Chapter 7

APPLICATIONS TO EXTENDED LINEAR-QUADRATIC PROGRAMMING

In this chapter, forward-backward splitting methods are applied to problems of extended linear-quadratic programming.

7.1. Introduction

A problem of *extended linear-quadratic programming* can be described abstractly as consisting of the minimization of a piecewise linear-quadratic convex function subject to a system of linear constraints. Mathematically, a *standard form* for a problem in extended linear-quadratic programming, as we described in (2.2.11) of Section 2.2, is

$$\underset{u \in U}{\text{minimize}} \quad f(u) = \left\{ p^T u + \tfrac{1}{2} u^T P u + \rho_{V,Q}(q - Du) \right\}, \tag{7.1.1}$$

where $P \in \mathbb{R}^{k \times k}$, $Q \in \mathbb{R}^{l \times l}$ are symmetric and positive semidefinite, U and V are polyhedral in $\mathbb{R}^k$ and $\mathbb{R}^l$ respectively, $\rho_{V,Q}$ is the monitoring function defined in (2.2.2).

Problems and theory of extended linear-quadratic programming have been recently introduced by Rockafellar and co-workers in describing various large-scale problems in stochastic programming and multistage optimization. The first users for the concept are Rockafellar and Wets in their papers of two-stage stochastic programming in 1986 and 1987 ([296], [297]). It was developed further in Rockafellar [270], [271], [273], [274], [275], Rockafellar and Wets [295], Zhu and Rockafellar [331], and so on. These problems have received more and more attention because not only they are more flexible in modeling and application in optimization than problems of conventional quadratic programming, but also the exact constraints that are present, besides the constraints formulated with penalty terms, may be quite simple.

Problem (7.1.1) includes many problems of optimization as special cases such as problems of linear programming, quadratic programming, piecewise linear programming and piecewise quadratic programming. Discrete-time problems of optimal control, both deterministic and stochastic (i.e., multistage stochastic programming) were also analyzed as extended linear-quadratic programming problems in Rockafellar and Wets [295] and shown to have a remarkable property of Lagrangian decomposability in the primal and dual arguments, both of which can be high dimensional. These models raise new computational challenges and possibilities.

From the Lagrangian point of view, a problem of extended linear-quadratic programming is to find a saddle point $(\bar{u}, \bar{v})$ of a function

$$L(u,v) = p^T u + \tfrac{1}{2}u^T Pu + q^T v - \tfrac{1}{2}v^T Qv - v^T Du \quad \text{on} \quad U \times V, \tag{7.1.2}$$

where P, Q, D, U, V etc. are the same as those in Problem (7.1.1). Problem (7.1.1) is actually the primal problem associated with L, U and V. The corresponding dual problem is similarly

$$\underset{v \in V}{\text{maximize}} \quad g(v) = \left\{ q^T v - \tfrac{1}{2}v^T Qv - \rho_{U,P}(D^T v - p) \right\}. \tag{7.1.3}$$

A main kind of numerical methods for the general saddle point problem of extended linear-quadratic programming is called *finite-envelope methods*, which use duality in generating envelope approximations to the primal and dual objective functions through a finite sequence of separate minimizations or maximizations of the Lagrangian. These methods are originally introduced by Rockafellar and Wets (1986, 1987). Rockafellar (1990, 1991) set the theory and foundation for numerical schemes for these methods in large-scale extended linear-quadratic programming. Zhu and Rockafellar (1992) combined finite-envelope representation with nonlinear gradient projection to obtain some alternative algorithms called *primal-dual projected gradient algorithms.* All of these methods do use the decomposability of the problem. However, they do not decompose the problem at each iteration into a set of low dimension subproblems that could be solved simultaneously, unlike the forward-backward splitting methods.

We call the *fully quadratic* case of (7.1.1) and (7.1.3) when the matrices are actually positive definite.

Theorem 2.2.2 in Section 2.2 gives the duality and optimality of Problem (7.1.1) and Problem (7.1.3), and Theorem 2.2.1 exploited some properties of the monitoring functions. While Theorem 3.1.1 in Section 3.1 reviewed some properties of the objective functions f and g in (7.1.1) and (7.1.3).

An important case that helps to clarify the formulation of Problem (7.1.1) and the motivations behind the parallel computation through the EFBS method is the *box-diagonal case.* Then P and Q are diagonal matrices,

$$P = \mathrm{diag}[\alpha_1, \cdots, \alpha_k], \qquad Q = \mathrm{diag}[\beta_1, \cdots, \beta_l], \tag{7.1.4}$$

and U and V are boxes representing upper and lower bounds on the components of $u = (u_1, \cdots, u_k)$ and $v = (v_1, \cdots, v_l)$:

$$U = [u_1^-, u_1^+] \times \cdots \times [u_k^-, u_k^+], \qquad V = [v_1^-, v_1^+] \times \cdots \times [v_l^-, v_l^+]. \tag{7.1.5}$$

Hence, (7.1.1) specializes to the problem of minimizing

$$\sum_{j=1}^{k} [p_j u_j + \tfrac{1}{2}\alpha_j u_j^2] + \sum_{i=1}^{l} \rho_{v_i^-, v_i^+, \beta_i} \left(q_i - \sum_{j=1}^{k} d_{ij} u_j \right) \tag{7.1.6}$$

subject to $u_j^- \le u_j \le u_j^+$ for $j = 1, \cdots, k$, where

$$\rho_{v_i^-, v_i^+, \beta_i}(w_i) = \underset{v_i^- \le v_i \le v_i^+}{\mathrm{maximize}} \quad [w_i v_i - \tfrac{1}{2}\beta_i v_i^2] \tag{7.1.7}$$

is an *one-dimensional* subproblem for every i and has very simple closed-form solution:

$$\rho_{v_i^-, v_i^+, \beta_i}(w_i) = \begin{cases} v_i^+ w_i - \frac{1}{2}\beta_i (v_i^+)^2 & \text{if } w_i/\beta_i > v_i^+, \\ \frac{1}{2} w_i^2/\beta_i & \text{if } v_i^- \le w_i/\beta_i \le v_i^+, \\ v_i^- w_i - \frac{1}{2}\beta_i (v_i^-)^2 & \text{if } w_i/\beta_i < v_i^-. \end{cases} \tag{7.1.8}$$

We shall adopt several assumptions of Rockafellar (1990) in this chapter:

Finite Monitoring Assumption. *The functions $\rho_{Q,V}$ and $\rho_{P,U}$ are finite everywhere; they do not take on ∞. Thus the objective functions f in (7.1.1) and*

g in (7.1.3) are finite everywhere, and there are no constraints implicit in these problems beyond $u \in U$ and $v \in V$.

This is definitely satisfied in the fully quadratic case of problems (7.1.1) and (7.1.3) (where P and Q are both nonsingular). It is also satisfied, regardless of the nature of P and Q, when the sets U and V are bounded.

The Finite Monitoring Assumption merely acts as a weak normalization of the mathematical model being used in a given application. If it were not satisfied, we could just either truncate the sets U and V or add strictly quadratic "proximal terms" to L. Both ways are harmless and have considerable theoretic background. Please refer to Rockafellar [271], [291], [273].

By virtue of the Finite Monitoring Assumption, we can write

$$f(u) = \max_{v \in V} L(u, v), \qquad g(v) = \min_{u \in U} L(u, v). \tag{7.1.9}$$

Correspondingly, we have *nonempty* point sets

$$F(u) = \operatorname*{argmax}_{v \in V} L(u, v), \qquad G(v) = \operatorname*{argmin}_{u \in U} L(u, v), \tag{7.1.10}$$

that depend on u and v. In the fully quadratic case where L is strongly convex-concave, $F(u)$ and $G(v)$ are uniquely determined vectors.

Double Decomposability Assumption. *For each $u \in U$ it is readily possible to calculate the value $f(u)$ in (7.1.9) and at least one of the elements of the set $F(u)$ in (7.1.10). Similarly, for each $v \in V$ it is readily possible to calculate the value $g(v)$ and at least one of the elements of $G(v)$.*

This assumption is important in finite-envelope methods and primal-dual projected gradient algorithms [295], [331]. It is also convenient for us since we need to compute the $f(u^k)$ and $g(v^k)$ in our algorithms. However, we only need the assumptions on $f(u)$ and $g(v)$, which is very reasonable, and we do not need the assumptions about $F(u)$ and $G(v)$ at all in our algorithms.

Line Searchability Assumption. *It is possible to minimize $f(u)$ over any line segment joining two points in U, and likewise, to maximize $g(v)$ over any line segment joining two points in V.*

This assumption is already implied, more or less, by the two preceding assumptions, but Rockafellar introduced it as an assumption for clarity. This is very useful in our modified EFBS algorithm with line searches.

7.2. Properties about Bounded Linear Operators

When an operator is bounded linear, it has some good properties such as Lipschitz continuity, etc.. Here we shall give some.

Proposition 7.2.1. *If T is a bounded linear operator on $\mathcal{H}$ and is strongly monotone, then T^{-1} is also strongly monotone.*

Proof. This is a corollary of Proposition 4.6.6. $\square$

Proposition 7.2.2. *If T is a linear bounded operator on $\mathcal{H}$, then T is (strongly) monotone if and only if T^* is (strongly) monotone.*

Proof. Since $\langle x, T(x)\rangle = \langle T^*(x), x\rangle = \langle x, T^*(x)\rangle$, thus T is (strongly) monotone if and only if T^* is (strongly) monotone. $\square$

Theorem 7.2.3. *If T is a symmetric bounded linear operator and is monotone, then T^{-1} is strongly monotone.*

Proof. Because T is symmetric and positive semi-definite, there exists an unique symmetric and positive semi-definite bounded linear operator E on $\mathcal{H}$ such that $T = E^2$ from the theory of Functional Analysis. So if $T \neq 0$,

$$\begin{aligned}\langle x, T(x)\rangle &= \langle E(x), E(x)\rangle \\ &= \|E(x)\|^2 \\ &\leq \frac{1}{\|E\|^2}\|T(x)\|^2 \\ &= \frac{1}{\|T\|}\|T(x)\|^2.\end{aligned}$$

That is, T^{-1} is strongly monotone. If $T = 0$, the conclusion is obviously true. $\square$

Theorem 7.2.3 is very important since we can obtain a convergent theorem about the FBS methods later.

Now we shall introduce another concept.

Definition 7.2.1. *A bounded linear operator R on $\mathcal{H}$ is said as **invertible** if there is a bounded linear operator G such that $RG = GR = I$ and we write $R^{-1} = G$.*

Lemma 7.2.4. *Suppose R and G are bounded linear operators. If $\langle x, Ry\rangle = \langle x, Gy\rangle$ for any $x, y \in \mathcal{H}$, then $R = G$.*

Proof. For any $x \in \mathcal{H}$, let $y = Rx - Gx$. Then

$$\langle y, y\rangle = \langle y, Rx\rangle - \langle y, Gx\rangle = 0$$

i.e., $y = 0$. Or $Rx = Gx$. Hence $R = G$. □

Lemma 7.2.5. *If R and G are bounded linear operators, then $(RG)^* = G^*R^*$. If they are also invertible, then $(RG)^{-1} = G^{-1}R^{-1}$ and $(R^*)^{-1} = (R^{-1})^*$.*

Proof. Since for any $x, y \in \mathcal{H}$ we have

$$\langle y, (RG)^*x\rangle = \langle RGy, x\rangle = \langle Gy, R^*x\rangle = \langle y, (G^*R^*)x\rangle$$

So Lemma 7.2.4 implies that $(RG)^* = G^*R^*$.

$$(RG)^{-1}(RG) = I \Longrightarrow (RG)^{-1}R = G^{-1} \Longrightarrow (RG)^{-1} = G^{-1}R^{-1}$$

$$(R^{-1})^*R^* = (RR^{-1})^* = I^* = I \Longrightarrow (R^*)^{-1} = (R^{-1})^*$$

□

Lemma 7.2.6. *Suppose R is a invertible bounded linear operator. Then*
(a) R is symmetric if and only if so is R^{-1},
(b) R is monotone if and only if so is R^{-1},
(c) R is strongly monotone if and only if so is R^{-1}.

Proof. (a): Since $RR^{-1} = R^{-1}R = I$ and $R^*(R^{-1})^* = I$, so when $R^* = R$, $R(R^{-1})^* = I \Longrightarrow (R^{-1})^* = R^{-1}$, and when $(R^{-1})^* = R^{-1}$, $R^*R^{-1} = I \Longrightarrow R^* = R$.
(b): Since $\langle x, Rx\rangle = \langle R^{-1}y, y\rangle$ for any $x, y \in \mathcal{H}$ with $y = Rx$ or $x = R^{-1}y$, hence R is monotone if and only if R^{-1} is monotone.
(c): Similar to (b), from $\langle x, Rx\rangle = \langle R^{-1}y, y\rangle$ and the boundedness of R and R^{-1} with $y = Rx$ or $x = R^{-1}y$, we have the conclusion easily. □

Lemma 7.2.7. *If R is a bounded linear operator, then R is invertible if and only if $\mathcal{R}(R) = \mathcal{H}$ and $\mathcal{N}(R) := \{x \,|\, Rx = 0\} = \{0\}$.*

Proof. (Necessity):
For any $x \in \mathcal{H}$, let $y = R^{-1}x$. Then $Ry = x$. Hence $im(R) = \mathcal{H}$.
If $Rx = 0$, then $x = R^{-1}(0) = 0$. Hence $\mathcal{N}(R) = \{0\}$.
(Sufficiency):
If $im(R) = \mathcal{H}$ and $\mathcal{N}(R) = \{0\}$, then for any $x_1 \neq x_2$, $Rx_1 \neq Rx_2$ and for any $y \in \mathcal{H}$, there is a $x \in \mathcal{H}$ such that $y = Rx$. Now we define a new operator G with $Gy = x$ if $y = Rx$. We know that G is a bounded linear operator from Functional Analysis and $GRx = x$, $RGy = y$ for any $x, y \in \mathcal{H}$. Thus $GR = RG = I$. Hence $G = R^{-1}$. $\square$

From Lemma 7.2.7, it is simple to prove the following.

Proposition 7.2.8. *If a bounded linear operator T is strongly monotone, then T is invertible.*

Proof. Let $\alpha > 0$ be the modulus. Then $\langle x, Tx\rangle \geq \alpha\langle x, x\rangle$ implies that $Tx \neq 0$ if $x \neq 0$. That is, $\mathcal{N}(T) = \{0\}$. Similarly, since T^* is also strongly monotone, $\mathcal{N}(T^*) = \{0\}$. Therefore, $im(T) = \mathcal{N}(T^*)^\perp = \{0\}^\perp = \mathcal{H}$. $\square$

Proposition 7.2.9. *If a monotone bounded linear operator T is symmetric and invertible, then T and T^{-1} are strongly monotone.*

Proof. Theorem 7.2.3 implies that T^{-1} is strongly monotone. But T is invertible, so Lemma 7.2.6 implies that T is also strongly monotone. $\square$

Without symmetry, Proposition 7.2.9 may be wrong. For example, if T is invertible and $T^* = -T$, then T is monotone but T and T^{-1} are both not strongly monotone.

Now let consider the convergence of the FBS algorithm when T_1 is a bounded linear operator.

Theorem 7.2.10 (FBS Convergence). *Suppose T_1 is boundedly linear and $\{z^k\}$ is generated by the FBS algorithm of Section 4.1. Then*

(1) if $\{z^k\}$ *strongly converges to a point* z^* *and* $\inf_k\{\lambda_k\} > 0$*, we have* $z^* \in T^{-1}(0)$*;*

(2) if T_1 *is also symmetric and* $\{\lambda_k\}$ *satisfies*

$$0 < \inf_{k\geq 0}\{\lambda_k\} \leq \sup_{k\geq 0}\{\lambda_k\} < 2\|T_1\|^{-1}, \tag{7.2.1}$$

then $\|z^k\| \to \infty$ *when* $T^{-1}(0) = \emptyset$*, otherwise* $\{z^k\}$ *strongly converges to a zero point of* T*.*

Proof. Part (1) is a direct corollary of Proposition 4.5.1 and Part (2) for $T^{-1}(0) \neq \emptyset$ is a corollary of Theorem 7.2.3 and Theorem 4.5.4. □

7.3. A Simplification of the Continuous-Time Models

Now we consider continuous time optimal control problems. Here we only emphasize on the linear-quadratic ones in Section 3.2.

Problem $(\mathcal{P}_c)$ in Section 3.2 was introduced in Rockafellar (1987). It was shown in that paper that $(\mathcal{P}_c)$ is the primal problem associated with the Lagrangian $\mathcal{J}_0(u, u_e; v, v_e)$ on $\mathcal{U} \times \mathcal{V}$, while the dual problem is $(\mathcal{Q}_c)$ of Section 3.2.

From Section 3.2, solving the saddle point problem associated with Lagrangian $\mathcal{J}(u, u_e; v, v_e)$ on $\mathcal{U} \times \mathcal{V}$ is equivalent to solving its primal problem $(\mathcal{P}_c)$ and its dual problem $(\mathcal{Q}_c)$.

Now let us consider the simplification of the saddle point problem. Let $x^0(t)$ and $y^0(t)$ are the solutions of equations (3.2.1) and (3.2.6) associated with $u(t) = 0$ a.e. , $v(t) = 0$ a.e. , $u_e = 0$ and $v_e = 0$. i.e.

$$\begin{aligned} \dot{x}^0(t) &= A(t)x^0(t) + b(t) \quad \text{a.e.,} \quad x^0(t_0) = b_e, \\ -\dot{y}^0(t) &= A^*(t)y^0(t) + c(t) \quad \text{a.e.,} \quad y^0(t_1) = c_e \end{aligned} \tag{7.3.1}$$

We may get (3.2.1) and (3.2.6) by combining (7.3.1) and the following ODE's

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \quad \text{a.e.,} \quad x(t_0) = B_e u_e, \\ -\dot{y}(t) &= A^*(t)y(t) + C^*(t)v(t) \quad \text{a.e.,} \quad y(t_1) = {C_e}^* v_e. \end{aligned} \tag{7.3.2}$$

Proposition 7.3.1. *Let $x^0(t)$ and $y^0(t)$ defined as the above. Suppose $x(t)$ and $y(t)$ are the solutions of equations (7.3.2). Then $\bar{x}(t) = x(t) + x^0(t)$ and $\bar{y}(t) = y(t) + y^0(t)$ are the solutions of ode's (3.2.1) and (3.2.6).*

Proof. This is elementary to verify them by differentiating $\bar{x}(t)$ and $\bar{y}(t)$ from (7.3.1) and (7.3.2). □

Lemma 7.3.2. *Under the assumptions of Proposition 7.3.1, the $[(u,u_e),(v,v_e)]$ in (3.2.5) satisfies*

$$[(u,u_e),(v,v_e)] = [(u,u_e),(v,v_e)]_0 + [u,u_e]_u + [v,v_e]_v + [b,c]_c. \tag{7.3.3}$$

where

$$\begin{aligned}
[u,u_e]_u &= \int_{t_0}^{t_1} y^0(t)\cdot B(t)u(t)\,dt + y^0(t_0)\cdot B_e u_e \\
&= \int_{t_0}^{t_1} x(t)\cdot c(t)\,dt + x(t_1)\cdot c_e, \\
[v,v_e]_v &= \int_{t_0}^{t_1} y(t)\cdot b(t)\,dt + y(t_0)\cdot b_e \\
&= \int_{t_0}^{t_1} x^0(t)\cdot C^*(t)v(t)\,dt + x^0(t_1)\cdot {C_e}^* v_e, \\
[b,c]_c &= \int_{t_0}^{t_1} y^0(t)\cdot b(t)\,dt + y^0(t_0)\cdot b_e \\
&= \int_{t_0}^{t_1} x^0(t)\cdot c(t)\,dt + x^0(t_1)\cdot c_e, \\
[(u,u_e),(v,v_e)]_0 &= \int_{t_0}^{t_1} y(t)\cdot B(t)u(t)\,dt + y(t_0)\cdot B_e u_e \\
&= \int_{t_0}^{t_1} x(t)\cdot C^*(t)v(t)\,dt + x(t_1)\cdot {C_e}^* v_e.
\end{aligned} \tag{7.3.4}$$

Proof. From the definition of $[(u,u_e),(v,v_e)]$ in (3.2.5) and Proposition

7.3.1, one has the following:

$$\begin{aligned}[(u,u_e),(v,v_e)] &= \int_{t_0}^{t_1} [y(t)+y^0(t)]\cdot[B(t)u(t)+b(t)]\,dt \\ &\quad + [y(t_0)+y^0(t_0)]\cdot[B_e u_e + b_e] \\ &= \int_{t_0}^{t_1} [x(t)+x^0(t)]\cdot[C^*(t)v(t)+c(t)]\,dt \\ &\quad + [x(t_1)+x^0(t_1)]\cdot[C_e{}^* v_e + c_e].\end{aligned} \tag{7.3.5}$$

Let $u(t)\equiv 0$, $v(t)\equiv 0$, $u_e = 0$ and $v_e = 0$. Then $x(t) = 0$ and $y(t) = 0$ from (7.3.2) and (7.3.5) gives the common value $[b,c]_c$. When we only take $u(t)\equiv 0$ and $u_e = 0$, (7.3.2) implies $x(t) = 0$ and hence (7.3.5) and $[b,c]_c$ give the common value $[v,v_e]_v$. Likely, taking $v(t)\equiv 0$ and $v_e = 0$ implies the common value $[u,u_e]_u$ from (7.3.5) and $[b,c]_c$. Finally, deleting $[b,c]_c$, $[u,u_e]_u$ and $[v,v_e]_v$ from (7.3.5) gives the common value $[(u,u_e),(v,v_e)]_0$. Obviously, adding the common values in (7.3.4) together gives (7.3.5). That is, (7.3.3) is true. □

Now we may simplify the saddle point problem associated with Lagrangian $\mathcal{J}(u,u_e;v,v_e)$ on $\mathcal{U}\times\mathcal{V}$ from the above theorem. Consider the following problem:

$$(\mathcal{SP})\qquad \begin{aligned} &\underset{(u,u_e),(v,v_e)}{\text{minimax}}\ \mathcal{J}_0(u,u_e;v,v_e) = \bar{J}_e(u_e,v_e) \\ &\qquad\qquad + \int_{t_0}^{t_1} \bar{J}(t,u(t),v(t))\,dt - [(u,u_e),(v,v_e)]_0, \\ &\text{subject to}\quad (u,u_e)\in\mathcal{U}, (v,v_e)\in\mathcal{V},\end{aligned}$$

where $x(t)$ and $y(t)$ satisfy (7.3.2) and

$$\begin{aligned}\bar{J}(t,u,v) &= \bar{p}(t)\cdot u + \bar{q}(t)\cdot v + \tfrac{1}{2}u\cdot P(t)u - \tfrac{1}{2}v\cdot Q(t)v - v\cdot D(t)u, \\ \bar{J}_e(u_e,v_e) &= \bar{p}_e\cdot u_e + \bar{q}_e\cdot v_e + \tfrac{1}{2}u_e\cdot P_e u_e - \tfrac{1}{2}v_e\cdot Q_e v_e - v_e\cdot D_e u_e.\end{aligned} \tag{7.3.6}$$

Theorem 7.3.3. *The saddle point problem $(\mathcal{SP})$ is equivalent to the saddle point problem associated with $\mathcal{J}(u,u_e;v,v_e)$ on $\mathcal{U}\times\mathcal{V}$ under the conditions*

$$\begin{aligned}\bar{p}(t) &= p(t) - B^*(t)y^0(t), \qquad \bar{p}_e = p_e - B_e{}^* y^0(t_0), \\ \bar{q}(t) &= q(t) - C(t)x^0(t), \qquad \bar{q}_e = q_e - C_e x^0(t_1)\end{aligned} \tag{7.3.7}$$

.

Proof. It is elementary to obtain from (3.2.5), (7.3.6), (7.3.3), (7.3.4) and the conditions (7.3.7) that

$$\mathcal{J}(u, u_e; v, v_e) = \mathcal{J}_0(u, u_e; v, v_e) - [b, c]_c$$

$[b, c]_c$ is a constant since $x^0(t)$ and $y^0(t)$ are independent of $(u, u_e; v, v_e)$. Hence the solutions are the same to both problems. i.e., the two problems are equivalent. $\square$

From the expressions and the dynamical systems of both problems, $(\mathcal{SP})$ is simpler and has the 'homogeneous' or 'linear' property in (u, u_e) or (v, v_e) which is also useful later on. Actually, the both problems in Theorem 7.3.3 are the same when $b(t) = 0, b_e = 0, c(t) = 0$ and $c_e = 0$ which imply $x^0(t) = 0$ and $y^0(t) = 0$. Therefore we may ignore $b(t), c(t), b_e$ and c_e when we analyze and solve the problem $(\mathcal{P})$ later on.

7.4. A Simplification of the Multistage Models

The extended linear quadratic discrete time problems in Rockafellar and Wets (1990) include two classes: deterministic and stochastic controls. Here we consider the deterministic or multistage problems which are described in Section 3.1. For these problems, the dynamics can be expressed by the system (3.1.2) and the model problem is $(\mathcal{P}_d)$ of Section 3.1.

Problem $(\mathcal{P}_d)$ was introduced in the paper of Rockafellar and Wets. It was shown in that paper that $(\mathcal{P}_d)$ is the primal problem associated with the Lagrangian $L(u, v)$ on $U \times V$, while $(\mathcal{Q}_d)$ is the dual problem of $(\mathcal{P}_d)$.

Except some primal-dual properties, Rockafellar and Wets also proved that the saddle point problem associated with L on $U \times V$ has some decomposability properties. But the main troubles for solving it are caused by $l(u, v)$ and the dynamical systems (3.1.2) and (3.1.7) in practice. As we did in Section 7.3, we shall try to simplify $l(u, v)$ in (3.1.6) and the dynamics equations in (3.1.2) and (3.1.7) in the following.

Similar to Sec. 7.3, assume that $x^0 = (x_0^0, \cdots, x_N^0)$ and $y^0 = (y_1^0, \cdots, y_{N+1}^0)$ are always the solutions of the ODE's, which are independent of u and v, as follows:

$$\begin{aligned} &x_t = A_t x_{t-1} + b_t \text{ for } t = 1, \ldots, N, \text{ with } x_0 = b_0; \\ &y_t = A_t^* y_{t+1} + c_t \text{ for } t = N, \ldots, 1, \text{ with } y_{N+1} = c_{N+1}. \end{aligned} \tag{7.4.1}$$

Consider the following ODE's which depend on u and v linearly and homogeneously.

$$\begin{aligned} &x_t = A_t x_{t-1} + B_t u_t \text{ for } t = 1, \ldots, N, \text{ with } x_0 = B_0 u_0; \\ &y_t = A_t^* y_{t+1} + C_t^* v_t \text{ for } t = N, \ldots, 1, \text{ with } y_{N+1} = C_{N+1}^* v_{N+1}. \end{aligned} \tag{7.4.2}$$

Proposition 7.4.1. *$x = (x_0, \ldots, x_N)$ and $y = (y_1, \ldots, y_{N+1})$ solve the equations (7.4.2) if and only if $\bar{x} = x + x^0$ solves (3.1.2) and $\bar{y} = y + y^0$ solves (3.1.7).*

Proof. It is elementary from (3.1.2), (3.1.7) and (7.4.2) by (7.4.1). □

From the above proposition, we may decompose $l(u, v)$ into 4 parts.

Lemma 7.4.2. *Suppose that x and y are given by (7.4.2). Then $l(u, v)$ in (3.1.6) can be decomposed into 4 parts:*

$$l(u, v) = l_0(u, v) + l_1(u) + l_2(v) + l_c \tag{7.4.3}$$

where

$$\begin{aligned} l_c &= \sum_{t=0}^{N} b_t \cdot y_{t+1}^0 = \sum_{t=1}^{N+1} c_t \cdot x_{t-1}^0, \\ l_0(u, v) &= \sum_{t=1}^{N+1} x_{t-1} \cdot C_t^* v_t = \sum_{t=0}^{N} y_{t+1} \cdot B_t u_t, \\ l_1(u) &= \sum_{t=0}^{N} y_{t+1}^0 \cdot B_t u_t = \sum_{t=1}^{N+1} x_{t-1} \cdot c_t, \\ l_2(v) &= \sum_{t=0}^{N} y_{t+1} \cdot b_t = \sum_{t=1}^{N+1} x_{t-1}^0 \cdot C_t^* v_t. \end{aligned} \tag{7.4.4}$$

Proof. Let us replace x and y in $l(u, v)$ of (3.1.6) by $\bar{x} = x + x^0$ and $\bar{y} = y + y^0$ from Proposition 7.4.1. If $u = 0$ and $v = 0$, then $\bar{x} = x^0$ and $\bar{y} = y^0$ from (7.4.2), and hence $l(u, v)$ gives the common value l_c of (7.4.4). If only $u = 0$, one gets $x = 0$ from (7.4.2). Thus $l(u,v)$ becomes the common value $l_2(v) + l_c$. Deleting l_c gets the common value $l_2(v)$. Likewise, $v = 0$ implies the common value $l_1(u)$. Directly deleting l_c, $l_1(u)$ and $l_2(v)$ from $l(u,v)$ gives the common value $l_0(u,v)$. □

Consider the following saddle point problem:

$$(\mathcal{SP}_d)\qquad \begin{aligned} \text{minimax}\quad L_0(u,v) := {} & \sum_{t=0}^{N}\left[\bar{p}_t{\cdot}u_t + \tfrac{1}{2}u_t{\cdot}P_t u_t\right] - \sum_{t=1}^{N} v_t{\cdot}D_t u_t \\ & + \sum_{t=1}^{N+1}\left[\bar{q}_t{\cdot}v_t - \tfrac{1}{2}v_t{\cdot}Q_t v_t\right] - l_0(u,v), \\ \text{subject to}\quad & u \in U, v \in V \text{ and (7.4.2)} . \end{aligned}$$

Theorem 7.4.3. *If $\bar{p}_t = p_t - B_t^* y_{t+1}^0$ for $t = 0,\ldots,N$ and $\bar{q}_t = q_t - C_t x_{t-1}^0$ for $t = 1,\ldots,N+1$, then Problem $(\mathcal{SP}_d)$ and the saddle point problem associated with $L(u,v)$ on $U \times V$ with (3.1.2) and (3.1.7) have the same solutions (u,v).*

Proof. Under the conditions of this theorem, it is easy to see that

$$L(u,v) = L_0(u,v) - l_c. \tag{7.4.5}$$

But both have the same constraint $U \times V$ and l_c is a constant here. Thus both problems have the same solutions (u,v). □

Theorem 7.4.3 tells us that solving Problem $(\mathcal{P}_d)$ and Problem $(\mathcal{Q}_d)$ is equivalent to solving Problem $(\mathcal{SP}_d)$ with the ODE's (7.4.1). When $b_t = 0$ for $t = 0,\ldots,N$ and $c_t = 0$ for $t = 1,\ldots,N+1$, one gets $x^0 = 0$ and $y^0 = 0$ from (7.4.1) and hence $p = \bar{p}$, $q = \bar{q}$. We shall ignore the b_t and c_t if we do further analysis.

7.5. EFBS and Modified EFBS Algorithms for Saddle Point Problems

In this section, we shall discuss two algorithms for solving a saddle point problem by the EFBS methods, especially solving the extended linear-quadratic programming problem (2.2.7), which includes Problem $(\mathcal{SP})$ and $(\mathcal{SP}_d)$ as special cases, when L is given by (7.1.2).

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be real Hilbert spaces and $L : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{R}$ be such that $L(u, v)$ is convex in $u \in \mathcal{H}_1$, concave in $v \in \mathcal{H}_2$, and is differentiable, closed over $\mathcal{H}$.

Consider the saddle point problem (2.2.7):

$$(\mathcal{G}) \qquad \underset{u\in U,\, v\in V}{\text{minimax}} \quad L(u, v)$$

where *minimax* has the same meaning as that in Section 2.2 and $U \subset \mathcal{H}_1$ and $V \subset \mathcal{H}_2$ are nonempty closed convex sets.

From Example 2.3.3, Problem $(\mathcal{G})$ can be transformed into the fundamental problem $(\mathcal{P}_0)$ with T as follows:

$$T = \begin{pmatrix} \nabla_u L + \mathcal{N}_U \\ -\nabla_v L + \mathcal{N}_V \end{pmatrix} = \begin{pmatrix} \nabla_u L \\ -\nabla_v L \end{pmatrix} + \mathcal{N}_U \times \mathcal{N}_V. \tag{7.5.1}$$

Hence, operator splitting techniques can be applied to solve the saddle point problem $(\mathcal{G})$. Here, we only consider the FBS method.

Since $\mathcal{N}_U \times \mathcal{N}_V$ is difficult to implement through the forward step but relatively easy to run by the backward step, it is put into T_2 when $T = T_1 + T_2$ is split. Now, suppose $L = L_1 + L_2$ is given with L_1 and L_2 convex-concave. Then, a corresponding splitting $T = T_1 + T_2$ is also given with T_1, T_2 maximal monotone through

$$T_1 = \begin{pmatrix} \nabla_u L_1 \\ -\nabla_v L_1 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} \nabla_u L_2 + \mathcal{N}_U \\ -\nabla_v L_2 + \mathcal{N}_V \end{pmatrix}. \tag{7.5.2}$$

For our consideration, we only discuss the special $L(u, v)$

$$L(u, v) = h(u) - s(v) - \langle v, Du\rangle_2 \tag{7.5.3}$$

here $h : \mathcal{H}_1 \to \mathbb{R}$ and $s : \mathcal{H}_2 \to \mathbb{R}$ are differentiable convex functions. $\langle .,.\rangle_1$ and $\langle .,.\rangle_2$ are the inner products over $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. $D : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ is a bounded linear operator and

$$\langle v, Du\rangle_2 = \langle u, D^* v\rangle_1 \quad \text{for any} \quad u \in \mathcal{H}_1, v \in \mathcal{H}_2. \tag{7.5.4}$$

where D^* is the conjugate operator of D, which is also a bounded linear operator.

In this chapter, we mainly focus to an important special case of (7.5.3), i.e., L is the Lagrangian representation (7.1.2) of extended linear-quadratic programming problems with

$$h(u) = \tfrac{1}{2}u{\cdot}Pu + p{\cdot}u, \qquad s(v) = \tfrac{1}{2}v{\cdot}Qv - q{\cdot}v. \tag{7.5.5}$$

Note that the space $\mathcal{L}^2\big([t_0, t_1], \mathbb{R}^k\big) \times \mathbb{R}^m$ is a Hilbert space associated with the inner product

$$\langle w_1, w_2\rangle = \int_{t_0}^{t_1} u_1(t){\cdot}u_2(t)\,dt + v_1{\cdot}v_2, \tag{7.5.6}$$

for $w_i = (u_i, v_i) \in \mathcal{L}^2\big([t_0, t_1], \mathbb{R}^k\big) \times \mathbb{R}^m$, $i = 1, 2$. Therefore, Problem $(\mathcal{SP})$ and Problem $(\mathcal{SP}_d)$ are the special cases of Problem $(\mathcal{G})$ with L given by (7.1.2) from the homogeneous properties of $[(u, u_e), (v, v_e)]_0$ and $l_0(u, v)$.

For the saddle point problem $(\mathcal{G})$, a computing extended forward-backward splitting algorithm can be described as follows:

EFBS Algorithm:

Step 0. Initial input: Give an initial point (u^0, v^0), initial $\lambda_0 > 0$, initial metric H_0, initial approximation $T^0 = T_1^0 + T_2^0$, and the error bound ε. Let $k = 0$.

Step 1. EFBS iteration: Use the EFBS scheme to get (u^{k+1}, v^{k+1}). That is,

$$(u^{k+1}, v^{k+1}) = (H_k + \lambda_k T_2^k)^{-1}\big[H_k(u^k, v^k) - \lambda_k T_1^k(u^k, v^k)\big]. \tag{7.5.7}$$

In other words, (u^{k+1}, v^{k+1}) is gotten through the following:

$$\hat{u}^k = H_k u^k - \lambda_k \nabla_u L_1^k(u^k, v^k), \quad \hat{v}^k = H_k v^k + \lambda_k \nabla_v L_1^k(u^k, v^k), \tag{7.5.8}$$

$$\begin{aligned}(u^{k+1}, v^{k+1}) = \operatorname*{argminimax}_{u\in U,\, v\in V} \Big\{ & L_2^k(u, v) \\ & + \frac{1}{2\lambda_k}\big[\|u - (H_k')^{-1}\hat{u}^k\|_{H_k'}^2 - \|v - (H_k'')^{-1}\hat{v}^k\|_{H_k''}^2\big]\Big\},\end{aligned} \tag{7.5.9}$$

where

$$H_k = \begin{pmatrix} H'_k & \\ & H''_k \end{pmatrix}. \tag{7.5.10}$$

Step 2. Primal-dual gap check: Compute $f_k = \bar{f}(u^{k+1})$ and $g_k = \bar{g}(v^{k+1})$, where

$$\bar{f}(u) = \sup_{v \in V} L(u, v), \qquad \bar{g}(v) = \inf_{u \in U} L(u, v). \tag{7.5.11}$$

If

$$f_k - g_k \leq \varepsilon, \tag{7.5.12}$$

then stop iteration and output the ε-optimal solution $(\hat{u}, \hat{v}) = (u^{k+1}, v^{k+1})$. Otherwise, go to next step.

Step 3. Distance check: If

$$\|(u^{k+1}, v^{k+1}) - (u^k, v^k)\|_{H_k} \leq \varepsilon,$$

or the iteration number arrives a maximal iteration number requirement, stop iteration and output the ε-optimal solution $(\hat{u}, \hat{v}) = (u^{k+1}, v^{k+1})$. Otherwise, choose suitable $\lambda_{k+1} > 0$, metric H_{k+1}, operator approximation $T^{k+1} = T_1^{k+1} + T_2^{k+1}$, and let $k := k + 1$. Go back to Step 1. □

From the above algorithm, we can notice that the step 2 can not be deduced from the EFBS methods in Chapter 5. Hence, we need to know its influence on the convergence of (u^k, v^k). For this purpose, we cite here from Rockafellar (1990, 1993) several fundamental properties on which our algorithmic developments depend.

Proposition 7.5.1 (equilibrium, [268]). *Let L be an arbitrary real-valued function on a product of any nonempty sets U and V. Then, it is always true that*

$$\inf_{u \in U} \bar{f}(u) \geq \sup_{v \in V} \bar{g}(v). \tag{7.5.13}$$

A pair $(\bar{u}, \bar{v})$ furnishes a saddle point of $L(u,v)$ on $U \times V$ if and only if

$$\begin{cases} \bar{f}(\bar{u}) = \inf_{u\in U} \bar{f}(u), \\ \bar{g}(\bar{v}) = \sup_{v\in V} \bar{g}(v), \\ \quad \bar{f}(\bar{u}) = \bar{g}(\bar{v}). \end{cases} \tag{7.5.14}$$

□

Proposition 7.5.1 implies that all of the cluster points of (u^k, v^k) will be the saddle points of $(\mathcal{G})$ if $\bar{f}(u^k) - \bar{g}(v^k) \to 0$ and $\bar{f}$ and $-\bar{g}$ are lower semicontinuous. Hence, the step 2 in the EFBS algorithm makes sense. Furthermore, if $(\mathcal{G})$ has a unique saddle point and L is convex-concave, then (u^k, v^k) will actually converge to the saddle point strongly.

Note that $\bar{f}(u) \equiv f(u)$ and $\bar{g}(v) \equiv g(v)$ when L in Problem $(\mathcal{G})$ is the one in (7.1.2).

Proposition 7.5.2 (optimality estimates, [271]). *Let L be given by (7.1.2), and be strongly convex-concave. Suppose $(\bar{u}, \bar{v})$ is the unique saddle point of L on $U \times V$. Suppose $\hat{u}$ and $\hat{v}$ are elements of U and V satisfying $f(u) - g(v) \le \varepsilon$ for a certain $\varepsilon \ge 0$. Then, $\hat{u}$ and $\hat{v}$ are ε-optimal in the sense that $|f(\hat{u}) - f(\bar{u})| \le \varepsilon$ and $|g(\hat{u}) - g(\bar{u})| \le \varepsilon$. Moreover, $\|\hat{u} - \bar{u}\|_P \le \sqrt{2\varepsilon}$ and $\|\hat{v} - \bar{v}\|_Q \le \sqrt{2\varepsilon}$.* □

Proposition 7.5.1 and Proposition 7.5.2 together tell us that (u^k, v^k) will converge to the saddle point of $(\mathcal{G})$ in the EFBS algorithm with $\varepsilon = 0$ if $f_k - g_k \to 0$ when L is given by (7.1.2). That is, the step 2 in the algorithm is logical for L in (7.1.2).

Proposition 7.5.3 (regularity properties, [271]). *Let L be given by (7.1.2), and be strongly convex-concave. The functions f and g are continuously differentiable everywhere:*

$$\begin{aligned} \nabla f(u) &= \nabla_u L(u, F(u)) = p + Pu - D^T F(u), \\ \nabla g(v) &= \nabla_v L(G(v), v) = q - Qv - DG(v), \end{aligned} \tag{7.5.15}$$

where $F(u)$ and $G(v)$ are defined in (7.1.10), while in terms of the constant

$$\gamma(P, Q, D) := \|Q^{-\frac{1}{2}} D P^{-\frac{1}{2}}\|^2 \tag{7.5.16}$$

the mappings F and G are Lipschitz continuous:

$$\begin{aligned} \|F(u') - F(u)\|_Q &\le \gamma(P,Q,D)\|u'-u\|_P \quad \text{for all} \quad u \quad \text{and} \quad u', \\ \|G(v') - G(v)\|_P &\le \gamma(P,Q,D)\|v'-v\|_Q \quad \text{for all} \quad v \quad \text{and} \quad v'. \end{aligned} \tag{7.5.17}$$

□

In the following, we shall discuss some properties of $\bar{f}$ and $\bar{g}$ in (7.5.11).

Theorem 7.5.4 (properties of $\bar{f}$ and $\bar{g}$). *Let L be given by (7.5.3) and h and s be strongly convex. Suppose $\mathcal{H}$ is actually an Euclidean space. Then,*

(a) $\bar{f}$ and $\bar{g}$ in (7.5.11) are strongly convex and strongly concave respectively, and they can be expressed by

$$\bar{f}(u) = \max_{v\in V} L(u,v) = L(u,\bar{F}(u)), \quad \bar{g}(v) = \min_{u\in U} L(u,v) = L(\bar{G}(v),v), \tag{7.5.18}$$

where

$$\bar{F}(u) := \operatorname*{argmax}_{v\in V} L(u,v), \qquad \bar{G}(v) := \operatorname*{argmin}_{u\in U} L(u,v) \tag{7.5.19}$$

are single-valued and continuous.

(b) $\bar{f}$ and $\bar{g}$ have continuous first derivatives, and one has

$$\begin{aligned} \nabla\bar{f}(u') &= \nabla_u L(u',\bar{F}(u')) = \nabla h(u') - D^T\bar{F}(u') \quad \text{at any} \quad u', \\ \nabla\bar{g}(v') &= \nabla_v L(\bar{G}(v'),v') = -D\bar{G}(v') - \nabla s(v') \quad \text{at any} \quad v'. \end{aligned} \tag{7.5.20}$$

Proof. The strong convexity of h and s ensures that $L(u,v)$ is strictly convex in u and strictly concave in v, so the sets $\bar{F}(u)$ and $\bar{G}(v)$ in (7.5.19) are always singletons. That is, the max and min in (7.5.18) are uniquely attained. Indeed, $\bar{F}(u)$ is the solution to a strictly convex (smooth) programming problem in v in which only the linear part of the objective depends on u, from which it follows further that the mapping $\bar{F}$ is continuous; similar for $\bar{G}(v)$. Now let $\bar{s}^*$ be the conjugate of the closed, proper, convex function

$$\bar{s}(v) = \begin{cases} s(v) & \text{if } v \in V, \\ \infty & \text{if } v \notin V. \end{cases} \tag{7.5.21}$$

Then

$$\bar{s}^*(w) = \sup_v \left\{ v^T w - \bar{s}(v) \right\}. \tag{7.5.22}$$

The function $\bar{s}$ being strictly convex on its effective domain V, its conjugate is closed, convex, proper [, Theorem 12.2], and differentiable [, Theorem 26.3], hence by convexity continuously differentiable[, Corollary 25.5.1]. However

$$\bar{f}(u) = h(u) + \bar{s}^*(-Du). \tag{7.5.23}$$

Thus, $\bar{f}$ is strongly convex and has continuous first derivative. The argument for strong concavity and the continuous differentiability of $\bar{g}$ is parallel. □

Now let us consider another algorithm. In the EFBS algorithm, no line searches are included. hence, we shall add line search techniques to the algorithm. Then, a modified extended forward-backward splitting algorithm can be described in the following:

MEFBS Algorithm:

Step 0 (Initial input). Give an initial point $(u^0, v^0) = (\tilde{u}^0, \tilde{v}^0)$, initial $\lambda_0 > 0$, initial metric H_0, initial approximation $T^0 = T_1^0 + T_2^0$, and the error bound ε. Let $k = 0$.

Step 1 (EFBS iteration). Use the EFBS scheme to get $(\tilde{u}^{k+1}, \tilde{v}^{k+1})$. That is,

$$(\tilde{u}^{k+1}, \tilde{v}^{k+1}) = (H_k + \lambda_k T_2^k)^{-1}(H_k - \lambda_k T_1^k)(\tilde{u}^k, \tilde{v}^k). \tag{7.5.24}$$

In other words, $(\tilde{u}^{k+1}, \tilde{v}^{k+1})$ is gotten through the following:

$$\hat{u}^k = H_k \tilde{u}^k - \lambda_k \nabla_u L_1^k(\tilde{u}^k, \tilde{v}^k), \quad \hat{v}^k = H_k \tilde{v}^k + \lambda_k \nabla_v L_1^k(\tilde{u}^k, \tilde{v}^k), \tag{7.5.25}$$

$$\begin{aligned}(\tilde{u}^{k+1}, \tilde{v}^{k+1}) = \operatorname*{argminimax}_{u\in U,\, v\in V} \Big\{ & L_2^k(u,v) \\ & + \frac{1}{2\lambda_k}\big[\|u - (H_k')^{-1}\hat{u}^k\|^2_{H_k'} - \|v - (H_k'')^{-1}\hat{v}^k\|^2_{H_k''}\big] \Big\},\end{aligned} \tag{7.5.26}$$

where H_k is the same as that in (7.5.10) for every $k \geq 0$.

Step 2 (Primal-dual gap check). Compute $f_k = \bar{f}(\tilde{u}^{k+1})$ and $g_k = \bar{g}(\tilde{v}^{k+1})$, where $\bar{f}$ and $\bar{g}$ are given by (7.5.11). If

$$f_k - g_k \leq \varepsilon,$$

then stop iterating and output the ε-optimal solution $(\hat{u}, \hat{v}) = (\tilde{u}^{k+1}, \tilde{v}^{k+1})$. Otherwise, go to next step.

Step 3 (Distance check). If

$$\|(\tilde{u}^{k+1}, \tilde{v}^{k+1}) - (\tilde{u}^k, \tilde{v}^k)\|_{H_k} \leq \varepsilon,$$

or the iteration number arrives a maximal iteration number requirement, stop iterating and output the ε-optimal solution $(\hat{u}, \hat{v}) = (\tilde{u}^{k+1}, \tilde{v}^{k+1})$. Otherwise, go to next step.

Step 4 (Line search). Minimize $\bar{f}(u)$ over the line segment $[u^k, \tilde{u}^{k+1}]$ to get u^{k+1}, and maximize $\bar{g}(v)$ over the line segment $[v^k, \tilde{v}^{k+1}]$ to get v^{k+1}.

Step 5 (Primal-dual gap check again). If $\bar{f}(u^{k+1}) - \bar{g}(v^{k+1}) \leq \varepsilon$, stop iterating and output the ε-optimal solution $(\hat{u}, \hat{v}) = (u^{k+1}, v^{k+1})$. Otherwise, choose suitable $\lambda_{k+1} > 0$, metric H_{k+1}, operator approximation $T^{k+1} = T_1^{k+1} + T_2^{k+1}$, and let k be increased by 1. Return then to Step 1. □

For the Lagrangian L in (7.1.2) of extended linear-quadratic programming, Line Search Assumption in Section 7.1 guarantees the implementation of the step 4 of the MEFBS algorithm. In fact, Proposition 7.5.3 and Theorem 7.5.4 are important for the step 4.

7.6. Strong Monotonicity of T_1 and T_2 in Problem (G)

Before we go to convergence of the two algorithms, we consider what the conditions are for the strong monotonicity of T_1 and T_2 in (7.5.2) with L given in (7.5.3). These conditions are important from Chapter 5 for convergence discussion of the algorithms of the preceding section.

Proposition 7.6.1. *If a projection splitting is applied, i.e., T_1 is given as follows:*

$$T_1(u,v) = \begin{pmatrix} \nabla_u L(u,v) \\ -\nabla_v L(u,v) \end{pmatrix} = \begin{pmatrix} \nabla h(u) - D^* v \\ \nabla s(v) + Du \end{pmatrix}, \tag{7.6.1}$$

then, T_1 is strongly monotone if and only if h and s are strongly convex. Furthermore, T_1^{-1} being strongly monotone implies that ∇h^{-1} and ∇s^{-1} are strongly monotone.

Proof. Since for any $(u',v'),(u,v) \in \mathcal{H}$, one has

$$\begin{aligned}
&\langle (u',v')-(u,v), T_1(u',v')-T_1(u,v)\rangle \\
&= \langle u'-u, \nabla h(u')-\nabla h(u) - D^*(v'-v)\rangle_1 \\
&\quad + \langle v'-v, \nabla s(v')-\nabla s(v) + D(u'-u)\rangle_2 \qquad (7.6.2)\\
&= \langle u'-u, \nabla h(u')-\nabla h(u)\rangle_1 + \langle v'-v, \nabla s(v')-\nabla s(v)\rangle_2 \\
&= \langle (u'-u, v'-v), (\nabla h(u')-\nabla h(u), \nabla s(v')-\nabla s(v))\rangle.
\end{aligned}$$

Obviously, if h and s are strongly convex, T_1 is strongly monotone from (7.6.2).

Conversely, if T_1 is strongly monotone, i.e., there is a $\alpha > 0$ such that

$$\langle (u',v')-(u,v), T_1(u',v')-T_1(u,v)\rangle \ge \alpha \|(u',v')-(u,v)\|^2 \qquad (7.6.3)$$

If we let $u'=u$ and $v'=v$ respectively, we obtain from (7.6.2) and (7.6.3)

$$\begin{aligned}
&\langle v'-v, \nabla s(v')-\nabla s(v)\rangle_2 \ge \alpha\|v'-v\|^2, \quad \forall v', v \in \mathcal{H}_2, \\
&\langle u'-u, \nabla h(u')-\nabla h(u)\rangle_1 \ge \alpha\|u'-u\|^2, \quad \forall u', u \in \mathcal{H}_1.
\end{aligned}$$

These proved that h and s are strongly convex.

Now, if T_1^{-1} is strongly monotone, that is, there is a $\beta > 0$ such that

$$\begin{aligned}
&\langle (u',v')-(u,v), T_1(u'v')-T_1(u,v)\rangle \\
&\ge \beta\|T_1(u',v')-T_1(u,v)\|^2, \quad \forall (u',v'),(u,v) \in \mathcal{H}.
\end{aligned} \qquad (7.6.4)$$

Let $u'=u$. Then $T_1(u,v')-T_1(u,v) = (-D^*(v'-v), \nabla s(v')-\nabla s(v))$, and hence (7.6.2) and (7.6.4) imply that

$$\langle v'-v, \nabla s(v')-\nabla s(v)\rangle_2 \ge \beta\|T_1(u,v')-T_1(u,v)\|^2 \ge \beta\|\nabla s(v')-\nabla s(v)\|^2.$$

Therefore, ∇s^{-1} is strongly monotone. Similarly $v'=v$ implies that ∇h^{-1} is strongly monotone. ◻

The strong monotonicity of ∇h^{-1} and ∇s^{-1} may not imply the strong monotonicity of T_1^{-1} generally except $D=0$.

Counter example: $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{R}$ and $h(u) = c_1 u$, $s(v) = c_2 v$. Then ∇h^{-1} and ∇s^{-1} are strongly monotone since

$$\nabla h(u')-\nabla h(u) \equiv 0, \quad \nabla s(v')-\nabla s(v) \equiv 0, \quad \forall\ u', u \in \mathcal{H}_1, v', v \in \mathcal{H}_2.$$

If we let $D = 1$, then $T_1(u', v') - T_1(u, v) = (v - v', u' - u)$. Thus there is no $\beta > 0$ such that (7.6.4) is true. i.e., T_1^{-1} from (7.6.1) is not strongly monotone.

The above example tell us that we need additional conditions to guarantee the inverse strong monotonicity of T_1 defined in (7.6.1).

Proposition 7.6.2. *If h and s are strongly convex and ∇h^{-1} and ∇s^{-1} are strongly monotone, then T_1^{-1} from (7.6.1) is strongly monotone.*

Proof. We want to prove that there is a $\beta > 0$ such that (7.6.4) is true. Here we should note that

$$T_1(u', v') - T_1(u, v) = \begin{pmatrix} \nabla h(u') - \nabla h(u) - D^*(v' - v), \\ \nabla s(v') - \nabla s(v) + D(u' - u) \end{pmatrix}. \tag{7.6.5}$$

Since D and D^* are bounded linear operators, there exists a $M > 0$ such that

$$\begin{aligned} &\|D^*(v' - v)\|^2 \le M\|v' - v\|^2, \quad \forall v', v \in \mathcal{H}_2, \\ &\|D(u' - u)\|^2 \le M\|u' - u\|^2, \quad \forall u', u \in \mathcal{H}_1. \end{aligned} \tag{7.6.6}$$

Hence

$$\begin{aligned} &\|\nabla h(u') - \nabla h(u) - D^*(v' - v)\|^2 \\ &= \|\nabla h(u') - \nabla h(u)\|^2 + \|D^*(v' - v)\|^2 \\ &\qquad - 2\langle \nabla h(u') - \nabla h(u), D^*(v' - v)\rangle_1 \\ &\le 2\|\nabla h(u') - \nabla h(u)\|^2 + 2\|D^*(v' - v)\|^2 \\ &\le 2\|\nabla h(u') - \nabla h(u)\|^2 + 2M\|v' - v\|^2. \end{aligned} \tag{7.6.7}$$

Likewise,

$$\|\nabla s(v') - \nabla s(v) + D(u' - u)\|^2 \le 2\|\nabla s(v') - \nabla s(v)\|^2 + 2M\|u' - u\|^2. \tag{7.6.8}$$

If h and s are strongly convex and ∇h^{-1} and ∇s^{-1} are strongly monotone, then there exist $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$ and $\beta_2 > 0$ such that

$$\begin{aligned} &\langle u' - u, \nabla h(u') - \nabla h(u)\rangle_1 \ge \alpha_1\|u' - u\|^2, \\ &\langle v' - v, \nabla s(v') - \nabla s(v)\rangle_2 \ge \alpha_2\|v' - v\|^2, \\ &\langle u' - u, \nabla h(u') - \nabla h(u)\rangle_1 \ge \beta_1\|\nabla h(u') - \nabla h(u)\|^2, \\ &\langle v' - v, \nabla s(v') - \nabla s(v)\rangle_2 \ge \beta_2\|\nabla s(u') - \nabla s(u)\|^2. \end{aligned}$$

Let $4\beta = \min\{\alpha_1, \alpha_2, \beta_1/M, \beta_2/M\}$.

Then the above inequalities and (7.6.2) give that

$$\begin{aligned}&\langle (u'-u, v'-v), T_1(u',v') - T_1(u,v)\rangle \\ &\qquad \geq 4\beta[\|\nabla h(u') - \nabla h(u)\|^2 + \|\nabla s(v') - \nabla s(v)\|^2],\\ &\langle (u'-u, v'-v), T_1(u',v') - T_1(u,v)\rangle \\ &\qquad \geq 4M\beta[\|u'-u\|^2 + \|v'-v\|^2].\end{aligned}$$

Adding the both sides together respectively and dividing them by 2, one obtains (7.6.4) from (7.6.5), (7.6.7) and (7.6.8). □

If $D = 0$ in (7.6.1), then T_1^{-1} is strongly monotone if and only if ∇h^{-1} and ∇s^{-1} are strongly monotone.

For the special case of (7.1.2), we have the following:

Proposition 7.6.3. *Suppose h and s are given by (7.5.5), and $P \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{m\times m}$. Then, h, or s, is strongly convex if and only if the matrix P, or the matrix Q, is positive definite. Furthermore, h and s being strongly convex implies ∇h^{-1} and ∇s^{-1} being strongly monotone respectively.*

Proof. Since $\nabla h(u) = Pu$, one has h strongly convex if and only if there exists $\alpha > 0$ such that $u{\cdot}Pu \geq \alpha u{\cdot}u = \alpha\|u\|^2$ for all $u \in \mathbb{R}^m$. This is true if and only if P is positive definite.

Furthermore, h is strongly convex if and only if P is positive definite, which is true if and only if P^{-1} is positive definite. This implies there exists $\beta > 0$ such that

$$\langle u, \nabla h(u)\rangle = \langle u, Pu\rangle = \langle P^{-1}Pu, Pu\rangle \geq \beta\|Pu\|^2.$$

This implies that ∇h^{-1} is strongly monotone.

It is similar to prove the conclusions for s and Q. □

Proposition 7.6.4. *Suppose h is given by (7.5.5) and $P \in \mathbb{R}^{n\times n}$. If λ_1 and λ_m are the smallest and the largest eigenvalues of P respectively, and $\lambda_1 > 0$, then λ_1 is a modulus for h being strongly convex and λ_m^{-1} is a modulus for ∇h^{-1} being strongly monotone.*

Proof. They are obvious from the proof of Lemma 7.6.3. □

Now let us consider another splitting strategy of $T = T_1 + T_2$ as follows:

Proposition 7.6.5. *If $U \neq \emptyset$, $V \neq \emptyset$, and*

$$T_1(u,v) = \begin{pmatrix} -D^*v \\ Du \end{pmatrix}, \qquad T_2(u,v) = \begin{pmatrix} \nabla h(u) + \mathcal{N}_U(u) \\ \nabla s(v) + \mathcal{N}_V(v) \end{pmatrix}, \tag{7.6.9}$$

then, T_1 satisfies the equality (4.2.1), and T_2 is strongly monotone in $U \times V$ if and only if h and s are strongly convex in U and V respectively.

Proof. Since for any $(u', v'), (u, v) \in U \times V$, one has

$$\begin{aligned} &\langle (u', v') - (u, v), w' - w\rangle \\ &= \langle u' - u, \nabla h(u') - \nabla h(u)\rangle_1 + \langle v' - v, \nabla s(v') - \nabla s(v)\rangle_2 \\ &= \langle (u' - u, v' - v), (\nabla h(u') - \nabla h(u), \nabla s(v') - \nabla s(v))\rangle. \end{aligned} \tag{7.6.10}$$

Obviously, if h and s are strongly convex in U and V, T_2 is strongly monotone from (7.6.10). Where w' and w are any elements of $T_2(u', v')$ and $T_2(u, v)$ respectively.

Conversely, if T_2 is strongly monotone in $U \times V$, i.e., there is a $\alpha > 0$ such that

$$\langle (u', v') - (u, v), w' - w\rangle \geq \alpha \|(u', v') - (u, v)\|^2, \tag{7.6.11}$$

then, when $w' = \big(\nabla h(u'), \nabla s(v')\big)$ and $w = \big(\nabla h(u), \nabla s(v)\big)$, (7.6.11) becomes

$$\begin{aligned} &\langle (u', v') - (u, v), \big(\nabla h(u'), \nabla s(v')\big) - \big(\nabla h(u), \nabla s(v)\big)\rangle \\ &\qquad \geq \alpha \|(u', v') - (u, v)\|^2. \end{aligned} \tag{7.6.12}$$

If we let $u' = u$ and $v' = v$ respectively in (7.6.12), we obtain

$$\langle v' - v, \nabla s(v') - \nabla s(v)\rangle_2 \geq \alpha \|v' - v\|^2, \forall v', v \in V,$$

$$\langle u' - u, \nabla h(u') - \nabla h(u)\rangle_1 \geq \alpha \|u' - u\|^2, \forall u', u \in U.$$

These proved that h and s are strongly convex in U and V respectively. $\square$

Corollary 7.6.6. *Let h and s be given by (7.5.5). Then, under the splitting strategy of (7.6.9), T_2 being strongly monotone in $U \times V$ is the same as P and Q being strongly monotone, or positive definite in an Euclidean space, in U and V respectively.*

Proof. It is elementary from (7.5.5) and Proposition 7.6.5. $\square$

7.7. Convergence of the Algorithms and Rate of Convergence

Let us consider the convergence of the algorithms in Section 7.5 in this section. We only consider the special case of $T_1^k \equiv T_1$ and $T_2^k \equiv T_2$ here. Of course, we are only interested in the case of $U \neq \emptyset$ and $V \neq \emptyset$.

Lemma 7.7.1. *Let $h : \mathcal{H}_1 \to \mathbb{R}$ be in $\mathcal{C}^2$ (twice continuously differentiable). Then, h is strongly convex with modulus α on $U \subset \mathcal{H}_1$ if and only if linear operators in $\{\nabla^2 h(u) \,|\, u \in U\}$ are uniformly strongly monotone with modulus α.*

Proof. (Necessity:) If h is strongly convex on U with α, then

$$\langle \nabla h(u') - \nabla h(u), u' - u \rangle \geq \alpha \|u' - u\|^2, \quad \forall u, u' \in U. \tag{7.7.1}$$

Considering the Taylor expansion of $\nabla h(u')$ at u, we obtain

$$\nabla h(u') = \nabla h(u) + \nabla^2 h(u)(u' - u) + o\big(\|u' - u\|\big)e \tag{7.7.2}$$

where e satisfies $e{\cdot}e = 1$ and $e{\cdot}(u' - u) = O\big(\|u' - u\|\big)$. Hence, (7.7.1) and (7.7.2) imply

$$\langle u' - u, \nabla^2 h(u)(u' - u) \rangle + o\big(\|u' - u\|^2\big) \geq \alpha \|u' - u\|^2. \tag{7.7.3}$$

So for any $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\frac{|o\big(\|u' - u\|^2\big)|}{\|u' - u\|^2} \leq \varepsilon, \quad \text{when} \quad \|u' - u\| < \delta.$$

That is,

$$\langle u' - u, \nabla^2 h(u)(u' - u) \rangle \geq (\alpha - \varepsilon)\|u' - u\|^2, \quad \text{if} \quad \|u' - u\| < \delta. \tag{7.7.4}$$

When $\|u' - u\| \geq \delta$, (7.7.4) is still true since $\nabla^2 h(u)$ is a linear operator for any $u \in U$ and there always is a constant N such that $\|u' - u\| < N\delta$. Thus, $\nabla^2 h(u)$ is strongly monotone with modulus $\alpha - \varepsilon$ for any $u \in U$ and $\varepsilon > 0$. Therefore, $\{\nabla^2 h(u) \,|\, u \in U\}$ is uniformly strongly monotone with modulus α.

(Sufficiency:) If $\{\nabla^2 h(u) \,|\, u \in U\}$ is uniformly strongly monotone with modulus α, then

$$\langle v, \nabla^2 h(u) v \rangle \geq \alpha \|v\|^2, \qquad \text{for all} \quad u, v \in U. \tag{7.7.5}$$

From the above, the mean value theorem and Taylor expansion, we have for any $u' \in U$ and $u \in U$ that

$$\begin{aligned} h(u') - h(u) &= \langle \nabla h(u), u' - u\rangle + \frac{1}{2}\langle u' - u, \nabla^2 h(u'')(u' - u)\rangle \\ &\geq \langle \nabla h(u), u' - u\rangle + \frac{\alpha}{2}\|u' - u\|^2, \\ h(u) - h(u') &= \langle \nabla h(u'), u - u'\rangle + \frac{1}{2}\langle u - u', \nabla^2 h(\tilde{u})(u - u')\rangle \\ &\geq \langle \nabla h(u'), u - u'\rangle + \frac{\alpha}{2}\|u' - u\|^2, \end{aligned} \tag{7.7.6}$$

where u'' and $\tilde{u}$ are two points in the segment $[u', u]$. We add the two inequalities in (7.7.6) to get

$$0 \geq \langle \nabla h(u') - \nabla h(u), u - u'\rangle + \alpha\|u' - u\|^2. \tag{7.7.7}$$

This implies (7.7.1). Therefore, h is strongly convex with modulus α. □

Theorem 7.7.2 (convergence comparison). *Let sequences $\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ be generated from the MEFBS algorithm. Then,*

$$0 \quad \leq \quad \bar{f}(u^k) - \bar{g}(v^k) \quad \leq \quad \bar{f}(\tilde{u}^k) - \bar{g}(\tilde{v}^k) \tag{7.7.8}$$

for all $k \geq 0$. Suppose L in $(\mathcal{G})$ is twice continuously differentiable in u and v, and is also strongly convex-concave on $U \times V$. Then, the sequences $\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ will strongly converge to the unique solution of the $(\mathcal{G})$ if $\gamma_k = \bar{f}(\tilde{u}^k) - \bar{g}(\tilde{v}^k) \to 0$. Moreover, the convergence rates of the sequences will be linear if $\gamma_k \to 0$ linearly.

Proof. The step 4 of the MEFBS algorithm implies

$$\bar{f}(u^k) \leq \min\{\bar{f}(u^{k-1}), \bar{f}(\tilde{u}^k)\}, \qquad \bar{g}(v^k) \geq \max\{\bar{g}(v^{k-1}), \bar{g}(\tilde{v}^k)\}. \tag{7.7.9}$$

Thus, the above and (7.5.11) give (7.7.8).

Now the strong convexity-concavity of L implies the strong monotonicity of T in (7.5.1). Hence there is a unique solution $(\bar{u}, \bar{v})$ of $(\mathcal{G})$. That is, $\bar{f}(\bar{u}) = L(\bar{u}, \bar{v}) = \bar{g}(\bar{v})$. Since

$$\begin{aligned} &\bar{f}(u^k) - \bar{f}(\bar{u}) \geq L(u^k, \bar{v}) - L(\bar{u}, \bar{v}) \\ &= \nabla_u L(\bar{u}, \bar{v}) \cdot (u^k - \bar{u}) + \frac{1}{2}(u^k - \bar{u}) \cdot \nabla^2_{uu} L(u', \bar{v})(u^k - \bar{u}), \end{aligned} \tag{7.7.10}$$

where u' is a point of segment $[u^k, \bar{u}]$. Inasmuch as $\bar{u}$ minimizes $L(u, \bar{v})$ over $u \in U$, we have $\nabla_u L(\bar{u}, \bar{v}) \cdot (u^k - \bar{u}) \geq 0$ for all $u \in U$. If $\alpha > 0$ is the modulus of the strong convexity-concavity of L on $U \times V$, then applying Lemma 7.7.1 in (7.7.10) we obtain

$$\bar{f}(u^k) - \bar{f}(\bar{u}) \geq \frac{1}{2}(u^k - \bar{u}) \cdot \nabla^2_{uu} L(u', \bar{v})(u^k - \bar{u}) \geq \frac{\alpha}{2}\|u^k - \bar{u}\|^2.$$

By a parallel argument,

$$\bar{g}(v^k) - \bar{g}(\bar{v}) \leq -\frac{\alpha}{2}\|v^k - \bar{v}\|^2.$$

We subtract this inequality from the preceding one to get

$$\frac{\alpha}{2}\|u^k - \bar{u}\|^2 + \frac{\alpha}{2}\|v^k - \bar{v}\|^2 \leq \bar{f}(u^k) - \bar{g}(v^k). \tag{7.7.11}$$

Through a same argument,

$$\frac{\alpha}{2}\|\tilde{u}^k - \bar{u}\|^2 + \frac{\alpha}{2}\|\tilde{v}^k - \bar{v}\|^2 \leq \bar{f}(\tilde{u}^k) - \bar{g}(\tilde{v}^k). \tag{7.7.12}$$

(7.7.8), (7.7.11) and (7.7.12) yield the strong convergence of two sequences if $\gamma_k \to 0$.

If $\gamma_k \to 0$ linearly, then there are a scalar $\theta \in [0, 1)$, an index $\bar{k}$ and a constant $C_{\bar{k}}$ from Definition 4.2.1 such that

$$|\gamma_k| \leq \theta^{k-\bar{k}} C_{\bar{k}} \qquad \text{for all} \quad k \geq \bar{k}. \tag{7.7.13}$$

(7.7.8), (7.7.11), (7.7.12) and the above imply

$$\|(u^k, v^k) - (\bar{u}, \bar{v})\| \leq \bar{\theta}^{k-\bar{k}} \bar{C}_{\bar{k}}, \quad \|(\tilde{u}^k, \tilde{v}^k) - (\bar{u}, \bar{v})\| \leq \bar{\theta}^{k-\bar{k}} \bar{C}_{\bar{k}}, \quad k \geq \bar{k}, \tag{7.7.14}$$

where

$$\bar{\theta} = \sqrt{\theta} \in [0, 1), \qquad \bar{C}_{\bar{k}} = \sqrt{\frac{2}{\alpha} C_{\bar{k}}}.$$

(7.7.14) yields the linear convergence of the two sequences to $(\bar{u}, \bar{v})$. □

From now on in this chapter, we shall always assume that $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$ is an Euclidean space.

Theorem 7.7.3 (global convergence). *Let $H_k \equiv H$. Suppose L in (7.5.3) is strongly convex-concave and one of the three conditions in Part (b) of Theorem 5.4.2 is satisfied for T_1, T_2 and $\{\lambda_k\}$. Then, in EFBS Algorithm (hence also in MEFBS Algorithm) with optimality bound $\varepsilon > 0$, termination must come with ε-optimal solution $(\hat{u}, \hat{v})$ in just a finite number of iterations. With $\varepsilon = 0$, unless the procedure happens to terminate with the exact optimal solution $\bar{u}$ and $\bar{v}$ in a finite number of iterations, the sequences generated will be such that $u^k \to \bar{u}$ and $v^k \to \bar{v}$ (hence $\tilde{u}^k \to \bar{u}$ and $\tilde{v}^k \to \bar{v}$ as well in MEFBS Algorithm) as $k \to \infty$ and the rate is at least linear.*

Proof. Theorem 5.4.2 implies that the above conclusions are true evidently for the EFBS algorithm.

As for the MEFBS algorithm, we only need to consider $\varepsilon = 0$. Since the sequence $\{(\tilde{u}^k, \tilde{v}^k)\}$ is generated directly from the EFBS algorithm, its Q-linear convergence to $(\bar{u}, \bar{v})$ is obvious from Theorem 5.4.2. Hence γ_k in Theorem 7.7.2 converges to 0 as $k \to \infty$. Now, let $H = \text{diag}\{H_1, H_2\}$. Then from Theorem 7.5.4 and Taylor expansion theory,

$$\bar{f}(\tilde{u}^k) - \bar{f}(\bar{u}) = \nabla \bar{f}(\bar{u})^T(\tilde{u}^k - \bar{u}) + o(\|\tilde{u}^k - \bar{u}\|_{H_1}). \tag{7.7.15}$$

Thus, for any constant $M' > 0$, there exists an index I_1 such that

$$o(\|\tilde{u}^k - \bar{u}\|_{H_1}) \leq M'\|\tilde{u}^k - \bar{u}\|_{H_1} \quad \text{when} \quad k \geq I_1.$$

Hence,

$$\bar{f}(\tilde{u}^k) - \bar{f}(\bar{u}) \leq (\|\nabla \bar{f}(\bar{u})\|_{H_1^{-1}} + M')\|\tilde{u}^k - \bar{u}\|_{H_1} \quad \text{when} \quad k \geq I_1. \tag{7.7.16}$$

In parallel, for some $M'' > 0$ and index I_2, one has

$$\bar{g}(\bar{v}) - \bar{g}(\tilde{v}^k) \leq (\|\nabla \bar{g}(\bar{v})\|_{H_2^{-1}} + M'')\|\tilde{v}^k - \bar{v}\|_{H_2} \quad \text{when} \quad k \geq I_2. \tag{7.7.17}$$

$\bar{f}(\bar{u}) = \bar{g}(\bar{v})$ implies

$$\gamma_k = \bar{f}(\tilde{u}^k) - \bar{g}(\tilde{v}^k) \leq \bar{M}\|(\tilde{u}^k, \tilde{v}^k) - (\bar{u}, \bar{v})\|_H \quad \text{when} \quad k \geq I_0, \tag{7.7.18}$$

where $I_0 = \max\{I_1, I_2\}$ and

$$\bar{M} = 2\max\{\|\nabla\bar{f}(\bar{u})\|_{H_1^{-1}} + M', \quad \|\nabla\bar{g}(\bar{v})\|_{H_2^{-1}} + M''\}.$$

However, from the proof of Theorem 5.4.2 there is a $\theta \in [0,1)$ such that

$$\|(\tilde{u}^k, \tilde{v}^k) - (\bar{u}, \bar{v})\|_H \leq \theta^k \|(\tilde{u}^0, \tilde{v}^0) - (\bar{u}, \bar{v})\|_H \quad \text{for all} \quad k \geq 0. \tag{7.7.19}$$

This and (7.7.18) imply that $\gamma_k \to 0$ at least linearly as $k \to \infty$. Therefore, Theorem 7.7.2 yields the linear convergence rates of $\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ to $(\bar{u}, \bar{v})$. □

Now let us consider the two splitting strategies in Section 7.6 and the Lagrangian L of extended linear-quadratic programming. Firstly, we discuss the projection splitting (7.6.1).

Theorem 7.7.4 (projection splitting). *Assume that L is given by (7.1.2), which is at the fully quadratic case, and the projection splitting (7.6.1) is used. Let $H = \operatorname{diag}\{P, Q\} = S^2$.*

(a) The least Lipschitz constant M and the largest strong monotone modulus α of $S^{-1}T_1S^{-1}$ are respectively

$$\alpha = 1, \qquad M = \left\| \begin{pmatrix} I & -\bar{D}^T \\ \bar{D} & I \end{pmatrix} \right\| = \sqrt{1 + \|\bar{D}\|^2}, \tag{7.7.20}$$

where $\bar{D} = Q^{-\frac{1}{2}} D P^{-\frac{1}{2}}$.

(b) If $\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ are generated from the MEFBS algorithm with $\varepsilon = 0$, then when

$$0 < \inf_k\{\lambda_k\} \leq \sup_k\{\lambda_k\} < \frac{2}{1 + \|\bar{D}\|^2}, \tag{7.7.21}$$

$\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ will at least linearly converge to the unique saddle point $(\bar{u}, \bar{v})$ of $(\mathcal{G})$. Furthermore, there exists a constant $\tilde{M}$ such that for all $k \geq 0$,

$$\begin{aligned} &\frac{1}{2}\|(u^k, v^k) - (\bar{u}, \bar{v})\|_H^2 \leq f(u^k) - g(v^k) \\ \leq &f(\tilde{u}^k) - g(\tilde{v}^k) \leq \tilde{M}\|(\tilde{u}^k, \tilde{v}^k) - (\bar{u}, \bar{v})\|_H. \end{aligned} \tag{7.7.22}$$

Proof. At the case of extended linear-quadratic programming, we have

$$(S^{-1}T_1S^{-1})(u',v') = \begin{pmatrix} I & -\bar{D}^T \\ \bar{D} & I \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} + S^{-1} \begin{pmatrix} p \\ -q \end{pmatrix}. \tag{7.7.23}$$

Hence,

$$\begin{aligned} &(S^{-1}T_1S^{-1})(u,v) - (S^{-1}T_1S^{-1})(u',v') \\ &= \begin{pmatrix} I & -\bar{D}^T \\ \bar{D} & I \end{pmatrix} \left[\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} u' \\ v' \end{pmatrix} \right]. \end{aligned} \tag{7.7.24}$$

Therefore, the largest strong monotone modulus and the least Lipschitz constant of $S^{-1}T_1S^{-1}$ are respectively $\alpha = 1$ and

$$M = \left\| \begin{pmatrix} I & -\bar{D}^T \\ \bar{D} & I \end{pmatrix} \right\|.$$

That is, for $z^T = (u^T, v^T)$

$$\begin{aligned} M^2 &= \max_{\|z\|=1} \left\| \begin{pmatrix} I & -\bar{D}^T \\ \bar{D} & I \end{pmatrix} z \right\|^2 \\ &= \max_{\|z\|=1} \left\{ \|u - \bar{D}^T v\|^2 + \|\bar{D}u + v\|^2 \right\} \\ &= \max_{\|z\|=1} \left\{ 1 + \|\bar{D}^T v\|^2 + \|\bar{D}u\|^2 \right\} \\ &= 1 + \max_{\|z\|=1} \left\| \begin{pmatrix} & \bar{D}^T \\ -\bar{D} & \end{pmatrix} z \right\|^2 \\ &= 1 + \|\bar{D}\|^2. \qquad (\text{ Proposition 6.4.4 }) \end{aligned}$$

Thus, Part (a) is proved.

As to Part (b), when $\{\lambda_k\}_{k=0}^{\infty}$ satisfies (7.7.21), the assumptions in the theorem and (7.7.20) imply that the condition (b.1) in Theorem 5.4.2 is satisfied. Hence, it is elementary from Theorem 7.7.3 for the linear convergence of the two sequences generated by the MEFBS algorithm. For the inequalities in (7.7.22), they will be easily proved by similar ideas in the proofs of Theorem 7.7.2 and Theorem 7.7.3. Firstly, follow the proof of Theorem 7.7.3 until (7.7.18) (now $\bar{f} \equiv f$ and $\bar{g} \equiv g$). For all $k < I_0$, suppose $(\tilde{u}^k, \tilde{v}^k) \neq (\bar{u}, \bar{v})$ (otherwise the algorithm terminated already). Let

$$\tilde{M} = \max \left\{ \bar{M}, \frac{f(\tilde{u}^k) - g(\tilde{v}^k)}{\|(\tilde{u}^k, \tilde{v}^k) - (\bar{u}, \bar{v})\|_H} \middle| k = 0, 1, \cdots, I_0 - 1 \right\}.$$

Therefore, (7.7.18) and the above give

$$f(\tilde{u}^k) - g(\tilde{v}^k) \le \tilde{M}\|(\tilde{u}^k, \tilde{v}^k) - (\bar{u}, \bar{v})\|_H \tag{7.7.25}$$

for all $k \ge 0$. On the other hand, follow the proof of Theorem 7.7.2 until (7.7.10). Since $\nabla^2_{uu} L(u', v) = P$ for any u' and v now, we have

$$f(u^k) - f(\bar{u}) \ge \nabla_u L(\bar{u}, \bar{v})^T (u^k - \bar{u}) + \frac{1}{2}\|u^k - \bar{u}\|_P^2 \ge \frac{1}{2}\|u^k - \bar{u}\|_P^2.$$

Similarly,

$$g(v^k) - g(\bar{v}) \le -\frac{1}{2}\|v^k - \bar{v}\|_Q^2.$$

Subtract this inequality from the preceding one to obtain

$$\begin{aligned} \frac{1}{2}\|(u^k, v^k) - (\bar{u}, \bar{v})\|_H^2 &= \frac{1}{2}\|u^k - \bar{u}\|_P^2 + \frac{1}{2}\|v^k - \bar{v}\|_Q^2 \\ &\le f(u^k) - g(v^k). \end{aligned} \tag{7.7.26}$$

This and (7.7.25), (7.7.8) yield (7.7.22). Part (b) is proved. □

We call a splitting *basic decomposition splitting* if it has a form of (7.6.9).

Theorem 7.7.5 (basic decomposition splitting). *Assume that L is given by (7.1.2), which is at the fully quadratic case, and the splitting (7.6.9) is used. Let $H = \text{diag}\{P, Q\} = S^2$.*

(a) *The least Lipschitz constant M of $S^{-1}T_1S^{-1}$ and the largest strong monotone modulus α of $S^{-1}T_2S^{-1}$ are respectively*

$$M = \|\bar{D}\| = \left\|\sqrt{Q^{-1}}D\sqrt{P^{-1}}\right\|, \qquad \alpha = 1. \tag{7.7.27}$$

(b) *If $\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ are generated from the MEFBS algorithm with $\varepsilon = 0$, then $\{(u^k, v^k)\}$ and $\{(\tilde{u}^k, \tilde{v}^k)\}$ will at least linearly converge to the unique saddle point $(\bar{u}, \bar{v})$ of $(\mathcal{G})$ when*

$$0 < \inf_k\{\lambda_k\} \le \sup_k\{\lambda_k\} < \begin{cases} \frac{2}{M^2-1} & \text{if } M > 1, \\ \infty & \text{if } M \le 1, \end{cases} \tag{7.7.28}$$

where M is given by (7.7.27). Furthermore, there exists a constant $\tilde{M}$ such that (7.7.22) holds for all $k \geq 0$.

Proof. At the case of (7.6.9) in extended linear-quadratic programming, we have

$$(S^{-1}T_1S^{-1})(u',v') = \begin{pmatrix} & -\bar{D}^T \\ \bar{D} & \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} \tag{7.7.29}$$

and

$$(S^{-1}T_2S^{-1})(u',v') = I\begin{pmatrix} u' \\ v' \end{pmatrix} + S^{-1}\begin{pmatrix} p \\ -q \end{pmatrix} + \begin{pmatrix} \mathcal{N}_{U'}(u') \\ \mathcal{N}_{V'}(v') \end{pmatrix} \tag{7.7.30}$$

because

$$\begin{aligned} R^{-1}\mathcal{N}_W(w) &= \{R^{-1}y \big| y^T(x-w) \leq 0 \quad \forall \;\; x \in W\} \\ &= \{R^{-1}y \big| (R^{-1}y)^T(Rx - Rw) \leq 0 \quad \forall \;\; x \in W\} \\ &= \{y' \big| y'^T(x'-w') \leq 0 \quad \forall \;\; x' \in RW\} \\ &= \mathcal{N}_{RW}(w'). \end{aligned} \tag{7.7.31}$$

Where

$$U' = \sqrt{P}U = \{\sqrt{P}u \big| u \in U\}, \qquad V' = \sqrt{Q}V = \{\sqrt{Q}v \big| v \in V\}.$$

Equation (7.7.29) and Proposition 6.4.4 imply the least Lipschitz constant M of $S^{-1}T_1S^{-1}$ is given in (7.7.27), while (7.7.30) and the monotonicity of $\mathcal{N}_{U'}$ and $\mathcal{N}_{V'}$ give largest strong monotone modulus $\alpha = 1$ of $S^{-1}T_2S^{-1}$. Hence, Part (a) is proved.

As for Part (b), when $\{\lambda_k\}_{k=0}^{\infty}$ satisfies (7.7.28), the assumptions in the theorem and (7.7.27) imply that the condition (b.2) in Theorem 5.4.2 is satisfied. Hence, it is elementary from Theorem 7.7.3 for the linear convergence of the two sequences generated by the MEFBS algorithm. For the inequalities in (7.7.22), the exact same argument as that in the proof of Theorem 7.7.4 will work because the proof does not depend on the splitting $T = T_1 + T_2$. Therefore, Part (b) is proved. □

7.8. Details of the Algorithm Implementation for the Multistage Models

In this section, we shall discuss some details of the MEFBS algorithm implementation for the multistage models in Section 7.4.

As for the continuous-time models in Section 7.3, we can discretize them into the multistage models through the following: an Euler discretization of problem $(\mathcal{SP})$ can be obtained by dividing the time interval into N equal intervals, with $t^i = t_0 + t_1(i-1)/N$, $i = 1, \cdots, N+1$, and finding $x_i \approx x(t^i)$, $y_i \approx y(t^i)$, $u_i \approx u(t^i)$, $v_i \approx v(t^i)$, $A_i \approx A(t^i)$, $P_i \approx P(t^i)$, and so on, by solving

$$\begin{aligned}
&\operatorname*{minimax}_{u_i, v_i} \left\{ \triangle t \sum_{i=1}^{N} J_i(u_i, v_{i+1}) + J_e(u_0, v_{N+2}) - [u, v]_0 \right\}, \\
&\text{subject to} \\
&u_0 \in U_e, u_i \in U_i \approx U(t^i), \quad v_i \in V_i \approx V(t^i), \ v_{N+2} \in V_e, \\
&x_{i+1} = x_i + \triangle t \big[A_i x_i + B_i u_i\big], \quad i = 1, \cdots, N, \quad x_1 = B_e u_0, \\
&y_{i-1} = y_i + \triangle t \big[A_i^T y_i + C_i^T v_i\big], \quad i = N+1, \cdots, 2, \quad y_{N+1} = C_e^T v_{N+2},
\end{aligned} \tag{7.8.1}$$

where $\triangle t = (t_1 - t_0)/N$ and

$$\begin{aligned}
J_i(u, v) &= \bar{p}_i^T u + \bar{q}_{i+1}^T v + \tfrac{1}{2} u^T P_i u - \tfrac{1}{2} v^T Q_{i+1} v - v^T D_i u, \\
J_e(u, v) &= \bar{p}_e^T u + \bar{q}_e^T v + \tfrac{1}{2} u^T P_e u - \tfrac{1}{2} v^T Q_e v - v^T D_e u, \\
[u, v]_0 &= \triangle t \sum_{i=0}^{N} y_{i+1}^T B_i u_i = \triangle t \sum_{i=2}^{N+2} v_i^T C_i x_{i-1}
\end{aligned} \tag{7.8.2}$$

and $u_0 = u_e$, $v_{N+2} = v_e$, $B_0 = B_e$, $C_{N+2} = C_e$. A solution that is accurate to $O(N^{-1})$ is obtained. Sometimes we use other discretizations which are accurate to $O(N^{-1})$ or $O(N^{-2})$ for some examples. When $D_e = 0$, or $u_e = 0$, or $v_e = 0$, problem (7.8.1) is a multistage problem similar to $(\mathcal{SP}_d)$.

Now let consider the multistage models in Section 7.4. That is, we only need to consider the following Lagrangian:

$$\begin{aligned}
L(u, v) = &\sum_{t=0}^{N} \big[p_t \cdot u_t + \tfrac{1}{2} u_t \cdot P_t u_t\big] - \sum_{t=1}^{N} v_t \cdot D_t u_t \\
&+ \sum_{t=1}^{N+1} \big[q_t \cdot v_t - \tfrac{1}{2} v_t \cdot Q_t v_t\big] - l_0(u, v),
\end{aligned} \tag{7.8.3}$$

subject to the dynamical systems in (7.4.2) and $u = (u_0, \cdots, u_N) \in U$, $v = (v_1, \cdots, v_{N+1}) \in V$ with

$$U = U_0 \times U_1 \times \cdots \times U_N, \qquad V = V_1 \times V_2 \times \cdots \times V_{N+1}, \tag{7.8.4}$$

where

$$l_0(u, v) = \sum_{t=1}^{N+1} x_{t-1} \cdot C_t^T v_t = \sum_{t=0}^{N} y_{t+1} \cdot B_t u_t. \tag{7.8.5}$$

At the fully quadratic case, P_t and Q_t are positive definite. Thus, from the discussion in Chapter 6, we may choose $H_k \equiv H = \mathrm{diag}\{P_0, \cdots, P_N, Q_1, \cdots, Q_{N+1}\}$ and $T = T_1 + T_2$ in (7.5.2) as the following basic decomposition splitting:

$$L_1(u, v) = -\sum_{t=1}^{N} v_t \cdot D_t u_t - l_0(u, v); \quad L_2(u, v) = L(u, v) - L_1(u, v). \tag{7.8.6}$$

Then, T_1 and $T_2 = \hat{T}_2 + \mathcal{N}_U \times \mathcal{N}_V$ are given by

$$T_1(u, v) = \begin{pmatrix} -B_0^T y_1 \\ -D_1^T v_1 - B_1^T y_2 \\ \vdots \\ -D_N^T v_N - B_N^T y_{N+1} \\ D_1 u_1 + C_1 x_0 \\ \vdots \\ D_N u_N + C_N x_{N-1} \\ C_{N+1} x_N \end{pmatrix}, \quad \hat{T}_2(u, v) = \begin{pmatrix} P_0 u_0 + p_0 \\ P_1 u_1 + p_1 \\ \vdots \\ P_N u_N + p_N \\ Q_1 v_1 - q_1 \\ \vdots \\ Q_N v_N - q_N \\ Q_{N+1} v_{N+1} - q_{N+1} \end{pmatrix}. \tag{7.8.7}$$

Theorem 7.8.1 (parallel structure of the algorithms). *Under the choices of the above, at each iteration, the step 1 of the MEFBS algorithm (hence EFBS algorithm) can be implemented by the following parallel structure:*

$$\begin{aligned} \tilde{u}_t^{k+1} &= \operatorname*{argmin}_{u \in U_t} \left[p_t^T u + \tfrac{1}{2} u^T P_t u + \frac{1}{2\lambda_k} \| u - P_t^{-1} \hat{u}_t^k \|^2 \right], \\ \hat{u}_t^k &= P_t \tilde{u}_t^k + \lambda_k [D_t^T \tilde{v}_t^k + B_t^T \tilde{y}_{t+1}^k], \quad D_0 = 0, \quad t = 0, 1, \ldots, N; \end{aligned} \tag{7.8.8}$$

and

$$\begin{aligned} \tilde{v}_t^{k+1} &= \operatorname*{argmin}_{v \in V_t} \left[\tfrac{1}{2} v^T Q_t v + \frac{1}{2\lambda_k} \| v - Q_t^{-1} \hat{v}_t^k \|^2 - q_t^T v \right], \\ \hat{v}_t^k &= Q_k \tilde{v}_t^k - \lambda_k [D_t \tilde{u}_t^k + C_t \tilde{x}_{t-1}^k], \quad D_{N+1} = 0, \quad t = 1, \ldots, N+1. \end{aligned} \tag{7.8.9}$$

Where $\{\tilde{y}_t^k\}$ and $\{\tilde{x}_t^k\}$ are the solutions of the dynamical systems in (7.4.2) associated with given $\tilde{u}^k$ and $\tilde{v}^k$. Moreover, the equalities in (7.8.8) and (7.8.9) will become a set of one-dimensional subproblems if (7.8.3) and (7.8.4) are box-diagonal.

Proof. They are elementary from (7.5.25), (7.5.26), (7.8.3), (7.8.5) and (7.8.7). □

We have discussed the choices of H_k and $T = T_1 + T_2$. We shall discuss the choice of the step sizes $\{\lambda_k\}$ for the MEFBS algorithm. From the results of Proposition 6.3.1, we need to estimate the Lipschitz constant M of $\sqrt{H^{-1}}T_1\sqrt{H^{-1}}$ and the strong monotone modulus α of $\sqrt{H^{-1}}T_2\sqrt{H^{-1}}$.

Theorem 7.8.2 (step size choice and constant estimate). *Under the choices of (7.8.6) and the above $H_k \equiv H$, if $S = \sqrt{H}$, then*
(a) $S^{-1}T_1S^{-1}$ and $S^{-1}T_2S^{-1}$ are respectively given by (7.7.29) and (7.7.30) with

$$\bar{D} = \sqrt{Q}^{-1} D \sqrt{P}^{-1} + \sqrt{Q}^{-1} C \bar{G}^{-1} B \sqrt{P}^{-1}, \tag{7.8.10}$$

where $Q = \mathrm{diag}\{Q_1, \cdots, Q_{N+1}\}$, $P = \mathrm{diag}\{P_0, \cdots, P_N\}$, $B = \mathrm{diag}\{B_0, \cdots, B_N\}$, $C = \mathrm{diag}\{C_1, \cdots, C_{N+1}\}$, $\bar{G}$ is given in (3.1.4) and

$$D = \begin{pmatrix} 0 & D_1 & & & \\ & 0 & D_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & D_N \\ & & & & 0 \end{pmatrix}; \tag{7.8.11}$$

(b) an estimate M for the Lipschitz constant $\|\bar{D}\|$ of $S^{-1}T_1S^{-1}$ and the largest strong monotone modulus α of $S^{-1}T_2S^{-1}$ are respectively $\alpha = 1$ and

$$\begin{aligned} \|\bar{D}\| \leq M = &\left[\max_{1\leq i\leq N} \{\|\sqrt{Q_i}^{-1} D_i \sqrt{P_i}^{-1}\|\}\right] + \Theta(A_N)\cdots\Theta(A_1)\cdot \\ &\left[\max_{1\leq i\leq N+1} \{\|\sqrt{Q_i}^{-1} C_i\|\}\right]\left[\max_{0\leq i\leq N} \{\|B_i \sqrt{P_i}^{-1}\|\}\right], \end{aligned} \tag{7.8.12}$$

where

$$\Theta(A) = \sqrt{1 + \frac{\|A\|^2}{2} + \frac{\|A\|\sqrt{4+\|A\|^2}}{2}}; \tag{7.8.13}$$

(c) *the best choices for* $\{\lambda_k\}$ *in Proposition 6.3.1 is* $\lambda_k \equiv \bar{\lambda} = 1/\|\bar{D}\|^2 = \theta/M^2$*, where* $\theta \geq 1$.

Proof. (7.7.30) is obvious from Theorem 7.7.5 and its proof. Hence, $\alpha = 1$. As for (7.7.29), (7.8.7) implies

$$T_1(u, v) = \begin{pmatrix} -B^T y - D^T v \\ Cx + Du \end{pmatrix}. \tag{7.8.14}$$

However, (7.4.2) can imply

$$\begin{aligned} \bar{G}x &= Bu \qquad \text{or} \qquad x = \bar{G}^{-1}Bu, \\ \bar{G}^T y &= C^T v \qquad \text{or} \qquad y = (\bar{G}^{-1})^T C^T v \end{aligned} \tag{7.8.15}$$

similar to the discussion in Section 3.1. Puting (7.8.15) into (7.8.14) gets

$$T_1(u, v) = \begin{pmatrix} & -B^T(\bar{G}^{-1})^T C^T - D^T \\ C\bar{G}^{-1}B + D & \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{7.8.16}$$

Applying Theorem 7.7.5 and its proof to the above T_1 will get (7.7.29) with $\bar{D}$ given by (7.8.10). Hence, the least Lipschitz constant of $S^{-1}T_1S^{-1}$ is $\|\bar{D}\|$. From (7.8.10),

$$\begin{aligned} \|\bar{D}\| &\leq \|\sqrt{Q}^{-1} D \sqrt{P}^{-1}\| + \|\sqrt{Q}^{-1} C\bar{G}^{-1}B\sqrt{P}^{-1}\| \\ &\leq \max_{1\leq i\leq N} \{\|\sqrt{Q_i}^{-1} D_i \sqrt{P_i}^{-1}\|\} + \|\sqrt{Q}^{-1}C\|\|\bar{G}^{-1}\|\|B\sqrt{P}^{-1}\| \\ &\leq \max_{1\leq i\leq N} \{\|\sqrt{Q_i}^{-1} D_i \sqrt{P_i}^{-1}\|\} + \\ &\quad \|\sqrt{Q}^{-1}C\|\|B\sqrt{P}^{-1}\|\Theta(A_N)\cdots\Theta(A_1) \\ &\leq \left[\max_{1\leq i\leq N} \{\|\sqrt{Q_i}^{-1} D_i \sqrt{P_i}^{-1}\|\}\right] + \Theta(A_N)\cdots\Theta(A_1) \\ &\quad \left[\max_{1\leq i\leq N+1} \{\|\sqrt{Q_i}^{-1} C_i\|\}\right] \left[\max_{0\leq i\leq N} \{\|B_i\sqrt{P_i}^{-1}\|\}\right] \end{aligned} \tag{7.8.17}$$

since

$$
\begin{aligned}
\|\bar{G}^{-1}\| &= \left\| \begin{pmatrix} I & & & \\ A_1 & I & & \\ \vdots & \vdots & \ddots & \\ A_N \cdots A_1 & A_N \cdots A_2 & \dots & I \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & A_N & I \end{pmatrix} \cdots \begin{pmatrix} I & & & \\ A_1 & I & & \\ & & \ddots & \\ & & & I \end{pmatrix} \right\| \\
&\leq \left\| \begin{pmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & A_N & I \end{pmatrix} \right\| \cdots \left\| \begin{pmatrix} I & & & \\ A_1 & I & & \\ & & \ddots & \\ & & & I \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} I & \\ A_N & I \end{pmatrix} \right\| \cdots \left\| \begin{pmatrix} I & \\ A_1 & I \end{pmatrix} \right\| \\
&= \Theta(A_N) \cdots \Theta(A_1).
\end{aligned} \tag{7.8.18}
$$

In fact, the last equality in (7.8.18) is from Proposition 6.4.2. Therefore, (7.8.17) gives (7.8.12). Part (c) is elementary now from the preceding two parts and Proposition 6.3.1. □

For the choice of θ in the part (c) of the theorem, there is no a general way to determine it. However, we suggest to use $\theta \in [1.2, 1.5]$ for a well-posed problem depending on our numerical experiments.

In the algorithms, we need to compute $f(u)$ and $g(v)$ at each iteration. For given saddle point problem associated with Lagrangian L of (7.8.3) on $U \times V$ in (7.8.4), we have

$$
\begin{aligned}
f(u) &= p_0^T u_0 + \tfrac{1}{2} u_0^T P_0 u_0 + \rho_{V_{N+1},Q_{N+1}}(q_{N+1} - C_{N+1} x_N) \\
&\quad + \sum_{t=1}^{N} \left[p_t^T u_t + \tfrac{1}{2} u_t^T P_t u_t + \rho_{V_t,Q_t}(q_t - C_t x_{t-1} - D_t u_t) \right], \\
g(v) &= q_{N+1}^T v_{N+1} - \tfrac{1}{2} v_{N+1}^T Q_{N+1} v_{N+1} - \rho_{U_0,P_0}(B_0^T y_1 - p_0) \\
&\quad + \sum_{t=1}^{N} \left[q_t^T v_t - \tfrac{1}{2} v_t^T Q_t v_t - \rho_{U_t,P_t}(B_t^T y_{t+1} + D_t^T v_t - p_t) \right].
\end{aligned} \tag{7.8.19}
$$

At the box-diagonal case, all ρ_{U_t,P_t} and ρ_{V_t,Q_t} specialize to a set of one-dimensional subproblems similar to (7.1.7) with similar closed-form solutions in (7.1.8). At this case, $f(u)$ and $g(v)$ are simple to compute.

7.9. A Numerical Example and Computing Results

The following example is originally a central bank model for stochastic programming. Here we only consider its deterministic case.

Example 7.1:

In Problem $(\mathcal{P}_d)$ and the dynamics (3.1.2) of Section 3.1, let

$$u_t \in \mathbb{R}^2, \qquad v_t \in \mathbb{R}, \qquad C_t = (0,1), \qquad D_t = (0,0), \tag{7.9.1}$$

$$U_t = [-2,5] \times [6,11], \quad V_t = [0,\infty), \quad q_t = 0, \quad Q_t = \frac{1}{2\delta^t}, \tag{7.9.2}$$

$$P_t = \begin{pmatrix} \delta^t\gamma & \\ & \delta^t \end{pmatrix}, \quad p_t = \begin{pmatrix} 0 \\ -\delta^t a \end{pmatrix}, \quad c_t = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad b_t = \begin{pmatrix} 0 \\ 6.5 \end{pmatrix}, \tag{7.9.3}$$

$$A_t = \begin{pmatrix} 1.016 & 0 \\ 0.13208 & 1 \end{pmatrix}, \qquad B_t = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \tag{7.9.4}$$

where

$$\delta = 0.99, \qquad \gamma = 0.01, \qquad a = 11 \tag{7.9.5}$$

and there are no u_0, P_0, C_{N+1}, and v_{N+1}.

We use the EFBS algorithm of Section 7.5 to do the numerical test. Two splitting strategies in (7.5.2), with $L(u,v)$ given by (3.1.5), are used:

(1) **BS:** the basic splitting with $L_1(u,v) = -l(u,v) - \sum_{t=1}^N v_t \cdot D_t u_t$;

(2) **PS:** the projection splitting with $L_1(u,v) = L(u,v)$.

We also test two different choices of H_k: $H_k \equiv I$ and $H_k \equiv H_2 = \operatorname{diag}\{P, Q\}$. Two sets of the starting points are freely chosen and produced by the following:

(a). If $\bar{u}_0 = (1,1)^T$ and $\bar{v}_0 = -1$, then the starting point (u^0, v^0) is

$$u_t^0 = \bar{P}_{U_t}(\bar{u}_0) = \begin{pmatrix} 1 \\ 6 \end{pmatrix}, \qquad v_t^0 = \bar{P}_{V_t}(\bar{v}_0) = 0. \tag{7.9.6}$$

(b). If $\bar{u}_0 = (0, 8)^T$ and $\bar{v}_0 = 0$, then the starting point (u^0, v^0) is

$$u_t^0 = \bar{P}_{U_t}(\bar{u}_0) = \begin{pmatrix} 0 \\ 8 \end{pmatrix}, \qquad v_t^0 = \bar{P}_{V_t}(\bar{v}_0) = 0. \tag{7.9.7}$$

Where $\bar{P}$ represent the projection operator of an appropriate dimension.

In the above example, P_t, Q_t and p_t are time-variant matrices. We choose $N = 5$ and $N = 10$ for different step sizes $\lambda_k \equiv \lambda$. The results are computed on the DEC 3100 workstation of applied mathematics computer laboratory at the University of Washington. The stopping error bound $\varepsilon = 10^{10}$ and the stopping criterion is $f_k - g_k \leq \varepsilon$ in the EFBS algorithm.

(1). Computational results for different H and T_1

For $N = 5$, the optimal objective value is $f(u^*) = g(v^*) = -290.654756138$.

Table 7.1. Comparison of EFBS Algorithm for N=5 and Point (a)

BS:$H = I$ λ/iter/time	PS:$H = I$ λ/iter/time	BS:H_2 λ/iter/time	PS:H_2 λ/iter/time
0.1/9299/6.7	0.2/4643/3.2	0.01/1397/-	0.01/1384/-
0.15/6201/4.0	0.25/3714/3.0	0.03/507/-	0.03/494/-
0.2/4652/2.9	0.28/3315/2.6	0.05/335/-	0.05/322/-
0.25/3722/2.4	0.3/3094/2.5	0.08/244/-	0.08/233/-
0.3/3103/2.0	0.32/2900/2.4	0.1/216/-	0.1/210/-
0.35/2660/1.6	0.33/6153/4.7	0.13/201/-	0.11/206/-
0.4/2328/1.4	0.35/div.	0.14/186/-	0.12/193/-
0.42/2715/1.8		0.15/201/-	0.13/202/-
0.45/div.		0.2/249/-	0.15/205/-
		0.23/355/-	0.2/498/-
		0.27/868/-	0.22/1707/-
		0.3/div.	0.23/div.

where BS means basic splitting, PS means projection splitting, $H = I$ and H_2 express $H_k \equiv I$ and $H_k \equiv \text{diag}\{P, Q\}$, respectively, *iter/time* represents the convergent iteration number and the system time used, *p-d gap*$\times 10^{-11}$ is the primal-dual gap of the Lagrangian at the end of iteration, 'div.' represents the algorithm diverges for given λ, and '-' means the time is too short to record.

(2). Computational results for different H and T_1

Table 7.2. Comparison of EFBS Algorithm for N=5 and Point (b)			
BS:$H = I$	PS:$H = I$	BS:H_2	PS:H_2
λ/iter/time	λ/iter/time	λ/iter/time	λ/iter/time
0.1/250/-	0.1/241/-	0.01/1350/-	0.01/1337/-
0.15/210/-	0.18/196/-	0.03/489/-	0.03/477/-
0.2/202/-	0.2/204/-	0.05/324/-	0.05/312/-
0.25/216/-	0.25/261/-	0.08/236/-	0.08/225/-
0.3/266/-	0.3/545/-	0.1/210/-	0.1/205/-
0.35/391/-	0.32/1316/-	0.13/199/-	0.11/196/-
0.4/953/-	0.35/div.	0.14/195/-	0.12/197/-
0.42/2714/2.2		0.15/191/-	0.13/191/-
0.45/div.		0.2/248/-	0.15/216/-
		0.25/476/-	0.2/476/-
		0.27/843/-	0.22/1660/-
		0.3/div.	0.23/div.

(3). Computational results for different T_1 and starting points:

The optimal objective value is $f(u^*) = g(v^*) = -541.0099541667$ for $N = 10$.

Table 7.3. Comparison of EFBS Algorithm for N=10 and H=I			
BS: (a)	PS: (a)	BS: (b)	PS: (b)
λ/iter/time	λ/iter/time	λ/iter/time	λ/iter/time
0.03/32491/25.9	0.03/32479/25.9	0.01/4016/3.0	0.01/4006/3.2
0.04/24372/19.8	0.04/24359/19.8	0.02/2106/1.6	0.02/2004/1.7
0.05/19495/15.8	0.05/19487/16.0	0.04/1064/-	0.04/1051/-
0.06/16248/12.4	0.06/16237/13.2	0.05/856/-	0.05/844/-
0.07/13928/10.8	0.07/13919/11.5	0.055/780/-	0.055/812/-
0.08/12186/10.0	0.08/12177/10.3	0.0565/761/-	0.0565/796/-
0.09/10834/8.7	0.083/11737/9.6	0.06/756/-	0.058/812/-
0.091/10715/9.0	0.085/11462/9.2	0.062/771/-	0.06/823/-
0.092/12490/9.6	0.087/22242/17.9	0.07/914/-	0.07/1042/-
0.095/div.	0.09/div.	0.08/1395/-	0.08/2013/-
		0.09/4809/3.7	0.085/5293/4.4
		0.092/11484/9.2	0.087/20485/16.7
		0.095/div.	0.09/div.

(4). Computational results for different T_1 and starting points:

Table 7.4. Comparison of EFBS Algorithm for N=10 and H=H_2				
λ	BS: (a) iter/time	PS: (a) iter/time	BS: (b) iter/time	PS: (b) iter/time
0.001	13685/10.5	13672/10.7	13213/10.0	13200/10.2
0.002	7075/5.2	7063/5.4	6827/5.1	6815/5.3
0.003	5170/3.9	5160/4.1	4987/3.9	4978/4.1
0.004	4466/3.6	4459/3.5	4324/3.5	4318/3.3
0.005	4246/3.4	4245/3.3	4131/3.2	4131/3.4
0.0051	4244/3.3	4244/3.3	4131/3.2	4131/3.3
0.0052	4245/3.3	4246/3.3	4134/3.2	4135/3.2
0.006	4387/3.4	4399/3.4	4288/3.3	4301/3.3
0.007	4984/3.7	5031/3.7	4896/3.6	4944/4.0
0.008	6544/4.8	6711/5.1	6461/4.9	6629/5.2
0.009	11910/9.4	12911/10.4	11839/8.8	12844/10.1
0.0095	24077/19.3	30224/23.7	24008/18.7	30148/23.2
0.01	div.	div.	div.	div.

where '(a)' and '(b)' express the starting points (a) and (b) respectively.

From the above tables and figures on the next several pages for this example, we can easily observe the following:

1. The different splitting strategies here do not make much difference in the convergence for the same starting points and stage N. (This observation coincides with our conclusions in Corollary 6.5.2 and Corollary 6.5.3.)
2. For $H_k \equiv I$, the EFBS algorithm is sensitive to the starting points (especially when the condition number of $\text{diag}\{P, Q\}$ is large).
3. For $H_k \equiv \text{diag}\{P, Q\}$, the EFBS algorithm is not sensitive to the two sets of starting points at all and the convergences appear at almost the same time.
4. The convergence of the EFBS algorithm depends on the step size λ more sensitively for $H_k \equiv I$ than $H_k \equiv \text{diag}\{P, Q\}$. Also, the convergence intervals for λ seem symmetric to the numerically best choice of λ for $H_k \equiv \text{diag}\{P, Q\}$, but no any symmetry for $H_k \equiv I$.
5. The convergence interval of λ for $H_k \equiv \text{diag}\{P, Q\}$ heavily depends on the eigenvalues of P and Q. Since the smallest eigenvalue of P is less than 10^{-2} [<< 1] (hence the Lipschitz constant of $H^{-\frac{1}{2}} T_1 H^{-\frac{1}{2}}$ is larger than T_1's),

the convergence interval of λ for $H_k \equiv \text{diag}\{P, Q\}$ should be smaller than $H_k \equiv I$. This is true from the results.

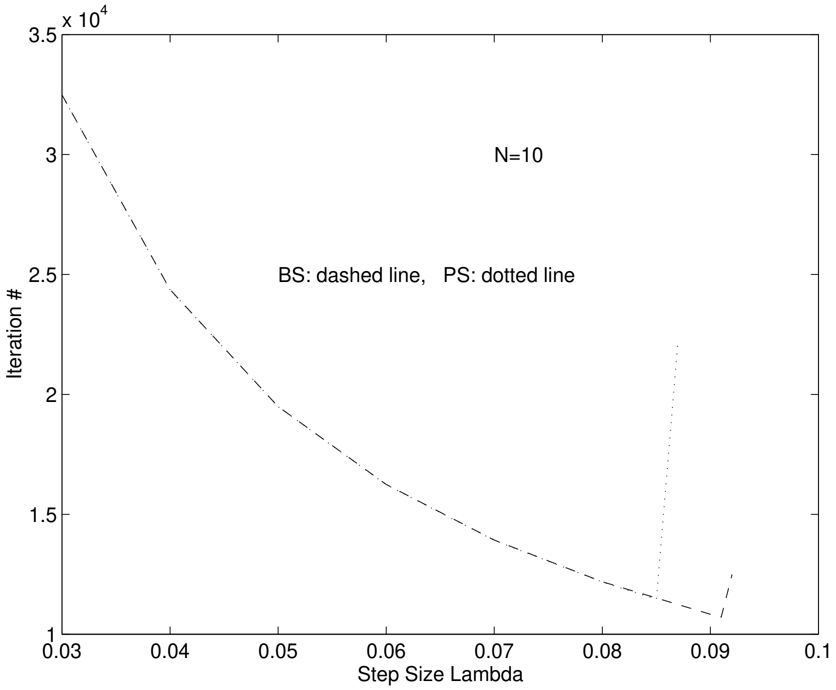

Figure 7.1: Comparison of BS and PS Methods for Point (a) & H=I

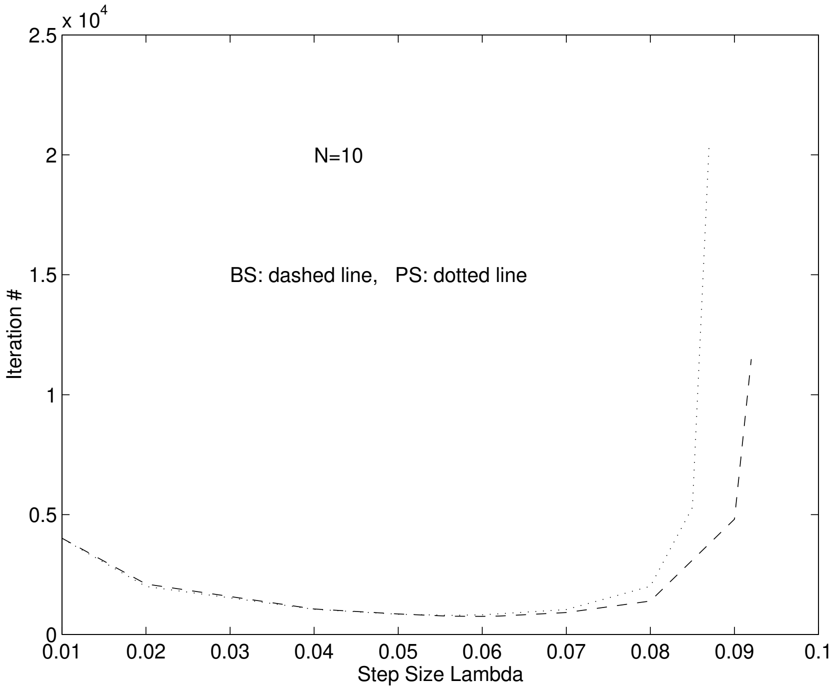

Figure 7.2: Comparison of BS and PS Methods for Point (b) & H=I

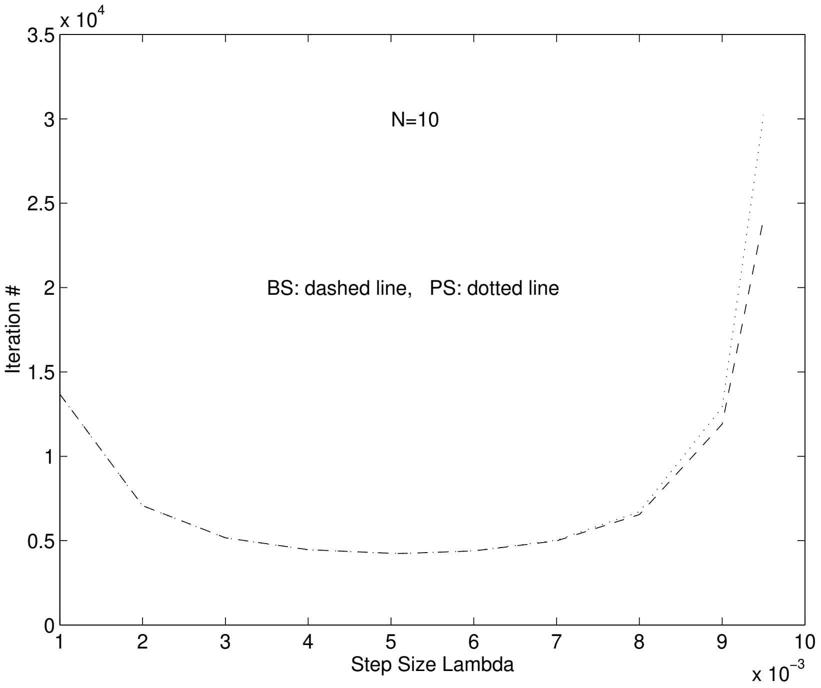

Figure 7.3: Comparison of BS and PS Methods for Point (a) & H=diag$\{P, Q\}$

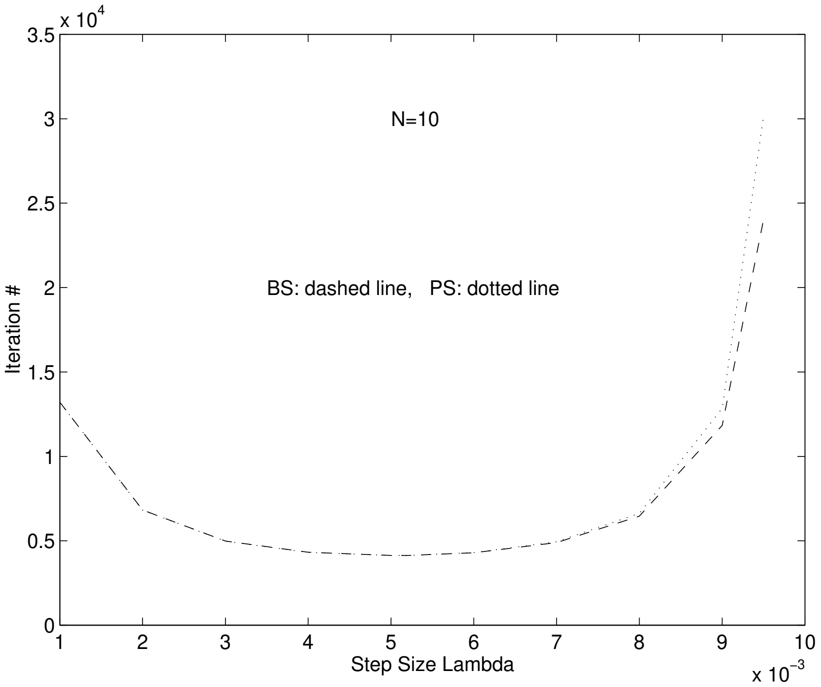

Figure 7.4: Comparison of BS and PS Methods for Point (b) & H=diag$\{P, Q\}$

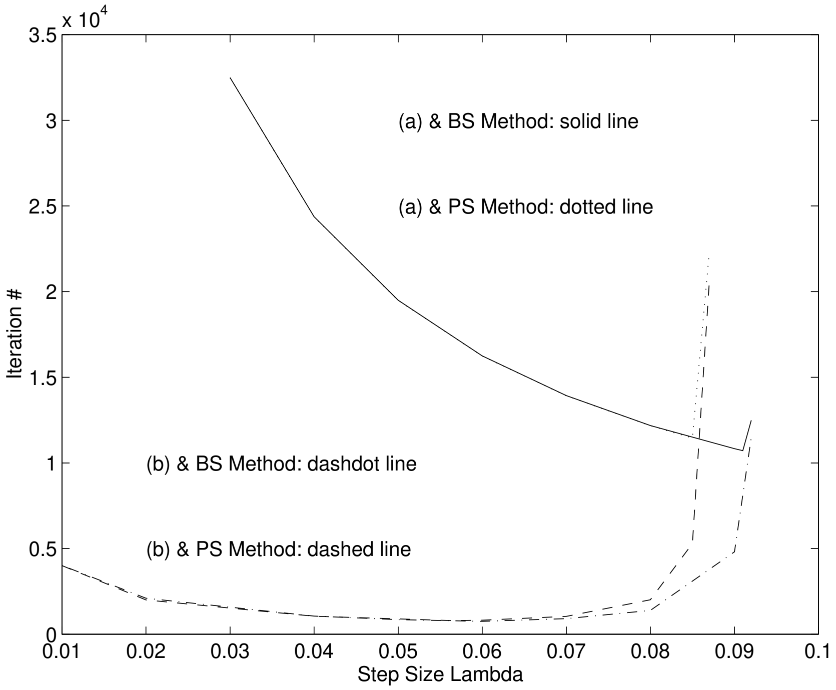

Figure 7.5: Comparison of Different Starting Points for H=I & N=10

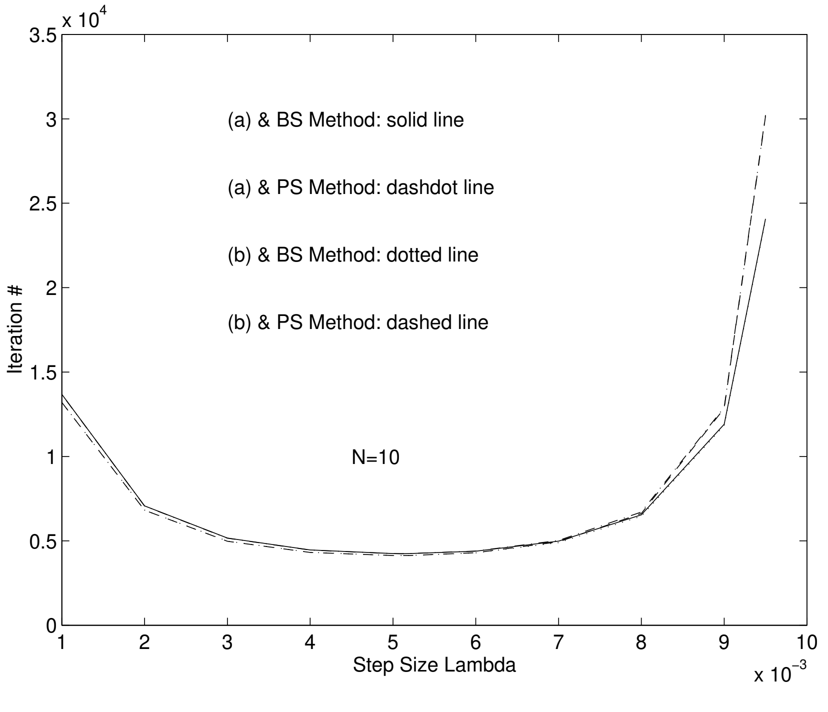

Figure 7.6: Comparison of Different Starting Points for H=diag$\{P, Q\}$

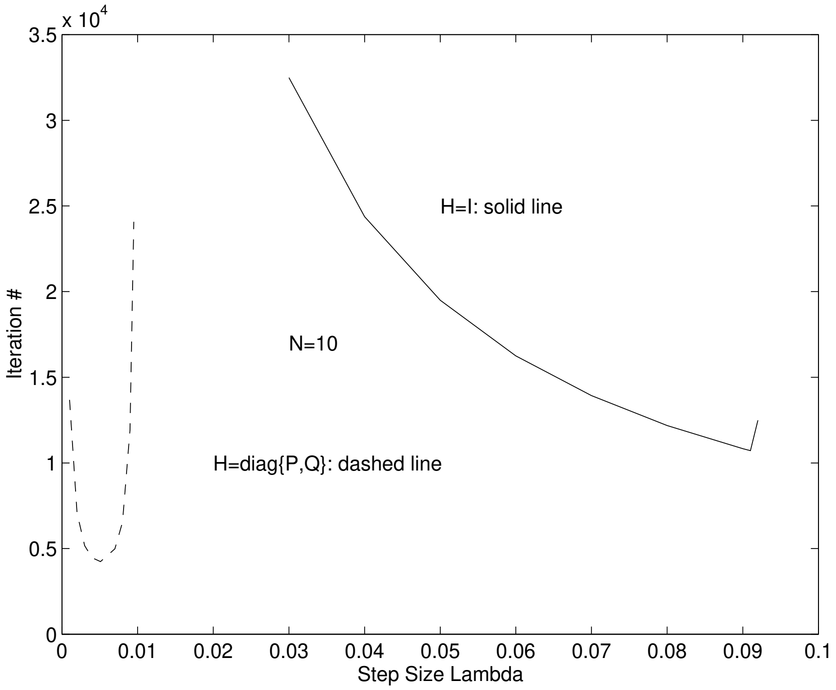

Figure 7.7: Comparison of Different Matrices H for (a) & BS Method

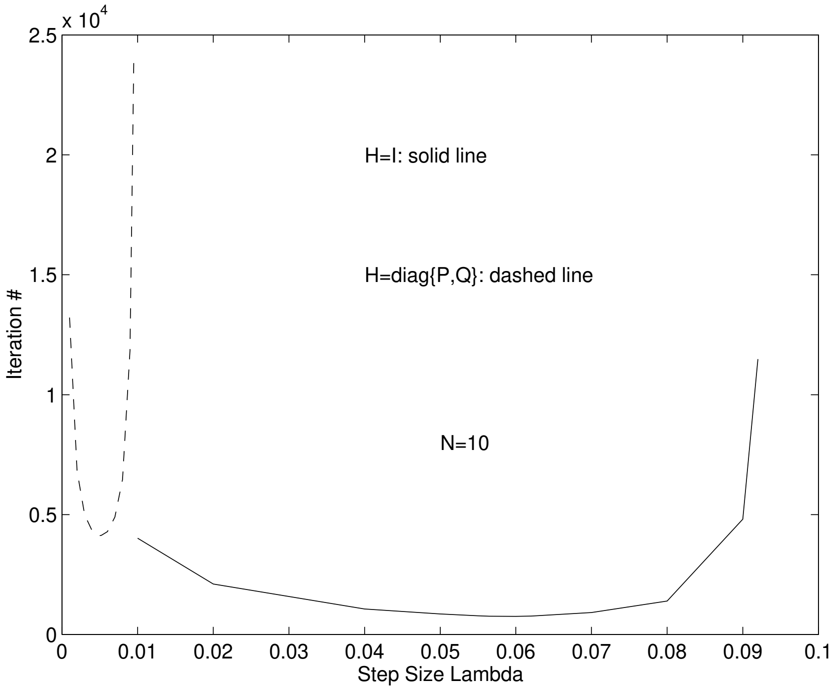

Figure 7.8: Comparison of Different Matrices H for (b) & BS Method

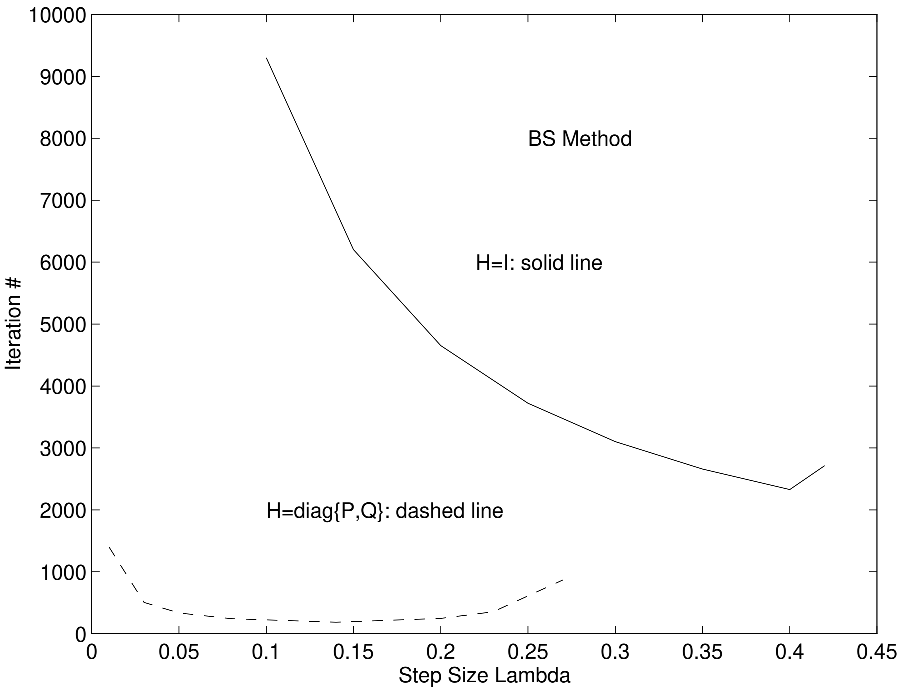

Figure 7.9: Comparison of Different Matrices H for (a) & N=5

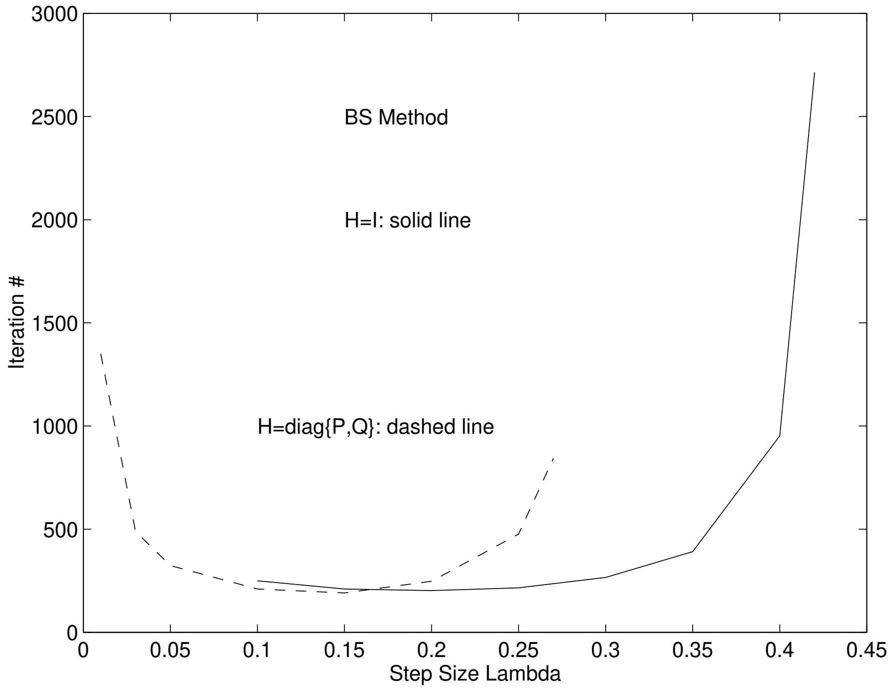

Figure 7.10: Comparison of Different Matrices H for (b) & N=5

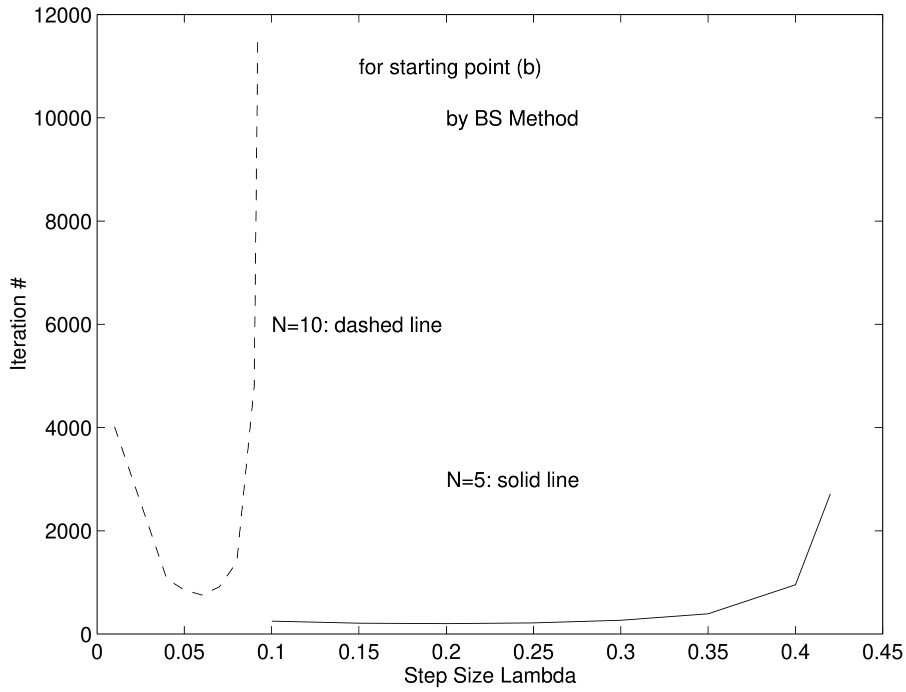

Figure 7.11: Comparison of Different Stages N for H=I

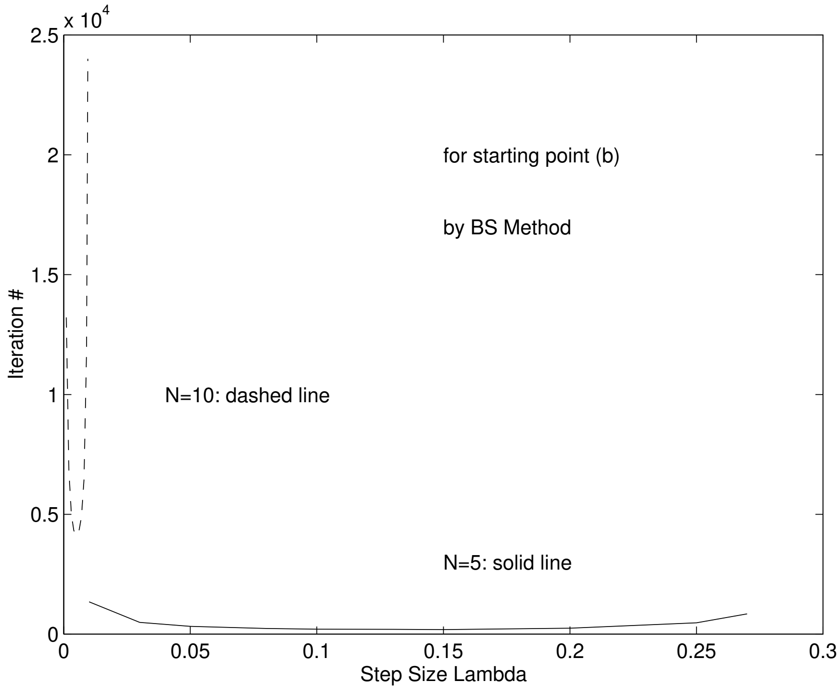

Figure 7.12: Comparison of Different Stages N for H=diag$\{P, Q\}$