# A Proximal Algorithm with Convergence Guarantee for a Nonconvex Minimization Problem Based on Reproducing Kernel Hilbert Space 

Hong-Xia Dou ${ }^{1, *}$ and Liang-Jian Deng ${ }^{2}$<br>1 School of Science, Xihua Univercity, Chengdu 610039, China<br>2 School of Mathematical Science, University of Electronic and Technology of China, Chengdu 611731, China; liangjian.deng@uestc.edu.cn<br>* Correspondence: hongxia.dou@mail.xhu.edu.cn

Citation: Dou, H.-X.; Deng, L.-J. A Proximal Algorithm with Convergence Guarantee for a Nonconvex Minimization Problem Based on Reproducing Kernel Hilbert Space. Symmetry 2021, 13, 2393.
https:/ /doi.org/10.3390/sym13122393

Academic Editors: Leszek Gasinski and Ioan Raşa

Received: 3 November 2021
Accepted: 6 December 2021
Published: 12 December 2021

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

Copyright: © 2021 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 (https:/ / creativecommons.org/licenses/by/ 4.0/).


#### Abstract

The underlying function in reproducing kernel Hilbert space (RKHS) may be degraded by outliers or deviations, resulting in a symmetry ill-posed problem. This paper proposes a nonconvex minimization model with $\ell_{0}$-quasi norm based on RKHS to depict this degraded problem. The underlying function in RKHS can be represented by the linear combination of reproducing kernels and their coefficients. Thus, we turn to estimate the related coefficients in the nonconvex minimization problem. An efficient algorithm is designed to solve the given nonconvex problem by the mathematical program with equilibrium constraints (MPEC) and proximal-based strategy. We theoretically prove that the sequences generated by the designed algorithm converge to the nonconvex problem's local optimal solutions. Numerical experiment also demonstrates the effectiveness of the proposed method.


Keywords: degraded function; nonconvex minimization; convergence analysis; RKHS; symmetry ill-posed problem

## 1. Introduction

The reproducing kernel Hilbert space (RKHS, denote as $\mathcal{H}$ ) has been widely studied in many studies [1-7]. Its most critical property is that the functions in RKHS can be linearly represented by reproducing kernel function. In addition, many studies have analyzed the properties of unitary or binary functions in RKHS. These functions usually can be regarded as signals or images in discrete form, so as to build optimization models and solve some application problems, such as image super-resolution and image restoration.

In general, the Hilbert space can be considered as $\mathcal{H} \subset L^{2}(\mathbb{P})$ which $\mathbb{P}$ is a probability measure on the subset $\mathcal{X} \subset \mathbb{R}, \mathcal{H}$ is complete for a class of real valued functions $f: \mathcal{X} \rightarrow \mathbb{R}$ with $\|f\|_{L^{2}(\mathbb{P})}<\infty$. Moreover, the reproducing kernel $\kappa$ of $\mathcal{H}$ can be defined as: (1) for any $x \in \mathcal{X}$, the function $\kappa(\cdot, x)$ blongs to $\mathcal{H}$. (2) the function $\kappa$ has so-called reproducing property, that is $f(x)=\langle f, \kappa(\cdot, x)\rangle_{\mathcal{H}}$ for all $f \in \mathcal{H}$, and $\langle\cdot, \cdot\rangle_{\mathcal{H}}$ represents an associated inner product. By this relation, we could get the Gram matrix $K \in \mathbb{R}^{n \times n}$ by the discretization of reproducing kernel $\kappa(\cdot, \mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^{n}$, thus it is easy to get the following discrete formulation by considering the bias $\mathbf{e} \in \mathbb{R}^{n}$ :

$$
\begin{equation*}
\mathbf{g}=\mathbf{K} \alpha+\mathbf{e}, \tag{1}
\end{equation*}
$$

where $\alpha \in \mathbb{R}^{n}$ is the coefficient we need to estimate, besides, $K$ is a real symmetry matrix.
In the real world, the underlying function $\mathbf{g}$ generally will be polluted by outliers and Gaussian perturbation, which gets the following symmetry ill-posed problem:

$$
\begin{equation*}
\mathbf{y}=\mathbf{K} \alpha+\mathbf{e}+\mathbf{u}+\sigma, \tag{2}
\end{equation*}
$$

where $\mathbf{u} \in \mathbb{R}^{n}$ can be viewed as outliers and $\sigma \in \mathbb{R}^{n}$ stands for Gaussian perturbation. Our final goal is to accurately estimate the coefficients $\alpha$ and $\mathbf{e}$ from the known $\mathbf{y}$ and $\mathbf{K}$. After obtaining $\alpha$ and $\mathbf{e}$, we could calculate the underlying function $\mathbf{g}$ by (1). Note that, solving the problem (2) is quite a chanllenging task, since the variables $\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u}$ and $\sigma$ in (2) are all unknown, which leads to a ill-posed problem.

In [8], Papageorfiou et al, considered that the function $g$ can be linearly represented by coefficient $\alpha$ and constant $c$ as $\mathbf{g}=\mathbf{K} \alpha+c \mathbf{1}$, and proposed a kernel regularized orthogonal maching pursuit (KROMP) method to solve the nonconvex problem. However, the KROMP method has two weakness: one is that the constant $c$ is not general and flexible; the other is that the convergence of its algorithm is not guaranteed theoretically. Therefore, this paper mainly establishes the nonconvex optimization model for the degraded problem (2) in RKHS, and gives the designed algorithm whose convergence can be guaranteed, finally shows the effectiveness of the proposed method in some simulation experiments.

Regularized modeling is a promising way to deal with ill-posed problems. Actually, the variable $\mathbf{u}$ representing outliers is generally sparse, which motivates us to formulate a sparsity-based regularization model. Especially, $\ell_{0}$-quasi norm that counts the nonzero elements in a vector is an ideal metric to depict the sparse property. Therefore, the nonconvex minimization problem for solving the ill-posed problem (2) can be simply shown as follows,

$$
\begin{equation*}
\min _{\alpha, \mathbf{e}, \mathbf{u} \in \mathbb{R}^{n}} \Phi(\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u})=\frac{\mu}{2}\|\mathbf{K} \alpha+\mathbf{e}+\mathbf{u}-\mathbf{y}\|_{2}^{2}+\frac{\lambda_{1}}{2}\|\boldsymbol{\alpha}\|_{2}^{2}+\frac{\lambda_{2}}{2}\|\mathbf{e}\|_{2}^{2}+\|\mathbf{u}\|_{0} \tag{3}
\end{equation*}
$$

where $\mu, \lambda_{1}$ and $\lambda_{2}$ are positive parameters. The first term in (3) is deduced from the Gaussian perturbation of $\sigma$ under the framework of maximum a posteriori (MAP) estimation. The second and third terms are two regularized terms to depict the underlying prior for $\alpha$ and $\mathbf{e}$. The last term is a $\ell_{0}$-quasi norm to depict the sparse prior of outlier $\mathbf{u}$. Note that the given regularization model (3) is a nonconvex minimization problem due to the nonconvex property of $\ell_{0}$ term. In general, the $\ell_{0}$ term will be replaced approximately by some other convex terms (e.g., $\ell_{1}$ term or hard threshold [9-12]) for simpler computation and convergence guarantee. However, if taking this way, it will lose the accuracy of depicting sparsity, which may result in unreasonable outcomes.

However, the nonconvex minimization problems usually have the following difficulties which are encountered to solve: (1) Whether the designed algorithm can effectively solve the minimization model? (2) Whether the convergence analysis of the designed algorithm can be guaranteed? (3) Whether the initial value affects the convergence of the designed algorithm? Thus, many studies have been devote to conquer these weaknesses of nonconvex problems.

Recently, the nonconvex problem can be reformulated as an equivalent minimization problem based on the mathematical program with equilibrium constraints (MPEC) which can be effectively solved by the classical algorithms [13-15]. For instance, Yuan et al. [14] have proposed an equivalent biconvex MPEC formulation for $\ell_{0}$-quasi norm of the nonconvex minimization problem. Additionally, the proximal alternating based algorithm has widely been used to solve the nonconvex and nonsmooth problems [16-24]. In [18], Bolte et al. propose a proximal alternating linear minimization algorithm (PALM) framework to solve nonconvex and nonsmoothing minimization problem, and give the convergence analysis of the algorithm.

In this paper, we mainly focus on the above mentioned difficulties of nonconvex minimization problem (3) to design an efficient algorithm with convergence guarantee theoretically. A simple and representative example is employed to verify the effectiveness of the proposed method. Besides, the contributions of this work can be summarized as: (1) New nonconvex minimization modeling based on RKHS; (2) Convergence guarantee of the designed algorithm for the nonconvex problem.

The outline of this paper is as follows. Section 2 shows the detailed algorithm for the nonconvex minimization problem (3). In Section 3, the convergence analysis of the given
algorithm is given. Numerical results are reported in Section 4. Finally, conclusions are drawn in Section 5.

## 2. The Solution for the Nonconvex Minimization Problem

Based on the MPEC lemma of $\ell_{0}$-quasi norm (see more details from [14]), the nonconvex minimization problem (3) can be equivalently reformulated the following model:

$$
\begin{align*}
\min _{\alpha, \mathbf{e}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n}} \Phi(\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u}, \mathbf{v}) & =\frac{\mu}{2}\|\mathbf{K} \alpha+\mathbf{e}+\mathbf{u}-\mathbf{y}\|_{2}^{2}+\frac{\lambda_{1}}{2}\|\boldsymbol{\alpha}\|_{2}^{2}+\frac{\lambda_{2}}{2}\|\mathbf{e}\|_{2}^{2} \\
& +\langle\mathbf{1}, \mathbf{1}-\mathbf{v}\rangle+\iota_{[0,1]}(\mathbf{v})  \tag{4}\\
\text { s.t. } \mathbf{v} \odot|\mathbf{u}| & =0
\end{align*}
$$

where $\odot$ represents point-wise multiplication, and $\iota_{[0,1]}(\mathbf{v})$ is indicator function projecting the elements of $\mathbf{v}$ into $[0,1]$. The constrained minimization problem (4) can be rewritten as the following unconstrained minimization problem by the penalty strategy:

$$
\begin{equation*}
\Phi(\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u}, \mathbf{v})=H(\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u}, \mathbf{v})+f_{1}(\boldsymbol{\alpha})+f_{2}(\mathbf{e})+f_{3}(\mathbf{u})+f_{4}(\mathbf{v}) \tag{5}
\end{equation*}
$$

where $H(\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u}, \mathbf{v})=\frac{\mu}{2}\|\mathbf{K} \boldsymbol{\alpha}+\mathbf{e}+\mathbf{u}-\mathbf{y}\|_{2}^{2}+\frac{\beta}{2}\|\mathbf{v} \odot|\mathbf{u}|\|_{2}^{2}, f_{1}(\boldsymbol{\alpha})=\frac{\lambda_{1}}{2}\|\boldsymbol{\alpha}\|_{2}^{2}, f_{2}(\mathbf{e})=\frac{\lambda_{2}}{2}\|\mathbf{e}\|_{2}^{2}$, $f_{3}(\mathbf{u})=0, f_{4}(\mathbf{v})=\langle\mathbf{1}, \mathbf{1}-\mathbf{v}\rangle+\iota_{[0,1]}(\mathbf{v})$.

We utilize the proximal-based algorithm to effectively deal with the unconstrained problem (5) by alternatingly solving each variable, which leads to the following subproblems:

$$
\begin{gather*}
\boldsymbol{\alpha}^{k+1} \in \arg \min _{\boldsymbol{\alpha}}\left\langle\boldsymbol{\alpha}-\boldsymbol{\alpha}^{k}, \partial_{\boldsymbol{\alpha}} H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)\right\rangle+f_{1}(\boldsymbol{\alpha})+\frac{\tau_{1}}{2}\left\|\boldsymbol{\alpha}-\boldsymbol{\alpha}^{k}\right\|_{2}^{2}  \tag{6}\\
\mathbf{e}^{k+1} \in \arg \min _{\mathbf{e}}\left\langle\mathbf{e}-\mathbf{e}^{k}, \partial_{\mathbf{e}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)\right\rangle+f_{2}(\mathbf{e})+\frac{\tau_{2}}{2}\left\|\mathbf{e}-\mathbf{e}^{k}\right\|_{2}^{2}  \tag{7}\\
\mathbf{u}^{k+1} \in \arg \min _{\mathbf{u}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}, \mathbf{v}^{k}\right)+f_{3}(\mathbf{u})+\frac{\delta_{1}}{2}\left\|\mathbf{u}-\mathbf{u}^{k}\right\|_{2}^{2}  \tag{8}\\
\mathbf{v}^{k+1} \in \arg \min _{\mathbf{v}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}\right)+f_{4}(\mathbf{v})+\frac{\delta_{2}}{2}\left\|\mathbf{v}-\mathbf{v}^{k}\right\|_{2}^{2} \tag{9}
\end{gather*}
$$

where the related parameters are all nonnegative, i.e., $\delta_{1}>0, \delta_{2}>0$, and $\tau_{1}=\gamma_{1}\left\|\mathbf{K}^{T} \mathbf{K}\right\|_{F}$, $\tau_{2}=\gamma_{2}\left\|\mathbf{I}_{m}^{T} \mathbf{I}_{m}\right\|_{F}$, here $\mathbf{I}_{m} \in \mathbb{R}^{m \times m}$ is identity matrix, and $\|\cdot\|_{F}$ represents Frobenius norm, $\gamma_{1}>1, \gamma_{2}>1$.

In particular, above subproblems are all convex functions whose closed-form solutions can be easily calculated as follows:

$$
\begin{gather*}
\boldsymbol{\alpha}^{k+1}=\left(\lambda_{1}+\tau_{1}\right)^{-1}\left(\mu \mathbf{K}^{T}\left(\mathbf{y}-\mathbf{K} \boldsymbol{\alpha}^{k}-\mathbf{e}^{k}-\mathbf{u}^{k}\right)+\tau_{1} \boldsymbol{\alpha}^{k}\right)  \tag{10}\\
\mathbf{e}^{k+1}=\left(\lambda_{2}+\tau_{2}\right)^{-1}\left(\mu\left(\mathbf{y}-\mathbf{K} \boldsymbol{\alpha}^{k+1}-\mathbf{e}^{k}-\mathbf{u}^{k}\right)+\tau_{2} \mathbf{e}^{k}\right)  \tag{11}\\
\mathbf{u}^{k+1}=\frac{\mu\left(\mathbf{y}-\mathbf{K} \boldsymbol{\alpha}^{k+1}-\mathbf{e}^{k+1}\right)+\delta_{1} \mathbf{u}^{k}}{\mu+\beta \mathbf{v}^{k} \odot \mathbf{v}^{k}+\delta_{1}}  \tag{12}\\
\mathbf{v}^{k+1}=\min \left(1, \max \left(0, \frac{\mathbf{1}+\delta_{2} \mathbf{v}^{k}}{\beta \mathbf{u}^{k+1} \odot \mathbf{u}^{k+1}+\delta_{2}}\right)\right) \tag{13}
\end{gather*}
$$

We iteratively and alternatingly update $\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}$ and $\mathbf{v}^{k+1}$ according to (10)-(13). The final algorithm for the nonconvex minimization problem (3) is summarized in Algorithm 1.

In Algorithm 1, "Max iter $^{\prime \prime}$ means the maximum iterations and the "rho" represents the relative error between adjacent iterations. When the iteration stops, the final underlying function can be estimated by the relation of (1).

```
Algorithm 1: The algorithm to minimize the problem (5)
    Input: blurred matrix \(K\), positive parameters \(\lambda_{1}, \lambda_{2}, \delta_{1} \delta_{2}\),
        and \(\gamma_{1}>1, \gamma_{2}>1\).
    Initialize: start with any \(\left(\boldsymbol{\alpha}^{0}, \mathbf{e}^{0}, \mathbf{u}^{0}, \mathbf{v}^{0}\right)\), rho \(=1, k=0\).
    While \(k<\) Max \(_{i \text { ter }}\) or \(r\) ho \(>10^{-4}\)
        (1) Compute \(\boldsymbol{\alpha}^{k+1}\) solved by Equation (10).
        (2) Compute \(\mathbf{e}^{k+1}\) by Equation (11).
        (3) Compute \(\mathbf{u}^{k+1}\) by Equation (12).
        (4) Compute \(\mathbf{v}^{k+1}\) by Equation (13).
        (5) Update penalty parameter \(\beta\) by \(\beta^{k+1}=1.1 \beta^{k}\).
        (6) Calculate the relative error \(r h o=\frac{\left\|\mathbf{g}^{k+1}-\mathbf{g}^{k}\right\|_{2}}{\left\|\mathbf{g}^{k}\right\|_{2}}\),
                where \(\mathbf{g}^{k+1}=\mathbf{K} \boldsymbol{\alpha}^{k+1}+\mathbf{e}^{k+1}\), and \(\mathbf{g}^{k}=\mathbf{K} \boldsymbol{\alpha}^{k}+\mathbf{e}^{k}\).
            (7) \(k \leftarrow k+1\).
    Endwhile
    Output: \(\alpha^{*}\), and \(\mathbf{e}^{*}\)
```


## 3. Convergence Analysis

For the sake of notational simplicity, we uniform expression $\|*\|$ as: (1) Frobenius norm $\|*\|_{F}$ if $*$ is a matrix; (2) $\ell_{2}$-norm $\|*\|_{2}$ if $*$ is a vector. Denote that $\mathbf{z}^{k}=\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)$ and its domain field is $\mathbb{R}^{n} \times \mathbb{R}^{n} \times \mathbb{R}^{n} \times \mathbb{R}^{n}$.

Lemma 1. Let the bounded sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ is generated by designed algorithm. Then the sequence $\left\{\Phi\left(\mathbf{z}^{k}\right)\right\}_{k \in \mathbb{N}}$ sufficiently decreases as follows:

$$
\begin{equation*}
\Phi\left(\mathbf{z}^{k}\right)-\Phi\left(\mathbf{z}^{k+1}\right) \geq \frac{\rho_{1}}{2}\left\|\mathbf{z}^{k+1}-\mathbf{z}^{k}\right\|^{2} \tag{14}
\end{equation*}
$$

where $\rho_{1}=\min \left\{\frac{\tau_{1}-L_{1}}{2}, \frac{\tau_{2}-L_{2}}{2}, \frac{\delta_{1}}{2}, \frac{\delta_{2}}{2}\right\}$. Note that $\tau_{1}-L_{1}>0, \tau_{1}-L_{1}>0$, and $L_{1}$ is Lipschitz constant of $\partial_{\alpha} H$ for $\boldsymbol{\alpha}$, and $L_{2}$ is that of $\partial_{\mathbf{e}} H$ for $\mathbf{e}$.

Proof. Since $H(\boldsymbol{\alpha}, \mathbf{e}, \mathbf{u}, \mathbf{v})$ is Lipschitz differential for variables $\boldsymbol{\alpha}$ and $\mathbf{e}$, respectively. There exist positive constants $L_{1}$ and $L_{2}$ which satisfy:

$$
\begin{align*}
H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right) \leq & H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+\left\langle\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}, \partial_{\boldsymbol{\alpha}} H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)\right\rangle \\
& +\frac{L_{1}}{2}\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\|^{2},  \tag{15}\\
H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k}, \mathbf{v}^{k}\right) \leq & H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+\left\langle\mathbf{e}^{k+1}-\mathbf{e}^{k}, \partial_{\mathbf{e}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)\right\rangle \\
& +\frac{L_{2}}{2}\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\|^{2} . \tag{16}
\end{align*}
$$

$\alpha$-subproblem: Based on the designed algorithm, $\boldsymbol{\alpha}^{k+1}$ is the minimum solution of $\alpha$-subproblem in $(k+1)$-th iteration, then we have:

$$
\begin{equation*}
\left\langle\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}, \partial_{\boldsymbol{\alpha}} H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)\right\rangle+f_{1}\left(\boldsymbol{\alpha}^{k+1}\right)+\frac{\tau_{1}}{2}\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\|^{2} \leq f_{1}\left(\boldsymbol{\alpha}^{k}\right) \tag{17}
\end{equation*}
$$

Combine (15) and (17), we have:

$$
\begin{align*}
H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+f_{1}\left(\boldsymbol{\alpha}^{k+1}\right) \leq & H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+f_{1}\left(\boldsymbol{\alpha}^{k}\right) \\
& +\frac{L_{1}-\tau_{1}}{2}\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\|^{2} \tag{18}
\end{align*}
$$

$\mathbf{e}$-subproblem: Based on the designed algorithm, $\mathbf{e}^{k+1}$ is the minimum solution of e-subproblem in $(k+1)$-th iteration, then we have:

$$
\begin{equation*}
\left\langle\mathbf{e}^{k+1}-\mathbf{e}^{k}, \partial_{\mathbf{e}} H\left(\boldsymbol{a}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}\right)\right\rangle+f_{2}\left(\mathbf{e}^{k+1}\right)+\frac{\tau_{2}}{2}\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\|^{2} \leq f_{2}\left(\mathbf{e}^{k}\right) \tag{19}
\end{equation*}
$$

Combine (16) and (19), we have:

$$
\begin{align*}
H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+f_{2}\left(\mathbf{e}^{k+1}\right) \leq & H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+f_{2}\left(\mathbf{e}^{k}\right) \\
& +\frac{L_{2}-\tau_{2}}{2}\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\|^{2} \tag{20}
\end{align*}
$$

$\mathbf{u}$-subproblem: Based on the designed algorithm, $\mathbf{u}^{k+1}$ is the minimum solution of $\mathbf{u}$-subproblem in $(k+1)$-th iteration, then we have:

$$
\begin{align*}
\left.H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)\right\rangle+f_{3}\left(\mathbf{u}^{k}\right) \geq & \left.H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k}\right)\right\rangle+f_{3}\left(\mathbf{u}^{k+1}\right) \\
& +\frac{\delta_{1}}{2}\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|^{2} \tag{21}
\end{align*}
$$

$\mathbf{v}$-subproblem: Based on the designed algorithm, $\mathbf{v}^{k+1}$ is the minimum solution of $\mathbf{v}$-subproblem in $(k+1)$-th iteration, then we have:

$$
\begin{align*}
H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k}\right)+f_{4}\left(\mathbf{v}^{k}\right) \geq & H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)+f_{4}\left(\mathbf{v}^{k+1}\right) \\
& +\frac{\delta_{2}}{2}\left\|\mathbf{v}^{k+1}-\mathbf{v}^{k}\right\|^{2} \tag{22}
\end{align*}
$$

According to (18), (20)-(22), we have:

$$
\begin{align*}
& \Phi\left(\mathbf{z}^{k}\right)-\Phi\left(\mathbf{z}^{k+1}\right) \\
= & H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+f_{1}\left(\boldsymbol{\alpha}^{k}\right)-H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right) \\
& -f_{1}\left(\boldsymbol{\alpha}^{k+1}\right)+H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+f_{2}\left(\mathbf{e}^{k}\right) \\
& -H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)-f_{2}\left(\mathbf{e}^{k+1}\right)+H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k}, \mathbf{v}^{k}\right) \\
& +f_{3}\left(\mathbf{u}^{k}\right)-H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k}\right)-f_{3}\left(\mathbf{u}^{k+1}\right) \\
& +H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k}\right)+f_{4}\left(\mathbf{v}^{k}\right)  \tag{23}\\
& -H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)-f_{4}\left(\mathbf{v}^{k+1}\right) \\
\geq & \frac{\tau_{1}-L_{1}}{2}\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\|^{2}+\frac{\tau_{2}-L_{2}}{2}\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\|^{2} \\
& +\frac{\delta_{1}}{2}\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|^{2}+\frac{\delta_{2}}{2}\left\|\mathbf{v}^{k+1}-\mathbf{v}^{k}\right\|^{2} \\
\geq & \rho_{1}\left\|\mathbf{z}^{k+1}-\mathbf{z}^{k}\right\|^{2},
\end{align*}
$$

where $\rho_{1}=\min \left\{\frac{\tau_{1}-L_{1}}{2}, \frac{\tau_{2}-L_{2}}{2}, \frac{\delta_{1}}{2}, \frac{\delta_{2}}{2}\right\}$. Thus, the consequence of Lemma 1 obviously holds.

Lemma 2. Let the bounded sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ be generated by the designed algorithm. Then:

$$
\begin{equation*}
\left\|\mathbf{d}^{k+1}\right\| \leq \rho_{2}\left\|\mathbf{z}^{k+1}-\mathbf{z}^{k}\right\| \tag{24}
\end{equation*}
$$

where $\mathbf{d}^{k+1} \in \partial \Phi\left(\mathbf{z}^{k+1}\right)$, and $\rho_{2}=4 \max \left\{\mu\left\|\boldsymbol{K}^{T} \mathbf{K}\right\|+\tau_{1}, \mu\|\mathbf{K}\|+\mu+\tau_{2}, \mu\|\mathbf{K}\|+\mu+\delta_{1}\right.$, $\left.2 \beta M^{2}+\delta_{2}\right\}$, which the contant $M>0$ is bounded-value of the sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ and $\partial$ is subdifferential operator.

Proof. Obviously, $\alpha^{k+1}$ satisfies the first-order optimal condition of $\alpha$-subproblem because of $\boldsymbol{\alpha}^{k+1}$ is the $k$-th solution of the $\alpha$-subproblem. Similar to the other variables, then we have:

$$
\begin{align*}
& 0 \in \partial_{\boldsymbol{\alpha}} H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+\partial f_{1}\left(\boldsymbol{\alpha}^{k+1}\right)+\tau_{1}\left(\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right) \\
& 0 \in \partial_{\mathbf{e}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)+\partial f_{2}\left(\mathbf{e}^{k+1}\right)+\tau_{2}\left(\mathbf{e}^{k+1}-\mathbf{e}^{k}\right),  \tag{25}\\
& 0 \in \partial_{\mathbf{u}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k}\right)+\partial f_{3}\left(\mathbf{u}^{k+1}\right)+\delta_{1}\left(\mathbf{u}^{k+1}-\mathbf{u}^{k}\right), \\
& 0 \in \partial_{\mathbf{v}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)+\partial f_{4}\left(\mathbf{v}^{k+1}\right)+\delta_{2}\left(\mathbf{v}^{k+1}-\mathbf{v}^{k}\right) .
\end{align*}
$$

Since the object function $\Phi$ is continuous and differentiable for each variable, we have:

$$
\begin{align*}
& \partial_{\boldsymbol{\alpha}} \Phi\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)=\partial_{\alpha} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)+\partial f_{1}\left(\boldsymbol{\alpha}^{k+1}\right), \\
& \partial_{\mathbf{e}} \Phi\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)=\partial_{\mathbf{e}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)+\partial f_{2}\left(\mathbf{e}^{k+1}\right), \\
& \partial_{\mathbf{u}} \Phi\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)=\partial_{\mathbf{u}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)+\partial f_{3}\left(\mathbf{u}^{k+1}\right),  \tag{26}\\
& \partial_{\mathbf{v}} \Phi\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)=\partial_{\mathbf{v}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)+\partial f_{4}\left(\mathbf{v}^{k+1}\right) .
\end{align*}
$$

As the sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ is bounded, there exists $M>0$ such that $\left\|\mathbf{z}^{k}\right\|<M$ holds. Then, combine (25) and (26), we have:

$$
\begin{align*}
& \left\|\mathbf{d}^{k+1}\right\| \\
\leq & \left\|\partial_{\alpha} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)-\partial_{\alpha} H\left(\boldsymbol{\alpha}^{k}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)-\tau_{1}\left(\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right)\right\| \\
& +\left\|\partial_{\mathbf{e}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)-\partial_{\mathbf{e}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k}, \mathbf{u}^{k}, \mathbf{v}^{k}\right)-\tau_{2}\left(\mathbf{e}^{k+1}-\mathbf{e}^{k}\right)\right\| \\
& +\left\|\partial_{\mathbf{u}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k+1}\right)-\partial_{\mathbf{u}} H\left(\boldsymbol{\alpha}^{k+1}, \mathbf{e}^{k+1}, \mathbf{u}^{k+1}, \mathbf{v}^{k}\right)-\delta_{1}\left(\mathbf{u}^{k+1}-\mathbf{u}^{k}\right)\right\| \\
& +\left\|-\delta_{2}\left(\mathbf{v}^{k+1}-\mathbf{v}^{k}\right)\right\| \\
\leq & \left\|\mu \mathbf{K}^{T}\left(\mathbf{K} \boldsymbol{\alpha}^{k+1}+\mathbf{e}^{k+1}+\mathbf{u}^{k+1}\right)-\mu \mathbf{K}^{T}\left(\mathbf{K} \boldsymbol{\alpha}^{k}+\mathbf{e}^{k}+\mathbf{u}^{k}\right)-\tau_{1}\left(\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right)\right\| \\
& +\left\|\mu\left(\mathbf{K} \boldsymbol{\alpha}^{k+1}+\mathbf{e}^{k+1}+\mathbf{u}^{k+1}\right)-\mu\left(\mathbf{K} \boldsymbol{\alpha}^{k+1}+\mathbf{e}^{k}+\mathbf{u}^{k}\right)-\tau_{2}\left(\mathbf{e}^{k+1}-\mathbf{e}^{k}\right)\right\| \\
& +\left\|\beta \mathbf{v}^{k+1} \odot \mathbf{v}^{k+1} \odot \mathbf{u}^{k+1}-\beta \mathbf{v}^{k} \odot \mathbf{v}^{k} \odot \mathbf{u}^{k+1}-\delta_{1}\left(\mathbf{u}^{k+1}-\mathbf{u}^{k}\right)\right\| \\
& +\left\|-\delta_{2}\left(\mathbf{v}^{k+1}-\mathbf{v}^{k}\right)\right\|  \tag{27}\\
\leq & \mu\left\|\mathbf{K}^{T} \mathbf{K}\right\|\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\|+\mu\|\mathbf{K}\|\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\|+\mu\|\mathbf{K}\|\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\| \\
& +\tau_{1}\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\| \\
& +\mu\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\|+\mu\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|+\tau_{2}\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\| \\
& +\beta\left\|\mathbf{v}^{k+1} \odot \mathbf{v}^{k+1} \odot \mathbf{u}^{k+1}-\mathbf{v}^{k} \odot \mathbf{v}^{k} \odot \mathbf{u}^{k+1}\right\| \\
& +\delta_{1}\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|+\delta_{2}\left\|\mathbf{v}^{k+1}-\mathbf{v}^{k}\right\| \\
\leq & \left(\mu\left\|\mathbf{K}^{T} \mathbf{K}\right\|+\tau_{1}\right)\left\|\boldsymbol{\alpha}^{k+1}-\boldsymbol{\alpha}^{k}\right\|+\left(\mu\|\mathbf{K}\|+\mu+\tau_{2}\right)\left\|\mathbf{e}^{k+1}-\mathbf{e}^{k}\right\| \\
& +\left(\mu\|\mathbf{K}\|+\mu+\delta_{1}\right)\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|+\left(2 \beta M^{2}+\delta_{2}\right)\left\|\mathbf{v}^{k+1}-\mathbf{v}^{k}\right\| \\
\leq & \rho_{2}\left\|\mathbf{z}^{k+1}-\mathbf{z}^{k}\right\|,
\end{align*}
$$

where $\rho_{2}=4 \max \left\{\mu\left\|\mathbf{K}^{T} \mathbf{K}\right\|+\tau_{1}, \mu\|\mathbf{K}\|+\mu+\tau_{2}, \mu\|\mathbf{K}\|+\mu+\delta_{1}, 2 \beta M^{2}+\delta_{2}\right\}$. Thus, Equation (24) holds.

Lemma 3. Let the bounded sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ be generated by designed algorithm and the initial variable $\mathbf{z}^{0}$ be bounded. Then the sequence $\left\{\Phi\left(\mathbf{z}^{k}\right)\right\}_{k \in \mathbb{N}}$ is bounded.

Proof. Obviously, the continuous function $\Phi$ is proper and coercive, since there exists $\Phi \rightarrow \infty$ if and only if $\|z\| \rightarrow \infty$. Thus, the sequence $\Phi\left(\mathbf{z}^{k}\right)$ is bounded because the sequence $\left\{\mathbf{z}^{k}\right\}$ generated by the designed algorithm is bounded.

Lemma 4. The function $\Phi(\mathbf{z})$ in Equation (3) is a Kurdyka-Łojasiewicz (KŁ) function (The definition of a KŁ function and some examples can be found in [18]).

Proof. According to the definition and some examples for the KŁ function in [18,25], $f_{1}, f_{2}, f_{3}, f_{4}, H$ are polynomial functions which are obviously real analytic functions. Thus, $\Phi(\mathbf{z})$ is a $K \notin$ function.

Theorem 1. Let the sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ be bounded which is generated by the designed algorithm. Then the sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ converges to a critical point.

Proof. Since the sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ is bounded, so it must exist a subsequence $\left\{\mathbf{z}^{k_{q}}\right\}_{q \in \mathbb{N}}$ which converges to a critical point $\tilde{\mathbf{z}}$ satisfying $\lim _{q \rightarrow \infty} \mathbf{z}^{k_{q}}=\tilde{\mathbf{z}}$. Since the function $\Phi(\mathbf{z})$ is continuous, then the sequence $\left\{\Phi\left(\mathbf{z}^{k_{q}}\right)\right\}_{q \in \mathbb{N}}$ is converged to $\Phi(\tilde{\mathbf{z}})$, i.e., $\lim _{q \rightarrow \infty} \Phi\left(\mathbf{z}^{k_{q}}\right)=$ $\Phi(\tilde{\mathbf{z}})$.

According to Lemmas 1-3, it means that the sequence $\left\{\Phi\left(\mathbf{z}^{k}\right)\right\}_{k \in \mathbb{N}}$ is also converged. Thus, the sequence $\left\{\Phi\left(\mathbf{z}^{k}\right)\right\}_{k \in \mathbb{N}}$ and the subsequence $\left\{\Phi\left(\mathbf{z}^{k_{q}}\right)\right\}_{q \in \mathbb{N}}$ will converge to the same function value as follows:

$$
\begin{equation*}
\lim _{k \rightarrow \infty} \Phi\left(\mathbf{z}^{k}\right)=\Phi(\tilde{\mathbf{z}}) \tag{28}
\end{equation*}
$$

If there is an index $\tilde{k}$ such that $\Phi\left(\mathbf{z}^{\tilde{k}}\right)=\Phi(\tilde{\mathbf{z}})$. It is obviously that the sequence $\left\{\mathbf{z}^{k}\right\}_{k \in \mathbb{N}}$ up to a stationary point and the corresponding function value does also not change based on Lemma 1, i.e., $\Phi\left(\mathbf{z}^{\tilde{k}+1}\right)=\Phi\left(\mathbf{z}^{\tilde{k}}\right)$. Thus, it is the critical point of $\left\{\mathbf{z}^{k}\right\}$ and the conclusion of Theorem 1 is obviously established.

Next part will prove that Theorem 1 still holds when the index $\tilde{k}$ is nonexistent.
Based on Lemma 1, it implies that for any $k>0$, we have:

$$
\begin{equation*}
\Phi(\tilde{\mathbf{z}})<\Phi\left(\mathbf{z}^{k}\right) \tag{29}
\end{equation*}
$$

From (28), it implies that for any $\eta>0$, there exists $k_{1} \in \mathbb{N}$, so we have the following inequation when $k>k_{1}$,

$$
\begin{equation*}
\Phi\left(\mathbf{z}^{k}\right)<\Phi(\tilde{\mathbf{z}})+\eta \tag{30}
\end{equation*}
$$

Denote a set of limit points of the sequence $\left\{\mathbf{z}^{k}\right\}$ as $\vartheta\left(\mathbf{z}^{0}\right)$, and let $\operatorname{dist}\left(\tilde{\mathbf{z}}, \vartheta\left(\mathbf{z}^{0}\right)\right)$ express the minimum distance between one point $\tilde{\mathbf{z}}$ and a set $\vartheta\left(\mathbf{z}^{0}\right)$, i.e.,

$$
\operatorname{dist}\left(\mathbf{z}^{k}, \vartheta\left(\mathbf{z}^{0}\right)\right) \triangleq \min \left\{\left\|\mathbf{z}^{k}-\widehat{\mathbf{z}}\right\|: \widehat{\mathbf{z}} \in \vartheta\left(\mathbf{z}^{0}\right)\right\}
$$

Based on Lemma 2, $\Phi(\mathbf{z})$ is a continuous and differential function, and we have $0 \in \partial \Phi(\tilde{\mathbf{z}})$ since $\left\|\mathbf{d}^{k+1}\right\| \rightarrow 0$ as $k \rightarrow \infty$. It implies that for any $\epsilon>0$, there exists a positive index $k_{2} \in \mathbb{N}$, when $k>k_{2}$, then:

$$
\begin{equation*}
\operatorname{dist}\left(\mathbf{z}^{k}, \vartheta\left(\mathbf{z}^{0}\right)\right)<\epsilon \tag{31}
\end{equation*}
$$

Set $l=\max \left\{k_{1}, k_{2}\right\}$, and for each $k>l$, there exists:

$$
\begin{equation*}
\Omega \triangleq\left\{\mathbf{z}: \operatorname{dist}\left(\mathbf{z}^{k}, \vartheta\left(\mathbf{z}^{0}\right)\right)<\epsilon\right\} \cap\{\mathbf{z}: \Phi(\tilde{\mathbf{z}})<\Phi(\mathbf{z})<\Phi(\tilde{\mathbf{z}})+\eta\} \tag{32}
\end{equation*}
$$

From the Lemma $4, \Phi(\mathbf{z})$ is a $K \notin$ function in domain $\Omega$. Thus, there is a concave function $\varphi$ such that:

$$
\begin{equation*}
\varphi^{\prime}\left(\Phi\left(\mathbf{z}^{k}\right)-\Phi(\tilde{\mathbf{z}})\right) \operatorname{dist}\left(0, \partial \Phi\left(\mathbf{z}^{k}\right)\right) \geq 1 \tag{33}
\end{equation*}
$$

Moreover, based on the concavity of $\varphi$, it has:

$$
\varphi\left(\Phi\left(\mathbf{z}^{k+1}\right)-\Phi(\tilde{\mathbf{z}})\right) \leq \varphi\left(\Phi\left(\mathbf{z}^{k}\right)-\Phi(\tilde{\mathbf{z}})\right)+\varphi^{\prime}\left(\Phi\left(\mathbf{z}^{k}\right)-\Phi(\tilde{\mathbf{z}})\right)\left(\Phi\left(\mathbf{z}^{k+1}\right)-\Phi\left(\mathbf{z}^{k}\right)\right)
$$

Denote $\Theta_{p, q}=\varphi\left(\Phi\left(\mathbf{z}^{p}\right)-\Phi(\tilde{\mathbf{z}})\right)-\varphi\left(\Phi\left(\mathbf{z}^{q}\right)-\Phi(\tilde{\mathbf{z}})\right)$ for all nonnegative integers $p$, and $q \in \mathbb{N}$. Combine Equation (33), Lemmas 1 and 2, we have:

$$
\begin{align*}
\Theta_{k, k+1} & \geq \varphi^{\prime}\left(\Phi\left(\mathbf{z}^{k}\right)-\Phi(\tilde{\mathbf{z}})\right)\left(\Phi\left(\mathbf{z}^{k}\right)-\Phi\left(\mathbf{z}^{k+1}\right)\right) \\
& \geq \varphi^{\prime}\left(\Phi\left(\mathbf{z}^{k}\right)-\Phi(\tilde{\mathbf{z}})\right) \rho_{1}\left\|\mathbf{z}^{k}-\mathbf{z}^{k+1}\right\|^{2}  \tag{34}\\
& \geq \frac{1}{\rho_{2}}\left\|\mathbf{z}^{k}-\mathbf{z}^{k-1}\right\|^{-1} \rho_{1}\left\|\mathbf{z}^{k}-\mathbf{z}^{k+1}\right\|^{2}
\end{align*}
$$

which can be rewritten as follows:

$$
\begin{equation*}
\left\|\mathbf{z}^{k}-\mathbf{z}^{k+1}\right\|^{2} \leq \frac{\rho_{2}}{\rho_{1}} \Theta_{k, k+1}\left\|\mathbf{z}^{k}-\mathbf{z}^{k-1}\right\| \tag{35}
\end{equation*}
$$

It is a well-known inequality $2 \sqrt{\alpha \beta} \leq \alpha+\beta$ for any $\alpha, \beta>0$, thus, we have:

$$
\begin{equation*}
2\left\|\mathbf{z}^{k}-\mathbf{z}^{k+1}\right\| \leq 2 \sqrt{\frac{\rho_{2}}{\rho_{1}} \Theta_{k, k+1}\left\|\mathbf{z}^{k}-\mathbf{z}^{k-1}\right\|} \leq\left\|\mathbf{z}^{k}-\mathbf{z}^{k-1}\right\|+\frac{\rho_{2}}{\rho_{1}} \Theta_{k, k+1} \tag{36}
\end{equation*}
$$

Taking summation to (36) for $i=l+1, \cdots, k$, it has the following inequation:

$$
\begin{align*}
2 \sum_{i=l+1}^{k}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\| & \leq \sum_{i=l+1}^{k}\left\|\mathbf{z}^{i}-\mathbf{z}^{i-1}\right\|+\frac{\rho_{2}}{\rho_{1}} \sum_{i=l+1}^{k} \Theta_{i, i+1}  \tag{37}\\
& \leq \sum_{i=l+1}^{k}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\|+\left\|\mathbf{z}^{l+1}-\mathbf{z}^{l}\right\|+\frac{\rho_{2}}{\rho_{1}} \Theta_{l+1, k+1}
\end{align*}
$$

that is:

$$
\begin{equation*}
\sum_{i=l+1}^{k}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\| \leq\left\|\mathbf{z}^{l+1}-\mathbf{z}^{l}\right\|+\frac{\rho_{2}}{\rho_{1}} \Theta_{l+1, k+1} \tag{38}
\end{equation*}
$$

According to the definition of $\Theta$, we have $\lim _{k \rightarrow \infty} \Theta_{l+1, k+1}=\varphi\left(\Phi\left(\mathbf{z}^{l+1}\right)-\Phi(\tilde{\mathbf{z}})\right)$. Thus,

$$
\begin{align*}
\lim _{k \rightarrow \infty} \sum_{i=1}^{k}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\| & =\sum_{i=1}^{l}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\|+\lim _{k \rightarrow \infty} \sum_{i=l+1}^{k}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\| \\
& \leq \sum_{i=1}^{l}\left\|\mathbf{z}^{i}-\mathbf{z}^{i+1}\right\|+\left\|\mathbf{z}^{l+1}-\mathbf{z}^{l}\right\|+\frac{\rho_{2}}{\rho_{1}} \varphi\left(\Phi\left(\mathbf{z}^{l+1}\right)-\Phi(\tilde{\mathbf{z}})\right)  \tag{39}\\
& <+\infty
\end{align*}
$$

Thus, $\left\{\mathbf{z}^{k}\right\}$ is a Cauchy sequence, and has a finite length. Because of the completeness of Hilbert space, the Cauchy sequence $\left\{\mathbf{z}^{k}\right\}$ is also certainly a convergence sequence. Thus, the sequence $\left\{\mathbf{z}^{k}\right\}$ generated by designed algorithm converges to a critical point $\mathbf{z}^{*}=\left(\boldsymbol{\alpha}^{*}, \mathbf{e}^{*}, \mathbf{u}^{*}, \mathbf{v}^{*}\right)$. Moreover, the convergence of the sequence generated by the designed algorithm can be guaranteed for any initial value.

## 4. Numerical Result

In this section, we conduct some simple simulation examples to show the effectiveness of the proposed method. We choose $f$ as the ground-truth function (discrete form as $\mathbf{f}$ ) and add Gaussian noise and sparse outliers to generate the observation. The proposed method is compared with the kernel-based regression using orthogonal matching pursuit
(KROMP) method [8], and the parameters of KROMP method are selected according to the range mentioned in the literature. The parameters of the proposed method in the experiment set empirically to: $\mu=10, \lambda_{1} \in\left[10^{-3}, 10^{2}\right], \lambda_{2} \in\left[10^{-3}, 10\right], \beta=0.01, \gamma_{1}=1.1$, $\gamma_{2}=10 \gamma_{1}$, and it should be noted that better visual and numerical results can be obtained by fine-tuning the parameters more carefully.

The related error (ReErr) for the quantitative evaluation, which is a commonly index to measure the effect of restoration, and it is defined as:

$$
\begin{equation*}
\operatorname{ReErr}=\frac{\|\mathbf{f}-\mathbf{g}\|_{2}}{\|\mathbf{f}\|_{2}} \tag{40}
\end{equation*}
$$

where $\mathbf{g}$ is the restoration result which is estimated by different methods. Experiments are implemented in MATLAB (R2016a) on a desktop with 16Gb RAM and Inter(R) Core(TM) CPU i5-4590: @3.30GHz.

Example 1. The binary continuous function $f$ is a given as follows:

$$
\begin{equation*}
f(x, y)=50 \operatorname{sinc}\left(\pi \sqrt{x^{2}+y^{2}}\right) \tag{41}
\end{equation*}
$$

where $x \in[-1,1], y \in[-1,1]$ (two dimensions, respectively, take 21 discrete points), $\operatorname{sinc}(x)=$ $\frac{\sin (\pi x)}{\pi x}$. After discretization of $f, 20 d B$ Gaussian noise and $10 \%$ outlier noise were added to obtain the final degraded data.

In order to show the experimenta results of different examples, we will show the ground-truth data, the degraded data polluted by noise and outliers, and the restored outcome calculated by the proposed method and the KROMP method, respectively.

Example 2. The binary continuous function $f$ is a given as follows:

$$
\begin{equation*}
f(x, y)=\frac{\sin \sqrt{x^{2}+y^{2}}}{\sqrt{x^{2}+y^{2}}} \tag{42}
\end{equation*}
$$

where $x \in[-8,8], y \in[-8,8]$ (two dimensions, respectively, take 31 discrete points), $\operatorname{sinc}(x)=$ $\frac{\sin (\pi x)}{\pi x}$. After discretization of $f$ which is a binary continuous function, 10 dB Gaussian noise and $10 \%$ outlier noise were added to obtain the final degraded data.

From Figures 1 and 2, although the shape of ground-truth in Example 2 is similar to that of Example 1 (see Figures 1a and 2a), in fact, the degree of degradation of Example 2 is much greater than that of Example 1, which the function value and the degree of noise pollution are different (see Figures 1b and 2b). The proposed method has obvious restorated outcomes, and can effectively recover original data (see Figure 1c, however the restorated outcomes of the KROMP method still have obvious noise residual in Figures 1d and 2d. It also shows the effectiveness of the proposed method.

Example 3. The binary continuous function $f$ is a given as follows

$$
\begin{equation*}
f(x, y)=x \exp \left(-x^{2}-y^{2}\right) \tag{43}
\end{equation*}
$$

where $x \in[-2,2], y \in[-2,2]$ (two dimensions, respectively, take 21 discrete points). After discretization of $f$ which is a binary continuous function, 10 dB Gaussian noise and $5 \%$ outlier noise were added to obtain the final degraded data.

It can be seen that even in the case of extremely large external pollution, such as Example 3 in Figure 3b, the proposed method can obtain more accurate recovery data
(Figure 3c) than the KROMP method (Figure 3d). In addition, the relative error (ReErr) results from Example 1 to Example 3 are shown in Table 1, and the better results have been bolded. It is obvious that the proposed method has smaller relative errors compared with the KROMP method, which verifies the effectiveness of the proposed method.

Table 1. The ReErr results from Example 1 to Example 3 (Bold: the best).

|  | KROMP | Proposed |
| :---: | :---: | :---: |
| Example 1 | 0.128 | $\mathbf{0 . 1 1 6}$ |
| Example 2 | 0.257 | $\mathbf{0 . 1 5 9}$ |
| Example 3 | 0.178 | $\mathbf{0 . 1 1 7}$ |



Figure 1. In Example 1, the ground-truth $\mathbf{f}$ was degraded by 20 dB Gaussian noise and $10 \%$ outliers, and the visual results of restoration by each method were obtained. (a) Ground-truth; (b) degraded; (c) the restored outcome by the proposed method, (d) the restored outcome by the KROMP method.


Figure 2. Cont.


Figure 2. In Example 2, the real data f was degraded by 10 dB Gaussian noise and $10 \%$ outliers, and the visual results of restoration by each method were obtained. (a) Ground-truth; (b) degraded; (c) the restored outcome by the proposed method, (d) the restored outcome by the KROMP method.


Figure 3. In Example 3, the real data $\mathbf{f}$ was degraded by 10 dB Gaussian noise and $5 \%$ outliers, and the visual results of restoration by each method were obtained. (a) Ground-truth; (b) degraded; (c) the restored outcome by the proposed method, (d) the restored outcome by the KROMP method.

## 5. Conclusions

In this paper, we proposed a new nonconvex modeling to deal with a challenging symmetry ill-posed problem. An efficient algorithm for solving the given nonconvex problem is designed by MPEC and the proximal-based regularization. Theoretically, the bounded sequence generated by the designed algorithm can be guaranteed to converge to the nonconvex problem's local optimal solution. Furthermore, the convergence of the designed algorithm can be guaranteed for any initial value. The numerical experiments show that the proposed method can achieve better restoration results. For example, our method could obtain smaller relative error comparing with benchmark KROMP approach, besides, could interpolate more mesh points and significantly reduce noise.

Author Contributions: Conceptualization, software and writing-original draft preparation, H.-X.D. and L.-J.D.; methodology and validation, formal analysis and writing-review and editing, H.-X.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by Scientific Research Startup Fund of Xihua University (RZ2000002862).
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: Data in the study is generated by ourself.
Acknowledgments: The authors thank the reviewers for their comments, which have improved the content of this manuscripts.
Conflicts of Interest: The authors declare no conflict of interest.

## References

1. Tanaka, K. Generation of point sets by convex optimization for interpolation in reproducing kernel Hilbert spaces. Numer. Algorithms 2020, 84, 1049-1079. [CrossRef]
2. Mo, Y.; Qian, T.; Mai, W.X.; Chen, Q.H. The AFD methods to compute Hilbert transform. Appl. Math. Lett. 2015, 45, 18-24. [CrossRef]
3. Karvonen, T.; Särkkü, S.; Tanaka, K. Kernel-based interpolation at approximate Fekete points. Numer. Algorithms 2020, 84, 1049-1079.
4. Silalahi, D.D.; Midi, H.; Arasan, J.; Mustafa, M.S.; Caliman, J.P. Kernel Partial Least Square Regression with High Resistance to Multiple Outliers and Bad Leverage Points on Near-Infrared Spectral Data Analysis. Symmetry 2021, 13, 547. [CrossRef]
5. Deng, L.J.; Guo, W.H.; Huang, T.Z. Single-image super-resolution via an iterative reproducing kernel Hilbert space method. IEEE Trans. Circuits Syst. Video Technol. 2015, 26, 2001-2014. [CrossRef]
6. Li, X.Y.; Wu, B.Y. A new reproducing kernel collocation method for nonlocal fractional boundary value problems with non-smooth solutions. Appl. Math. Lett. 2018, 86, 194-199. [CrossRef]
7. Wu, Q.; Li, Y.; Xue, W. A Kernel Recursive Maximum Versoria-Like Criterion Algorithm for Nonlinear Channel Equalization. Symmetry 2019, 11, 1067. [CrossRef]
8. Papageorgiou, G.; Bouboulis, P.; Theodoridis, S. Robust kernel-based regression using orthogonal matching pursuit. In Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing, Southampton, UK, 22-25 September 2013; pp. 1-6.
9. Chen, S.S.; Donoho, D.L.; Saunders, M.A. Atomic decomposition by basis pursuit. SIAM Rev. 2001, 43, 129-159. [CrossRef]
10. Donoho, D.L. For most large underdetermined systems of linear equations the minimial $\ell_{1}$-norm solution is also the sparsest solution. Commun. Pure Appl. Math. 2006, 59, 797-829. [CrossRef]
11. Dong, B.; Zhang, Y. An efficient algorithm for $\ell_{0}$ minimization in wavelet frame based image restoration. J. Sci. Comput. 2013, 54, 350-368. [CrossRef]
12. Zuo, W.M.; Meng, D.Y.; Zhang, L.; Feng, X.C.; Zhang, D. A generalized iterated shrinkage algorithm for non-convex sparse coding. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1-8 December 2013; pp. 217-224.
13. Ye, J.J.; Zhu, D. New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches. SIAM J. Optim. 2010, 20, 1885-1905. [CrossRef]
14. Yuan, G.Z.; Ghanem, B. $\ell_{0}$ TV: A sparse optimization method for impulse noise image restoration. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 41, 352-364. [CrossRef]
15. Yuan, G.Z.; Ghanem, B. An exact penalty method for binary optimization based on MPEC formulation. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4-9 February 2017.
16. Attouch, H.; Bolte, J.; Redont, P.and Soubeyran, A. Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka-Lojasiewicz inequality. Math. Oper. Res. 2010, 35, 435-457. [CrossRef]
17. Wang, Y.J.; Dong, M.M.; Xu, Y. A sparse rank-1 approximation algorithm for high-order tensors. Appl. Math. Lett. 2020, 102, 106-140. [CrossRef]
18. Bolte, J.; Sabach, S.; Teboulle, M. Proximal alternating linearized minimization or nonconvex and nonsmooth problems. Math. Program. 2014, 146, 459-494. [CrossRef]
19. Sun, T.; Barrio, R.; Jiang, H.; Cheng, L.Z. Convergence rates of accelerated proximal gradient algorithms under independent noise. Numer. Algorithms 2019, 81, 631-654. [CrossRef]
20. Hu, W.; Zheng, W.; Yu, G. A Unified Proximity Algorithm with Adaptive Penalty for Nuclear Norm Minimization. Symmetry 2019, 11, 1277. [CrossRef]
21. An, Y.; Zhang, Y.; Guo, H.; Wang, J. Compressive Sensing Based Three-Dimensional Imaging Method with Electro-Optic Modulation for Nonscanning Laser Radar. Symmetry 2020, 12, 748. [CrossRef]
22. Ma, F. Convergence study on the proximal alternating direction method with larger step size. Numer. Algorithms 2020, 85, 399-425. [CrossRef]
23. Pham, Q.M.; Lachmund, D.; Hào, D.N. Convergence of proximal algorithms with stepsize controls for non-linear inverse problems and application to sparse non-negative matrix factorization. Numer. Algorithms 2020, 85, 1255-1279. [CrossRef]
24. Tiddeman, B.; Ghahremani, M. Principal Component Wavelet Networks for Solving Linear Inverse Problems. Symmetry 2021, 13, 1083. [CrossRef]
25. Xu, Y.Y.; Yin, W.T. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. SIAM J. Imaging Sci. 2013, 6, 1758-1789. [CrossRef]
