delays (n_i and a_i for $i = 1, \ldots, M$) as follows:

$$p(n_1, n_2, \ldots, n_M, a_1, a_2, \ldots, a_M / r(n)) =$$
$$K \frac{1}{N^M} \frac{1}{(\sqrt{2\pi})^N \sigma^N}$$
$$\exp(-\frac{1}{2\sigma^2} \sum_{n=1}^{N} (r(n) - \sum_{i=1}^{M} a_i s(n - n_i))^2). \tag{4}$$

Quantity K is a constant. Also $\frac{1}{N^M} \frac{1}{(\sqrt{2\pi})^N \sigma^N}$ is a constant, being independent of all unknowns. All constants can be consolidated into one; Equation 4 becomes:

$$p(n_1, n_2, \ldots, n_M, a_1, a_2, \ldots, a_M / r(n)) =$$
$$C \exp(-\frac{1}{2\sigma^2} \sum_{n=1}^{N} (r(n) - \sum_{i=1}^{M} a_i s(n - n_i))^2), \tag{5}$$

Once the joint posterior distribution of all unknowns is described, MAP estimates of those parameters can be obtained through its maximization. Maximizing the distribution in Equation 5 over the unknown amplitudes and delays is equivalent to maximization of the likelihood function of Equation 5

in [12] in the discrete case. This problem seems to require a search in a $2 \times M$-dimensional space. As shown in Ref. [12], however, amplitude estimates can be analytically obtained, and subsequently a search in an M-dimensional space is required for delay estimation.

In order to justify results that will follow, it is here essential to point out what is involved in the maximum likelihood (or MAP, in this case) amplitude and time-delay estimation. By obtaining derivatives of the likelihood function with respect to amplitudes and borrowing notation from [12], estimates a_i are:

$$a_i = \Lambda^{-1} \phi_i, \tag{6}$$

$i = 1, \ldots, M$, where $\phi_i = \sum_{j=1}^{N} s(n - n_j) \ r(n)$. Matrix Λ is defined as:

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \ldots & \lambda_{1N} \\ \lambda_{21} & \lambda_{22} & \ldots & \lambda_{2N} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \lambda_{N1} & \lambda_{N2} & \ldots & \lambda_{NN} \end{pmatrix}. \tag{7}$$

Estimates for the time delays can be subsequently obtained by maximizing over time delays function $f(n_1, n_2, \ldots, n_M)$, where

$$f(n_1, n_2, \ldots, n_M) = \Phi^T \Lambda^{-1} \Phi. \tag{8}$$

Vector Φ is defined as follows:

$$\Phi^T = [\phi_1 \phi_2 \ldots \phi_M] \tag{9}$$

For a problem involving M arrivals, the estimation process involves a search in an M-dimensional space.

III   Building the Gibbs Sampler

Gibbs Sampling is an iterative Monte Carlo sampling process where realizations from a joint distribution are obtained by cycling through conditional

distributions that are typically easier to sample from than the joint distribution [21]. The first step is, hence, to derive the necessary conditional distributions. This will be achieved using teh distributions of Section II; the analysis of Section II is now extended to include noise variance as an unknown.

A non-informative prior distribution is considered for the variance as well, as in Ref. [24]:

$$p(\sigma^2) \propto \frac{1}{\sigma^2}, \tag{10}$$

typical of a variable taking only positive values.

Including the prior for the unknown variance, and consolidating constants, the joint posterior distributions is as follows:

$$p(n_1, n_2, \ldots, n_M, a_1, a_2, \ldots, a_M, \sigma^2/r(n)) =$$
$$C \frac{1}{\sigma^{N+2}} \exp(-\frac{1}{2\sigma^2} \sum_{n=1}^{N} (r(n) - \sum_{i=1}^{M} a_i s(n - n_i))^2), \tag{11}$$

The conditional posterior distribution for the variance is identified as an inverse χ^2 distribution:

$$p(\sigma^2/n_1, n_2 \ldots, n_M, a_1, a_2 \ldots, a_M, r(n)) = \frac{1}{\sigma^{N+2}} \exp(-\frac{1}{2\sigma^2} \sum_{n=1}^{N} (r(n) - \sum_{i=1}^{M} a_i s(n-n_i))^2). \tag{12}$$

Samples from such a distribution can be drawn readily.

From the joint posterior function of Equation 11, the marginal posterior probability distributions for time delays and amplitudes of arrivals are obtained. Assuming that all amplitudes a_j, $j = 1, \ldots, M$ and $j \neq i$, and delays n_k, $k = 1, \ldots, M$ are known, we can derive the following conditional distribution for amplitude a_i:

$$p(a_i, /n_1, n_2, \ldots, n_M, a_1, a_2, \ldots, a_M, \sigma^2, r(n)) =$$
$$C \frac{1}{\sigma^{N+2}} \exp(-\frac{1}{2\sigma^2}(a_i - (\sum_{n=1}^{N} r(n)s(n - n_i))$$
$$- \sum_{j=1(j \neq i)}^{M} a_j s(n - n_i)s(n - n_j)))^2). \tag{13}$$

7

The argument of the exponential of Equation 13 reveals a Normal distribution for amplitude a_i with mean

$\sum_{n=1}^{N} r(n)s(n-n_i) - \sum_{j=1(j\neq i)}^{M} a_j s(n-n_i)s(n-n_j))$

and variance σ^2.

The marginal posterior distributions for delays n_i, $i = 1, ..., M$, are obtained on a grid (between 1 and N with unit spacing, where N is the length of the received sequence). Using the distribution of Equation 11, the conditional posterior distribution of n_i for known a_j, $j = 1, \ldots, M$, n_k, $k = 1, \ldots, M$, $k \neq i$, and σ^2 is:

$$p(n_i/n_1, n_2 \ldots, n_M, a_1, a_2 \ldots, a_M, \sigma^2, r(n)) =$$
$$G \exp(-\frac{1}{2\sigma^2}$$
$$\sum_{n=1}^{N} (r(n) - \sum_{i=1}^{M} a_i s(n-n_i))^2). \tag{14}$$

The conditional distributions derived above will be used as building blocks in the Gibbs Sampler for the estimation of the posterior probability distribution of time delays and amplitudes.

In the present context we are concerned with obtaining the joint posterior distribution for amplitudes a_i, time delays n_i, $i = 1\ldots, M$, and variance σ^2. Conditional on all time delays and amplitudes of the arrivals as well as noise variance, the marginal conditional posterior distribution of each amplitude is analytically tractable in a closed form as shown in Equation 13. So is the marginal conditional posterior distribution of σ^2 (Equation 12). The marginal posterior distributions for time delays are not analytically tractable; we thus proceed with a grid–based approximation using the distributions of Equation 14 (griddy Gibbs [25]).

Gibbs Sampling begins with a set of initial conditions for all $2M + 1$ unknown parameters (a_i and n_i, $i = 1, \ldots, M$, and σ^2). The process as implemented here first draws a sample from the inverse χ^2 distribution of Equation 12; this is the new, updated value of the variance for the first iteration. Subsequently, a sample is drawn from the Normal marginal conditional posterior of a_1 (Equation 13). Given the new values of σ^2, a_1, and the initial values for $a_3, \ldots, a_M, n_1, n_2, \ldots, n_M$, a sample is then drawn in the same way for a_2. We continue this procedure, drawing samples for all unknown pa-
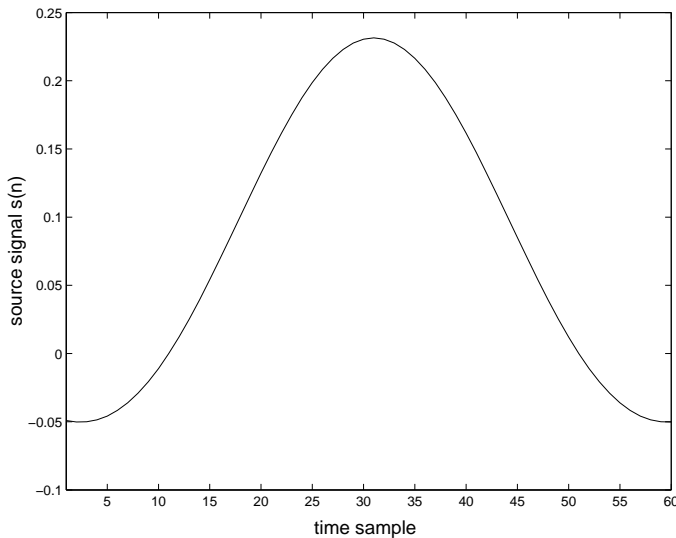
Figure 1: Transmitted signal.

rameters from their respective marginal conditional posterior distributions. For a large number of iterations, the obtained sample sequences eventually converge to the true joint posterior distribution of a_i, n_i, $i = 1, \ldots, M$, and σ^2 [21, 26, 27].

IV   Performance Evaluation

To evaluate the proposed Gibbs Sampling MAP approach, a two-arrival problem was simulated. The Gibbs Sampler was evaluated through a comparison to the analytical calculation of the maximum of the joint posterior distribution. The transmitted signal $s(n)$ for the simulations is shown in Figure 1. The first arrival was always considered to be at the 50th sample. The second arrival varied between the 52nd and 160th sample. The amplitudes were always 100 and -90 for the two arrivals, respectively.

Figures 2 and 3 present the rms errors for time delays and amplitudes, respectively, for noise variance $\sigma^2 = 0.01 \, ||s(n)||^2$ as obtained from the analytical MAP estimator and the Gibbs Sampler. The results were initially surprising; the Gibbs Sampler appears to outperform the analytical proces-
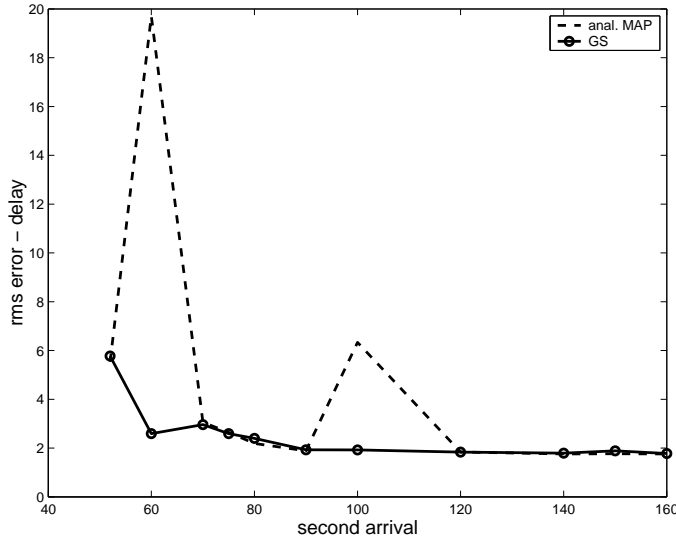
9

Figure 2: Rms errors for time delays: $\sigma^2 = 0.01 \, ||s(n)||^2$.

sor, yielding smaller errors for both delays and amplitudes. The analytical MAP process was seen as a benchmark for good performance. The Gibbs Sampler, estimating the posterior distribution that the analytical processor maximizes, was expected to approach the performance of that but not to exceed it (as, theoretically, it cannot). Equations 6 and 8 reveal the source of the discrepancy. These equations make use of matrix inversion for calculation of the estimates; when the matrix has a large condition number, the estimation becomes less robust with highly varying results from one case to the next (large variance in the estimates).

The analytical processing was repeated using diagonal loading for matrix Λ. Diagonal loading stabilizes the inversion and, consequently, estimation of time delays and amplitudes; it reduces the variance in the estimates but also introduces biases [28]. The comparison between the new (with loading) analytical MAP estimator and the Gibbs Sampler is illustrated in Figures 4 and 5; the results show a very good agreement between the analytical MAP process and the Gibbs Sampling approximation to the analytical approach. A practically important result is that, because the Gibbs Sampler does not make use of any matrix inversions, it does not suffer from instabilities.

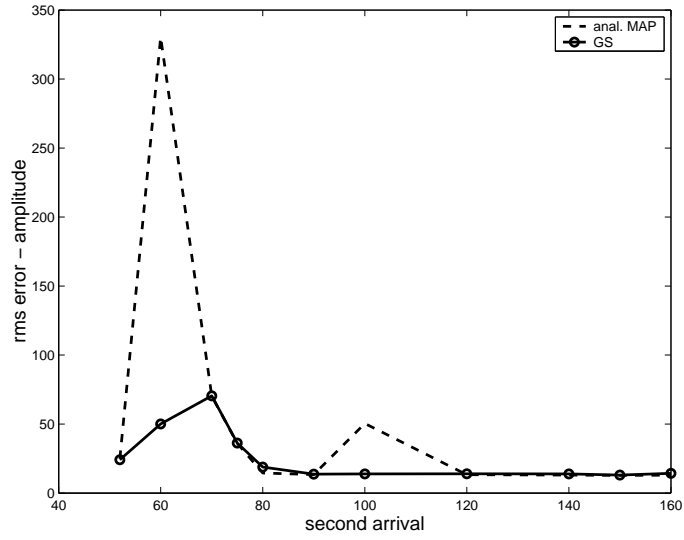Simulations were also run for noise variance $\sigma^2 = 0.05 \, ||s(n)||^2$. Fig-

10

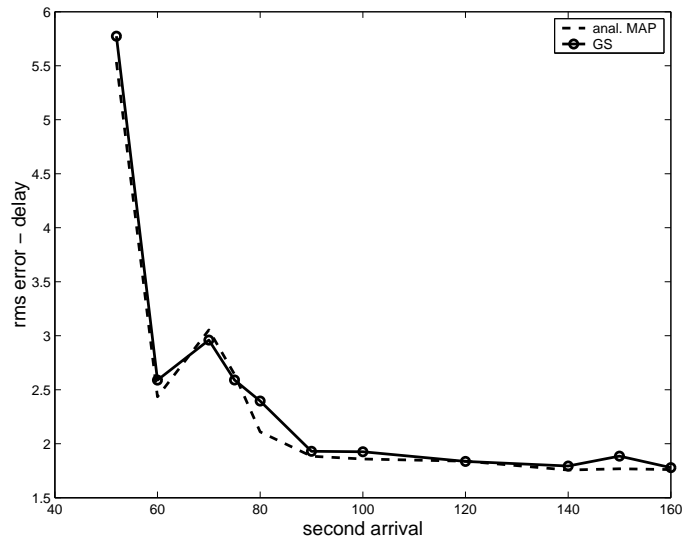Figure 3: Rms errors for amplitudes: $\sigma^2 = 0.01 \; ||s(n)||^2$.

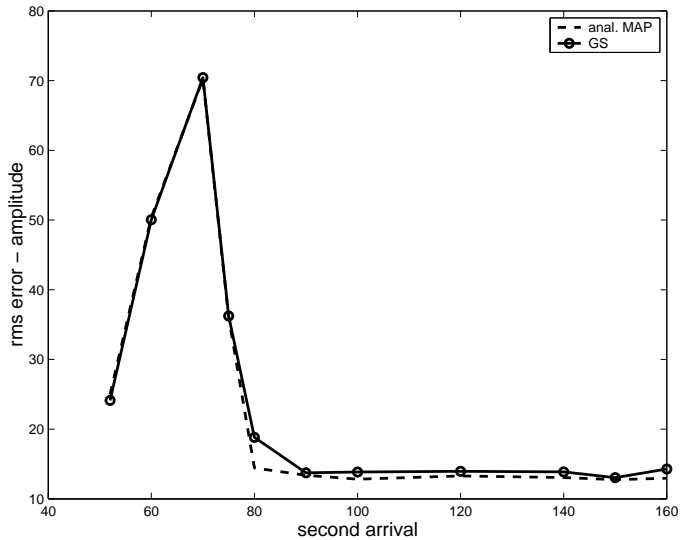Figure 4: Rms errors for time delays: $\sigma^2 = 0.01 \; ||s(n)||^2$.

11

Figure 5: Rms errors for amplitudes: $\sigma^2 = 0.01 \, ||s(n)||^2$.

ures 6 and 7 present the delay and amplitude results for the analytical MAP processor and the Gibbs Sampler; no diagonal loading has been performed. Again, substantial variance characterizes the analytically obtained results. A closer match between the two processors is achieved with diagonal loading for matrix Λ (Figures 6 and 7); especially in the case of time delays, the two processors perform very similarly. It should be pointed out here that the variance of the analytical processor estimates (and, consequently, the error) is artificially reduced with the loading process.

V   Gibbs Sampling vs. Expectation-Maximization

To circumvent the calculations required for the analytical MAP (or maximum-likelihood) process, Feder and Weinstein applied the EM method for time delay and amplitude estimation [20]. The method, being elegant and fast, became an important tool for time delay estimation.

EM is a two step process: starting from randomly picked initial values for the unknowns, the expectation of the log-likelihood is formed (expectation step). The expectation is, subsequently, maximized (maximization step) over
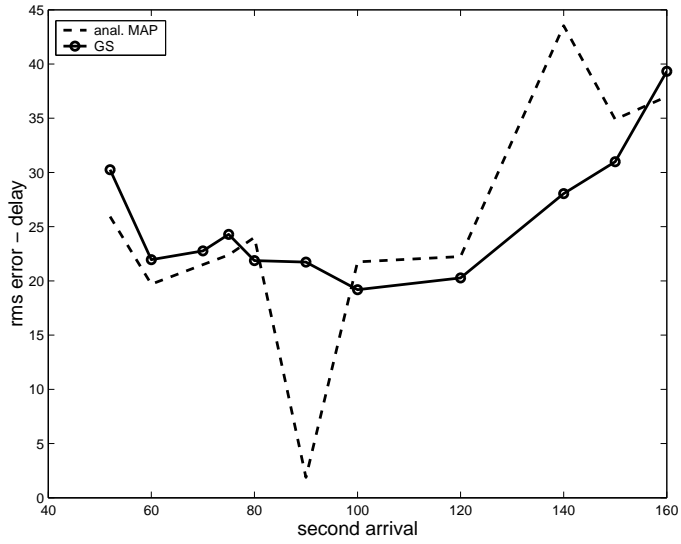
12

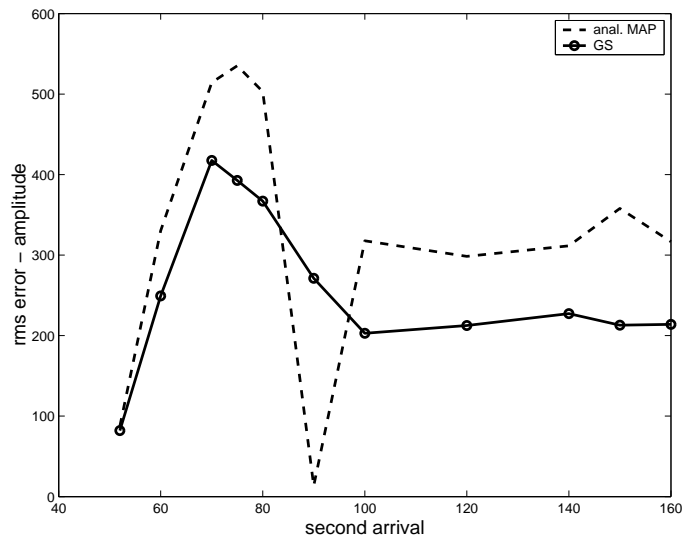Figure 6: Rms errors for time delays: $\sigma^2 = 0.05 \, ||s(n)||^2$.

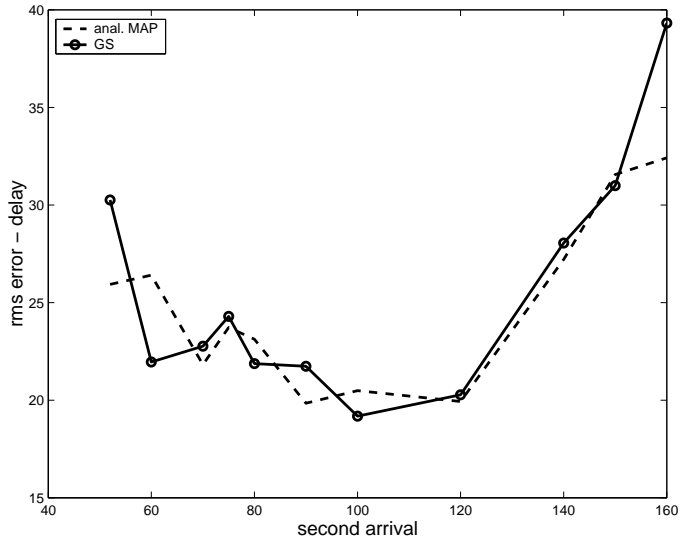Figure 7: Rms errors for amplitudes: $\sigma^2 = 0.05 \, ||s(n)||^2$.

13

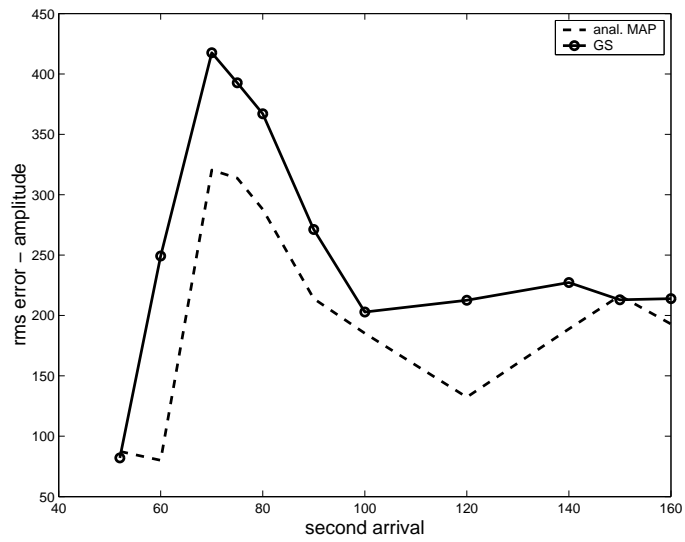Figure 8: Rms errors for time delays: $\sigma^2 = 0.05 \, ||s(n)||^2$.

Figure 9: Rms errors for amplitudes: $\sigma^2 = 0.05 \, ||s(n)||^2$.

14

Table 1: Medians of estimates from 200 realizations for the analytical MAP method, Gibbs Sampling (GS), and EM for noise variance $\sigma^2 = 0.01 \, ||s(n)||^2$.

true delay	anal. MAP	GS	EM
50	50	50	50
80	80	80	51
true amplitude	anal. MAP	GS	EM
100	101.13	102.72	54.21
-90	-92.50	-95.43	54.37

the unknowns and estimates of those are produced. The two-step procedure is repeated for a few iterations (typically, less than ten) until convergence to a maximum is achieved. This maximum, however, could be a local extremum, since, in many cases, likelihood functions (and posterior distributions) are multi-modal. EM performs a local search and the estimates depend on the initial conditions. As discussed in [20], it is best that the process is run with several sets of initial conditions for inferences to be made on whether convergence to the global maximum has been achieved.

Table 1 presents the median of the time delay and amplitude estimates from 200 realizations for one of the examined cases. For the estimation, we used analytical MAP estimation (with diagonal loading), the proposed Gibbs Sampler, and EM. Gibbs Sampling and EM required selection of initial conditions. For both methods and all runs the same initial conditions were selected: 10 and 20 for the two delays, and 30 and 30 for the two amplitudes. The poor EM performance, although startling at first sight, is not surprising. As mentioned earlier in the paper, EM is a local, hill-climbing technique. As such, its performance is highly dependent on initial conditions. All results reported in the previous section were generated with a single set of initial conditions. For EM to explore the search space more globally, as mentioned previously, it is usually recommended that several sets of initial conditions are employed and the process is applied several times. When such an approach is followed, EM typically gives good results in time-delay estimation.
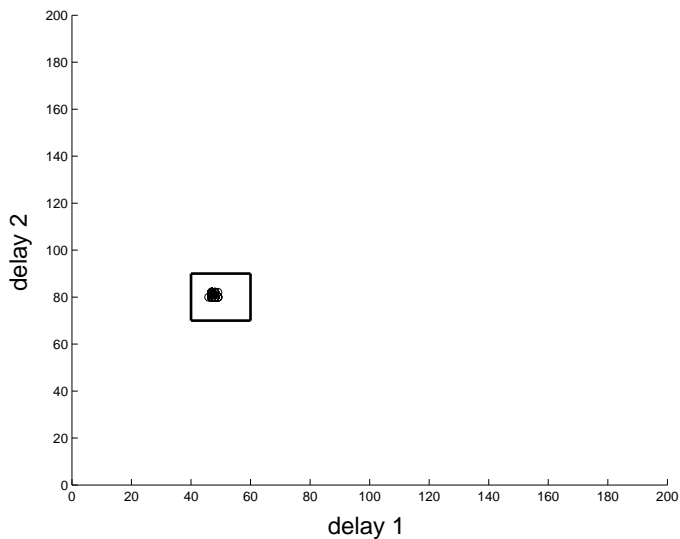
Figure 10: Scatter plot for delays obtained via Gibbs Sampling for 100 different initial conditions.

To illustrate this point, we selected a single noisy realization for the two-arrival case. We performed time delay and amplitude estimation for a single realization employing EM for 100 different sets of randomly selected initial conditions. At the same time, we applied the Gibbs Sampler to the same realization with the same set of initial conditions. The true delays were at 50 and 80; the amplitudes were 100 and -90. The Gibbs Sampler results, regardless of the initial conditions, yielded MAP estimates very close to the true values; those are shown in Figures 10 and 11, demonstrating small deviations.

Figure 12 shows samples for the second amplitude vs. iteration for two different initial conditions. The samples differ during the first few iterations but concentrate around the value of -90 after approximately 40 iterations.

Figures 13 and 14 show scatter plots of time delay and amplitude estimates for the first and second arrivals for the different initial conditions. The boxes in the plots demonstrate areas around the true parameter values. Only 14 sets of estimates fall inside the boxes (within 10 units of the true delays and 20 units of the true amplitudes).

A further test demonstrated a weakness of EM for closely spaced arrivals, which was not present in Gibbs Sampling. A case with arrivals at samples
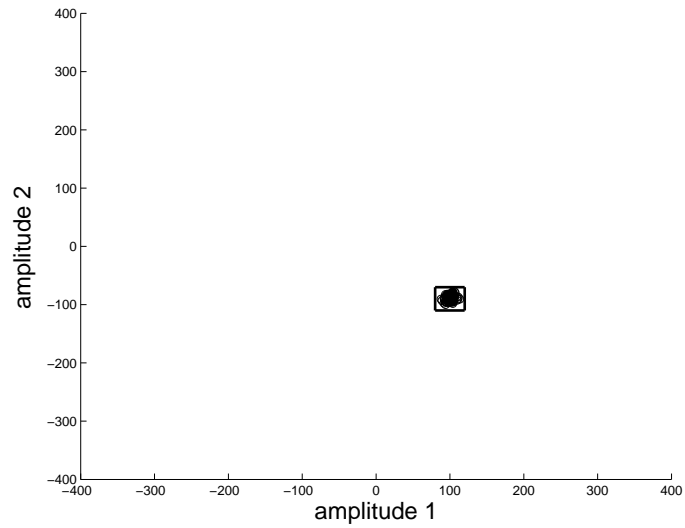
16

Figure 11: Scatter plot for amplitudes obtained via Gibbs Sampling for 100 different initial conditions.
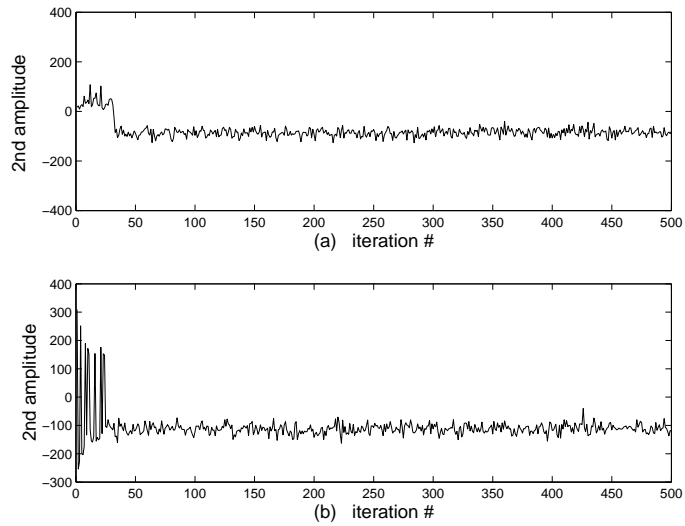

Figure 12: Samples obtained from the second amplitude for different initial conditions.
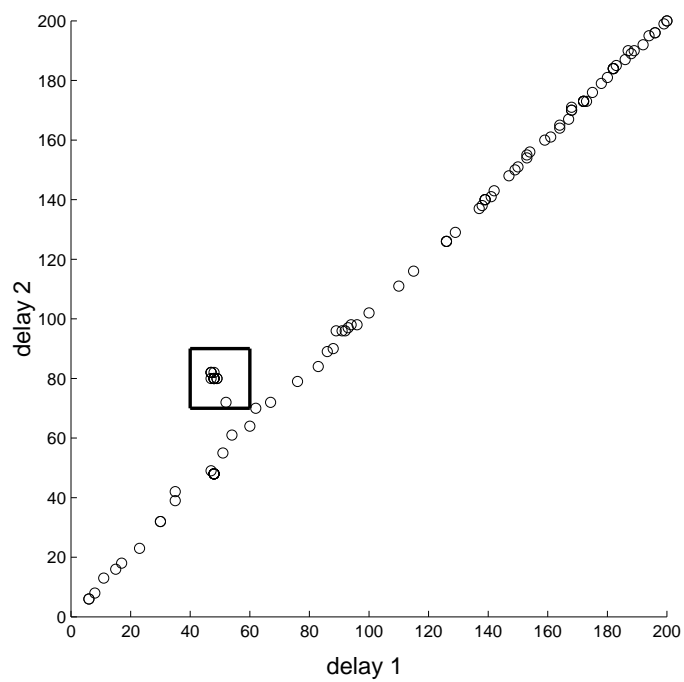
17

Figure 13: Scatter plot for delays obtained via EM for 100 different initial conditions.
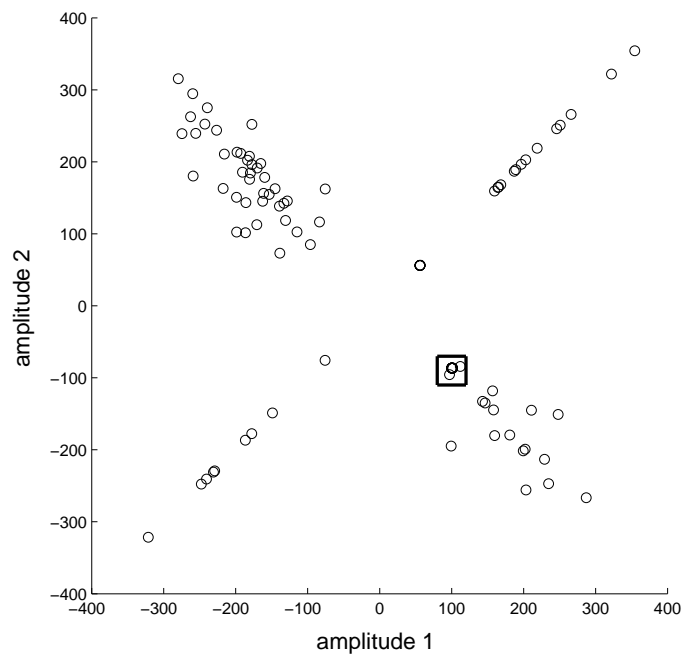
Figure 14: Scatter plot for amplitudes obtained via EM for 100 different initial conditions.
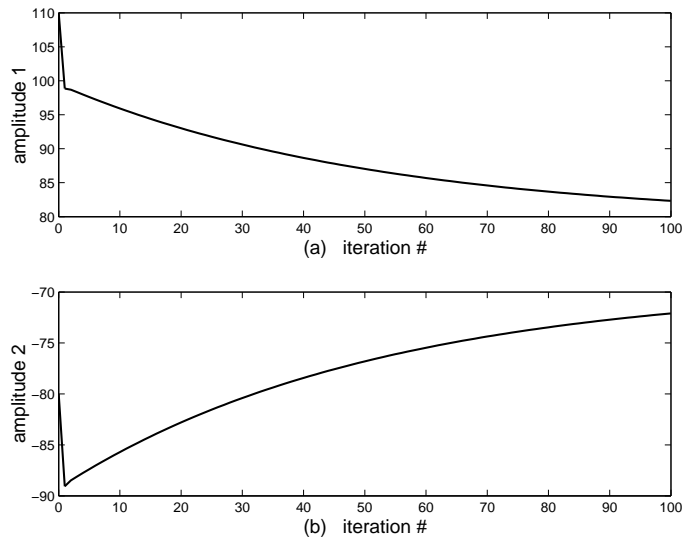
Figure 15: EM: Amplitudes vs. iteration for closely spaced arrivals (at samples 50 and 52).

50 and 52 was selected; the amplitudes were 100 and -90. Initial conditions for EM were set at 49 and 53 for delays and 100 and -80 for amplitudes (all very close to the true values). EM was run on one noisy realization for 100 iterations. The delay estimates did not change from the initial conditions (49 and 53). The amplitudes vs. iteration are shown in Figure 15. The figure demonstrates that there is a divergence in the amplitude estimates; as the iterations progress, the amplitude estimates deviate further away from the true values. Such a behavior is not present in EM results obtained for delays with a wider separation. Figure 16 shows amplitude vs. iteration for a case with delays at 50 and 80; true amplitudes were 100 and -90. Initial conditions were 49 and 83 for delays, and 110 and -80 for amplitudes. The amplitudes converged within less than ten iterations to values 101 and -86 (the delays stabilized to values 48 and 82 in one iteration). The divergence issue for the closely spaced arrivals did not appear in the Gibbs Sampling results.

EM and several other methods applied to the task of time delay and amplitude estimation yield point estimates (single values) for time delays and amplitudes, while our approach provides estimates of full posterior probability distributions of the parameters. These are particularly useful, since they
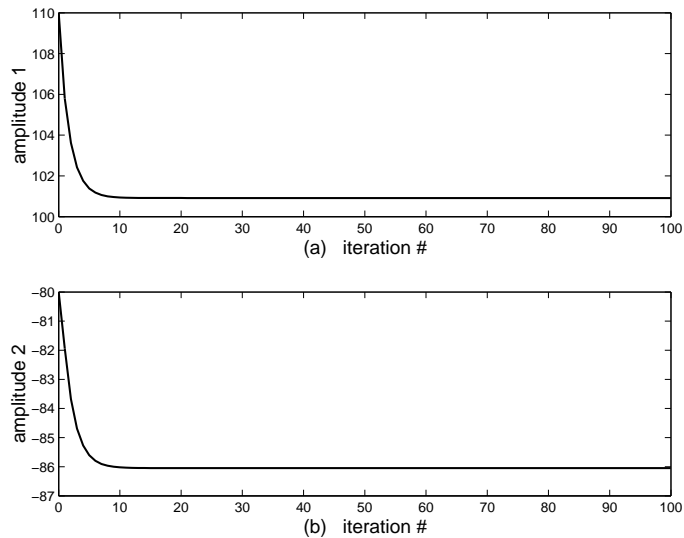
20

Figure 16: EM: Amplitudes vs. iteration for closely spaced arrivals (at samples 50 and 80).

include a substantial amount of information that is naturally absent from point estimates; point estimates do not offer insight into variance structure and multimodality, which would explain the vulnerability of point-estimators in finding a global maximum.

As an example, Figure 17 illustrates the two-dimensional posterior probability distributions over (a) delay 1 and amplitude 1 and (b) delay 2 and amplitude 2, respectively, for a two-arrival problem. The true arrival times are at samples 50 and 80; the true amplitudes are 100 and -90. The Gibbs Sampler gives the following estimates: 44 and 100 for the first delay and amplitude, and 193 and 82 for the second delay and amplitude. Although the first arrival is quite accurately characterized, the second arrival is erroneously estimated. However, observing Figure 17(b), we can see that there is some significant probability concentration around the correct values of 80 and -90 (for delay and amplitude) for the second arrival. Thus, although in terms of point estimates, the second arrival is not correctly recovered by the Gibbs Sampler (or other point estimators), the uncertainty introduced by the second peak of the posterior distribution strongly suggests that an alternative set of estimates may be relevant, and that further exploration is necessary for a better identification of the second arrival.
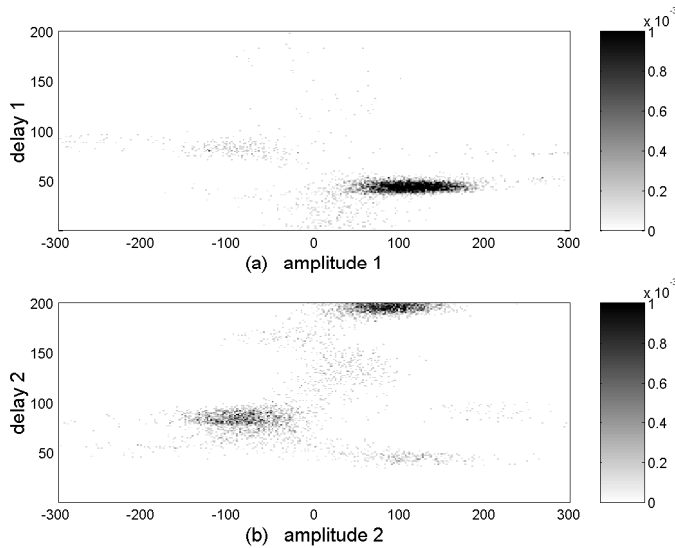
21

Figure 17: Estimated posterior probability distributions of time delays and amplitudes via Gibbs Sampling: (a) first arrival, (b) second arrival.

VI   Unknown number of arrivals

In the preceding section, the number of arrivals was assumed known and equal to two. Typically, there is no precise information on the number of arrivals, which depends on the propagation medium. Here, the assumption of a known arrival number is relaxed; along with the amplitudes, delays, and noise variance, the number arrivals is estimates.

According to the analysis presented in this paper, the joint probability distribution we have so far estimated with Gibbs Sampling is in reality a conditional distribution, conditioning being on the number of arrivals M: $p(a_1, \ldots, a_M, n_1, \ldots, n_M, \sigma^2/r(n), M)$. Following the Bayesian paradigm, a prior distribution can be specified for M. In the absence of specific information on M, we select a uniform prior:

$$p(M) = \frac{1}{M_2 - M_1 + 1}, \quad M_1 \le M \le M_2, \tag{15}$$

where M_1 and M_2 are lower and upper bounds for the expected arrival number.

For the uniform prior of Equation 15, estimation of the number of arrivals, M, can be achieved using the Schwartz-Rissanen criterion for model selection [29]. According to the criterion, M is chosen in order to minimize

$$-\log p(r(n)/\hat{a}_i(i=1,\ldots,M), \hat{n}_i(i=1,\ldots,M), \hat{\sigma^2}) =$$
$$(-\frac{1}{2\sigma^2} \sum_{n=1}^{N} (r(n) - \sum_{i=1}^{M} \hat{a}_i s(n - \hat{n}_i)^2) - M\log(N), \qquad (16)$$

where N is the length of the received signal, and $\hat{a}_i$ and $\hat{n}_i$ are the amplitude and delay estimates obtained from the Gibbs Sampler for a selected value of M between M_1 and M_2; $\hat{\sigma^2}$ is the estimate of the unknown variance. The Schwartz-Rissanen criterion uses estimates of amplitudes, delays, and variance to eventually calculate M. The criterion can be suitably altered in case of priors for M other than the uniform one of Equation 15.

Two-hundred runs were generated to test the estimation of the number of arrivals; the variance was set to $0.01||s(n)||^2$. Three arrivals were present (at samples 50, 75, and 150 with amplitudes 100, -80, and 60, respectively). Using prior knowledge, it was assumed that M could vary between 2 and 5. Using the Schwartz-Rissanen criterion, M was estimated correctly to be 3 138 out of the 200 times.

VII   Convergence of the Gibbs Sampler

There is not a straightforward manner with which to choose an optimal number of samples necessary for the Gibbs Sampler to converge. Monitoring convergence is a topic of open research. Several approaches are recommended for testing the convergence of the estimated distribution to the true joint posterior [21]. In this work, the Gibbs Sampler was originally tested by initializing the process with different parameter values. This is a standard procedure used to test convergence of Monte Carlo methods in general [21, 30]. As shown in Section V (Figures 12 and 11), the Gibbs Sampler results were insensitive to initial conditions, indicating convergence of the process to the true posterior distribution.

Running several parallel Gibbs Samplers initialized in different ways is an effective monitor for convergence but is computationally demanding. As an

Table 2: Modes of the posterior distributions for the amplitudes vs. iteration groups.

Iterations	a_1	a_2	a_3
1001-2000	455	-405	75
2001-3000	135	-125	75
3001-4000	395	-475	75
4001-5000	1175	-1175	75
5001-6000	1255	-1215	75
6001-7000	85	-75	65
7001-8000	95	-85	65
8001-9000	105	-85	65
9001-10000	95	-85	65

alternative convergence test, we monitored the modes of the marginal posterior distributions of the parameters (which can be readily calculated from the Gibbs Sampling results [27]) for a single set of initial conditions. Table 2 demonstrates the process for a three-arrival problem. The table includes modes for the distributions of the three amplitudes (true values: 100, -80, 60) for groups of 1000 iterations. The modes for the first two amplitudes vary significantly for different groups of iterations up to iteration 6000. Following that, the modes stabilize close to 100 and -85. After stabilization has been observed, the Gibbs Sampler is stopped.

It was observed that, as expected, the number of necessary iterations for convergence increased with the number of arrivals M. The necessary number of iterations increased with noise variance as well. A closed-form relationship, however, relating convergence and factors affecting the estimation process was not derived. For the two arrival problem that we investigated in Section IV, we found that 5000 iterations were adequate for most realizations.

VIII   Conclusions

A novel approach for time delay and amplitude estimation in multipath environments was presented. The technique estimates joint posterior distributions using Gibbs Sampling; once an estimate of the posterior distribution of time delays and amplitudes is available, its maximum yields estimates for the unknowns. The proposed method differs from other approaches typically used in time delay estimation in that, in addition to point estimates, it offers full posterior distributions for the parameters of interest. Those distributions could potentially highlight information (such as a missed arrival), which would have otherwise remained obscure.

The method performs well compared to the analytical MAP estimator. It is also stable with respect to initial conditions and is not adversely affected in the case of closely spaced arrivals. Its convergence is monitored through an examination of the stability of statistics (modes, in our case) of marginal posterior distributions.

References

[1] A. Tolstoy, *Matched Field Processing for Underwater Acoustics*. Singapore: World Scientific, 1993.

[2] A. Baggeroer, W. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameters estimation problem," *J. Acoust. Soc. Am.*, vol. 83, pp. 571–587, 1988.

[3] H. Bucker, "Matched-field tracking in shallow water," *J. Acoust. Soc. Am.*, vol. 96 (6), 1994.

[4] S. E. Dosso, M. R. Fallat, B. J. Sotirin, and J. L. Newton, "Array element localization for horizontal arrays via Occam's inversion," *J. Acoust. Soc. Am.*, vol. 104 (2), pp. 846–859, 1998.

[5] S. E. Dosso, G. H. Brooke, S. J. Kilistoff, B. J. Sotirin, V. K. McDonald, M. R. Fallat, and N. E. Collison, "High-precision array element localization for vertical line arrays in the Arctic Ocean," *IEEE Journal of Oceanic Engineering*, vol. 23, no. 4, pp. 365–379, 1998.

[6] S. E. Dosso and B. Sotirin, "Optimal array element localization," *J. Acoust. Soc. Am.*, vol. 106, pp. 3445–3459, 1999.

[7] E. K. Westwood and D. P. Knobles, "Source track localization via multipath correlation matching," *J. Acoust. Soc. Am.*, vol. 102, no. 5, pp. 2645–2654, 1997.

[8] P. Pignot and R. Chapman, "Tomographic inversion for geoacoustic properties in a range dependent shallow water environment," *J. Acoust. Soc. Am.*, vol. 104 (3), pp. 1338–1348, 1998.

[9] L. Jaschke, "Geophysical inversion by the freeze bath method with an application to geoacoustic ocean bottom parameter estimation," *Master's Thesis, University of Victoria*, 1997.

[10] L. Jaschke and R. Chapman, "Matched field inversion of broadband data using the freeze bath method," *J. Acoust. Soc. Am.*, vol. 106 (4), pp. 1838–1851, October 1999.

[11] X. Ma, *Efficient inversion methods in underwater acoustics*. PhD thesis, New Jersey Institute of Technology, May 2001.

[12] J. E. Ehrenberg, T. E. Ewart, and R. D. Morris, "Signal processing techniques for resolving individual pulses in a multipath signal," *J. Acoust. Soc. Am.*, vol. 63, pp. 1861–1865, June 1978.

[13] G. C. Carter, "Time delay estimation for passive sonar signal processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, pp. 463–470, June 1981.

[14] G. C. Carter, ed., *Coherence and time delay estimation*. IEEE Press, 1993.

[15] R. J. Tremblay, G. C. Carter, and D. W. Lytle, "A practical approach to the estimation of amplitude and time delay parameters of a composite signal," *IEEE Journal of Oceanic Engineering*, vol. OE-12, pp. 273–278, January 1987.

[16] I. Kirsteins, "High resolution time delay estimation," in *ICASSP-87*, pp. 451–453, 1987.

[17] S. Umesh and D. W. Tufts, "Estimation of parameters of exponentially damped sinusoids using fast maximum likelihood estimation with application to NMR spectroscopy data," *IEEE Transactions on Signal Processing*, vol. 44, pp. 2245–2259, September 1996.

[18] A. D. Blackowiak and S. D. Rajan, "Multi-path arrival estimates using simulated annealing: application to crosshole tomography experiment," *ieeejoe*, pp. 1861–1865, July 1995.

[19] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B*, vol. 39, no. 1, pp. 1–38, 1977.

[20] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.

[21] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman and Hall, CRC, first ed., 1996.

[22] Z.-H. Michalopoulou and M. Picarelli, "A Gibbs sampling approach to maximum a posteriori time delay and amplitude estimation," in *ICASSP-02*, 2002.

[23] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer Verlag, second ed., 1985.

[24] G. Box and G. Tiao, *Bayesian Inference in Statistical Analysis*. Addison Wesley, 1973.

[25] M. A. Tanner, *Lecture Notes in Statistics: Tools for Statistical Inference (Observed Data and Data Augmentation Methods)*, vol. 67. Springer Verlag, second ed., 1992.

[26] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, pp. 721–741, November 1984.

[27] A. E. Gelfand and A. F. Smith, "Sampling–based approaches to calculating marginal densities," *Journal of the American Statistical Association*, vol. 85, pp. 398–409, November 1990.

[28] R. Aster, B. Borchers, and C. Thurber, *Parameter Estimation and Inverse Problems*. Elsevier, 2004.

[29] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, pp. 387–392, April 1985.

[30] S. Dosso, "Quantifying uncertainty in matched field inversion. I A fast Gibbs sampler approach.," *J. Acoust. Soc. Am.*, vol. 111, no. 1, pp. 129–142, 2002.