Sparse Blind Deconvolution with Nonconvex Optimization for Ultrasonic NDT Application

Xuyang Gao¹, Yibing Shi¹, Kai Du¹, Qi Zhu² and Wei Zhang¹,⁎

¹ School of Automation Engineering, University of Electronic Science and Technology of China, No.2006, Xiyuan Avenue, West Hi-tech Zone, Chengdu 611731, China; xuyanggao@std.uestc.edu.cn (X.G.); ybshi@uestc.edu.cn (Y.S.); 202022060402@std.uestc.edu.cn (K.D.)
² College of Mechatronic Engineering, Southwest Petroleum University, No.8, Xindu Road, Xindu District, Chengdu 610500, China; 201231010016@swpu.edu.cn
⁎ Correspondence: weizhang@uestc.edu.cn; Tel.: +86-131-9486-9424

Received: 9 November 2020; Accepted: 3 December 2020; Published: 4 December 2020

Abstract: In the field of ultrasonic nondestructive testing (NDT), robust and accurate detection of defects is a challenging task because of the attenuation and noising of the ultrasonic wave from the structure. For determining the reflection characteristics representing the position and amplitude of ultrasonic detection signals, sparse blind deconvolution methods have been implemented to separate overlapping echoes when the ultrasonic transducer impulse response is unknown. This letter introduces the $\ell_1/\ell_2$ ratio regularization function to model the deconvolution as a nonconvex optimization problem. The initialization influences the accuracy of estimation and, for this purpose, the alternating direction method of multipliers (ADMM) combined with blind gain calibration is used to find the initial approximation to the real solution, given multiple observations in a joint sparsity case. The proximal alternating linearized minimization (PALM) algorithm is embedded in the iterate solution, in which the majorize-minimize (MM) approach accelerates convergence. Compared with conventional blind deconvolution algorithms, the proposed methods demonstrate the robustness and capability of separating overlapping echoes in the context of synthetic experiments.

Keywords: ultrasonic detection; sparse blind deconvolution; nonconvex optimization; blind gain calibration

1. Introduction

Ultrasonic nondestructive testing (NDT) is an established detection technology, which can accurately characterize multilayer or composite materials [1]. The primary ultrasonic NDT technology mostly considered in practical systems is the pulse-echo method, which evaluates the surface and defects based on the estimation of the time of arrival (TOA), time of flight (TOF), time of flight diffraction (TOFD) and the time difference of arrival (TDOA) of received signals [2]. However, extracting the implicit information is difficult with the distortion and attenuation of the captured signals under noisy environments. Hence, the inspection aims to correctly estimate the parameter values, or separate superimposed waveforms, representing the physical properties.

The conventional method to improve the detection resolution is to increase the transducer frequency, leading to a decrease in acoustic penetration [3]. The other accepted practice estimates the waveform parameters by the Gaussian echo model (or other modified models) and prior information from the transducer [4]. These approaches have some limitations, such as reliance on hardware design and finite propagation distortion. Generally, the signal sample of ultrasonic pulse-echo testing is regarded as the convolution result between the reflection function over the propagation path and the system pulse response. Therefore, deconvolution signal processing methods can be implemented to
improve temporal resolution, which significantly affects revealing any small defect object masked by superimposed echoes. The existing mainstream methods are divided into the parametric method and optimization theory method.

Parametric methods depend on prior knowledge of waveform and design specific structures to obtain the estimated parameters. With the assumption that the reflection sequence was sparse, Chang [5] applied the Gauss-Newton (GN) algorithm to parameter estimation of the basic waveform and employed the split variable augmented Lagrangian shrinkage algorithm (SALSA) in the reflection sequence solution. Jin [6] determined the reference echo signal via maximum likelihood estimation (MLE) and achieved sparse deconvolution results by the orthogonal matching pursuit (OMP) algorithm. Bossmann [7] deconvolved the superimposed echo signal by removing the additive noise with a particular matching pursuit (MP) algorithm. The variational mode decomposition (VMD) algorithm was applied to deconvolute the signal by reducing the noise level, which increased the resolution of several defects in different positions [8]. Another class of methods is the minimum entropy deconvolution (MED) technique without acquiring prior assumptions, and the main idea is to find the inverse filter to make the output as sparse as possible [9,10]. In the research conducted by Li [11], $\ell_0$ norm transformation is applied to enhance the sparsity of MED outputs and eliminate the iteration overflow caused by the long inverse filter. Morphological filtering with the sparse MED algorithm was proposed in [12], where the scale of the signal characteristics can be adaptively selected according to morphological filtering results. Under the premise of mastering the relevant information about the instrument, these algorithms were efficient [3]. Moreover, transforming the ultrasonic parameter estimation into sparse deconvolution was proved to be potentially possible.

However, due to distortion of the waveform in the transmission process, the received signal has a high probability that only the morphological hypothesis (e.g., the sparse prior) can satisfy. Hence, many researchers in recent years have directly abstracted ultrasonic pulse-echo deconvolution into nonconvex optimization with sparsity assumptions, where both the transmitted wavelet and the reflectivity function are unknown. Li [13] constructed a reliable two-step computing framework and applied phase lifting to initialize the blur function. Under statistical assumptions of the signal, Qu [14] studied the multichannel sparse blind deconvolution (MCS-BD) problem via the gradient descent (GD) algorithm. When multiple input signals are present, finding sparse vectors in a subspace can be directly converted into blind calibration [15]. In order to improve the stability of initialization, $\ell_2$ norm projection [16], least-squares (LS) [17], and pruned tree search [18] are proposed. Zhang [19] explored the influence of symmetric structure on the iterative solution and recovered the shift truncation over the kernel sphere. In simulation experiments, methods based on optimization theory have better antinoise performance than parametric methods.

This letter mainly focuses on sparse blind deconvolution for pulse-echo application, which is an essential branch of ultrasonic NDT. The current challenges of solving this problem with optimization theory are as follows:

- Although accurate estimation parameters can be obtained through finite iterations, utilizing the optimization framework is time-consuming [20].
- Without prior waveform information, the optimization process is nonconvex, which means that a robust initialization algorithm is inevitable.

We construct a two-step algorithm based on the nonconvex optimization inspired by [13] to solve these problems. The nonconvex model for multichannel ultrasonic signal deconvolution using the $\ell_1/\ell_2$ ratio is proposed, the solution process of which is divided into two parts. First, with multiple observations, the convolution process is transformed into a sparse blind correction problem in the sparsity case, and the initial estimations of system response and sparse sequence are obtained by the alternating direction method of multipliers (ADMM) [21]. The proximal alternating linearized minimization (PALM) algorithm [22] is applied to solve the nonconvex and nonsmooth optimization problems over each individual received sample. The majorize-minimize (MM) approximation
is introduced with symmetric positive definite (SPD) matrix compensation [23] for convergence accelerating purposes. The performance of the proposed algorithms is investigated through simulation analysis and numerical experiment, which demonstrates that the superimposed signals can be significantly faster deconvoluted with reasonable initial estimations.

The remainder of this letter is organized as follows. Section 2 models the multiple blind deconvolution problem and introduces the initialization method and iteration algorithm. Section 3 evaluates the proposed method with performance analysis and compares it with classical deconvolution methods. Section 4 concludes and summarizes this paper.

2. Optimization Models and Methods

2.1. Convolution Model of Ultrasonic Inspection

The ultrasonic excitation signal is usually generated by a high-voltage pulse generator, which stimulates a sensor made of piezoelectric material. The generated wave propagates in the measured object and forms a superimposed ultrasonic echo after reflection by defects and surfaces. We integrate into one signal the measurement model from TOFD and the A-scan experiments regarding the specific ultrasonic pulse-echo application, as illustrated in Figure 1.

Figure 1. Diagram for time of flight diffraction (TOFD) testing and example of A-scan signal: (a) TOFD arrangement for defect inspection; (b) observed receive signal sample by convolution; (c) hypothetical reflectivity sequence.

To analyze the inspection and achieve separation in the time domain, one can regularly assume that the received signal convolves results by the generated ultrasonic pulse-echo with the reflectivity sequence. The latter is a small combination of the reflection coefficients. Hence, the convolution model of the multiple measurement process can be represented as

$$y_l = h_l * x_l + w_l, \quad (1)$$

where \* denotes the convolution operation, \(1 \leq l \leq L\), \(y_l \in \mathbb{R}^n\) is the observation sample at the receiving end, \(h_l \in \mathbb{R}^s\) represents the impulse response (or the blurring function), \(x_l \in \mathbb{R}^n\) is the \(l\)th reflected sequence with the prior sparsity and \(w_l\) is the additive noise. We utilized the Gaussian echo model to
analyze the impulse response and assumed that the relevant parameters were stable over time in a complete measurement process. Thus:

\[ h_l := h(\theta, t) = \beta e^{-at^2} \cos(2\pi\omega t + \phi), \]  

(2)

with the parameter vector \( \theta = (\beta, a, \omega, \phi) \), which describes amplitude, bandwidth factor, center frequency and phase, respectively. In this letter, we are only interested in the waveform of \( h(\theta, t) \) without solving the specific \( \theta \). Hence by ignoring the subscript \( l \) in Equation (1), we propose the following blind deconvolution model to recover \( h \) and \( x \) from \( y \):

\[ F(x, h) = f(x, h) + g(x, h), \]

(3)

where

\[ f(x, h) = f_0(x, h) + \lambda \frac{\|x_1\|}{\|x_2\|}, \]

(4)

is the least-squares objective function \( f_0(x, h) = \frac{1}{2}\|h \ast x - y\|^2 \) with \( \ell_1/\ell_2 \) the ratio regularization function. \( \lambda \) is a regularization parameter. \( g(x, h) \) contains prior information about the optimization objects, such as amplitude range, and is the continuous convex function in the domain. The \( \ell_1/\ell_2 \) function is the normalized version of the \( \ell_1 \) function and holds a scale-invariant, which behaves correctly for deconvolution with irregular blurring functions [24] by considering the following nonconvex optimization algorithm:

\[ \hat{h}, \hat{x} := \arg\min_{h, x} F(x, h), \quad \text{for all} \quad 1 \leq l \leq L. \]

(5)

Obviously, the solution to this problem is not uniquely identifiable, and we can always find a solution from the equivalence class \( y_l = (\epsilon R x) \ast (\epsilon^{-1} R y), \epsilon \neq 0 \) without noise, where \( R \) is the shift symmetry matrix [13]. The convergence to the correct solution relies on the initial estimation \((x_0, h_0)\) of Equation (5). When the center frequency parameter is small, or the basic waveform is the centered Gaussian filter, the global optimal solution can be obtained by random initialization or wavelet initialization function [16]. Because of the constraint from the regularization function, the initialization of the reflected sequence is relatively relaxed. For avoiding the computational difficulty and local optimal solution of solving the model, a well-produced initialization \( h_0 \) is crucial to the success of Equation (5).

2.2. Initialization Based on Blind Gain Calibration

Without considering the noise, we modify Equation (1) as:

\[ y_l = C(g)\overline{x}_l, \]

(6)

where \( C(g) \) is the Toeplitz circulant matrix of \( g = [h_1, h_2, \ldots, h_s, 0, 0, \ldots, 0]^T \in \mathbb{R}^n \), and can be written as

\[ C(g) = \begin{bmatrix} g_1 & \cdots & g_2 \\ \vdots & \ddots & \vdots \\ g_s & \cdots & g_1 \end{bmatrix}. \]

(7)

By denoting that multiple measurements can be written in matrix form, i.e., \( Y = [y_2, y_2, \ldots, y_L] \in \mathbb{R}^{n \times L} \) and \( X = [x_1, x_2, \ldots, x_L] \in \mathbb{R}^{n \times L} \), we rewrite Equation (6) as:

\[ Y = C(g)X. \]

(8)
The circulant matrix $C(g)$ has the equivalent eigenvalue decomposition, thus:

$$C(g) := F^H \text{diag} (\tilde{g}) F,$$  

where $F$ is the $n \times n$ unitary discrete Fourier transform (DFT) matrix, and $\tilde{g} = \sqrt{n} F g \in \mathbb{C}^n$. By applying the DFT matrix $F$ to both sides of Equation (8), we have:

$$FY = \text{diag}(\tilde{g}) FX.$$  

Let $\tilde{Y} = FY$ and, the measurement process can be parameterizing in a linear representation:

$$\tau_i \tilde{y}_{i,l} = F_i^H X_l, \quad 1 \leq i \leq n \quad \text{and} \quad 1 \leq l \leq L,$$  

where $\tilde{y}_{i,l}$ is the element of the $i$th row and the $l$th column in $\tilde{Y}$, $F_i^H$ denotes the $i$th row of $F$.

The bilinear inverse problem of $(\tilde{g}, X)$ is transformed into a linear inverse problem of $(\tau, X)$. Even in the case of actual noise interference [25], we also can solve the problem:

$$\text{diag}(\tau) \tilde{Y} \approx FX.$$  

The proposed initialization algorithm becomes equivalent to blind gain calibration proposed in [26], which leads to solving:

$$\left(\hat{\tau}, \hat{X}\right) := \arg\min_{\tau, X} \|X\|_1, \quad s.t. \text{diag}(\tau) \tilde{Y} = FX.$$  

Evidently, it satisfies for the pair $(0, 0)$. To avoid this, we introduce an additional linear constraint on $\tau$, thus:

$$1^T \tau = c,$$  

where $1^T = [1, 1, \ldots, 1]$. The combination of Equations (13) and (14) is a convex problem, which can be solved using the existing solvers.

However, numerical simulations indicate that the direct use of this linear equality constraint complicates the problem. Therefore, we use the ADMM algorithm, which updates local subproblems to coordinate the global problem and solve Equation (13). An equivalent operator of the sum constraint is applied in the iteration process to simplify the initialization process.

First, the augmented Lagrangian function of Equation (13) with the expansion of Equation (11) is expressed as:

$$L_\rho(X, \tau, \Lambda) = \sum_{l=1}^{L} \|X_l\|_1 + \sum_{i=1}^{n} \sum_{l=1}^{L} a_{i,l} \left( F_i^H X_l - \tau_i \tilde{y}_{i,l} \right) + \frac{\rho}{2} \sum_{i=1}^{n} \sum_{l=1}^{L} \left( F_i^H X_l - \tau_i \tilde{y}_{i,l} \right)^2,$$  

where $\rho > 0$ is the penalty parameter and $\Lambda$ is the Lagrange multiplier matrix. The Equation (13) processes in following steps by updating parameters individually:

$$\begin{cases} 
L^{(k+1)} := \arg\min_{X} L_\rho(X, \tau^{(k)}, \Lambda^{(k)}) \\
\tau^{(k+1)} := \arg\min_{\tau} L_\rho(X^{(k+1)}, \tau, \Lambda^{(k)}) \\
\Lambda^{(k+1)} := \Lambda^{(k)} + \rho \left( FX^{(k+1)} - \text{diag}(\tau^{(k+1)}) \tilde{Y} \right)
\end{cases}.$$  

For update for $X$, $\tau^{(k)}$ can be regarded as constant values, and the minimization problem is reduced to:

$$\arg\min_{\tau} \mathcal{L}_p(X, \tau^{(k)}, \lambda^{(k)}) \arg\min_{\tau} \sum_{l=1}^{L} \|\tilde{x}_l\|_1 + \lambda^T_l (F\tilde{x}_l - \tilde{b}_l^{(k)}) + \frac{\rho}{2} \sum_{l=1}^{L} \|F\tilde{x}_l - \tilde{b}_l^{(k)}\|_2^2,$$  \hspace{1cm} (17)

where $\tilde{b}_l^{(k)}$ is the $l$th column in $\text{diag}(\tau^{(k+1)})\tilde{Y}$. The optimization problem for $\tilde{x}_j$ is stated as follows:

$$\arg\min_{\tilde{x}_j} \|\tilde{x}_j\|_1 + \lambda^T_j (F\tilde{x}_j - \tilde{b}_j^{(k)}) + \frac{\rho}{2} \|F\tilde{x}_j - \tilde{b}_j^{(k)}\|_2^2,$$  \hspace{1cm} (18)

Let $\zeta_0(x) = \lambda^T_j (Fx - \tilde{b}_j^{(k)}) + \frac{\rho}{2} \|Fx - \tilde{b}_j^{(k)}\|_2^2$, and thus:

$$\|\nabla \zeta_0(x') - \nabla \zeta_0(x)\|_2 = \rho \|F^T F(x' - x)\|_2 \leq \rho \|F^T F\|_2 \|x' - x\|_2 = \rho \|F\|_2^2 \|x' - x\|_2,$$  \hspace{1cm} (19)

which means an $\mathcal{L}$-Lipschitzian gradient for every $x' \in \mathbb{R}^n$ (see Appendix A). The Lipschitzian constants $\mathcal{L}_0$ of $\zeta_0(x)$ is $\rho \|F\|_2^2$. Hence, the update rule of $\tilde{x}_j$ is defined based on an approximation:

$$\tilde{x}_j^{(k+1)} := \arg\min_{\tilde{x}_j} \|\tilde{x}_j\|_1 + \frac{\mathcal{L}_0}{2} \|\tilde{x}_j - z^{(k)}\|_2^2,$$  \hspace{1cm} (20)

where:

$$z^{(k)} = \tilde{x}_j^{(k)} - \frac{1}{\mathcal{L}_0} \nabla \zeta_0(\tilde{x}_j^{(k)})$$  \hspace{1cm} (21)

Because $\|\tilde{x}_j\|_1$ is the $\ell_1$ norm function, the iterative shrinkage-thresholding (IST) method [27] can be used to find a closed-form solution:

$$\tilde{x}_j^{(k+1)} = \arg\min_{u} \left( \frac{u - z^{(k)}}{2} \right) + \frac{|u|}{\mathcal{L}_0} = \text{soft}\left(\frac{z^{(k)}}{\mathcal{L}_0}, \frac{1}{\mathcal{L}_0}\right),$$  \hspace{1cm} (22)

where $\tilde{x}_j^{(k+1)}$ means the $j$th element of $\tilde{x}_i^{(k)}$ (resp. $z^{(k)}$). soft($\cdot$) is the soft-thresholding function, which is defined as:

$$\text{soft}(u, v) = \text{sign}(u) \max[|u| - v, 0].$$  \hspace{1cm} (23)

Note that $X^{(k+1)}$ is separable and consists of $\tilde{x}_j^{(k+1)}$, the update of $X^{(k+1)}$ in Equation (13) can be rewritten as:

$$X^{(k+1)} := \text{soft}(X^{(k)} - \frac{1}{\mathcal{L}_0} F^T (\lambda^{(k)} + \rho (FX^{(k)} - \text{diag}(\tau^{(k)})\tilde{Y})), \frac{1}{\mathcal{L}_0}).$$  \hspace{1cm} (24)

For updating $\tau^{(k)}$, $X$ is regarded as constant values, and thus:

$$\tau_i^{(k+1)} := \arg\min_{\tau_i} - \sum_{l=1}^{L} \lambda^T_{i,l} \tilde{g}_{i,l} + \frac{\rho}{2} \sum_{l=1}^{L} F^T F_{i,l} (\tau_i^{(k+1)} - \tau_i^{(k)})^2 = \arg\min_{\tau_i} \mathcal{L}_p^{(k)}(X^{(k+1)}, \tau_i, \lambda^{(k)}),$$  \hspace{1cm} (25)

and when $\nabla \mathcal{L}_p^{(k)}(X^{(k+1)}, \tau_i, \lambda^{(k)}) = 0, \tau_i^{(k+1)}$ is represented as:

$$\tau_i^{(k+1)} := \frac{\sum_{l=1}^{L} \lambda_{i,l} \tilde{g}_{i,l} + \rho \sum_{l=1}^{L} F^T F_{i,l} \tilde{g}_{i,l}}{\rho \sum_{l=1}^{L} \tilde{g}_{i,l}^2} = \frac{\bar{U}_{i}^{(k+1)}}{\bar{V}_i},$$  \hspace{1cm} (26)
where \( U, V \in \mathbb{C}^n \), and:

\[
U^{(k+1)} = \text{diag}((A^{(k)} + \rho M_d)Y^T), \quad V = \text{diag}(\rho Y^T), \quad M = FX^{(k+1)}.
\]  

(27)

Similar to \( X^{(k+1)} \), we use \( U^{(k+1)}/V \) directly to update \( \tau^{(k+1)} \). Considering the constraints from formula Equation (14), we modify the update method of \( X^{(k+1)} \) as:

\[
\tau_i^{(k+1)} := \frac{U_i^{(k+1)} + \delta^{(k+1)}}{V_i}, \quad \delta^{(k+1)} = \frac{c - \sum_{i=1}^{n} \frac{U_i^{(k+1)}}{V_i}}{\sum_{i=1}^{n} \frac{1}{V_i}},
\]  

(28)

where \( \text{trace}(\text{diag}(\tau^{(k+1)})) = c \) satisfies the sum constraint.

Thus, through Equations (16), (24), and (28), the impulse response and sparse sequences from each measurement are estimated up to specific scaling. Despite the arbitrary selection of \( c \) and the assumption without noise, the initialized pair \( (x_0, h_0) \) is close enough to the ground truth, and guarantees Equation (5) can converge to the global optimal solution approximately.

2.3. Alternating Optimization Method Based on PALM

The basic PALM method is proposed for minimization of a sum of finite functions and requires the smooth function (or partially Lipschitz properties). However, the \( \ell_1/\ell_2 \) ratio is both nonconvex and nonsmooth, which means the smoothed \( \ell_1/\ell_2 \) ratio proposed in [16] can replace the \( \ell_1/\ell_2 \) ratio in \( f(x,h) \). Hence, we rewrite Equation (4) as \( f(x,h) = f_0(x,h) + \phi(x) \), and:

\[
\phi(x) = \lambda \log\left(\frac{\ell_1(x)}{\ell_2(x)}\right), \quad \ell_1(x) = \sum_{i=1}^{n} \sqrt{x_i^2 + \alpha^2 - \alpha}, \quad \ell_2(x) = \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \eta^2,
\]  

(29)

which can be smooth approximations for sparse representation. Besides, we assume that \( g(x,h) \) is separable and is a proper, lower semicontinuous, convex function. Thus:

\[
g(x,h) = g_1(x) + g_2(h),
\]  

(30)

where \( g_1 \) and \( g_2 \) are the indicator functions with prior knowledge. Equation (5) is represented as:

\[
\arg\min_{h,x} F(x,h) := f_0(x,h) + \phi(x) + g_1(x) + g_2(h).
\]  

(31)

By starting with the initial estimates \((x^{(0)}, h^{(0)})\) obtained from the initialization, we generate the sequence via the scheme:

\[
\begin{align*}
  x^{(k+1)} &= \text{argmin}_x F(x, h^{(k)}) \\
  h^{(k+1)} &= \text{argmin}_h F(x^{(k+1)}, h).
\end{align*}
\]  

(32)

The PALM for solution of Equation (32) are defined respectively by:

\[
\begin{align*}
  x^{(k+1)} &= \text{prox}_{\delta_1,\ell_1}(x^{(k)} - \frac{1}{\delta_1} \nabla_x F(x^{(k)}, h^{(k)})) \\
  h^{(k+1)} &= \text{prox}_{\delta_2,\ell_2}(h^{(k)} - \frac{1}{\delta_2} \nabla_h F(x^{(k+1)}, h^{(k)})).
\end{align*}
\]  

(33)

where the proximal operator is defined by the proximal map, which is associated to a specific function, i.e.,

\[
\text{prox}_{\delta,\ell}(z) = \arg\min_x \alpha(x) + \frac{\ell}{2}||x - z||_2^2.
\]  

(34)

\( c^{(k)} \) and \( d^{(k)} \) are the partial Lipschitz constants of \( V f \) (see Appendix A).
The solution of the problem is converted to determine the appropriate \( c^{(k)} \) and \( d^{(k)} \), so that the function Equation (33) converges. From idea of gradient descent, \( c^{(k)} \) and \( d^{(k)} \) can be selected as:

\[
\begin{align*}
    c^{(k)} &= \gamma_1 \mathcal{L}_1(h^{(k)}), \\
    d^{(k)} &= \gamma_2 \mathcal{L}_2(x^{(k+1)}),
\end{align*}
\]

where \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are the Lipschitz constants for \( \nabla_x f \) and \( \nabla_h f \) [22]. Similar to the algorithm in [28], \( x^{(k+1)} \) can be considered constants when updating \( h^{(k+1)} \), and:

\[
\mathcal{L}_2(x^{(k+1)}) \approx \| (T_{x^{(k+1)}})^T T_{x^{(k+1)}} \|_2,
\]

where \( T_x : \mathbb{H}^n \to \mathbb{H}^{n \times n} \) is the Toeplitz matrix operator. For \( x^{(k+1)} \), we rewrite the update step as:

\[
\begin{align*}
    x^{(k+1)} &\propto x^{(k)} - \frac{1}{\gamma_1} A(x^{(k)}, h^{(k)})^{-1} \nabla_x f(x^{(k)}, h^{(k)}), \\
    \tilde{x}^{(k)} &= x^{(k)} - \frac{1}{\gamma_1} A(x^{(k)}, h^{(k)})^{-1} \nabla_x f(x^{(k)}, h^{(k)}),
\end{align*}
\]

and \( A(x^{(k)}, h^{(k)}) \) is the symmetric positive definite (SPD) matrix which satisfies the MM principle [23], which can be obtained by building majorizing approximations for \( f(x, h) \) [29]. Thus:

\[
A(x^{(k)}, h^{(k)}) = \left( \mathcal{L}'_1 + \frac{9}{8\eta} I_n + \text{diag}_{\gamma_2 \omega} (x^{(k)}) \right),
\]

\[
\text{diag}_{\gamma_2 \omega} (x) = \text{diag}\left( \left[ (\alpha_l^2 + \sigma^2)^{-\frac{1}{2}} \right]_{1 \leq l \leq n} \right),
\]

where \( \mathcal{L}'_1 \) is the Lipschitzian constants of \( \nabla_x f(x, h^{(k)}) \). By Equation (8), we have:

\[
\nabla_x f(x, h^{(k)}) = \mathcal{M}^T (Mx - y), \quad \mathcal{M} = \mathcal{F}^T \text{diag}(Fh^{(k)}) \mathcal{F},
\]

and \( \mathcal{L}'_1 = \| M \|_2^2 \). The other items in \( A(x^{(k)}, h^{(k)}) \) are given by the proposition established in [16].

Thus far, for each measurement sample \( y_i \), we can obtain the corresponding estimated \( (\hat{x}_i, \hat{h}_i) \) by solving Equation (32) with a global initialization \( (x_0, h_0) \) from Equation (16).

3. Simulation Results

3.1. Stable Initialization with Phase Transitions

Phase diagrams, which demonstrate the empirical probability of success over a range of sparsity and samples for a fixed sampling window length, are implemented to evaluate the feasibility of the proposed initialization algorithm. We classify a success initialization \( h_0 \) of \( h \) if \( \| h_0 - h \|_2 / \| h \|_2 \leq 0.3 \). Figure 2 presents variations of the initialization success rate by phase transitions, concerning the \( K \)-sparsity of each \( \tilde{x}_i \) and the number of samples \( L \). The additive noise \( w \) is considered as white Gaussian noise with the variance \( \sigma^2 \). The length of the sequence \( n = 100 \) and the recovery success rate are calculated using 50 Monte Carlo simulations on every grid point. We rescale the estimated \( h_0 \) to have the same norm as the correct \( h \).
3. Simulation Results

3.1. Stable Initialization with Phase Transitions

Phase diagrams, which demonstrate the empirical probability of success over a range of sparsity and samples for a fixed sampling window length, are implemented to evaluate the feasibility of the proposed initialization algorithm. We classify a success initialization of \( h_0 \) if
\[
\|h_0 - h\|_2 / \|h\|_2 \leq 0.3
\]
Figure 2 presents variations of the initialization success rate by phase transitions, concerning the \( \beta \)-sparsity of each \( x \) and the number of samples \( m \).

The additive noise \( \mathbf{j} \) is considered as white Gaussian noise with the variance \( \sigma^2 \). The length of the sequence \( n = 100 \) and the recovery success rate are calculated using 50 Monte Carlo simulations on every grid point. We rescale the estimated \( h_0 \) to have the same norm as the correct \( h \).

\[
\begin{array}{ll}
(a) & \text{noise free; (b) } \sigma = 0.01; \\
(c) & \text{(c) } \sigma = 0.02; \\
(d) & \text{(d) } \sigma = 0.03. 
\end{array}
\]

Figure 2. Phase transitions of the proposed initialization algorithm with different noise levels: (a) noise free; (b) \( \sigma = 0.01 \); (c) \( \sigma = 0.02 \); (d) \( \sigma = 0.03 \). The gray bar denotes the success rate.

As shown in Figure 2a, the initialization algorithm can estimate \( h \) correctly when the sparsity and the number of samples are appropriate. To achieve successful initialization with Equation (16), \( L \) must scale linearly with \( K \). However, when \( K \) does not satisfy the sparse hypothesis, i.e., \( K \ll n \) does not hold, blind gain calibration cannot achieve convergence to the neighborhood of \( h \). In the case of noise interference, the proposed method has a high probability of obtaining \( h_0 \) with increased samples. Because the influence of noise is not introduced in Equation (11), few samples lead to additional errors for the iterative recovery. Meanwhile, as shown in Figure 2d, the probability of successful initialization decreases with serious noise interference.

3.2. Numerical Comparison for Deconvolution Evaluation

Figure 3 illustrates the comparison of conventional parametric deconvolution methods numerically with the same noise variance (\( \sigma = 0.02 \)). The regularization parameters and indicator functions of [12,29] are adjusted in this comparison. In the case of \( K = 3, L = 20, \) and \( n = 1000 \), offset and polarity reversal presents in the conventional fast deconvolution methods, such as MED [11] and LS [17]. The MED algorithm uses an inverse filter to solve sparse reflection sequence directly. However, the uncorrelation between the inverse filter and the blurring function results in deviation and fluctuation, as shown in Figure 3a. Besides, the MED algorithm extracts the arrival time by manually modulating the threshold, which reduces detection stability. The LS algorithm estimates the blurring function and analyzes the concentration of the error, where the relatively accurate time parameters can be obtained.
The regularized coefficient of noise is based on the prior information of the signal. Therefore, the deviation appears in the estimation of amplitude parameters. The reflection sequence, estimated by LS (Figure 2b), is sparser than that estimated by MED. Although MED uses sparse $\ell_0$ constraint, the noise also affects the deconvolution results.

Figure 3. Comparison of parametric methods: (a) $\ell_0$ norm MED (the threshold is used to determine detection points); (b) LS algorithm.

Compared with parametric methods, the method based on optimization theory has better adaptability to the superimposed echo. Therefore, to evaluate the performance of the proposed algorithm, the sparsity is increased appropriately. Further, when $K = 20$, $L = 20$, and $n = 1000$ ($\sigma = 0.02$), Figure 4 shows the initialization comparison of the Smoothed One-Over-Two (SOOT) [29] algorithm and the proposed method. The centered Gaussian filter from SOOT is different from the basic ultrasonic waveform, where converging on local optimum is inevitable. Although the blind gain calibration is implemented with time cost, the proposed algorithm provides a starting point close to the original blurring function for alternating convergence.

Figure 4. Comparison of initialization algorithms.
Figure 5 reveals the results of the initialization algorithm and alternating solution. Although the SOOT algorithm reduces the influence of initialization by variable metric strategies and increasing the number of subiterations, the spike interference still affects the solution where the initialization of the centered Gaussian filter is not applicable to the true ultrasonic response. Despite limited missing solutions and amplitude deviation, the proposed algorithm can accurately estimate the position and amplitude of the reflected sequence in the case of relative sparsity.

Figure 4. Comparison of initialization algorithms.

(a) Stopping criterion for iterative algorithms is set as \( \| f_f \| \) the norm distances between the estimated values and the correct values for different noise levels, respectively. Table 1 compares the residual deconvolution error, where \( \ell_1 \) and \( \ell_\infty \) represent the norm distances between the estimated values and the correct values for different noise levels, respectively. The running time is obtained by taking the average value from experiments, and the stopping criterion for iterative algorithms is set as \( \| x^{(k+1)} - x^{(k)} \|_1 \leq 10^{-3} \). Although the deconvolution

For a fair comparison, we used the same noise conditions and parameter settings as illustrated in [29], where the fixed noise standard variances (0.01, 0.02, and 0.03) and regularization parameters are considered. Table 1 compares the residual deconvolution error, where \( \ell_2 \) and \( \ell_1 \) represent the norm distances between the estimated values and the correct values for different noise levels, respectively. The running time is obtained by taking the average value from experiments, and the stopping criterion for iterative algorithms is set as \( \| x^{(k+1)} - x^{(k)} \|_1 \leq 10^{-3} \). Although the deconvolution

Figure 5. Deconvolution results of different methods with optimization theory: (a) SOOT; (b) proposed algorithm.

For a fair comparison, we used the same noise conditions and parameter settings as illustrated in [29], where the fixed noise standard variances (0.01, 0.02, and 0.03) and regularization parameters are considered. Table 1 compares the residual deconvolution error, where \( \ell_2 \) and \( \ell_1 \) represent the norm distances between the estimated values and the correct values for different noise levels, respectively. The running time is obtained by taking the average value from experiments, and the stopping criterion for iterative algorithms is set as \( \| x^{(k+1)} - x^{(k)} \|_1 \leq 10^{-3} \). Although the deconvolution
based on LS [17] is superior in speed, the cost is a large estimation error. MED-based methods [9–11] achieve relatively good sparse sequence estimation in less computing time. However, because of the characteristics of the inverse filter, the basic response of the signal is inaccessible. A solution using alternate optimization can obtain a better solution at the cost of computing time. The CVX toolbox without improvement causes additional time consumption due to the introduction of random initialization. Besides, because the optimization conditions do not make specific constraints on the ultrasonic waveform, the estimation error of the blurring function is larger than that of the reflection sequence. Meanwhile, the proposed algorithm is significantly faster than the other algorithms based on nonconvex optimization. The proposed initialization is closer to the ground truth, where the algorithm can achieve the stopping criterion more quickly.

**Table 1.** Comparison of algorithm deconvolution results. The interpretation platform used the Intel Core i5-9400f processor with 16GB memory.

| Algorithm   | $\sigma = 0.01$ | $\sigma = 0.02$ | $\sigma = 0.03$ | Running Time (s) |
|-------------|-----------------|-----------------|-----------------|-----------------|
| LS          | $\ell_2, h$     | $\ell_1, h$     | $\ell_2, h$     | $\ell_1, h$     | $\ell_2, h$     | $\ell_1, h$     | 3.89 |
| $\ell_0$ MED | $\ell_0$        | 0.8121          | $\ell_0$        | 1.1002          | $\ell_0$        | 1.3063          | 5.62 |
| M-S-MED     | $\ell_0$        | 0.7203          | $\ell_0$        | 0.8212          | $\ell_0$        | 1.0304          | 8.20 |
| OMED        | $\ell_0$        | 0.5548          | $\ell_0$        | 0.6167          | $\ell_0$        | 0.9033          | 10.34 |
| $\ell_1/\ell_2$ CVX | 1.5301       | 0.1023          | $\ell_0$        | 1.6623          | $\ell_0$        | 0.4180          | 2.0301       | 0.8902       | 65.29 |
| SOOT        | 0.8923          | 0.0221          | $\ell_0$        | 1.2340          | $\ell_0$        | 0.0936          | 1.6721       | 0.1645       | 34.50 |
| proposed    | 0.3622          | 0.0133          | $\ell_0$        | 0.4528          | $\ell_0$        | 0.0584          | 0.8973       | 0.1045       | 22.32 |

4. Conclusions

In this letter, a deconvolution method based on nonconvex optimization is proposed. By simplifying the problem to blind gain calibration, a two-step iterative initialization can effectively obtain the initial guesses for ultrasonic pulse-echo response and reflection sequence with the ADMM algorithm. Then, using $\ell_1/\ell_2$ ratio regularization, an exact algorithm based on the PALM framework is applied. Simulation and numerical analysis illustrate that the proposed method algorithm has the probability of successful initialization. In the algorithm comparison experiment, deconvolution results are obtained with high accuracy and acceptable time cost.

The proposed method has potential gains in ultrasonic NDT applications for defect detection and sensor parameter estimation. When the specific instrument parameters are ambiguous or the propagation is distorted, the parameters representing the measured object can be obtained accurately through sparse blind deconvolution. However, the time cost limits the application of the algorithm in other areas (such as ultrasonic imaging) unless the processing speed is further improved.

**Author Contributions:** Conceptualization of the paper, X.G. and K.D.; Design of the algorithm and experiment, X.G. and K.D.; Original draft preparation, X.G.; Review and editing, Q.Z.; Supervision and funding acquisition Y.S. and W.Z.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by China National Offshore Oil Corporation under the Grant numbered CNOOC-KJ ZDHXJSGG YF 2019-02.

**Acknowledgments:** X.G. would like to thank Q.Z. for providing the dataset and administrative support.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A**

For every $k \in \mathbb{N}$, the quadratic function defined as

$$
\forall x \in \mathbb{R}^N, \quad Q(x, x_k)F(x_k) + (x - x_k)^T \nabla F(x_k) + \frac{1}{2} \|x - x_k\|_A^2, \quad (A1)
$$
is a majorant function of \( F(x_k) \). Thus:

\[
\forall k \in \mathbb{N}, \quad v_1 I_N \leq A_k \leq v_2 I_N,
\]

for \( v_1, v_2 \in [0, +\infty)^2 \). If \( F \) is Lipschitz differentiable on \( \text{dom } R \), yielding:

\[
\forall x, y \in (\text{dom } R)^2, \quad F(x) < F(y) + (x - y)^T V F(x) + \frac{1}{2} \| x - y \|^2,
\]

for \( \text{dom } R \) is a convex set and \( A_k = L I_N \), where \( L > 0 \) is the Lipschitz constant of \( VF \).

References

1. Katunin, A.; Wronkowicz-Katunin, A.; Dragan, K. Impact Damage Evaluation in Composite Structures Based on Fusion of Results of Ultrasonic Testing and X-ray Computed Tomography. *Sensors* 2020, 20, 1867. [CrossRef] [PubMed]
2. Zhao, N.; Basarab, A.; Kouame, D.; Tourneret, J.-Y. Joint Segmentation and Deconvolution of Ultrasound Images Using a Hierarchical Bayesian Model Based on Generalized Gaussian Priors. *IEEE Trans. Image Process.* 2016, 25, 3736–3750. [CrossRef] [PubMed]
3. Park, Y.; Choi, A.; Kim, K. Monaural Sound Localization Based on Reflective Structure and Homomorphic Deconvolution. *Sensors* 2017, 17, 2189. [CrossRef] [PubMed]
4. Jeong, H.; Cho, S.; Zhang, S.; Li, X. Acoustic nonlinearity parameter measurements in a pulse-echo setup with the stress-free reflection boundary. *J. Acoust. Soc. Am.* 2018, 143, EL237–EL242. [CrossRef] [PubMed]
5. Chang, Y.; Zi, Y.; Zhao, J.; Yang, Z.; He, W.; Sun, H. An adaptive sparse deconvolution method for distinguishing the overlapping echoes of ultrasonic guided waves for pipeline crack inspection. *Meas. Sci. Technol.* 2017, 28, 035002. [CrossRef]
6. Jin, H.; Chen, J.; Yang, K. A blind deconvolution method for attenuative materials based on asymmetrical Gaussian model. *J. Acoust. Soc. Am.* 2016, 140, 1184–1191. [CrossRef]
7. Bossmann, F.; Plonka, G.; Peter, T.; Nemitz, O.; Schmitte, T. Sparse Deconvolution Methods for Ultrasonic NDT Application on TOFD and Wall Thickness Measurements. *J. Nondestruct. Eval.* 2012, 31, 225–244. [CrossRef]
8. Abdessalem, B.; Farid, C. Resolution Improvement of Ultrasonic Signals Using Sparse Deconvolution and Variational Mode Decomposition Algorithms. *Russ. J. Nondestruct. Test.* 2020, 56, 479–489. [CrossRef]
9. Buzzoni, M.; Antoni, J.; D’Elia, G. Blind deconvolution based on cyclostationarity maximization and its application to fault identification. *J. Sound Vib.* 2018, 432, 569–601. [CrossRef]
10. Cheng, Y.; Zhou, N.; Zhang, W.; Wang, Z. Application of an improved minimum entropy deconvolution method for railway rolling element bearing fault diagnosis. *J. Sound Vib.* 2018, 425, 53–69. [CrossRef]
11. Li, X.; Li, X.; Liang, W.; Chen, L. l(0)-norm regularized minimum entropy deconvolution for ultrasonic NDT & E. *NDT E Int.* 2012, 47, 80–87. [CrossRef]
12. Li, M.; Li, X.; Gao, C.; Song, Y. Acoustic microscopy signal processing method for detecting near-surface defects in metal materials. *NDT E Int.* 2019, 103, 130–144. [CrossRef]
13. Li, X.; Ling, S.; Strohmer, T.; Wei, K. Rapid, robust, and reliable blind deconvolution via nonconvex optimization. *Appl. Comput. Harmon. Anal.* 2019, 47, 893–934. [CrossRef]
14. Qu, Q.; Li, X.; Zhu, Z. Exact Recovery of Multichannel Sparse Blind Deconvolution via Gradient Descent. *SIAM J. Imaging Sci.* 2020, 13, 1630–1652. [CrossRef]
15. Wang, L.; Chi, Y. Blind Deconvolution From Multiple Sparse Inputs. *IEEE Signal Process. Lett.* 2016, 23, 1384–1388. [CrossRef]
16. Repetti, A.; Mai Quyen, P.; Duval, L.; Chouzenoux, E.; Pesquet, J.-C. Euclid in a Taxicab: Sparse Blind Deconvolution with Smoothed l(1)/l(2) Regularization. *IEEE Signal Process. Lett.* 2015, 22, 539–543. [CrossRef]
17. Guan, J.; Wang, X.; Wang, W.; Huang, L. Sparse Blind Speech Deconvolution with Dynamic Range Regularization and Indicator Function. *Circuits Syst. Signal Process.* 2017, 36, 4145–4160. [CrossRef]
18. Jing, S.; Hall, J.; Zheng, Y.R.; Xiao, C. Signal Detection for Underwater IoT Devices With Long and Sparse Channels. *IEEE Internet Things J.* 2020, 7, 6664–6675. [CrossRef]
19. Zhang, Y.; Kuo, H.-W.; Wright, J. Structured Local Optima in Sparse Blind Deconvolution. *IEEE Trans. Inf. Theory* 2020, 66, 419–452. [CrossRef]
20. Yang, H.; Su, X.; Chen, S. Blind Image Deconvolution Algorithm Based on Sparse Optimization with an Adaptive Blur Kernel Estimation. *Appl. Sci.* **2020**, *10*, 2437. [CrossRef]

21. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Found. Trends Mach. Learn.* **2011**, *3*, 1–122. [CrossRef]

22. Bolte, J.; Sabach, S.; Teboulle, M. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Math. Program.* **2014**, *146*, 459–494. [CrossRef]

23. Chouzenoux, E.; Pesquet, J.-C.; Repetti, A. Variable Metric Forward-Backward Algorithm for Minimizing the Sum of a Differentiable Function and a Convex Function. *J. Opt. Theory Appl.* **2014**, *162*, 107–132. [CrossRef]

24. Krishnan, D.; Tay, T.; Fergus, R. Blind Deconvolution Using a Normalized Sparsity Measure. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, 20–25 June 2011; pp. 233–240. [CrossRef]

25. Li, Y.; Lee, K.; Bresler, Y. Blind Gain and Phase Calibration via Sparse Spectral Methods. *IEEE Trans. Inf. Theory* **2019**, *65*, 3097–3123. [CrossRef]

26. Gribonval, R.; Chardon, G.; Daudet, L. Blind calibration for compressed sensing by convex optimization. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, 25–30 March 2012; pp. 2713–2716.

27. Combettes, P.L.; Wajs, V.R. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* **2005**, *4*, 1168–1200. [CrossRef]

28. Sun, Y.; Babu, P.; Palomar, D.P. Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning. *IEEE Trans. Signal Process.* **2017**, *65*, 794–816. [CrossRef]

29. Mai Quyen, P.; Oudompheng, B.; Nicolas, B.; Mars, J.I. Sparse deconvolution for moving-source localization. In *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing Proceedings*, Shanghai, China, 20–25 March 2016; pp. 355–359.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).