TODO NEWTITLE

Vitor Gelsleichter Probst Curtarelli, and Anderson Wedderhoff Spengler, *Fellow, IEEE*

***Abstract*—TODO NEWABSTRACT**

***Index Terms*—TODO KEYWORDS**

I. Introduction

Although the speed of sound c in water is considered constant at around $1500\mathrm{m/s}$, it can vary when considering environments of both small and large scales, respectively in small and large proportions [1–4]. Many properties can affect the speed of sound in water, such as pressure, temperature, salinity, and water currents [5–8]. As these properties vary over the body of water, so does the velocity, creating a sound speed field (SSF) in the body of water. Although the term sound speed profile (SSP) is also commonly used [3, 4, 9], it more fittingly describes the 1D estimation scenario.

Initial works in the topic of estimating the velocity field in water aimed at estimating these key properties, as well as the relationship between them and the velocity of water. Works such as [5–8, 10–12] operated under this framework, with some still being used today to model the behavior of water currents on a global scale. The pursuit of knowing the speed of sound at each point of a body of water – the SSF – permits estimating the properties of a wave that travels through this medium [**?**]. By Snell's law [**snells-law**] and Fermat's principle of least time [**fermats-least-time**], a change in speed of a wave also changes its direction. Therefore a wave that travels through a sufficiently large body of water will be also affected by it, resulting in curves, refractions, reflections, and formation/dispersion of beams. With this, we can estimate the path a unidirectional wave would course. By treating an omnidirectional wave as a cluster of unidirectional waves, we can calculate the travel-time between the source and any point within the medium, as well as its wavefront's profile. The knowledge of the path taken by an ensemble of rays is if great importance in the fields of ... and ... [**?**, **?**], and the estimation of the SSF within the environment is relevant for ..., ..., and ... [**?**, **?**, **?**]

While traditional techniques used the environment properties – pressure, temperature, and salinity – to estimate the SSF, most modern SSF estimation approaches avoid estimating these properties, and treat this SSF estimation as an inversion problem. That is, since the knowledge of the SSF allows calculating the path of a wave through the medium (this being denominated the direct problem), then the knowledge of these paths admits the estimation of the SSF in the medium. Different techniques tackle this inversion problem by using measurements, and employing the acquired information through different lenses. Such techniques encompass Travel-Time Tomography [13, 14], Matched Field Processing [15, 16] and Inversion [17], and Full Waveform Inversion [18, 19].

Among these, the travel-time tomography (TTT) is known for its simplicity and versatility [20–22]. It employs the difference in arrival time of the acoustic signals at the sensors, requiring a spatial distribution of sources and receivers for sufficient spatial resolution. In general, the SSF is discretized in layers and/or cells, where the sound speed is considered constant within each. The SSF's inversion iterates alternately between calculating the paths and lengths in each cell via solving the Eikonal equation [**fast-marching**, **fast-sweeping**, **ray-tracing**], and estimating the speed in each cell. Its simpleness, robustness, and computational efficiency are its main advantages, the later enabling real-time implementation. However, it depends on a precise positioning and syncing of the devices, and its discretization assumption results in limited resolution.

Many different approaches have been proposed for the TTT inversion, such as regularization [20, 23], blurring [22, 24], ray bending [25, 26], conjugate gradient minimization [27, 28], and others [25, 27]. Each of these aim to enhance the inverse problem by correcting different issues that may appear, like ill-posed problems, physics-defiant models, or practically incoherent solutions.

In this paper, we propose a new approach, employing the singular-value decomposition (SVD) and the Moore-Penrose pseudo-inverse (from now on called p-inverse) for the primary inverse problem, as well as a null-space exploiting step for achieving a minimum on a secondary metric. These three techniques can be employed jointly to achieve an inverse mapping function that minimizes a primary cost function, and a secondary cost function within the null-space of the primary. This procedure ensures an arbitrarily small primary cost function (within the limits of the forward map), and a minimum on the secondary metric that doesn't have an impact on the primary.

Through simulations emulating real scenarios, the practical results show that the proposed framework more precisely

Manuscript Info. Corresponding author: Vitor P. Curtarelli.
A. W. Spengler is with the Department of Electrical and Electronic Engineering at UFSC - Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil (email: anderson.spengler@ufsc.br).
V. G.P. Curtarelli is with the Department of Electrical and Electronic Engineering at UFSC - Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil (email: vitor.curtarelli@posgrad.ufsc.br).

models the observed SSF, as well as achieving smaller error values between observed and estimated travel times, and .... A computational complexity analysis shows that ....

This paper is organized as follows: in section II we present the inverse problem's mathematical model as well as its physical underlying interpretation, and we establish the standard literature technique to which ours will be compared. In section III we present the techniques to be used and develop the new proposed model, bringing a discussion on its theoretical advantages and disadvantages. In section IV we present the simulation conditions for comparing the presented techniques, as well as the results obtained through these simulations. section V concludes the paper.

Although the acronym SSF means sound speed field, mathematically it is more practical to model the problem via the slowness, the inverse of speed. Since both the sound speed field and sound slowness field would result in the same acronym, and both are simply inversely proportional, the SSF acronym will be used indistinguishably for both.

II. SSF Estimation and the TTT technique

As previously explained, the TTT uses the delay – travel-time – between the wave's generation at a source and its arrival at a receptor to estimate the path traveled, and therefore the wave's speed along the route taken.

Given a ray that follows a path $\vec{\mathbf{r}}$, and a sound speed field $c(\mathbf{p})$ as a function of space ($\mathbf{p} = [x, y, z]$); then the travel time along the curve $\vec{\mathbf{r}}$ is

$$t_{\vec{\mathbf{r}}} = \int_{\vec{\mathbf{r}}} \frac{1}{c(\mathbf{p})} \mathsf{d}\mathbf{p} \tag{1}$$

This is known as the forward map: given an SSF $c(\mathbf{p})$ and a path through this SSF, estimate the travel time. The inverse problem's objective is to obtain the SSF (and a path, if ray bending is considered), given a set of travel times between different sources and receivers – therefore different paths – within this environment.

Under the TTT technique, the approach most commonly found in the literature is to discretize the environment in cells, assuming each cell has a constant speed, and estimating this discrete SSF using the measurements. We assume that the space is divided into an $N_\mathsf{z} \times N_\mathsf{x}$ grid of cells (vertical $\times$ horizontal), with a total of $N = N_\mathsf{x} \cdot N_\mathsf{z}$ cells; and a total of M rays are produced, between any number of sources and receivers, and travel through this modeled environment.

Mathematically, this discretization and ray pathing can be written as

$$t_m \approx \sum_n r_{m;n} s_n \tag{2}$$

where $r_{m;n}$ is the length of the m-th ray within the n-th cell, s_n is the n-th cell's slowness, and t_m is the m-th ray's travel time. In vector form, we rewrite eq. (2) as

$$t_m(\mathbf{s}) = \mathbf{r}_m^\mathsf{T}(\mathbf{s})\mathbf{s} \tag{3}$$

in which $\mathbf{r}_m(\mathbf{s}) \in \mathbb{R}^{N\times 1}$ corresponds to the lengths in each cell of the m-th ray (this being usually a sparse vector), and $\mathbf{s} \in \mathbb{R}^{N\times 1}$ represents the modeled slowness in each cell. We explicit the dependency of $\mathbf{r}_m(\mathbf{s})$ on $\mathbf{s}$ since the ray bending effects on the m-th ray's path will depend on the slowness field, following Fermat's principle and Snell's law.

We now let $\mathbf{t}(\mathbf{s}) \in \mathbb{R}^{M\times 1}$ as a vector of travel-times for the current model, for all M considered rays; and $\mathbf{R}(\mathbf{s}) \in \mathbb{R}^{M\times N}$ is a vertical stacking of all $\mathbf{r}_m^\mathsf{T}(\mathbf{s})$. With this, we achieve that

$$\mathbf{t}(\mathbf{s}) = \mathbf{R}(\mathbf{s})\mathbf{s} \tag{4}$$

A. SSF estimation

Given a vector of observed traveled times $\mathbf{t}_\mathrm{o}$, the SSF estimation is given by minimizing the error between the modeled (eq. (4)) and the observed travel times, this being translated to

$${}^\star\mathbf{s} = \underset{\mathbf{s}}{\operatorname{argmin}}\, E(\mathbf{s}) = \|\mathbf{R}(\mathbf{s})\mathbf{s} - \mathbf{t}_\mathrm{o}\|^2 \tag{5}$$

Note that this problem is non-linear: the solution to the ray paths $\mathbf{R}(\mathbf{s})$ depends on the slowness field $\mathbf{s}$, and in turn the optimal $\mathbf{s}$ depends on the achieved ray path solution $\mathbf{R}(\mathbf{s})$. Given this, the most common approach is to iteratively solve this problem. That is, given the i-th iteration solution $\mathbf{s}_i$ and its respective $\mathbf{R}(\mathbf{s}_i)$ (which will now be called $\mathbf{R}_i$ for simplicity), we have that

$$\mathbf{s}_{i+1} = \underset{\mathbf{s}}{\operatorname{argmin}} \|\mathbf{R}_i\mathbf{s} - \mathbf{t}_\mathrm{o}\|^2 \tag{6}$$

and the $(i+1)$-th iteration of $\mathbf{R}$ is obtained by solving the Eikonal equation, given the slowness field $\mathbf{s}_{i+1}$.

One extra component that is usually added is a regularization step, turning eq. (6) into

$$\mathbf{s}_{i+1} = \underset{\mathbf{s}}{\operatorname{argmin}}\, E'(\mathbf{s}) = \|\mathbf{R}_i\mathbf{s} - \mathbf{t}_\mathrm{o}\|^2 + \alpha^2 \|\mathbf{D}\mathbf{s}\|^2 \tag{7}$$

in which $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a regularizing matrix, and α a regularization parameter. The proper choice of $\mathbf{D}$ ensures a well-posed problem, improving some quality or metric in the obtained solution $\mathbf{s}_{i+1}$ at the cost of some error in the modeled travel times.

The solution to the minimization problem is given by an inverse map function $\mathbf{G}_i \in \mathbb{R}^{N\times M}$, resulting in

$$\mathbf{s}_{i+1} = \mathbf{G}_i\mathbf{t}_\mathrm{o} \tag{8}$$

In order to achieve $\mathbf{s}_{i+1}$, the gradient of $E'(\mathbf{s})$ is taken,

$$\nabla E'(\mathbf{s}) = 2\mathbf{R}_i^\mathsf{T}(\mathbf{R}_i\mathbf{s} - \mathbf{t}_\mathrm{o}) + 2\alpha^2\mathbf{D}^\mathsf{T}\mathbf{D}\mathbf{s} \tag{9}$$

which, when set to $\mathbf{0}$, leads to

$$\mathbf{G}_i = \left[\mathbf{R}_i^\mathsf{T}\mathbf{R}_i + \alpha^2\mathbf{D}^\mathsf{T}\mathbf{D}\right]^{-1}\mathbf{R}_i^\mathsf{T} \tag{10}$$

This requires the inversibility of $\mathbf{R}_i^\mathsf{T}\mathbf{R}_i + \alpha^2\mathbf{D}^\mathsf{T}\mathbf{D}$. Given the construction of the problem, either $\mathbf{R}_i$ has to be full-rank with $M \geq N$, or $\mathbf{D}$ has to be full-rank, for this inverse to be possible. When these conditions aren't met, non-inverting schemes have to be used, such as the popularly employed conjugate gradient method [24, 26–28].

As explained, the regularizing matrix $\mathbf{D}$ (and α) have to be appropriately chosen to improve some desired metric. For example, choosing $\mathbf{D} = \mathbf{I}_N$ (the $N \times N$ identity) results in the minimization weighting in the mean-squared value of the slowness vector [22]; this is equivalent to considering the presence of white noise in the measurements, and trading model accurateness for noise rejection [**beamforming-maxwng**]. Alternatively, $\mathbf{D}$ can be chosen to (discretely) approximate the Laplacian of the underlying continous slowness field [24, 27], resulting in a minimization that penalizes the Laplacian of the achieved slowness field; that is, minimizes the SSF's curvature, searching for a smoother field.

III. SVD-Based Travel-Time Tomography

As stated, the kernel of the inverse problem is to iterate $\mathbf{s}_i$ and $\mathbf{R}_i$, searching for an optimal slowness field that most accurately describes the observed field, using the travel-time information. While the literature technique relies on the regularization being added to the cost function $E(\mathbf{s})$ to ensure a desired behavior (noise rejection or smoothness) and invertibility, other approaches are possible and available for study.

Here, we will present the SVD technique, its applicability for a pseudo-inverse (p-inverse), and how this enables the regularization step to be applied in a separate step to that of the primary inverse.

A. Singular value decomposition and the p-inverse

Given the matrix $\mathbf{R}_i$, we can decompose it as

$$\mathbf{R}_i = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T} \tag{11}$$

in which $\mathbf{U} \in \mathbb{R}^{M\times M}$ and $\mathbf{V} \in \mathbb{R}^{N\times N}$ are the left- and right-singular matrices of $\mathbf{R}_i$, forming left and right orthonormal bases; and $\mathbf{\Sigma} \in \mathbb{R}^{M\times N}$ is a diagonal matrix, with diagonal entries λ_j and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_R > 0$ being the singular values of $\mathbf{R}_i$, in decreasing order; where $R = \operatorname{rank}(\mathbf{R}_i)$. If $N > M$, the last $N - M$ columns of $\mathbf{\Sigma}$ are 0, and if $M > N$, the last $M - N$ rows of $\mathbf{\Sigma}$ are 0. By definition, all $\mathbf{U}$, $\mathbf{\Sigma}$ and $\mathbf{V}$ are dependent on the iterator i, but this relation will be omitted for clarity.

Given R, if $R < N$ then $\operatorname{rank}\left(\mathbf{R}_i^\mathsf{T}\mathbf{R}_I\right) = R < N$, and therefore this matrix isn't invertible. We denote the Moore-Penrose p-inverse or $\mathbf{R}_i$ as $\mathbf{R}_i^\dagger$, it being given by

$$\begin{aligned}\mathbf{R}_i^\dagger &= \left[\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}\right]^\dagger \\ &= \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^\mathsf{T}\end{aligned} \tag{12}$$

where, for the diagonal matrix $\mathbf{\Sigma}$, its p-inverse $\mathbf{\Sigma}^\dagger \in \mathbb{R}^{N\times M}$ is its transpose, with all non-zero diagonal entries being reciprocated. With this p-inverse, the solution to the problem in eq. (4) (without regularization) is given by

$$\begin{aligned}\mathbf{G}_i &= \left[\mathbf{R}_i^\mathsf{T}\mathbf{R}_i\right]^\dagger\mathbf{R}_i^\mathsf{T} \\ &= \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^\mathsf{T} \\ &= \mathbf{R}^\dagger\end{aligned} \tag{13}$$

where, via the properties of the p-inverse, $\left[\mathbf{R}_i^\mathsf{T}\mathbf{R}_i\right]^\dagger\mathbf{R}_i^\mathsf{T} = \mathbf{R}_i^\dagger$.

It is easy to prove that $\nabla E(\mathbf{s}_{i+1}) = \mathbf{0}$ when using the p-inverse (Theorem 1, appendix A), showing that the obtained slowness field fulfills the desired condition on the cost function. We let $\mathbf{U}_2$ as the last $M - R$ columns of $\mathbf{u}$, connected to the left null-space of $\mathbf{R}_i$. With this, the cost function with this solution can be given by

$$E(\mathbf{s}_{i+1}) = \mathbf{t}_\mathrm{o}^\mathsf{T}\mathbf{U}_2\mathbf{U}_2^\mathsf{T}\mathbf{t}_\mathrm{o} \tag{14}$$

B. SVD truncation

In general, the smaller a singular value λ_n is, the lessened its impact will be on the reconstruction of $\mathbf{R}_i$, and thus in the minimization problem. However, through the p-inverse, its impact on the estimation of $\mathbf{s}_{i+1}$ will be increased, over-representing the corresponding bases.

This effect can be attenuated by setting an arbitrary number of the smallest diagonal entries of $\mathbf{\Sigma}$ to 0. This procedure reduces the effective condition number (largest-to- smallest-non-zero singular value ratio), improving stability in the inversion problem; and increases its practical null-space, which will be followingly exploited. In contrast, using this truncation introduces some minimization error, with the minimum found not having truly zero gradient.

We denote $\tilde{\mathbf{\Sigma}}$ as a truncated $\mathbf{\Sigma}$, with only its first $\tilde{R}$ diagonal entries being non-zero. Trivially, $R \geq \tilde{R}$ for this to be meaningful. Given this approximation, we define $\tilde{\mathbf{R}}_i = \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}^\mathsf{T}$, and the inverse map function can be approximated as

$$\begin{aligned}\mathbf{G}_i &= \tilde{\mathbf{R}}_i^\dagger \\ &= \mathbf{V}\tilde{\mathbf{\Sigma}}^\dagger\mathbf{V}^\mathsf{T}\end{aligned} \tag{15}$$

For this solution, given the truncation, $\nabla E(\mathbf{s}_{i+1}) \neq \mathbf{0}$. Instead, we can show that (Theorem 2, appendix A)

$$\nabla E(\mathbf{s}_{i+1}) = 2\left(\sum_{j=\tilde{R}+1}^{R}\lambda_j\mathbf{v}_j\mathbf{u}_j^\mathsf{T}\right)\mathbf{t}_\mathrm{o} \tag{16}$$

where the summation happens over the truncated singular values. Therefore, the smaller the zeroed singular values are, the smaller the gradient will be, and the closer the obtained solution would be to the space of optimal solutions ($\nabla E = \mathbf{0}$). Furthermore, the cost function with the truncated SVD can be computed as

$$E(\mathbf{s}_{i+1}) = \mathbf{t}_\mathrm{o}^\mathsf{T}\mathbf{U}_{\tilde{2}}\mathbf{U}_{\tilde{2}}^\mathsf{T}\mathbf{t}_\mathrm{o} \tag{17}$$

in which $\mathbf{U}_{\tilde{2}}$ is the left null-space of $\tilde{\mathbf{R}}_i$. Furthermore, the change in cost function compared to the non-truncated SVD can be written as

$$\Delta E(\mathbf{s}_{i+1}) = \mathbf{t}_\mathrm{o}^\mathsf{T}(\mathbf{U}_{\tilde{2}} - \mathbf{U}_2)(\mathbf{U}_{\tilde{2}} - \mathbf{U}_2)^\mathsf{T}\mathbf{t}_\mathrm{o} \tag{18}$$

Noticeably, this gain in cost function ΔE doesn't depend on the truncated singular values directly. Therefore, the choice of λ_n's to truncate isn't explicitly related to the increase in cost function.

C. Null-space exploitation

We define $\mathbf{V}_2$ as the last $N-R$ columns of $\mathbf{V}$, corresponding to the right null-space of $\mathbf{R}_i$; that is, the columns of $\mathbf{V}_2$ form an orthonormal basis for the null-space of $\mathbf{R}_i$.

Given any coordinate vector $\boldsymbol{\mu} \in \mathbb{R}^{(N-R)\times 1}$, any vector $\bar{\mathbf{s}} = \mathbf{V}_2\boldsymbol{\mu}$ will lie on the null-space of $\mathbf{R}_i$, and therefore

$$\begin{aligned}\mathbf{R}_i\bar{\mathbf{s}} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\mathsf{T}\mathbf{V}_2\boldsymbol{\mu} \\ &= 0\end{aligned} \tag{19}$$

from the orthonormality of $\mathbf{V}$ and the null diagonal-entries of $\boldsymbol{\Sigma}$. This means that, for any $\mathbf{s}_{i+1} = \mathbf{G}_i\mathbf{t}_\mathrm{o} + \mathbf{V}_2\boldsymbol{\mu}$, we have

$$E(\mathbf{G}_i\mathbf{t}_\mathrm{o} + \mathbf{V}_2\boldsymbol{\mu}) = E(\mathbf{G}_i\mathbf{t}_\mathrm{o}) \tag{20a}$$

$$\nabla E(\mathbf{G}_i\mathbf{t}_\mathrm{o} + \mathbf{V}_2\boldsymbol{\mu}) = \nabla E(\mathbf{G}_i\mathbf{t}_\mathrm{o}) \tag{20b}$$

With this, both the cost function and its gradient are unaffected by $\boldsymbol{\mu}$, for any $\boldsymbol{\mu}$. Note that this independence of cost function and gradient relative to $\boldsymbol{\mu}$ is only valid when using the non-truncated SVD. That is, if instead $\mathbf{V}_2$ are the last $N-\tilde{R}$ columns of $\mathbf{V}$, then the equations in eq. (20b) aren't valid, with there being some residual cost function and gradient related to the truncation. However, similarly to section III-B, the cost and gradient functions will depend on the truncated singular values, or their respective orthonormal vectors.

D. Secondary minimization

This information on the independence of $\boldsymbol{\mu}$ can be leveraged to find an optimal solution within a secondary metric, while keeping the primary cost function and its gradient intact, through eq. (20b). Given a regularization matrix $\mathbf{D}$ as before, we define a new cost function $F(\mathbf{G}_i\mathbf{t}_\mathrm{o}, \boldsymbol{\mu})$ to be minimized, with $\boldsymbol{\mu}_{i+1}$ being given by its minimization,

$$\boldsymbol{\mu}_{i+1} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}}\, F(\mathbf{G}_i\mathbf{t}_\mathrm{o}, \boldsymbol{\mu}) = \|\mathbf{D}(\mathbf{G}_i\mathbf{t}_\mathrm{o} + \mathbf{V}_2\boldsymbol{\mu})\|^2 \tag{21}$$

with $\boldsymbol{\mu}$ being iteratively calculated, similarly to $\mathbf{R}_i$ and $\mathbf{s}_{i+1}$. Given the column-orthogonality of $\mathbf{V}_2$, then $\operatorname{rank}(\mathbf{V}_2) = N - R$ (meaning it is full-rank), and therefore this problem can be directly solved as

$$\boldsymbol{\mu}_{i+1} = -\left[\mathbf{V}_2^\mathsf{T}\mathbf{D}^\mathsf{T}\mathbf{D}\mathbf{V}_2\right]^{-1}\mathbf{V}_2^\mathsf{T}\mathbf{D}^\mathsf{T}\mathbf{D}\mathbf{G}_i\mathbf{t}_\mathrm{o} \tag{22}$$

or, through the p-inverse's properties,

$$\boldsymbol{\mu}_{i+1} = -\left[\mathbf{D}\mathbf{V}_2\right]^\dagger\mathbf{D}\mathbf{G}_i\mathbf{t}_\mathrm{o} \tag{23}$$

Therefore, the complete solution when considering this secondary null-space exploiting minimization is

$$\mathbf{s}_{i+1} = \left(\mathbf{I}_n - \mathbf{V}_2\left[\mathbf{D}\mathbf{V}_2\right]^\dagger\mathbf{D}\right)\mathbf{G}_i\mathbf{t}_\mathrm{o} \tag{24}$$

where $\mathbf{G}_i$ can be either with eq. (13) or eq. (15).

With this solution, the primary cost function $E(\mathbf{s})$ is minimized globally (at most within the truncated SVD), and the secondary cost function $F(\mathbf{G}_i\mathbf{t}_\mathrm{o}, \boldsymbol{\mu})$ is minimized within the primary cost function's solution space. The main insight for this expansion by using $\boldsymbol{\mu}_{i+1}$ is, given that $\mathbf{G}_i$ isn't full-rank, there is an entire space of solutions that result in the same primary metric, and a secondary metric can be optimized within this space. Note that this procedure requires $\operatorname{rank}(\mathbf{G}_i) < N$ (either naturally or through truncation), otherwise there is no null-space to be exploited, and the second minimization can't be performed.

An important remark is that $\mathbf{V}_2\left[\mathbf{D}\mathbf{V}_2\right]^\dagger\mathbf{D}$ can't be further simplified. By the p-inverse's properties, this would require $\mathbf{V}_2$ to have linearly independent rows, which it can't given it is a tall matrix (either $N > N - R$, or $N > N - \tilde{R}$ for the truncated scenario). Another relevant attribute is that, if $\mathbf{D}$ is invertible, the solution for taking the gradient of $F(\mathbf{G}_i\mathbf{t}_\mathrm{o}, \boldsymbol{\mu})$ w.r.t. $\boldsymbol{\mu}$, $\mathbf{G}_i\mathbf{t}_\mathrm{o}$, or $\mathbf{G}_i\mathbf{t}_\mathrm{o} + \mathbf{V}_2\boldsymbol{\mu}$, all result in eq. (23).

E. Spatial under- over-modeling

Previously on the proposed method, no mention on the dimensions of $\mathbf{R}_i$ was given, or its impact on the minimization scheme. Given that $\mathbf{R}_i \in \mathbb{R}^{M\times N}$, with $\mathbf{G}_i$ having the transposed dimentionality, the relationship of M and N (as well as with the rank R) can be impactful on the inverse map's behavior.

As commented in section II-A, $M \geq N = R$ is sufficient to obtain the inverse map through a direct matrix inversion. For the proposed method, no such condition was imposed, but analyzing these conditions can bring some insight to the process. In particular, $M < N$ (labeled over-modeled scenario, where there are more variables than equations/information on the system) ensures that $R \leq M$ and therefore $N - R > 0$, guaranteeing the existance of a null-space for $\mathbf{R}_i$, and securing the possibility of a secondary minimization on $\boldsymbol{\mu}$.

Assuming $R = M < N$, then $\mathbf{U}_2 = []_{0\times M}$ is an empty vector,

IV. Simulation and Results

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

V. Conclusion

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Appendix

A. Proofs and Theorems

We let $E(\mathbf{s}) = \|\mathbf{J}_i\mathbf{s} - \mathbf{t}_\text{o}\|^2$, $\mathbf{J}_i = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}$ be the SVD of $\mathbf{J}_i$, and $[\cdot]^\dagger$ denote the Moore-Penrose pseudo-inverse.

Theorem 1. *The p-inverse solution has zero gradient.*

Given a solution $\mathbf{s} = \mathbf{J}_i^\dagger\mathbf{t}_\text{o}$, then the gradient of $E(\mathbf{s})$ for this solution will be

$$\begin{aligned}\nabla E(\mathbf{s}_{i+1}) &= 2\Big(\mathbf{J}_i^\mathsf{T}\mathbf{J}_i\mathbf{J}_i^\dagger\mathbf{t}_\text{o} - \mathbf{J}_i^\mathsf{T}\mathbf{t}_\text{o}\Big)\\ &= 2\Big(\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}\mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o} - \mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o}\Big)\\ &= 2\Big(\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{\Sigma}\mathbf{\Sigma}^\dagger\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o} - \mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o}\Big)\\ &= 2\Big(\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\left[\mathbf{I}_{M;R} - \mathbf{I}_M\right]\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o}\Big)\end{aligned} \tag{25}$$

where $\mathbf{I}_{M;R} \in \mathbb{R}^{M\times M}$ is a diagonal matrix with the first R entries being 1, and the remaining $M-R$ ones being 0; $\mathbf{I}_M \in \mathbb{R}^{M\times M}$ is the identity matrix; and $R = \operatorname{rank}(\mathbf{J}_i)$.

From this, $\mathbf{I}_{M;R} - \mathbf{I}_M$ has its first R diagonal entries 0, and the last ones -1. Since all rows of $\mathbf{\Sigma}^\mathsf{T}$ are in the null space of $\mathbf{I}_{M;R} - \mathbf{I}_M$ (only the first R entries of any row of $\mathbf{\Sigma}^\mathsf{T}$ are non-zero, exactly where $\mathbf{I}_{M;R} - \mathbf{I}_M$ is zero), then $\mathbf{\Sigma}^\mathsf{T}\left[\mathbf{I}_{M;R} - \mathbf{I}_M\right] = \mathbf{0}$. Therefore,

$$\begin{aligned}\nabla E(\mathbf{J}_i^\dagger\mathbf{t}_\text{o}) &= 2\Big(\mathbf{V}\,\mathbf{0}\,\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o}\Big)\\ &= \mathbf{0}\end{aligned} \tag{26}$$

■

Theorem 2. *The SVD truncated solution doesn't have zero gradient.*

Given $\tilde{\mathbf{\Sigma}}$ be a truncated $\mathbf{\Sigma}$ with only $\tilde{R}$ non-zero diagonal entries, then a solution to the minimization of $E(\mathbf{s})$ of the form $\mathbf{s} = \mathbf{V}\tilde{\mathbf{\Sigma}}^\dagger\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o}$ will have a gradient

$$\begin{aligned}\nabla E(\mathbf{s}_{i+1}) &= 2\Big(\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}\mathbf{V}\tilde{\mathbf{\Sigma}}^\dagger\mathbf{U}^\mathsf{T} - \mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}\Big)\mathbf{t}_\text{o}\\ &= 2\Big(\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{\Sigma}\tilde{\mathbf{\Sigma}}^\dagger\mathbf{U}^\mathsf{T} - \mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}\Big)\mathbf{t}_\text{o}\\ &= 2\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\Big(\mathbf{I}_{M;\tilde{R}} - \mathbf{I}_M\Big)\mathbf{U}^\mathsf{T}\mathbf{t}_\text{o}\\ &= 2\left(\sum_{i=\tilde{R}+1}^{R}\sigma_i\mathbf{v}_i\mathbf{u}_i^\mathsf{T}\right)\mathbf{t}_\text{o}\end{aligned} \tag{27}$$

which isn't identically $\mathbf{0}$.

■

References

[1] S. W. Kieffer, "Sound speed in liquid-gas mixtures: Water-air and water-steam," *Journal of Geophysical Research*, vol. 82, no. 20, Jul. 10, 1977.

[2] J. M. Trusler and E. W. Lemmon, "Determination of the thermodynamic properties of water from the speed of sound," *The Journal of Chemical Thermodynamics*, vol. 109, Jun. 2017.

[3] Z. Li, L. He, R. Zhang, F. Li, Y. Yu, and P. Lin, "Sound speed profile inversion using a horizontal line array in shallow water," *Science China Physics, Mechanics & Astronomy*, vol. 58, no. 1, Jan. 2015.

[4] S. Pinson and L. Guillon, "Sound speed profile characterization by the image source method," *The Journal of the Acoustical Society of America*, vol. 128, no. 4, Oct. 1, 2010.

[5] W. D. Wilson, "Speed of Sound in Distilled Water as a Function of Temperature and Pressure," *The Journal of the Acoustical Society of America*, vol. 31, no. 8, Aug. 1, 1959.

[6] H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Applications of Modern Acoustics). Academic Press, 1998.

[7] T. J. McDougall and P. M. Barker, *Getting Started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox*. Trevor J McDougall, 2011.

[8] V. A. Del Grosso, "New equation for the speed of sound in natural waters (with comparisons to other equations)," *The Journal of the Acoustical Society of America*, vol. 56, no. 4, Oct. 1, 1974.

[9] Q. Li, J. Shi, Z. Li, Y. Luo, F. Yang, and K. Zhang, "Acoustic sound speed profile inversion based on orthogonal matching pursuit," *Acta Oceanologica Sinica*, vol. 38, no. 11, Nov. 2019.

[10] C.-T. Chen and F. J. Millero, "Speed of sound in seawater at high pressures," *The Journal of the Acoustical Society of America*, vol. 62, no. 5, Nov. 1, 1977.

[11] G. S. K. Wong and S.-m. Zhu, "Speed of sound in seawater as a function of salinity, temperature, and pressure," *The Journal of the Acoustical Society of America*, vol. 97, no. 3, Mar. 1, 1995.

[12] F. Roquet, G. Madec, T. J. McDougall, and P. M. Barker, "Accurate polynomial expressions for the density and specific volume of seawater using the TEOS-10 standard," *Ocean Modelling*, vol. 90, Jun. 2015.

[13] P. Bormann, E. R. Engdahl, and R. Kind, "Seismic Wave Propagation and Earth models," *New Manual of Seismological Observatory Practice 2 (NMSOP2)*, P. Bormann, Ed., 2012.

[14] P. Stefanov, G. Uhlmann, A. Vasy, and H. Zhou, "Travel Time Tomography," *Acta Mathematica Sinica, English Series*, vol. 35, no. 6, Jun. 2019.

[15] K. L. Gemba, W. S. Hodgkiss, and P. Gerstoft, "Adaptive and compressive matched field processing," *The Journal of the Acoustical Society of America*, vol. 141, no. 1, Jan. 1, 2017.

[16] R. K. Brienzo and W. S. Hodgkiss, "Broadband matched-field processing," *The Journal of the Acoustical Society of America*, vol. 94, no. 5, Nov. 1, 1993.

[17] S. E. Dosso and M. J. Wilmut, "Bayesian multiple-source localization in an uncertain ocean environment," *The Journal of the Acoustical Society of America*, vol. 129, no. 6, Jun. 1, 2011.

[18] J. Virieux and S. Operto, "An overview of full-waveform inversion in exploration geophysics," *GEOPHYSICS*, vol. 74, no. 6, Nov. 2009.

[19] A. Fichtner et al., "Multiscale full waveform inversion," *Geophysical Journal International*, vol. 194, no. 1, Jul. 1, 2013.

[20] K. Aki, A. Christoffersson, and E. S. Husebye, "Determination of the three-dimensional seismic structure of the lithosphere," *Journal of Geophysical Research*, vol. 82, no. 2, Jan. 10, 1977.

[21] K. Dines and R. Lytle, "Computerized geophysical tomography," *Proceedings of the IEEE*, vol. 67, no. 7, 1979.

[22] W. S. Phillips and M. C. Fehler, "Traveltime tomography: A comparison of popular methods," *GEOPHYSICS*, vol. 56, no. 10, Oct. 1991.

[23] K. Aki and W. H. K. Lee, "Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model," *Journal of Geophysical Research*, vol. 81, no. 23, Aug. 10, 1976.

[24] R. Ali, S. Hsieh, and J. Dahl, "Open-source Gauss-Newton-based methods for refraction-corrected ultrasound computed tomography," in *Medical Imaging 2019: Ultrasonic Imaging and Tomography*, N. V. Ruiter and B. C. Byram, Eds., SPIE, Mar. 15, 2019.

[25] G. Jeong, F. Li, T. M. Mitcham, U. Villa, N. Duric, and M. A. Anastasio, "Investigating the Use of Traveltime and Reflection Tomography for Deep Learning-Based Sound-Speed Estimation in Ultrasound Computed Tomography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 11, Nov. 2024.

[26] A. Hormati, I. Jovanović, O. Roy, and M. Vetterli, "Robust ultrasound travel-time tomography using the bent ray model," J. D'hooge and S. A. McAleavey, Eds., Mar. 4, 2010.

[27] J. Zhang, U. S. Ten Brink, and M. N. Toksöz, "Nonlinear refraction and reflection travel time tomography," *Journal of Geophysical Research: Solid Earth*, vol. 103, no. B12, Dec. 10, 1998.

[28] H.-g. Tang, R.-j. Xie, Y. Wu, C.-s. Zhou, J.-l. Yuan, and W. Qin, "Travel time tomography by ray tracing using the fast sweeping method," *Applied Geophysics*, vol. 21, no. 4, Dec. 2024.