Mathematical and Computational Applications

MDPI

Article

On the $O(1/n)$ Convergence Rate of the Auxiliary Problem Principle for Separable Convex Programming and Its Application to the Power Systems Multi-Area Economic Dispatch Problem

Yaming Ren [1,*] and Zhongxian Chen [2]

1 School of Electrical and Information Engineering , Anhui University of Technology, Maanshan 243002, China
2 State Grid Hefei Power Supply Company, Hefei 230000, China; chenzhognxianseu@sina.com
* Correspondence: renyaming1981@gmail.com; Tel.: +86-0555-2316689

Academic Editor: Mehmet Ali Ilgın
Received: 16 April 2016; Accepted: 25 July 2016; Published: 29 July 2016

Abstract: The auxiliary problem principle has been widely applied in power systems to solve the multi-area economic dispatch problem. Although the effectiveness and correctness of the auxiliary problem principle method have been demonstrated in relevant literatures, the aspect connected with accurate estimate of its convergence rate has not yet been established. In this paper, we prove the $O(1/n)$ convergence rate of the auxiliary problem principle method.

Keywords: auxiliary problem principle; variational inequality; convergence rate

1. Introduction

The auxiliary problem principle (APP) [1], originally proposed by G. Cohen in [2], has a wide range of applications in the power systems field [3–8]. In fact, the mathematical formulation of multi-area economic dispatch problem can be expressed as follows.

$$\min \ \{f(x_1) + g(x_2) \,|\, Ax_1 + Bx_2 = b, x_1 \in \Omega_1, x_2 \in \Omega_2\} \tag{1}$$

where $f : R^m \to R$ and $g : R^n \to R$ are convex function. $\Omega_1 \subseteq R^m$ and $\Omega_2 \subseteq R^n$ are closed convex sets. $A \in R^{r\times m}$ and $B \in R^{r\times n}$ are given fixed matrices (not necessarily full rank). $b \in R^r$ is given constant.

For solving (1), the corresponding APP iterative scheme can be expressed as follows.

$$\begin{aligned} x_1^{k+1} = \arg\ \min \Big\{ & f(x_1) + \tfrac{\beta}{2}\|Ax_1\|^2 - \beta \left\langle Ax_1, Ax_1^k \right\rangle \\ & + \left\langle -\lambda^k + c(Ax_1^k + Bx_2^k - b), Ax_1 \right\rangle | x_1 \in \Omega_1 \Big\} \end{aligned} \tag{2}$$

$$\begin{aligned} x_2^{k+1} = \arg\ \min \Big\{ & g(x_2) + \tfrac{\beta}{2}\|Bx_2\|^2 - \beta \left\langle Bx_2, Bx_2^k \right\rangle \\ & + \left\langle -\lambda^k + c(Ax_1^k + Bx_2^k - b), Bx_2 \right\rangle | x_2 \in \Omega_2 \Big\} \end{aligned} \tag{3}$$

$$\lambda^{k+1} = \lambda^k - c\left(Ax_1^{k+1} + Bx_2^{k+1} - b\right) \tag{4}$$

where $\lambda \in R^r$ is the Lagrangian multiplier for the linear constraint $Ax_1 + Bx_2 = 0$ and $c > 0$ is a given fixed penalty parameter. $\langle\cdot,\cdot\rangle$ denotes the inner product, i.e., $\langle x, x\rangle = x^T x$. The superscript k denotes iteration index. $\beta > 2c$ is given fixed auxiliary problem principle parameter [7].

Although the APP iterative scheme is known to be an efficient approach for the convex problem with separable operators [9], the theoretical analysis of its convergence rate has not been established and applied in the literature.

In 2004, Nemirovski gave a proof to show that prox-type method has the $O(1/n)$ convergence rate for variational inequalities with Lipschitz continuous monotone operators, where n denotes the iteration number [10]. Then, for the same problem, the $O(1/n)$ convergence rate of the projection and contraction method was proved in [11]. Inspired by these literatures, taking advantage of the variational inequality approach, the accurate estimate of alternating direction method's convergence rate has made considerable headway in recent years. To be more exact, in 2012, Bingsheng He's analysis indicated that the Douglas-Rachford alternating direction method has the $O(1/n)$ convergence rate [12]. After that, in 2014, Yuan Shen and Minghua Xu studied the $O(1/n)$ convergence rate of Ye-Yuan's modified alternating direction method of multipliers [13].

In this paper, our aim is to investigate the convergence rate of the iterative scheme APP under the framework of variational inequality. In fact, problem (1) is equivalent to solving the following variational inequality (VI) problem: Find (x_1, x_2, λ) such that

$$f\left(x_1'\right) - f\left(x_1\right) + \left(x_1' - x_1\right)\left(-A^T\lambda\right) \geq 0 \quad \forall x_1' \in \Omega_1 \tag{5}$$

$$g\left(x_2'\right) - g\left(x_2\right) + \left(x_2' - x_2\right)\left(-B^T\lambda\right) \geq 0 \quad \forall x_2' \in \Omega_2 \tag{6}$$

$$\left(\lambda' - \lambda\right)\left(Ax_1 + Bx_2 - b\right) \geq 0 \quad \forall \lambda' \in R^r \tag{7}$$

Then, the compact form of (5)–(7) can be expressed as follows.

$$\theta\left(u'\right) - \theta\left(u\right) + \left(w' - w\right)^T F\left(w\right) \geq 0, \quad \forall w' \in W \tag{8}$$

where

$$u = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad w = \begin{pmatrix} u \\ \lambda \end{pmatrix}, \quad F\left(w\right) = \begin{pmatrix} -A^T\lambda \\ -B^T\lambda \\ Ax_1 + Bx_2 - b \end{pmatrix} \tag{9}$$

$$W = \Omega_1 \times \Omega_2 \times R^r, \ \theta\left(u\right) = f\left(x_1\right) + g\left(x_2\right) \tag{10}$$

and the mapping $F\left(w\right)$ is monotone.

2. The Convergence Analysis of APP

In this section, we give a convergence analysis of iterative scheme APP under the framework of variational inequality. Meanwhile, the analysis is useful for the accurate estimate of APP's convergence rate in thr next section. Throughout this paper, we assume the solution set of VI problem (8) is nonempty and denoted by W^*. w^* denotes an arbitrary (but fixed) point in the solution set W^*.

Lemma 1. *A single iteration of APP*

$$\begin{aligned} x_1^{k+1} = \arg\ \min \Big\{ & f\left(x_1\right) + \tfrac{\beta}{2}\|Ax_1\|^2 - \beta\left\langle Ax_1, Ax_1^k\right\rangle \\ & + \left\langle -\lambda^k + c(Ax_1^k + Bx_2^k - b), Ax_1\right\rangle | x_1 \in \Omega_1 \Big\} \end{aligned} \tag{11}$$

$$\begin{aligned} x_2^{k+1} = \arg\ \min \Big\{ & g\left(x_2\right) + \tfrac{\beta}{2}\|Bx_2\|^2 - \beta\left\langle Bx_2, Bx_2^k\right\rangle \\ & + \left\langle -\lambda^k + c(Ax_1^k + Bx_2^k - b), Bx_2\right\rangle | x_2 \in \Omega_2 \Big\} \end{aligned} \tag{12}$$

is equivalent to

$$
\begin{aligned}
x_1^{k+1} = \arg\ \min \Big\{ & f(x_1) + \tfrac{\beta-c}{2} \left\| Ax_1 - Ax_1^k \right\|^2 \\
& + \tfrac{c}{2} \left\| Ax_1 + Bx_2^k - b \right\|^2 + \left\langle -\lambda^k, Ax_1 \right\rangle | x_1 \in \Omega_1 \Big\}
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
x_2^{k+1} = \arg\ \min \Big\{ & g(x_2) + \tfrac{\beta-c}{2} \left\| Bx_2 - Bx_2^k \right\|^2 \\
& + \tfrac{c}{2} \left\| Ax_1^k + Bx_2 - b \right\|^2 + \left\langle -\lambda^k, Bx_2 \right\rangle | x_2 \in \Omega_2 \Big\}
\end{aligned}
\tag{14}
$$

Proof of Lemma 1. Adding a quadratic term $\frac{\beta}{2} \left\| Ax_1^k \right\|^2$ to the objective function (11) without changing its optimization result, then (11) can be expressed as follows.

$$
\begin{aligned}
x_1^{k+1} = \arg\ \min \Big\{ & f(x_1) + \tfrac{\beta}{2} \left\| Ax_1 - Ax_1^k \right\|^2 \\
& + \left\langle -\lambda^k + c(Ax_1^k + Bx_2^k - b), Ax_1 \right\rangle | x_1 \in \Omega_1 \Big\}
\end{aligned}
\tag{15}
$$

Considering the following equation

$$
\begin{aligned}
\left\langle c(Ax_1^k + Bx_2^k - b), Ax_1 \right\rangle &= c \left\langle Ax_1^k - Ax_1, Ax_1 \right\rangle + c \left\langle Ax_1 + Bx_2^k - b, Ax_1 \right\rangle \\
&= \tfrac{c}{2} \left(\left\| Ax_1 + Bx_2^k - b \right\|^2 - \left\| Ax_1 - Ax_1^k \right\|^2 \right) \\
&\quad + \tfrac{c}{2} \left(\left\| Ax_1^k \right\|^2 - \left\| Bx_2^k - b \right\|^2 \right)
\end{aligned}
\tag{16}
$$

Then, combing (15) and (16), we obtain

$$
\begin{aligned}
x_1^{k+1} = \arg\ \min \Big\{ & f(x_1) + \tfrac{\beta-c}{2} \left\| Ax_1 - Ax_1^k \right\|^2 + \tfrac{c}{2} \left\| Ax_1 + Bx_2^k - b \right\|^2 \\
& + \left\langle -\lambda^k, Ax_1 \right\rangle + \tfrac{c}{2} \left(\left\| Ax_1^k \right\|^2 - \left\| Bx_2^k - b \right\|^2 \right) | x_1 \in \Omega_1 \Big\}
\end{aligned}
\tag{17}
$$

Removing the constant term $\frac{c}{2} \left(\left\| Ax_1^k \right\|^2 - \left\| Bx_2^k - b \right\|^2 \right)$, we get

$$
\begin{aligned}
x_1^{k+1} = \arg\ \min \Big\{ & f(x_1) + \tfrac{\beta-c}{2} \left\| Ax_1 - Ax_1^k \right\|^2 + \tfrac{c}{2} \left\| Ax_1 + Bx_2^k - b \right\|^2 \\
& + \left\langle -\lambda^k, Ax_1 \right\rangle | x_1 \in \Omega_1 \Big\}
\end{aligned}
\tag{18}
$$

Analogously, we have

$$
\begin{aligned}
x_2^{k+1} = \arg\ \min \Big\{ & g(x_2) + \tfrac{\beta-c}{2} \left\| Bx_2 - Bx_2^k \right\|^2 + \tfrac{c}{2} \left\| Ax_1^k + Bx_2 - b \right\|^2 \\
& + \left\langle -\lambda^k, Bx_2 \right\rangle | x_2 \in \Omega_2 \Big\}
\end{aligned}
\tag{19}
$$

as we wanted to prove. Thus Lemma 1 is proved. □

Lemma 2. *Let sequence* $\left\{ w^k \right\}$ *is generated by the iterative scheme APP. We denote* $\|x\|_M = x^T M x$ *and* $\|x\| = x^T x$, *then we get*

$$
\left\| w^k - w^* \right\|_M^2 - \left\| w^{k+1} - w^* \right\|_M^2 \geq \left\| w^k - w^{k+1} \right\|_M^2 \tag{20}
$$

$$f\left(x_1\right)+g\left(x_2\right)-f\left(x_1^{k+1}\right)-g\left(x_2^{k+1}\right) \\ +\left(w-w^{k+1}\right)^T\left\{F\left(w^{k+1}\right)+M\left(w^{k+1}-w^k\right)\right\}\geq 0,\ \forall w\in W \tag{21}$$

where

$$M=\begin{pmatrix} (\beta-c)\,A^TA & -cA^TB & 0 \\ -cB^TA & (\beta-c)\,B^TB & 0 \\ 0 & 0 & \frac{1}{c}I_m \end{pmatrix},\ \beta>2c \tag{22}$$

Proof of Lemma 2. According to the description of Lemma 1 and using variational inequality approach, solving (11) and (12) is equivalent to solving $\left(x_1^{k+1},x_2^{k+1}\right)$ which satisfies following inequalities,

$$f\left(x_1\right)-f\left(x_1^{k+1}\right)+\left(x_1-x_1^{k+1}\right)^T \\ \left\{(\beta-c)\,A^T\left(Ax_1-Ax_1^k\right)+cA^T\left(Ax_1+Bx_2^k-b\right)-A^T\lambda^k\right\}\geq 0,\ \forall x_1\in\Omega_1 \tag{23}$$

$$g\left(x_2\right)-g\left(x_2^{k+1}\right)+\left(x_2-x_2^{k+1}\right)^T \\ \left\{(\beta-c)\,B^T\left(Bx_2-Bx_2^k\right)+cB^T\left(Ax_1^k+Bx_2-b\right)-B^T\lambda^k\right\}\geq 0,\ \forall x_2\in\Omega_2 \tag{24}$$

Considering

$$\lambda^{k+1}=\lambda^k-c\left(Ax_1^{k+1}+Bx_2^{k+1}-b\right) \tag{25}$$

Thus, the following result is given by utilizing (23)–(25)

$$f\left(x_1\right)+g\left(x_2\right)-f\left(x_1^{k+1}\right)-g\left(x_2^{k+1}\right) \\ +\left(w-w^{k+1}\right)^T\left\{F\left(w^{k+1}\right)+M\left(w^{k+1}-w^k\right)\right\}\geq 0,\ \forall w\in W \tag{26}$$

Setting $w=w^*$ in (26), we get

$$\left(w^*-w^{k+1}\right)^T M\left(w^{k+1}-w^k\right) \\ \geq f\left(x_1^{k+1}\right)+g\left(x_2^{k+1}\right)-f\left(x_1^*\right)-g\left(x_2^*\right)+\left(w^{k+1}-w^*\right)^T F\left(w^{k+1}\right) \tag{27}$$

Mapping F is monotone, we have

$$\left(w^{k+1}-w^*\right)^T F\left(w^{k+1}\right)\geq\left(w^{k+1}-w^*\right)^T F\left(w^*\right) \tag{28}$$

According to (8), we get

$$\left(w^{k+1}-w^*\right)^T F\left(w^*\right)\geq 0 \tag{29}$$

Combing (27)–(29), we get

$$\begin{aligned} &\left(w^*-w^{k+1}\right)^T M\left(w^{k+1}-w^k\right) \\ &\geq f\left(x_1^{k+1}\right)+g\left(x_2^{k+1}\right)-f\left(x_1^*\right)-g\left(x_2^*\right)+\left(w^{k+1}-w^*\right)^T F\left(w^{k+1}\right) \\ &\geq f\left(x_1^{k+1}\right)+g\left(x_2^{k+1}\right)-f\left(x_1^*\right)-g\left(x_2^*\right)+\left(w^{k+1}-w^*\right)^T F\left(w^*\right) \\ &\geq 0 \\ &\Rightarrow\left(w^*-w^k+w^k-w^{k+1}\right)^T M\left(w^{k+1}-w^k\right)\geq 0 \\ &\Rightarrow\left(w^*-w^k\right)^T M\left(w^{k+1}-w^k\right)\geq\left(w^k-w^{k+1}\right)^T M\left(w^k-w^{k+1}\right) \end{aligned} \tag{30}$$

Using (30), we obtain

$$\begin{aligned}
&\left\|w^k - w^*\right\|_M^2 - \left\|w^{k+1} - w^*\right\|_M^2 \\
&= \left\|w^k - w^*\right\|_M^2 - \left\|w^k - w^* - (w^k - w^{k+1})\right\|_M^2 \\
&= 2\left(w^* - w^k\right)^T M\left(w^{k+1} - w^k\right) - \left\|w^k - w^{k+1}\right\|_M^2 \\
&\geq 2\left\|w^k - w^{k+1}\right\|_M^2 - \left\|w^k - w^{k+1}\right\|_M^2 \\
&= \left\|w^k - w^{k+1}\right\|_M^2
\end{aligned} \tag{31}$$

Based on the above discussion, the proof of Lemma 2 is completed. □

If matrices A and B are full rank, for $\forall\ w \in W$, we can get $\|w\|_M = w^T M w = (\beta - 2c)\left(\|Ax_1\|^2 + \|Bx_2\|^2\right) + c\|Ax_1 - Bx_2\|^2 + \frac{1}{c}\|\lambda\|^2 \geq 0$, and the equality hold up if and only if $w = 0$. It is clear that matrix M is positive definite matrix and (20) is Fejér monotone. We get

$$\lim_{k\to\infty} w^k = w^* \tag{32}$$

Furthermore, for general matrices A and B, (20) can be rewritten as follows.

$$\left\|v^k - v^*\right\|_N^2 - \left\|v^{k+1} - v^*\right\|_N^2 \geq \left\|v^k - v^{k+1}\right\|_N^2 \tag{33}$$

where

$$v = \begin{pmatrix} Ax_1 \\ Bx_2 \\ \lambda \end{pmatrix},\ N = \begin{pmatrix} (\beta - c)\, I & -cI & 0 \\ -cI & (\beta - c)\, I & 0 \\ 0 & 0 & \frac{1}{c} I_m \end{pmatrix} \tag{34}$$

It is clear that (33) is Fejér monotone, so we get

$$\lim_{k\to\infty}\left\|Ax_1^{k+1} - Ax_1^k\right\| = 0,\ \lim_{k\to\infty}\left\|Bx_2^{k+1} - Bx_2^k\right\| = 0,\ \lim_{k\to\infty}\left\|\lambda^{k+1} - \lambda^k\right\| = 0 \tag{35}$$

Lemma 3. *Let sequence $\left\{w^k\right\}$ is generated by the iterative scheme APP. If*

$$\left\|Ax_1^{k+1} - Ax_1^k\right\| = 0,\ \left\|Bx_2^{k+1} - Bx_2^k\right\| = 0,\ \left\|\lambda^{k+1} - \lambda^k\right\| = 0 \tag{36}$$

then, w^{k+1} is the solution of VI problem (8).

Proof of Lemma 3. According to [14], solving (8) is equivalent to finding a zero point of e(w).

$$e(w) = \begin{bmatrix} e_{x_1}(w) \\ e_{x_2}(w) \\ e_\lambda(w) \end{bmatrix} = \begin{bmatrix} x_1 - P_{\Omega 1}\left\{x_1 - \left[\nabla f(x_1) - A^T\lambda\right]\right\} \\ x_2 - P_{\Omega 2}\left\{x_2 - \left[\nabla g(x_2) - B^T\lambda\right]\right\} \\ Ax_1 + Bx_2 - b \end{bmatrix} \tag{37}$$

where $P_\Omega(\cdot)$ denotes the projection on Ω. $\nabla f(\cdot)$ denotes the gradient of $f(\cdot)$.

Based on the iterative scheme APP and the projection equation, we obtain

$$\begin{aligned}
x_1^{k+1} = P_{\Omega 1}\Big\{x_1^{k+1} - \Big[&\nabla f\left(x_1^{k+1}\right) - A^T\left(\lambda^k - c\left(Ax_1^{k+1} + Bx_2^k - b\right)\right) \\
&+ (\beta - c)\, A^T\left(Ax_1^{k+1} - Ax_1^k\right)\Big]\Big\}
\end{aligned} \tag{38}$$

$$x_2^{k+1} = P_{\Omega 2}\left\{x_2^{k+1} - \left[\nabla g\left(x_2^{k+1}\right) - B^T\left(\lambda^k - c\left(Ax_1^k + Bx_2^{k+1} - b\right)\right) + (\beta - c)\, B^T\left(Bx_2^{k+1} - Bx_2^k\right)\right]\right\} \tag{39}$$

Recall (37), we get,

$$e\left(w^{k+1}\right) = \begin{bmatrix} e_{x_1}\left(w^{k+1}\right) \\ e_{x_2}\left(w^{k+1}\right) \\ e_{\lambda}\left(w^{k+1}\right) \end{bmatrix} = \begin{bmatrix} x_1^{k+1} - P_{\Omega 1}\left\{x_1^{k+1} - \left[\nabla f\left(x_1^{k+1}\right) - A^T\lambda^{k+1}\right]\right\} \\ x_2^{k+1} - P_{\Omega 2}\left\{x_2^{k+1} - \left[\nabla g\left(x_2^{k+1}\right) - B^T\lambda^{k+1}\right]\right\} \\ Ax_1^{k+1} + Bx_2^{k+1} - b \end{bmatrix} \tag{40}$$

and hence,

$$\left\|e\left(w^{k+1}\right)\right\| \le \left\|e_{x_1}\left(w^{k+1}\right)\right\| + \left\|e_{x_2}\left(w^{k+1}\right)\right\| + \left\|e_{\lambda}\left(w^{k+1}\right)\right\| \tag{41}$$

Replacing the first x_1^{k+1} in $e_{x_1}\left(w^{k+1}\right)$ by (38) and using

$$\|P_{\Omega}(x) - P_{\Omega}(y)\| \le \|x - y\| \tag{42}$$

We get

$$\begin{aligned} \left\|e_{x_1}\left(w^{k+1}\right)\right\| &= \left\|x_1^{k+1} - P_{\Omega 1}\left\{x_1^{k+1} - \left[\nabla f\left(x_1^{k+1}\right) - A^T\lambda^{k+1}\right]\right\}\right\| \\ &\le \Big\|A^T\left\{\left(\lambda^k - \lambda^{k+1}\right) - c\left(Ax_1^{k+1} + Bx_2^k - b\right)\right\} \\ &\quad - (\beta - c)\, A^T\left(Ax_1^{k+1} - Ax_1^k\right)\Big\| \\ &\le \left\|A^T\left\{\left(\lambda^k - \lambda^{k+1}\right) - c\left(Ax_1^{k+1} + Bx_2^{k+1} - b\right) + c\left(Bx_2^{k+1} - Bx_2^k\right)\right\}\right\| \\ &\quad + \left\|(\beta - c)\, A^T\left(Ax_1^{k+1} - Ax_1^k\right)\right\| \\ &\le \|A^Tc\|\left\|Bx_2^{k+1} - Bx_2^k\right\| + \|(\beta - c)\, A^T\|\left\|\left(Ax_1^{k+1} - Ax_1^k\right)\right\| \end{aligned} \tag{43}$$

Similarly, replacing the first x_2^{k+1} in $e_{x_2}\left(w^{k+1}\right)$ by (39) and using (42), we get

$$\begin{aligned} \left\|e_{x_2}\left(w^{k+1}\right)\right\| &= \left\|x_2^{k+1} - P_{\Omega 2}\left\{x_2^{k+1} - \left[\nabla g\left(x_2^{k+1}\right) - B^T\lambda^{k+1}\right]\right\}\right\| \\ &\le \Big\|B^T\left\{\left(\lambda^k - \lambda^{k+1}\right) - c\left(Ax_1^k + Bx_2^{k+1} - b\right)\right\} \\ &\quad - (\beta - c)\, B^T\left(Bx_2^{k+1} - Bx_2^k\right)\Big\| \\ &\le \left\|B^T\left\{\left(\lambda^k - \lambda^{k+1}\right) - c\left(Ax_1^{k+1} + Bx_2^{k+1} - b\right) + c\left(Ax_1^{k+1} - Ax_1^k\right)\right\}\right\| \\ &\quad + \left\|(\beta - c)\, B^T\left(Bx_2^{k+1} - Bx_2^k\right)\right\| \\ &= \left\|B^Tc\left(Ax_1^{k+1} - Ax_1^k\right)\right\| + \left\|(\beta - c)\, B^T\left(Bx_2^{k+1} - Bx_2^k\right)\right\| \\ &\le \|B^Tc\|\left\|Ax_1^{k+1} - Ax_1^k\right\| + \|(\beta - c)\, B^T\|\left\|Bx_2^{k+1} - Bx_2^k\right\| \end{aligned} \tag{44}$$

Combining (41), (43) and (44), we obtain

$$\begin{aligned} \left\|e\left(w^{k+1}\right)\right\| &\le \left\|e_{x_1}\left(w^{k+1}\right)\right\| + \left\|e_{x_2}\left(w^{k+1}\right)\right\| + \left\|e_{\lambda}\left(w^{k+1}\right)\right\| \\ &\le \left(\|A^Tc\| + \|(\beta - c)\, B^T\|\right)\left\|Bx_2^{k+1} - Bx_2^k\right\| \\ &\quad + \left(\|(\beta - c)\, A^T\| + \|B^Tc\|\right)\left\|Ax_1^{k+1} - Ax_1^k\right\| + \left\|Ax_1^{k+1} + Bx_2^{k+1} - b\right\| \end{aligned} \tag{45}$$

and using

$$\lambda^{k+1} = \lambda^k - c\left(Ax_1^{k+1} + Bx_2^{k+1} - b\right) \tag{46}$$

Hence, we need only prove that if w^{k+1} satisfies

$$\lim_{k\to\infty}\left\|Ax_1^{k+1}-Ax_1^k\right\|=0,\ \lim_{k\to\infty}\left\|Bx_2^{k+1}-Bx_2^k\right\|=0,\ \lim_{k\to\infty}\left\|\lambda^{k+1}-\lambda^k\right\|=0 \tag{47}$$

then, w^{k+1} is the solution of problem (8).

Therefore, the proof of lemma 3 is completed. □

3. The Convergence Rate Analysis of APP

In this section, we first introduce Lemma 4 which is originally described as Theorem 2.1 in [12]. Lemma 4 provides a basic property for the solution set of VI problem.

Lemma 4. *The solution set of VI problem is convex and can be characterized as,*

$$W^*=\bigcap_{w\in W}\left\{\tilde{w}\in W:\theta(u)-\theta(\tilde{u})+(w-\tilde{w})^T F(w)\geq 0\right\} \tag{48}$$

Lemma 4 demonstrates, for $\varepsilon=O(1/n)$, if there is a point $\tilde{w}\in W$ satisfying

$$\theta(\tilde{u})-\theta(u)+(\tilde{w}-w)^T F(w)\leq\varepsilon,\forall w\in W \tag{49}$$

then, iterative scheme APP has $O(1/n)$ convergence rate.

Lemma 5. *Let sequence $\left\{w^k\right\}$ be generated by APP algorithm, we get*

$$\theta(u)-\theta\left(u^{k+1}\right)+\left(w-w^{k+1}\right)^T F(w)+\frac{1}{2}\left\|w-w^k\right\|_M^2\geq\frac{1}{2}\left\|w-w^{k+1}\right\|_M^2,\forall w\in W \tag{50}$$

Proof of Lemma 5. Using the following equation [12],

$$(a-b)^T H(c-d)=\frac{1}{2}\left(\|a-d\|_H^2-\|a-c\|_H^2\right)+\frac{1}{2}\left(\|c-b\|_H^2-\|d-b\|_H^2\right) \tag{51}$$

where H is a symmetric and positive semidefinite matrix.

Here, setting $a=w,\ b=w^{k+1},\ c=w^k,\ d=w^{k+1}$, we get

$$\begin{aligned}
&\left(w-w^{k+1}\right)^T M\left(w^k-w^{k+1}\right)\\
&=\tfrac{1}{2}\left(\left\|w-w^{k+1}\right\|_M^2-\left\|w-w^k\right\|_M^2\right)+\tfrac{1}{2}\left(\left\|w^k-w^{k+1}\right\|_M^2-\left\|w^{k+1}-w^{k+1}\right\|_M^2\right)\\
&=\tfrac{1}{2}\left(\left\|w-w^{k+1}\right\|_M^2-\left\|w-w^k\right\|_M^2\right)+\tfrac{1}{2}\left\|w^k-w^{k+1}\right\|_M^2\\
&\geq\tfrac{1}{2}\left(\left\|w-w^{k+1}\right\|_M^2-\left\|w-w^k\right\|_M^2\right)
\end{aligned} \tag{52}$$

Combining Lemma 2 and (52), we obtain

$$\theta(u)-\theta\left(u^{k+1}\right)+\left(w-w^{k+1}\right)^T F\left(w^{k+1}\right)+\frac{1}{2}\left\|w-w^k\right\|_M^2\geq\frac{1}{2}\left\|w-w^{k+1}\right\|_M^2 \tag{53}$$

Based on the above discussion, the proof of Lemma 5 is completed. □

Lemma 6. *Let $\left\{w^k\right\}$ be generated by APP algorithm. For any integer $n>0$,*

$$\theta(\tilde{u}_n)-\theta(u)+(\tilde{w}_n-w)^T F(w)\leq\frac{1}{2(n+1)}\left\|w-w^0\right\|_M^2,\forall w\in W \tag{54}$$

where $\tilde{w}_n=\frac{1}{n+1}\sum_{k=0}^n w^{k+1}$, $\tilde{u}_n=\frac{1}{n+1}\sum_{k=0}^n u^{k+1}$, n is the iteration number, w^0 denotes the initial point.

Proof of Lemma 6. According to lemma 5 , we sum the inequality (50) over $k = 0, 1, \cdots, n$, we obtain

$$\begin{array}{l}\sum\limits_{k=0}^{n} \left(\theta(u) - \theta\left(u^{k+1}\right)\right) + \left((n+1)w - \sum\limits_{k=0}^{n} w^{k+1}\right)^T F(w) + \frac{1}{2}\left\|w - w^0\right\|_M^2 \geq \\ \frac{1}{2}\left\|w - w^{k+1}\right\|_M^2, \forall w \in W\end{array} \tag{55}$$

(55) can be rewritten as,

$$\frac{\left\|w - w^0\right\|_M^2}{2(n+1)} \geq \frac{1}{n+1}\sum_{k=0}^{n} \theta\left(u^{k+1}\right) - \theta(u) + \left(\frac{1}{n+1}\sum_{k=0}^{n} w^{k+1} - w\right)^T F(w), \forall w \in W \tag{56}$$

Because

$$\theta(u) = f(x_1) + \mathrm{g}(\mathrm{x}_2) \tag{57}$$

and $f(x_1), g(x_2)$ are convex functions, we have,

$$\theta(\tilde{u}_n) \leq \frac{1}{n+1}\sum_{k=0}^{n} \theta\left(u^{k+1}\right) \tag{58}$$

Combining (56) and (58), we obtain,

$$\frac{\left\|w - w^0\right\|_M^2}{2(n+1)} \geq \theta(\tilde{u}_n) - \theta(u) + (\tilde{w}_n - w)^T F(w), \forall w \in W \tag{59}$$

Based on above discussion, the proof of Lemma 6 is completed. □

According to Lemmas 4 and 5, it is found that iterative scheme APP has $O(1/n)$ convergence rate in an ergodic sense.

4. Numerical Experiments

In this section, we present the 40-unit test system to show the efficiency of the auxiliary problem principle. To be exact, the test system consists of two areas (area 1 and area 2). There are 25 units and 15 units in area 1 and area 2 respectively. The corresponding mathematical formulation can be expressed as follows.

$$\min \ \{f(x_1) + g(x_2) \,|\, Ax_1 + Bx_2 = b, x_1 \in \Omega_1, x_2 \in \Omega_2\} \tag{60}$$

where

$$A = \left(\underbrace{0, \cdots, 0}_{25}, 1\right), B = \left(\underbrace{0, \cdots, 0}_{15}, -1\right), b = 0 \tag{61}$$

$$x_1 = (P_1, P_2, \cdots, P_{25}, P_{b1})^T, \ x_2 = (P_{26}, P_{27}, \cdots, P_{40}, P_{b2})^T \tag{62}$$

$$f(x_1) = \sum_{i=1}^{25}\left(a_i^2 P_i^2 + b_i P_i + c_i\right), \ \ g(x_2) = \sum_{i=26}^{40}\left(a_i^2 P_i^2 + b_i P_i + c_i\right) \tag{63}$$

$$\Omega_1 = \left\{x_1 \left|\sum_{i=1}^{25} P_i + P_{b1} = 8000;\ P_{i,\min} \leq P_i \leq P_{i,\max},\ 1 \leq i \leq 25;\ |P_{b1}| \leq 800\right.\right\} \tag{64}$$

$$\Omega_2 = \left\{x_2 \left|\sum_{i=26}^{40} P_i - P_{b2} = 2000;\ P_{i,\min} \leq P_i \leq P_{i,\max},\ 26 \leq i \leq 40;\ |P_{b2}| \leq 800\right.\right\} \tag{65}$$

P_i is the active output of unit i in this test system. Both P_{b1} and P_{b2} denote transfer power flow between two areas. $P_{i,\min}$, $P_{i,\max}$ are given variable upper and lower limits, and a_i, b_i, c_i are given fixed parameters for objective function as shown in Table 1 [15].

Table 1. Data for 40-Unit test sysem.

i	$P_{i,min}$	$P_{i,max}$	a_i	b_i	c_i
1	40	80	0.03073	8.336	170.44
2	60	120	0.02028	7.0706	309.54
3	80	190	0.00942	8.1817	369.03
4	24	42	0.08482	6.9467	135.48
5	26	42	0.09693	6.5595	135.19
6	68	140	0.01142	8.0543	222.33
7	110	300	0.00357	8.0323	287.71
8	135	300	0.00492	6.999	391.98
9	135	300	0.00573	6.602	455.76
10	130	300	0.00605	12.908	722.82
11	94	375	0.00515	12.986	635.2
12	94	375	0.00569	12.796	654.69
13	125	500	0.00421	12.501	913.4
14	125	500	0.00752	8.8412	1760.4
15	125	500	0.00708	9.1575	1728.3
16	125	500	0.00708	9.1575	1728.3
17	125	500	0.00708	9.1575	1728.3
18	220	500	0.00313	7.9691	647.85
19	220	500	0.00313	7.955	649.69
20	242	550	0.00313	7.9691	647.83
21	242	550	0.00313	7.9691	647.81
22	254	550	0.00298	6.6313	785.96
23	254	550	0.00298	6.6313	785.96
24	254	550	0.00298	6.6313	785.53
25	254	550	0.00298	6.6313	785.53
26	254	550	0.00277	7.1032	801.32
27	254	550	0.00277	7.1032	801.32
28	10	150	0.52124	3.3353	1055.1
29	10	150	0.52124	3.3353	1055.1
30	10	150	0.52124	3.3353	1055.1
31	20	70	0.25098	13.052	1207.8
32	20	70	0.16766	21.887	810.79
33	20	70	0.2635	10.244	1247.7
34	20	70	0.30575	8.3707	1219.2
35	18	60	0.18362	26.258	641.43
36	18	60	0.32563	9.6956	1112.8
37	20	60	0.33722	7.1633	1044.4
38	25	60	0.23915	16.339	832.24
39	25	60	0.23915	16.339	834.24
40	25	60	0.23915	16.339	1035.2

APP algorithm is employed to solve the problem. Here, parameters are selected as penalty parameter $c = 0.01$ and auxiliary problem principle parameter $\beta = 0.03$. Stop criterion is set to be

$$\max\left(\left\|Ax_1^{k+1} - Ax_1^k\right\|, \left\|Bx_2^{k+1} - Bx_2^k\right\|, \left\|\lambda^{k+1} - \lambda^k\right\|\right) \leq 10^{-4} \tag{66}$$

Figures 1 and 2 reflect the convergence characteristic of objective function and stop criterion for this test system, respectively. It is clear that objective function is converged to the optimal solution and the stop criterion is very close to zero when the number of iterations reaches 20. The effectiveness and correctness of the auxiliary problem principle have been demonstrated.

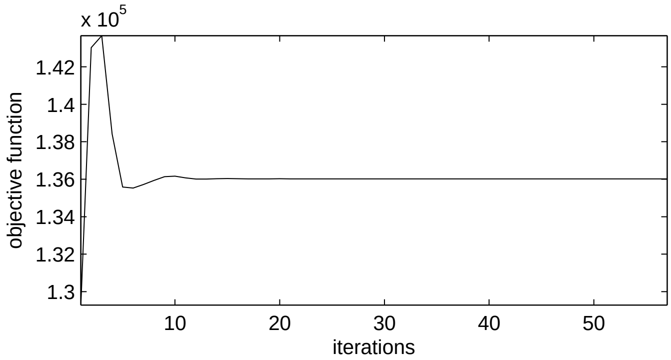

Figure 1. Convergence characteristic of objective function.

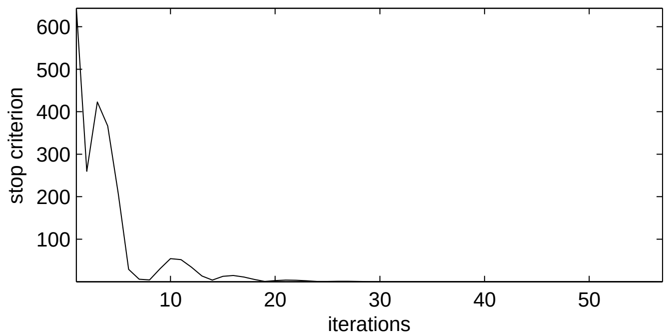

Figure 2. Convergence characteristic of stop criterion.

5. Conclusions

In this paper, taking advantage of special characterization of variational inequality solution set, we derive the $O(1/n)$ convergence rate of the auxiliary problem principle.

Acknowledgments: The authors would like to thank the reviewers for their valuable comments and suggestions to improve the present work.

Author Contributions: This research was carried out in collaboration among all two authors. Yaming Ren designed the algorithm, analyzed the data and wrote the paper. Zhongxian Chen obtained the numerical experiments data and carried out the experiments. All authors have read and approved the final manuscript.

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

References

1. Cohen, G.; Zhu., D.L. Decomposition coordination methods in large scale optimization problems: The non-differentiable case and the use of augmented lagrangians. *Adv. Large Scale Syst.* **1984**, *1*, 203–266.
2. Cohen, G. Auxiliary problem principle and decomposition of optimization problems. *J. Optim. Theory Appl.* **1980**, *32*, 277–305.
3. Jiang, Q.Y.; Zhou, B.R.; Zhang, M.Z. Parallel augment lagrangian relaxation method for transient stability constrained unit commitment. *IEEE Trans. Power Syst.* **2013**, *28*, 1140–1148.
4. Chung, K.H.; Kim, B.H.; Hur, D. Distributed implementation of generation scheduling algorithm on interconnected power systems. *Energy Convers. Manag.* **2011**, *52*, 3457–3464.
5. Liu, K.; Li, Y.; Sheng, W. The decomposition and computation method for distributed optimal power flow based on message passing interface (MPI). *Int. J. Electr. Power Energy Syst.* **2011**, *33*, 1185–1193.
6. Kim, B.H.; Baldick, R. A comparison of distributed optimal power flow algorithms. *IEEE Trans. Power Syst.* **2000**, *15*, 599–604.
7. Kim, B.H.; Baldick, R. Coarse-grained distributed optimal power flow. *IEEE Trans. Power Syst.* **1997**, *12*, 932–939.

8. Batut, J.; Renaud, A. Daily generation scheduling optimization with transmission constraints: A new class of algorithms. *IEEE Trans. Power Syst.* **1992**, *7*, 982–989.
9. Beltran, C.; Heredia, F.J. Unit commitment by augmented lagrangian relaxation: Testing two decomposition approaches. *J. Optim. Theory Appl.* **2002**, *112*, 295–314.
10. Nemirovski, A. Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM J. Optim.* **2004**, *15*, 229–251.
11. Cai, X.J.; Gu, G.Y.; He, B.S. On the $O(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators. *Comput. Optim. Appl.* **2014**, *57*, 339–363.
12. He, B.; Yuan, X. On the $O(1/n)$ convergence rate of the douglas-rachford alternating direction method. *SIAM J. Numer. Anal.* **2012**, *50*, 700–709.
13. Shen, Y.; Xu, M.H. On the $O(1/t)$ convergence rate of Ye-Yuan's modified alternating direction method of multipliers. *Appl. Math. Comput.* **2014**, *226*, 367–373.
14. He, B.S.; Yang, H.; Wang, S.L. Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. *J. Optim. Theory Appl.* **2000**, *106*, 337–356.
15. Chen, P.H.; Chang, H.C. Large-scale economic-dispatch by genetic algorithm. *IEEE Trans. Power Syst.* **1995**, *10*, 1919–1926.