A gradient-based optimization approach for underwater acoustic source localization `FREE`

Dariush Kari; Andrew C. Singer ; Hari Vishnu ; Amir Weiss

Check for updates

View Online   Export Citation

Articles You May Be Interested In

Development tools for deep learning models of acoustical signal processing

J Acoust Soc Am (April 2022)

A physics-informed neural network for sound propagation in the atmospheric boundary layer

Proc. Mtgs. Acoust. (March 2021)

Correlation of acoustic wavefront and signal time-base instabilities in the ocean

J Acoust Soc Am (June 1976)

184th Meeting of the Acoustical Society of America

Chicago, Illinois

8-12 May 2023

*Computational Acoustics: Paper 2pSP11

A gradient-based optimization approach for underwater acoustic source localization

Dariush Kari and Andrew C. Singer

Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA; dkari2@illinois.edu; acsinger@illinois.edu

Hari Vishnu

Acoustic Research Laboratory, National University of Singapore, Singapore, SINGAPORE; harivishnu@gmail.com

Amir Weiss

Massachusetts Institute of Technology, Cambridge, MA, USA; amirwei@mit.edu

In underwater acoustic localization via matched-field-processing, given a propagation model and a suitable environmental parameterization, one searches for the location (of the transmitter or receiver) whose replica field is closest to the observed one. The high computational complexity of such non-gradient-based optimization methods renders them infeasible for many real-time scenarios, especially when an accurate solution is desired, due to resolution of the search grid required, or as the search dimensionality increases (e.g., when it is necessary to optimize over uncertain environmental parameters such as sound speed or bathymetry). In this work, we propose a ray-based, differentiable model for acoustic propagation for the purpose of a gradient-based optimization for localization. For localization applications in which accurate times of arrivals might not be available, the proposed method can be adapted to work without requiring this information. In such a scenario, it seeks the location (and possibly environmental parameters) that minimize the squared-error between the observed signal and its estimate via the differentiable model. We leverage the PyTorch optimization and auto-differentiation tools for the implementation and demonstrate successful localization on synthetic data inspired from a real-world scenario in a dense multipath environment.

*POMA Student Paper Competition Winner

12 March 2025 08:57:25

1. INTRODUCTION

Conventional model-based underwater acoustic (UWA) localization algorithms such as matched field processing (MFP)[1] are highly dependent upon accurate environmental knowledge[2–4] that is rarely available. To perform localization using MFP, one needs to generate replica fields for different candidate locations using an underwater propagation model, such as Bellhop[5] or KRAKEN,[6] whose parameters are set based on the environment of interest. The accuracy of finding the best matching location is sensitive to the mismatch of these model parameters, which may not be known at all, or not known accurately.

Tackling the lack of prior knowledge and overcoming sensitivity to mismatch involves including an additional search over the unknown parameters to find the best set of parameters that match the measurements. However, this leads to a substantial increase of the search space dimensionality relevant for the optimization. The optimization involved becomes impractical when the dimensionality increases, in part because one cannot use gradient-based (GB) methods to find the optimum, due to non-differentiability of propagation models, like Bellhop. Tackling this typically requires using brute-force methods or non-gradient-based optimization methods such as simulated annealing (SA)[7] or genetic algorithm (GA),[8] which require a high computational effort, often rendering them infeasible for real-time applications. To address this computational challenge, we propose an alternative method by developing a differentiable model for acoustic wave propagation, which allows us to make use of GB optimization methods for localization. In this approach, the model parameters are tuned based on the gradient of the incurred cost, thus reducing the complexity involved in the previously-mentioned approaches.

This GB optimization comes with challenges of its own. One such challenge is the risk of getting stuck in local minima if the cost function is not well-behaved, i.e, if it is not locally convex in a neighborhood around the global minimum that encompasses the initial point. The cost functions used in this paper are non-convex over each of the spatial dimensions. Faced with such a cost function, we propose a smoothing technique to make the cost function more amenable to the GB optimization by broadening the local convexity interval around the ground-truth, thus making the method feasible.

While some MFP-based methods exploit hydrophone *arrays*,[2,3] deploying an array may not be possible in some cases, and the compact setup offered by a *single-hydrophone*[9–12] may be more attractive. Examples of such scenarios include applications that need to cover large areas (e.g. bioacoustics) or where the platforms are small,[10] and usually with broadband sources or in shallow waters where multipath information can be leveraged for source localization. We focus on the single-hydrophone case as a starting point, although the proposed approach can be extended to the array-based localization as well.

The multipath structure in shallow water environments leads to interference between arrivals that can make them unresolvable, which in addition to low signal-to-noise ratio (SNR)[9] can impede accurate estimation of the arrival times. Nevertheless, for successful single-hydrophone localization, it is not necessary to directly estimate the times of arrival; for instance localization can be performed by correlating the received waveform and a model predicted waveform as performed by Jesus, et al.[9] In Fig. 1, we show an example of such unresolvable arrivals in real data recorded in the Scripps Ocean Atmosphere Research Simulator (SOARS) at the Scripps Institution of Oceanography (SIO), where the transmitter is an ITC-1032 and the receiver is an HTI-96-Min hydrophone.

In this work, we develop a differentiable model for the received signal using a ray-based approach that takes the environmental parameters, transmitted signal waveform, and the transmitter and receiver's locations and synthesizes the received signal by combining raypaths. We then estimate the parameters of interest, e.g., the transmitter location, via GB optimization on a target cost function involving the difference between the model-synthesized and recorded signals at the receiver. Specifically, we use the Adam optimizer, implemented in PyTorch. Note that in single-hydrophone scenarios with multipath, the frequency diversity of the broadband sources can compensate for the spatial diversity provided by arrays.[9,10] Thus, we use a Gaussian pulse-shaped waveform as a broadband transmitted signal.

(a)   *(b)*
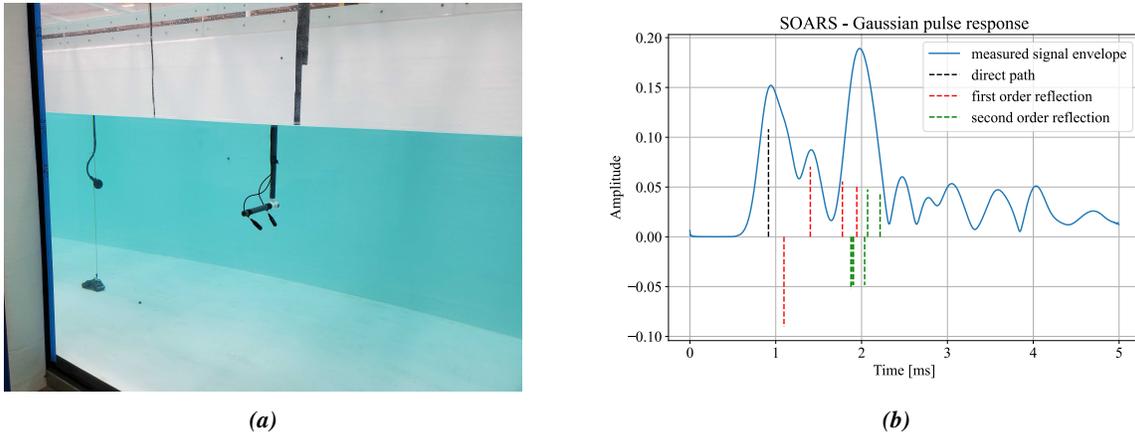
Figure 1: (a) Experiment setup in the SOARS tank. It consists of a transmitter (ITC-1032 transmitter on the left), and two HTI-96-Min receivers on the right of the picture. (b) The Gaussian pulse response of the channel in the frequency range of $26-38\ kHz$, along with model-predicted arrival times and amplitudes for different raypaths in the tank. A negative amplitude indicates a surface interaction in the arrival. The first peak, around $1\ ms$, corresponds to the interference of the direct path and the surface reflected arrival. Furthermore, the two sidewall reflections seem to have merged with each other and with higher-order reflections to give a larger peak at around $2\ ms$.

The rest of the paper is organized as follows. In Section 2 we present the problem formulation. We then propose a GB algorithm in Section 3, followed by simulation results in Section 4. We finish the paper with some concluding remarks in Section 5.

2.   PROBLEM FORMULATION

Notation: Vectors are denoted by bold letters. Also, subscripts "s" and "r" refer to "source" and "receiver", respectively. In addition, a subscript 0 indicates an initial estimate or belief of the value for an unknown parameter. For a temporal signal $x(t)$, t denotes the continuous-time variable and the corresponding discrete-time signal is $x[k] = x(k/Fs)$, where F_s is the sampling frequency.

In the current scenario, we use a single receiver to localize a source that emits a modulated Gaussian pulse $s(t) = e^{-\frac{B^2t^2}{2}}\cos(2\pi f_c t)$, where B is the pulse bandwidth and f_c is the central frequency. The received signal can be expressed as a sum of contributions from multiple raypaths[13] and the ambient noise $\eta(t)$ as

$$r(t, \boldsymbol{\theta}_s, \boldsymbol{\theta}_r) = \sum_{i=1}^{N} \beta_i(\boldsymbol{\theta}_s, \boldsymbol{\theta}_r)\, s\left(t - \tau_i(\boldsymbol{\theta}_s, \boldsymbol{\theta}_r)\right) + \eta(t), \tag{1}$$

where N is the number of paths in the multipath received signal, $\boldsymbol{\theta}_s = [x_s, y_s, z_s]$ is the source's location, $\boldsymbol{\theta}_r = [x_r, y_r, z_r]$ is the receiver's location, and β_i and τ_i indicate the i-th path attenuation coefficient and delay, respectively. β_i and τ_i can be represented[13] as

$$\beta_i = \frac{\rho_i e^{-\alpha \ell_i}}{\ell_i}, \qquad \tau_i = \ell_i/c, \tag{2}$$

where ℓ_i is the i-th path length, α is the (known) frequency-dependent volume attenuation, ρ_i is the product of all reflection coefficients corresponding to all reflecting surfaces encountered along the i-th path, and c

is the sound speed. Note that in well-mixed shallow waters and confined spaces like water tanks, c can be assumed constant.

For notational simplicity, since the receiver's location is fixed in forthcoming analysis, we represent $r(t, \boldsymbol{\theta}_s, \boldsymbol{\theta}_r)$, $\tau_i(\boldsymbol{\theta}_s, \boldsymbol{\theta}_r)$, and $\beta_i(\boldsymbol{\theta}_s, \boldsymbol{\theta}_r)$ by the shorter versions $r(t, \boldsymbol{\theta}_s)$, $\tau_i(\boldsymbol{\theta}_s)$, and $\beta_i(\boldsymbol{\theta}_s)$, respectively. Note that β_i can be positive or negative and depends on the transmit signal parameters, path length, and the reflection coefficients.[13] Given the transmitted signal $s(t)$, we can simplify the received signal model, and thereby the localization problem, by incorporating only the paths with significant contributions to the received signal; for instance Weiss, et al.[14] considers only the direct path, bottom reflection, and the surface reflection for short ranges in shallow ocean environments. Considering the N_p most significant paths ($N_p < N$), the noiseless received signal is modeled by

$$\hat{r}(t, \boldsymbol{\theta}_s') = \sum_{i=1}^{N_p} \beta_i\left(\boldsymbol{\theta}_s'\right) s\left(t - \tau_i(\boldsymbol{\theta}_s')\right). \tag{3}$$

Given the environmental parameters, the problem of localization reduces to finding the values $\boldsymbol{\theta}_s^*$ of $\boldsymbol{\theta}_s'$ for which $\hat{r}(t, \boldsymbol{\theta}_s')$ is the "most similar" to the measured signal $r(t, \boldsymbol{\theta}_s)$. To combat noise and perform pulse-compression, we matched-filter $r(t)$ and $\hat{r}(t)$ with the known $s(t)$, which yield $r_m(t)$ and $\hat{r}_m(t)$, respectively.

By defining a cost function $\mathcal{C}\left(\mathbf{r}_m, \hat{\mathbf{r}}_m\right)$, which characterizes the difference between the estimated and recorded received signal in some sense, the source localization problem can be formulated as

$$\boldsymbol{\theta}_s^* = \arg\min_{\boldsymbol{\theta}_s'} \mathcal{C}\left(\mathbf{r}_m, \hat{\mathbf{r}}_m\right), \tag{4}$$

where $\mathbf{r}_m = [r_m[k]]_{k=0}^{M-1}$ and $\hat{\mathbf{r}}_m = [\hat{r}_m[k]]_{k=0}^{M-1}$, and M is the total number of time samples considered, which is assumed long enough to cover N_p paths. In order to solve (4) via GB optimization, we introduce a GB localization algorithm in the following section.

3. GRADIENT-BASED APPROACH

Assuming that: (1) the model that generates $\hat{r}(t, \boldsymbol{\theta}_s')$ can be differentiated w.r.t. $\boldsymbol{\theta}_s'$; (2) the cost function $\mathcal{C}$ is convex in an interval around the global minimum $\boldsymbol{\theta}_s^*$; and (3) there is an initial estimate of the $\boldsymbol{\theta}_s$ that lies in that interval, we can use a GB algorithm to achieve $\boldsymbol{\theta}_s^*$, as illustrated in Fig. 2. In the rest of this section, we develop a feasible GB algorithm by ensuring that these assumptions hold.

We build our model using the differentiable programming framework provided by PyTorch. The proposed algorithm includes three stages. In the first stage, we rely heavily on the prior knowledge and assumptions (discussed next) to find an initial estimate. As the algorithm progresses, it uses finer information obtained from the observed signal to improve this estimate; in the second stage, it uses the signal envelope, and in the third stage it uses the temporal signal as it contains all available information (envelope and phase). Moreover, in stages 2 and 3, as the algorithm progresses, the step size in the optimizer is gradually reduced, allowing it to make fine-tuned changes to arrive at the cost function minimum.

A. STAGE 1: INITIAL ESTIMATE USING TIMES-OF-ARRIVALS

The initial stage involves using the times-of-arrivals (TOAs) extracted from the received signal. The *approximate* TOAs are obtained from the peaks in the envelope of $r_m(t)$ defined as

$$r_e(t) = \|r_m(t) + \jmath\mathcal{H}(r_m(t))\|_2\,,$$
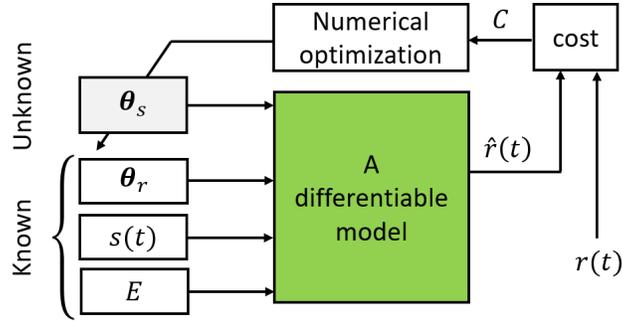
Figure 2: The block diagram of the proposed localization approach. The algorithm currently assumes prior knowledge of the source signal $s(t)$, receiver's location θ_r, and the set of environmental parameters E including the sound speed, geometry, etc. However, if there is uncertainty in any of the parameters, those can be additionally included in the set of parameters to be estimated via optimization.

where $\mathcal{H}(.)$ indicates the Hilbert transform operation and $r_e(t)$ is the envelope of $r(t)$. Note that there is not a bijection between the arrivals and the peaks because some arrivals are unresolvable. In addition, each peak must be labelled, which is a difficult problem in general. These significantly affect the initial estimate, and necessitate further refining of this estimate, which will be addressed in next stages.

Nevertheless, we are able to achieve a reasonable initial estimate using maximum a posteriori (MAP) estimation by assuming that for each θ'_s, the peak corresponding to the i-th path in the received signal will occur at $t = p_i$ and

$$p_i \sim \mathcal{N}(\tau_i(\theta'_s), \sigma^2),$$

where τ_i is calculated based on the geometry and σ depends on the signal bandwidth and the SNR (assuming σ is fixed, it becomes irrelevant in the MAP estimation). Although such assumptions might not hold in general, they still yield an acceptable initial estimate.[15] Consequently, our approach is as follows:

1. We assign the first peak to the direct path and determine a grid $\mathcal{A}$ of points that can yield that TOA. If there is a priori information about θ_s, it can be used in this stage to reduce the grid size. For instance, the prior information can be in the form of $z_s \in [z_{min,0}, z_{max,0}]$.

2. For each grid point, we obtain the τ_i's for all paths. Note that, we use only first-order reflections in this stage. For example, in a two-dimensional environment provided in Fig. 3, we calculate τ_i for the bottom reflected and surface reflected paths. Then, we use a nearest neighbor approach to assign each τ_i to the most-likely peak of $r_e(t)$. Based on the assignment, we calculate an error of $e_i(\theta'_s) = (\tau_i(\theta'_s) - p_i)^2$, and, $e(\theta'_s) = \sum_i e_i(\theta'_s)$.

3. Finally, the initial estimate is
$$\theta_{s,0} = \arg \min_{\theta'_s \in \mathcal{A}} e(\theta'_s).$$

Note that since our model is differentiable, even in this stage, we leverage the GB optimization with a target cost of $e(\theta'_s)$ to refine each grid point.

Figure 4 shows an example of this stage for localization in the 3-ray environment of Fig. 3. It is observed that although the set of TOAs approximately matches that of the peaks, the error in the peak assignment has resulted in a discrepancy between $r_e(t)$ and $\hat{r}_e(t)$ that will be addressed in next stages. The performance of this stage of the algorithm highly depends on the geometry; symmetries in the geometry adversely affect the performance, because the algorithm does not distinguish between different reflections.
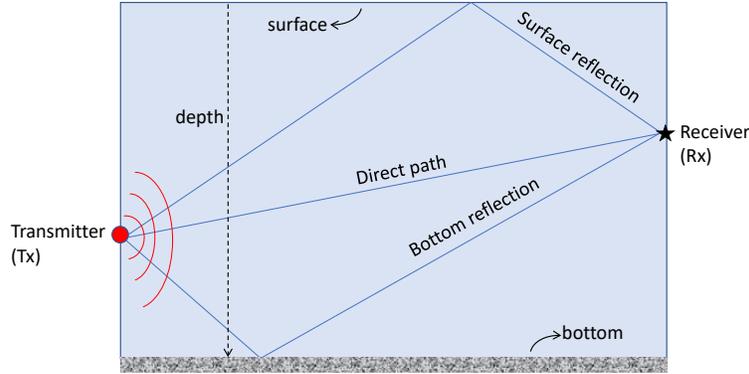
Figure 3: The 3 paths in the illustrative 3-ray model with flat surface and bottom and iso-velocity channel.
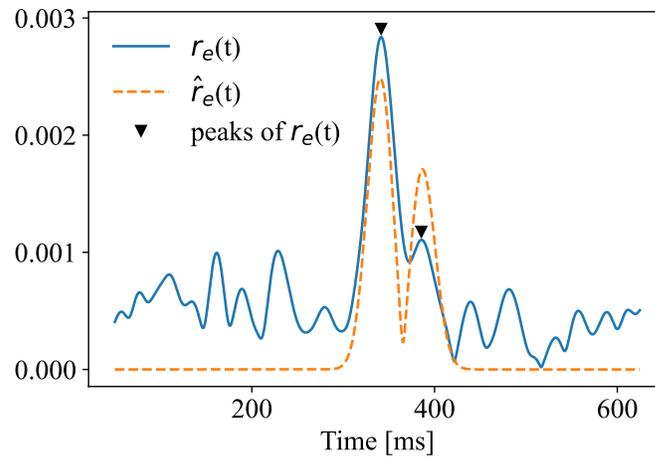
Figure 4: An example of the first stage in the 3-ray environment of Fig. 3, with SNR=20 dB. Although the peaks of $r_e(t)$ and $\hat{r}_e(t)$ coincide, there is a discrepancy between $r_e(t)$ and $\hat{r}_e(t)$.

B.   STAGE 2: ESTIMATE REFINING USING SIGNAL ENVELOPE

If the initial estimate of $\hat{\boldsymbol{\theta}}_s$ is less than half a wavelength apart from the true $\boldsymbol{\theta}_s$, one can use a squared error (SE) cost function to refine this estimate and achieve a more accurate solution. The SE cost is defined as

$$\mathcal{C}_{\text{SE}}\left(\mathbf{r}_m, \hat{\mathbf{r}}_m\right) = \|\mathbf{r}_m - \hat{\mathbf{r}}_m\|_2^2. \tag{5}$$

However, the initial estimate does not necessarily lie in the half-wavelength interval around the true solution (for example, if it is obtained by the TOA-based method of stage 1). Moreover, since the received signal is a passband signal, the SE cost function highly fluctuates around the true $\boldsymbol{\theta}_s$, making the problem non-convex. As an illustrative example, consider the two-dimensional 3-ray model provided in Fig. 3, and assume that the water depth is $D = 200$ m, $y_s = y_r = 0$, x_r, x_s and z_r are known, and the only unknown is z_s. Figure 5 shows the SE cost surface for this problem, where the ground truth is $z_s = 20$ m. Evidently, the SE cost function can fluctuate significantly.

To tackle the challenge posed by the fluctuations in cost function due to the passband signal, we propose to first refine $\boldsymbol{\theta}_{s,0}$ using the signal envelope $r_e(t)$ instead of $r_m(t)$ in the cost function. Therefore, we define another cost function based on the Kullback-Leibler (KL) divergence[16] between the envelopes of the
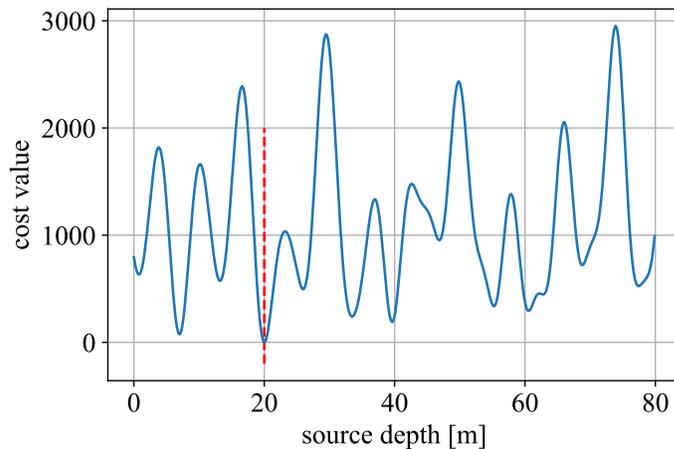
Figure 5: The SE cost surface, in the 3-ray environment in the presence of AWGN at an SNR=20 dB. The true source depth is 20 m.

estimated and received signals as

$$\mathcal{C}_{\text{KL}}\left(\mathbf{r}_e, \hat{\mathbf{r}}_e\right) = D_{\text{KL}}\left(\frac{\mathbf{r}_e}{C}, \frac{\hat{\mathbf{r}}_e}{\hat{C}}\right) + D_{\text{KL}}\left(\frac{\hat{\mathbf{r}}_e}{\hat{C}}, \frac{\mathbf{r}_e}{C}\right), \tag{6}$$

where $C = \sum_{k=0}^{M-1} r_e[k]$ and $\hat{C} = \sum_{k=0}^{M-1} \hat{r}_e[k]$ are the normalizing factors that make the areas under the envelopes equal to 1, and D_{KL} denotes the KL divergence. The intuition behind using the SE and KL divergence is that the SE corresponds to the maximum likelihood solution in the presence of additive white Gaussian noise, and the KL divergence (6) for Gaussian pulses takes into account the difference in the peak locations, i.e., time of arrivals, that is usually used for localization.[17, 18]

Although using the envelopes changes the optimization landscape, it does not completely remove the fluctuations, as observed in Fig. 6. Therefore, in this stage, we use a smoothing technique, where the cost is defined as

$$\mathcal{C}_{\mathcal{W}(\boldsymbol{\theta}_s')}\left(\mathbf{r}, \hat{\mathbf{r}}_{\boldsymbol{\theta}_s'}\right) = \sum_{\boldsymbol{\theta}_w \in \mathcal{W}(\boldsymbol{\theta}_s')} \mathcal{C}_{\text{KL}}\left(\mathbf{r}_e, \hat{\mathbf{r}}_{e,\boldsymbol{\theta}_w}\right), \tag{7}$$

where $\mathcal{W}(\boldsymbol{\theta}_s')$ denotes the smoothing window around the $\boldsymbol{\theta}_s'$. This means that at each round of the optimization, we evaluate the cost function not only for a single point, but also for its neighboring points, which in turn increases the complexity. However, as the algorithm converges, smaller window sizes can be used to reduce complexity. The results in Fig. 6 verify the efficacy of this approach in the problem of source depth estimation. Note that although this method works well for low-dimensional problems like source depth estimation (and with no uncertainties in the parameters), it is not as effective for higher dimensional problems.

Alternatively, one can perform a grid search around the initial estimate $\boldsymbol{\theta}_{s,0}$, $\mathcal{W}(\boldsymbol{\theta}_s')$, and for each grid point and run the GB algorithm to descend to the neighboring local minimum on the cost surface. Finally, an estimate of the global minimum can be obtained by comparing the local minima obtained from the grid. We use this method in our source localization problem with uncertainties in the parameters, which is presented as "KL + grid" or "SE + grid", according to the cost used.
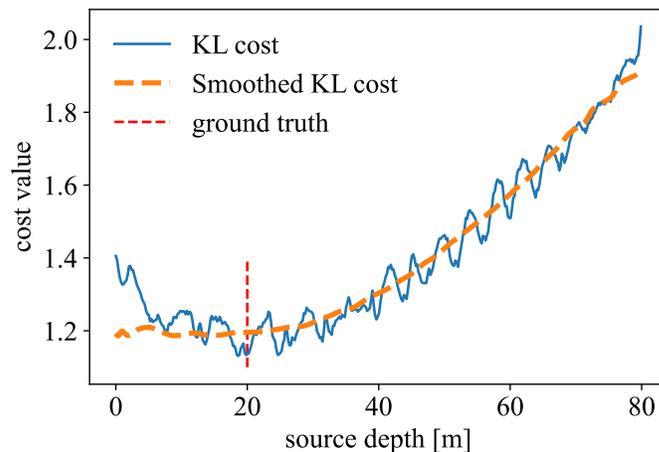
Figure 6: The KL cost surface, before and after smoothing, in the 3-ray environment in the presence of AWGN at the SNR=20 dB. For smoothing, we have used the window defined in (8) with $W = 40$ and $\delta = 0.2$.

C.   STAGE 3: FINE-TUNING USING TEMPORAL SIGNAL

We assume that the coarse estimate $\hat{\boldsymbol{\theta}}_s$, obtained via the envelope of the received waveform, is within the half-wavelength vicinity of the true estimate. We can then obtain a finer estimate of the source location by minimizing the SE cost on the temporal waveform (see Fig. 5) to exploit both the envelope and phase information and achieve a more accurate final estimate.

4.   SIMULATION RESULTS

In this section, to show the effectiveness of our algorithm through simulations, we compare the performance in terms of the root mean-squared-error (RMSE) of our GB localization (GBL) to that of the TOA-based method (stage 1) and the Cramér-Rao Lower Bound (CRLB). Specifically, we show that 1) given a good initialization, we can achieve the CRLB, and 2) the algorithm improves the performance of the simple TOA-based method. The CRLB derivations will be provided in an extended version of this paper. Note that for all experiments, we have simulated data and assumed an AWGN channel.

A.   THREE-RAY ENVIRONMENT

In this section, we generate simulations for the simple geometry shown in Fig. 3 with a depth of $D = 200 \ m$, $c = 1500 \ m/s$, $y_s = y_r = 0 \ m$, $x_s = 0 \ m$, $x_r = 600 \ m$, $z_s = 20 \ m$, $z_r = 120 \ m$, and a Gaussian pulse with $f_c = 500$ Hz and $B = 100$ Hz. We denote the step-size of the Adam optimizer by μ.

i.   Source Depth Estimation

In the first scenario, we assume the only unknown is z_s and an initial estimate of $z_{s,0} = 30 \ m$ is available. We begin by localizing using a smoothed KL cost and perform the optimization with the parameters provided in Table 1. In this experiment, the smoothing window is defined as

$$\mathcal{W}(z_s') = \{z_s' + \delta w \mid w \in [-W, -W + 1, ..., W], \delta > 0, W \in \mathbb{N}\}. \tag{8}$$

The results in Fig. 7 show that this method can achieve the CRLB for this setup when the SNR ≥ 15 dB.

Optimization Step	Parameters	Cost Function	Stage
$1 \leq i \leq 2$	$W = 50, \delta = 0.2, \mu = 4$	Smoothed KL	2
$3 \leq i \leq 4$	$\mu = 0.2$	KL	
$5 \leq i \leq 22$	$\mu = 0.1$	SE	3
$23 \leq i \leq 38$	$\mu = 0.01$	SE	
$39 \leq i \leq 100$	$\mu = 0.005/(i - 38)$	SE	

Table 1: Optimization parameters for the source depth estimation.
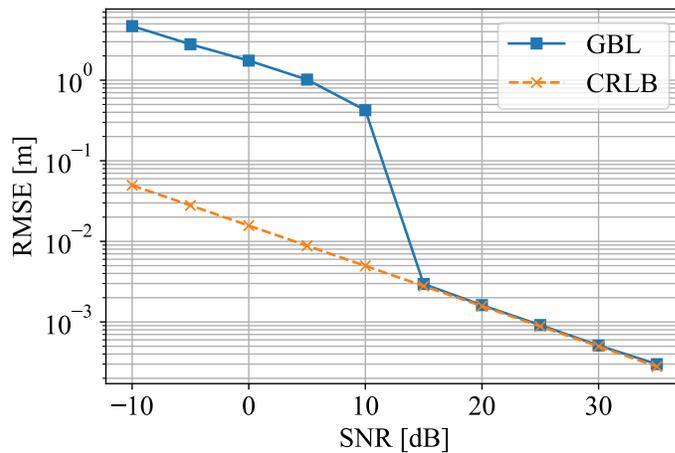
Figure 7: Depth estimation performance of the proposed approach in the presence of AWGN shows that it achieves the CRLB at high SNR.

ii.   Source Localization with Mismatch in The Sound Speed

In the second scenario, we assume that x_s and z_s are unknown, and our prior knowledge includes $c_0 = 1495 \ m/s$, $z_{min,0} = 5 \ m$, and $z_{max,0} = 35 \ m$. We use the TOA-based method to obtain an initial estimate of $\boldsymbol{\theta}_s = [x_s, z_s]$, then use a grid around the initial estimate defined similar to (8) as

$$\mathcal{W}(\boldsymbol{\theta}'_s) = \{\boldsymbol{\theta}'_s + \boldsymbol{\delta}w \mid w \in [-W, -W+1, ..., W], \boldsymbol{\delta} = [0, \delta], W \in \mathbb{N}\}, \tag{9}$$

and run the GB algorithm for those points to achieve a more accurate result, as detailed in Table 2.

Figure 8 shows that if the prior soundspeed c_0 is used without modification, GBL can improve on the TOA estimate. However, if we account for the uncertainty in c_0 and include this parameter along with the unknown location in the optimization procedure (with a corresponding step-size of μ_c in the optimizer), we can further improve the estimate.

iii.   Source Localization with Mismatch in The Water Depth

In this scenario, we assume that the sound speed is known accurately beforehand, but there is uncertainty in the water depth, with the prior belief of $D_0 = 198 \ m$, and the unknowns are x_s and z_s. Figure 9 shows that if the prior depth D_0 is used without modification, GBL can improve on the TOA estimate. However, by accounting for the uncertainty in D_0 and including this parameter along with the unknown location in the optimization procedure (with a corresponding step-size of μ_d in the optimizer), we can further improve the estimate. The implementation details are provided in Table 3.

Optimization Step	Parameters	Cost Function	Stage
$1 \leq i \leq 20$	$W = 7, \delta = 1.3, \mu = 0.04, \mu_c = 0.6$	KL + grid	
$21 \leq i \leq 50$	$W = 7, \delta = 1.3, \mu = 0.02, \mu_c = 0.12$	KL + grid	
$51 \leq i \leq 70$	$W = 7, \delta = 1.3, \mu = 0.01, \mu_c = 0.06$	SE + grid	2
$71 \leq i \leq 150$	$W = 7, \delta = 1.3, \mu = 0.005, \mu_c = 0.03$	SE + grid	
$151 \leq i \leq 350$	$\mu = 0.002, \mu_c = 0.12$	SE	3

Table 2: Optimization parameters for the localization with the mismatch in the sound speed.
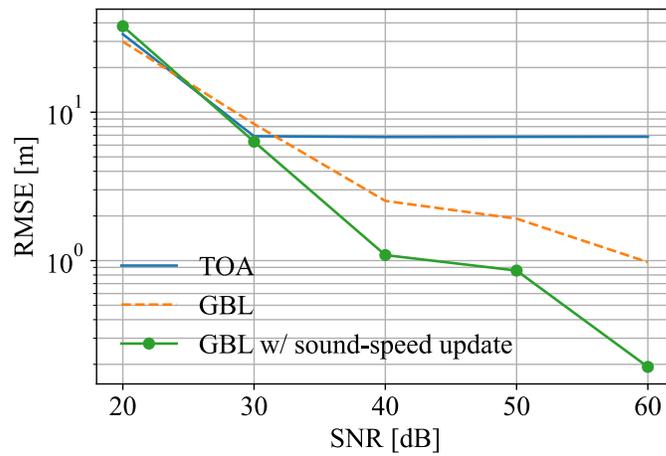
Figure 8: Localization performance in the presence of uncertainty in the sound speed; The true sound speed is $1500 \ m/s$*, while we have assumed a prior of* $1495 \ m/s$ *in the localization algorithms. The orange curve shows the performance of the GBL when not modifying the sound speed is* c_0 *during the optimization, while the green curve shows the performance of the GBL when we include* c *in the set of parameters to be optimized.*

Optimization Step	Parameters	Cost Function	Stage
$1 \leq i \leq 20$	$W = 7, \delta = 1.3, \mu = 0.2, \mu_d = 0.4$	KL + grid	
$21 \leq i \leq 100$	$W = 7, \delta = 1.3, \mu = 0.1, \mu_d = 0.4$	KL + grid	
$101 \leq i \leq 120$	$W = 7, \delta = 1.3, \mu = 0.1, \mu_d = 0.2$	SE + grid	2
$121 \leq i \leq 200$	$W = 7, \delta = 1.3, \mu = 0.05, \mu_d = 0.2$	SE + grid	
$201 \leq i \leq 400$	$\mu = 0.05, \mu_d = 0.08$	SE	3

Table 3: Optimization parameters for the localization with the mismatch in the water depth.

B.   SOARS TANK-BASED SIMULATION SCENARIO

In this experiment, we implement a simulation environment that approximates the localization experiment done in the SOARS tank. This scenario is a 3-dimensional localization problem; x_s, y_s, and z_s are unknowns. The transmit signal is a modulated Gaussian pulse with $f_c = 32$ kHz and $B = 12$ kHz and Table 4 provides the details of the geometry.

Figure 10 presents the performance of the GBL, when used with different number of reflections and initialized by $\boldsymbol{\theta}_{s,0} = [0.01, -0.03, 0.46]$, i.e., $0.0117 \ m$ away from the ground truth. The results show that for SNR ≥ 25 dB, the GBL achieves the CRLB. Note that if we use only the direct path and first-order
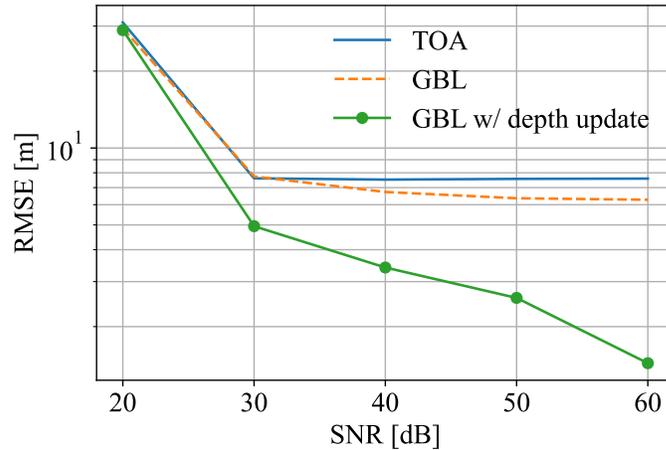
Figure 9: Localization performance in the presence of uncertainty in the water depth; The true depth is $200\ m$***, while we have assumed*** $198\ m$ ***in the localization algorithms. The orange curve shows the performance of the GBL when not modifying the depth*** D_0 ***during the optimization, while the green curve shows the performance of the GBL when we include*** c ***in the set of parameters to be optimized.***

$\boldsymbol{\theta}_s$	$[0.000, -0.0376, 0.4611]\ [m]$
$\boldsymbol{\theta}_r$	$[1.4239, -0.0224, 0.4440]\ [m]$
right wall	$y = 1.228\ [m]$
left wall	$y = -1.228\ [m]$
bottom	$z = 1.2630\ [m]$
surface	$z = 0\ [m]$
sound speed	$1513.12\ [m/s]$
wall reflection coefficient	0.93

Table 4: Details of the geometry of the simulation mimicking the SOARS tank geometry.

reflections, we cannot significantly improve this initial estimate.

5.   CONCLUSION

We have developed a single-sensor GB localization algorithm that leverages the automatic differentiation tools in PyTorch, along with a smoothing technique to make the cost function more amenable to optimization via GB methods. The proposed GB method is a generic framework for localization using differentiable models, that has the capacity to offer computationally faster localization by avoiding costly brute-force searches or GA/SA kind of approaches. Furthermore, the proposed method can be extended to similar problems such as wall or scatterer localization in a water tank. This technique is general enough to be used with any differentiable propagation model, which can also include machine-learning models trained from recorded data, for example.
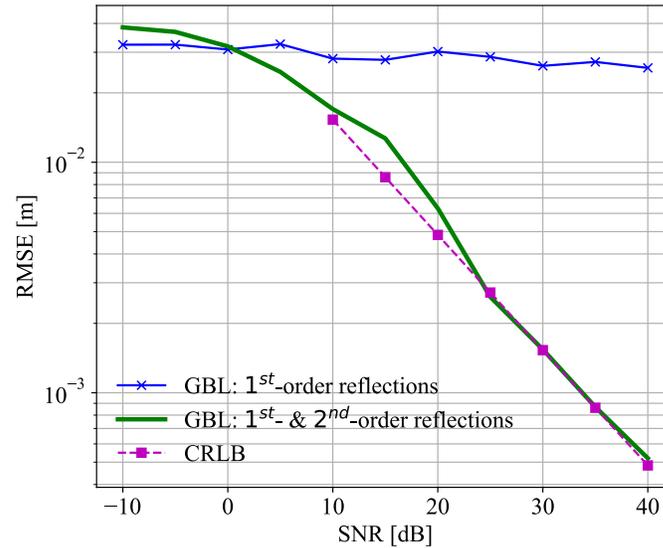
Figure 10: The RMSE of the proposed approach in the simulation environment modeled on SOARS tank shows that it achieves the CRLB at high SNR. Note that if we use both first-order and second-order reflections, we can greatly improve the initial estimate.

ACKNOWLEDGMENTS

REFERENCES

[1] A. B. Baggeroer, W. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," The Journal of the Acoustical Society of America **83**(2), 571–587 (1988).

[2] K. L. Gemba, W. S. Hodgkiss, and P. Gerstoft, "Adaptive and compressive matched field processing," The Journal of the Acoustical Society of America **141**(1), 92–103 (2017).

[3] P. Hursky, M. B. Porter, M. Siderius, and V. K. McDonald, "High-frequency (8–16 khz) model-based source localization," The Journal of the Acoustical Society of America **115**(6), 3021–3032 (2004).

[4] B. M. Worthmann, H. Song, and D. R. Dowling, "High frequency source localization in a shallow ocean sound channel using frequency difference matched field processing," The Journal of the Acoustical Society of America **138**(6), 3549–3562 (2015).

[5] M. B. Porter, "The bellhop manual and user's guide: Preliminary draft," Heat, Light, and Sound Research, Inc., La Jolla, CA, USA, Tech. Rep **260** (2011).

[6] M. B. Porter, "The kraken normal mode program," , Naval Research Lab Washington DC (1992).

[7] C. E. Lindsay and N. R. Chapman, "Matched field inversion for geoacoustic model parameters using adaptive simulated annealing," IEEE Journal of Oceanic Engineering **18**(3), 224–231 (1993).

[8] D. Whitley, "A genetic algorithm tutorial," Statistics and computing **4**(2), 65–85 (1994).

[9] S. M. Jesus, M. B. Porter, Y. Stéphan, X. Démoulin, O. C. Rodríguez, and E. M. F. Coelho, "Single hydrophone source localization," IEEE Journal of Oceanic Engineering **25**(3), 337–346 (2000).

[10] J. Bonnel, A. Thode, D. Wright, and R. Chapman, "Nonlinear time-warping made simple: A step-by-step tutorial on underwater acoustic modal separation with a single hydrophone," The journal of the acoustical society of America **147**(3), 1897–1926 (2020).

[11] K. W. Lo, "A matched-field processing approach to ranging surface vessels using a single hydrophone and measured replica fields," The Journal of the Acoustical Society of America **149**(3), 1466–1474 (2021).

[12] E. L. Ferguson, S. B. Williams, and C. T. Jin, "Convolutional neural network for single-sensor acoustic localization of a transiting broadband source in very shallow water," The Journal of the Acoustical Society of America **146**(6), 4687–4698 (2019).

[13] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational ocean acoustics* (Springer Science & Business Media, 2011).

[14] A. Weiss, T. Arikan, H. Vishnu, G. B. Deane, A. C. Singer, and G. W. Wornell, "A semi-blind method for localization of underwater acoustic sources," IEEE Transactions on Signal Processing **70**, 3090–3106 (2022).

[15] A. Weiss, A. Lancho, Y. Bu, and G. W. Wornell, "A bilateral bound on the mean-square error for estimation in model mismatch," arXiv preprint arXiv:2305.08207 (2023).

[16] T. M. Cover, *Elements of information theory* (John Wiley & Sons, 1999).

[17] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," IEEE transactions on acoustics, speech, and signal processing **24**(4), 320–327 (1976).

[18] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," IEEE Signal processing magazine **22**(4), 54–69 (2005).

12 March 2025 08:57:25