Chapter 8

OPTIMAL CONTROL PROBLEM APPLICATIONS

In this chapter, forward-backward splitting methods are applied to problems of optimal control.

8.1. Introduction

The problem of optimal control of an initial ordinary differential equation, with Bolza objectives and mixed constraints, has the form

$$(\mathcal{O}_c) \qquad \begin{aligned} &\min_{x,u} \int_0^T L\big(x(t),u(t),t\big)\,dt + \varphi\big(x(T)\big),\\ &\dot{x}(t) = f\big(x(t),u(t),t\big), \qquad x(0) = b_0,\\ &g\big(x(t),u(t),t\big) \le 0, \quad t \in [0,T], \quad g_f\big(x(T)\big) \le 0. \end{aligned}$$

Where, $x : [0,T] \to \mathbb{R}^{n_s}$, $u : [0,T] \to \mathbb{R}^{n_c}$, $L : \mathbb{R}^{n_s} \times \mathbb{R}^{n_c} \times [0,T] \to \mathbb{R}$, $\varphi : \mathbb{R}^{n_s} \to \mathbb{R}$, $g : \mathbb{R}^{n_s} \times \mathbb{R}^{n_c} \times [0,T] \to \mathbb{R}^{n_g}$, $g_f : \mathbb{R}^{n_s} \to \mathbb{R}^{n_f}$.

A discrete-time counterpart is the problem

$$(\mathcal{O}_d) \qquad \begin{aligned} &\min_{x_i,u_i} \sum_{i=1}^{N} L_i(x_i,u_i) + \varphi(x_{N+1}),\\ &x_{i+1} = f_i(x_i,u_i), \quad i = 1,\ldots,N, \qquad x_1 = b_0,\\ &g_i(x_i,u_i) \le 0, \quad i = 1,\ldots,N, \quad g_f(x_{N+1}) \le 0. \end{aligned}$$

Many efficient algorithms have been proposed for various special classes of the above problems. Some examples of those algorithms and special cases can be found in Polak (1970, 1973), Dunn and Bertsekas (1989), Edge and Powers (1976), Kelley and Sachs (1989), for unconstrained version of $(\mathcal{O}_c)$ or $(\mathcal{O}_d)$ and for Newton or Newton-like methods; in Dem'yanov and Rubinov (1970), Dunn (1981, 1987), Gafni and Bertsekas (1984), Pantoja and Mayne (1991), Wright (1990, 1991), for the control-only constrained case and the gradient projection class methods, the stagewise algorithms; and so on.

The most general cases of $(\mathcal{O}_c)$, $(\mathcal{O}_d)$ are known to be significantly more difficult to solve than the special cases in the above. Wright (1993) describes interior point methods for the general problems together with an inexact sequential quadratic programming algorithm. Miele (1975), Polak, Yang, and Mayne (1990), Evtushenko (1985, Chapter 6), Di Pillo, Grippo and Lampariello (1983) *etc.* have considered similar problems and proposed algorithms based on nonlinear programming techniques, such as gradient methods, penalty function methods, quasi-Newton methods, *etc.*

In this chapter, we focus on a linear-quadratic version of the discrete problem $(\mathcal{O}_d)$. This can be formulated as

$$
(\mathcal{O}) \qquad
\begin{aligned}
&\min_{x_i,u_i} \sum_{i=1}^{N} \left[p_i^T u_i + c_i^T x_i + \frac{1}{2}(x_i^T S_i x_i + 2x_i^T R_i u_i + u_i^T P_i u_i)\right] \\
&\qquad\qquad + c_{N+1}^T x_{N+1} + \frac{1}{2} x_{N+1}^T S_{N+1} x_{N+1}, \\
&x_{i+1} = A_i x_i + B_i u_i + b_i, \quad i = 1, \ldots, N, \qquad x_1 = b_0, \\
&\hat{G}_i u_i + \hat{F}_i x_i \le \hat{g}_i, \quad i = 1, \ldots, N, \qquad F_{N+1} x_{N+1} \le g_{N+1}.
\end{aligned}
$$

Where $x_i \in \mathbb{R}^{n_s}$, $u_i \in \mathbb{R}^{n_c}$, $\hat{g}_i \in \mathbb{R}^{n_g}$, S_i and P_i are symmetric and positive semidefinite. A_i, B_i, R_i, $\hat{G}_i$, and $\hat{F}_i$ are matrices in appropriate spaces.

Sequential quadratic programming algorithms, for the nonlinear problem $(\mathcal{O}_d)$, give rise to subproblems of this form at each iteration (Ref. Wright 1993, Pang 1986, Pantoja and Mayne 1991, *etc.*).

We assume that Problem $(\mathcal{O})$ is a convex problem, i.e.,

$$
\begin{aligned}
&\begin{bmatrix} P_i & R_i \\ R_i & S_i \end{bmatrix} \quad \text{is positive semidefinite for} \quad i = 1, \ldots, N; \\
&\qquad\quad S_{N+1} \quad \text{is positive semidefinite.}
\end{aligned}
\tag{8.1.1}
$$

We split the inequality constraints for $i = 1, \ldots, N$ into two parts:

$$
G_i u_i + F_i x_i \le g_i, \quad i = 1, \ldots, N, \qquad u_i \in U_i, \tag{8.1.2}
$$

here $g_i \in \mathbb{R}^m$ and U_i is a "box" in $\mathbb{R}^{n_c}$, i.e., U_i is the product of closed intervals, closed half lines, and whole lines $\mathbb{R}$.

Recently, Rockafellar and Wets (1986, 1987), Zhu and Rockafellar (1993) proposed finite generation algorithms and primal-dual gradient algorithms with decomposable properties, from Lagrangian and duality theory, for extended linear-quadratic programming problems which include our discrete-time linear-quadratic optimal control problems. In this chapter, forward-backward splitting methods are proposed for the above problem $(\mathcal{O})$ by applying the results of Chapter 7. In this chapter, however, we only consider using the special extended forward-backward splitting scheme:

$$z^{k+1} = (H + \lambda_k T_2)^{-1}(H - \lambda_k T_1)(z^k), \qquad \forall k \geq 0, \tag{8.1.3}$$

where T_1 and T_2 are maximal monotone operators on $\mathcal{H}$, with $T = T_1 + T_2$, T_1 is single-valued, $\{\lambda_k\}_{k=0}^{\infty}$ is a sequence of positive scalars, H is a given positive definite matrix from $\mathcal{H}$ to itself, and $\mathcal{H}$ is an Euclidean space.

8.2. Transformation of the Problem into a Saddle Point Problem

For using the EFBS algorithms in Chapter 7, we should transform the linear-quadratic discrete problem $(\mathcal{O})$ into a primal problem of a special convex-concave saddle point problem which is a special case of the problem $(\mathcal{G})$ of the section 7.5. Solving the problem $(\mathcal{O})$ can be transformed, hence, into solving a convex-concave saddle point problem $(\mathcal{G})$, which could utilize the results of the chapter 7.

Now let us consider the following linear-quadratic saddle point problem:

$$(\mathcal{D}) \qquad \operatorname{minimax} \mathcal{J}(u,v) \quad \text{subject to} \quad (u,v) \in U \times V \quad \text{and}$$

$$x_{i+1} = A_i x_i + B_i u_i + b_i, \quad i = 1, \ldots, N, \qquad x_1 = b_0, \tag{8.2.1}$$

$$y_i = A_i^T y_{i+1} + C_i^T v_i + c_i, \quad y_{N+1} = C_{N+1}^T v_{N+1} + c_{N+1}, \tag{8.2.2}$$

where

$$
\begin{aligned}
&U = U_1 \times U_2 \times \ldots \times U_N, \qquad u = (u_1, \ldots, u_N)\\
&V = V_1 \times V_2 \times \ldots \times V_{N+1}, \qquad v = (v_1, \ldots, v_{N+1})\\
&\mathcal{J}(u,v) = \sum_{i=1}^{N} J_i(u_i, v_i) + \varphi_N(v_{N+1}) - l(u,v),\\
&J_i(u_i, v_i) = p_i^T u_i + q_i^T v_i + \frac{1}{2} u_i^T P_i u_i - v_i^T D_i u_i - \frac{1}{2} v_i^T Q_i v_i,\\
&\varphi_N(v) = q_{N+1}^T v - \frac{1}{2} v^T Q_{N+1} v,\\
&l(u,v) = \sum_{i=1}^{N+1} x_i^T (C_i^T v_i + c_i) = \sum_{i=1}^{N} y_{i+1}^T (B_i u_i + b_i) + y_1^T b_0
\end{aligned}
$$

and $u_i \in \mathbb{R}^{n_c}$, $v_i \in \mathbb{R}^{n_v}$, $x_i \in \mathbb{R}^{n_s}$, $y_i \in \mathbb{R}^{n_s}$, all P_i for $i = 1, \ldots, N$, Q_j for $j = 1, \ldots, N+1$ are symmetric and positive semidefinite, U_i and V_j are closed boxes in $\mathbb{R}^{n_c}$ and $\mathbb{R}^{n_v}$. Obviously, Problem $(\mathcal{D})$ is a convex-concave saddle point problem.

Let

$$
\begin{aligned}
\mathcal{F}(u) &= \sup_v \mathcal{J}(u,v) \qquad \text{over} \quad v \in V,\\
\mathcal{G}(v) &= \inf_u \mathcal{J}(u,v) \qquad \text{over} \quad u \in U.
\end{aligned} \tag{8.2.3}
$$

Then, the primal problem and dual problem of $(\mathcal{D})$ are respectively

$$\inf_u \mathcal{F}(u) \quad \text{subject to} \quad u \in U \quad \text{and} \quad (8.2.1). \tag{8.2.4}$$

$$\sup_v \mathcal{G}(v) \quad \text{subject to} \quad v \in V \quad \text{and} \quad (8.2.2). \tag{8.2.5}$$

Theorem 8.2.1 (Rockafellar [274]).

(a) *A control sequence $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_N)$ solves the primal problem (8.2.4) if and only if there is a dual solution control sequence $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_{N+1})$ for (8.2.5) such that $(\bar{u}, \bar{v})$ is a saddle point of $\mathcal{J}(u,v)$ relative to $U \times V$.*

(b) *It is always true that $\inf_{u \in U} \mathcal{F}(u) \geq \sup_{v \in V} \mathcal{G}(v)$, and hence $\mathcal{F}(u) \geq \mathcal{G}(v)$ for any feasible (u,v).*

(c) *Problem (8.2.4) has a solution if and only if its objective $\mathcal{F}(u)$ is bounded below on U and is finite for at least one $u \in U$. Then* min *in (8.2.4)=* max *in (8.2.5).*

Next, let us discuss the relation between Problem $(\mathcal{O})$ and Problem $(\mathcal{D})$.

From the convex assumption, we can suppose $R_i \equiv 0$ in $(\mathcal{O})$, without loss of generality, from the following logic:

Proposition 8.2.2. *If R_i in $(\mathcal{O})$ are not all zero matrices, then Problem $(\mathcal{O})$ can be transformed into an equivalent problem which has no across term in its objective function.*

Proof. Since $(\mathcal{O})$ is a convex problem and all P_i, S_i $(i = 1, \dots, N)$ are symmetric, there exist unit lower triangular matrices X_i and diagonal matrix Λ_i such that

$$\Lambda_i = X_i \begin{bmatrix} P_i & R_i^T \\ R_i & S_i \end{bmatrix} X_i^T \quad \text{for} \quad i = 1, \dots, N.$$

Where

$$\Lambda_i = \begin{bmatrix} \bar{P}_i & \\ & \bar{S}_i \end{bmatrix}, X_i = \begin{bmatrix} X_{1i} & \\ X_{12i} & X_{2i} \end{bmatrix}.$$

Let

$$\begin{bmatrix} u_i \\ x_i \end{bmatrix} = X_i^T \begin{bmatrix} \bar{u}_i \\ \bar{x}_i \end{bmatrix} = \begin{bmatrix} X_{1i}^T \bar{u}_i + X_{12i}^T \bar{x}_i \\ X_{2i}^T \bar{x}_i \end{bmatrix}.$$

Then the objective function becomes

$$\sum_{i=1}^{N} [\bar{p}_i^T \bar{u}_i + \bar{c}_i^T \bar{x}_i + \frac{1}{2}(\bar{u}_i^T \bar{P}_i \bar{u}_i + \bar{x}_i^T \bar{S}_i \bar{x}_i)] + c_{N+1}^T x_{N+1} + \frac{1}{2} x_{N+1}^T S_{N+1} x_{N+1},$$

and the constraints for $i = 1, \dots, N$ are changed as

$$\begin{aligned} &\bar{x}_{i+1} = \bar{A}_i \bar{x}_i + \bar{B}_i \bar{u}_i + \bar{b}_i, \quad i = 1, \dots, N-1, \qquad x_1 = b_0, \\ &x_{N+1} = \bar{A}_N \bar{x}_N + \bar{B}_N \bar{u}_N + b_N, \\ &\bar{G}_i \bar{u}_i + \bar{F}_i \bar{x}_i \le \bar{g}_i, \qquad i = 1, \dots, N. \end{aligned}$$

Here

$$\begin{aligned} &\bar{p}_i = X_{1i} p_i, \quad \bar{c}_i = X_{12i} p_i + X_{2i} c_i, \quad \bar{b}_i = (X_{2(i+1)}^T)^{-1} b_i, \\ &\bar{A}_i = (X_{2(i+1)}^T)^{-1}(A_i X_{2i}^T + B_i X_{12i}^T), \quad i = 1, \dots, N-1, \\ &\bar{G}_i = G_i X_{1i}^T, \quad \bar{F}_i = G_i X_{12i}^T + F_i X_{2i}, \quad i = 1, \dots, N \\ &\bar{A}_N = A_n X_{2N}^T + B_N X_{12N}^T, \quad \bar{B}_N = B_N X_{1N}^T, \\ &\bar{B}_i = (X_{2(i+1)}^T)^{-1} B_i X_{1i}^T, \quad i = 1, \dots, N-1. \end{aligned}$$

This is what we want. $\square$

Therefore, we can consider using the following problem to replace Problem $(\mathcal{O})$ in our discussion without loss of generality.

$$
(\mathcal{O}') \qquad
\begin{aligned}
&\min_{x_i,u_i} \mathcal{F}'(u) := \sum_{i=1}^{N} \left[p_i^T u_i + c_i^T x_i + \frac{1}{2}(x_i^T S_i x_i + u_i^T P_i u_i) \right] \\
&\qquad\qquad + c_{N+1}^T x_{N+1} + \frac{1}{2} x_{N+1}^T S_{N+1} x_{N+1}, \\
&x_{i+1} = A_i x_i + B_i u_i + b_i, \qquad i = 1, \ldots, N, \qquad x_1 = b_0, \\
&G_i u_i + F_i x_i \le g_i, \quad u_i \in U_i, \quad i = 1, \ldots, N, \quad F_{N+1} x_{N+1} \le g_{N+1}.
\end{aligned}
$$

Theorem 8.2.3. *The linear-quadratic discrete problem $(\mathcal{O}')$ is the primal problem associated with $\mathcal{J}$ on $U \times V$ of Problem $(\mathcal{D})$ when P_i, U_i, p_i, A_i, B_i, b_i, and c_i are the same in both problems and the others in $(\mathcal{D})$ are*

$$
\begin{aligned}
&Q_i = \begin{bmatrix} Q_{1i} & \\ & 0 \end{bmatrix}, \quad v_i = \begin{bmatrix} v_{1i} \\ v_{2i} \end{bmatrix}, \quad q_i = \begin{bmatrix} 0 \\ -g_i \end{bmatrix}, \\
&C_i = \begin{bmatrix} C_{1i} \\ -F_i \end{bmatrix}, \quad D_i = \begin{bmatrix} 0 \\ -G_i \end{bmatrix}, \quad C_{1i} = Q_{1i}^{\frac{1}{2}} S_i^{\frac{1}{2}}, \\
&V_i = \mathbb{R}^{l_i} \times \mathbb{R}_+^m.
\end{aligned} \tag{8.2.6}
$$

$Q_{1i} \in \mathbb{R}^{l_i \times l_i}$ is a positive definite matrix and could be a preconditioner for C_{1i} and Q_i.

Proof. We only need to prove that the primal problem in (8.2.4) is the problem $(\mathcal{O}')$ under given conditions. However,

$$\max_v \mathcal{J}(u, v) \quad \text{over} \quad v \in V$$

is equivalent to finding a $\bar{v}$ such that

$$\nabla_v \mathcal{J}(u, \bar{v}) \in \mathcal{N}_V(\bar{v}),$$

where $\mathcal{N}_V(v)$ is the normal cone of V at v, that is,

$$
\begin{aligned}
&q_i - Q_i \bar{v}_i - D_i u_i - C_i x_i \in \mathcal{N}_{V_i}(\bar{v}_i), \quad i = 1, \ldots, N, \\
&q_{N+1} - Q_{N+1} \bar{v}_{N+1} - C_{N+1} x_{N+1} \in \mathcal{N}_{V_{N+1}}(\bar{v}_{N+1}).
\end{aligned}
$$

Hence, from the given relations between $(\mathcal{D})$ and $(\mathcal{O}')$,

$$\bar{v}_{1i} = Q_{1i}^{-1}(-C_{1i}x_i), \qquad \bar{v}_{2i} = \begin{cases} 0, & \text{when } G_iu_i + F_ix_i - g_i \le 0; \\ \infty, & \text{when } G_iu_i + F_ix_i - g_i > 0, \end{cases}$$

$$\bar{v}_{1(N+1)} = Q_{1(N+1)}^{-1}(-C_{1(N+1)}x_{N+1}),$$

$$\bar{v}_{2(N+1)} = \begin{cases} 0, & \text{when } F_{N+1}x_{N+1} - g_{N+1} \le 0; \\ \infty, & \text{when } F_{N+1}x_{N+1} - g_{N+1} > 0. \end{cases}$$

Therefore, if $(x_1, \ldots, x_{N+1})$ is determined by u through equality constraints (dynamic system), then

$$\mathcal{F}(u) = \max_v \mathcal{J}(u,v) = \begin{cases} \infty, & \text{when one of inequalities is false;} \\ \mathcal{F}'(u), & \text{otherwise.} \end{cases}$$

Since the primal problem is to find minimum of $\mathcal{F}(u)$, we only need to consider those u and corresponding x_i such that they satisfy the constraints of $(\mathcal{O}')$. That is, $(\mathcal{O}')$ is the primal problem of $(\mathcal{D})$ under given conditions. $\square$

The example 2.3.3 gives us a way to transform Problem $(\mathcal{D})$ into a fundamental problem $(\mathcal{P}_0)$ similar to that in the chapter 7. Let

$$T = \begin{pmatrix} \nabla_u \mathcal{J} + \mathcal{N}_U \\ -\nabla_v \mathcal{J} + \mathcal{N}_V \end{pmatrix}. \tag{8.2.7}$$

Then T is maximal monotone and solving the convex-concave saddle point problem $(\mathcal{D})$ is equivalent to solving the problem $(\mathcal{P}_0)$ with T given by (8.2.7).

Now we consider using forward-backward splitting methods to solve Problem $(\mathcal{D})$. As we discussed in Chapter 3 and Chapter 7, the FBS methods are especially suited to fit the structure of a saddle point problem of this nature. Since normal cone operators are hard to implement through forward steps but easy to do by backward steps and the operator corresponding to $l(u,v)$ is relatively difficult to implement through backward steps, we always split the box constraints into backward steps and $-l(u,v)$ into forward steps. Therefore, if we have a splitting of $T = T_1 + T_2$ in (8.2.7), there will be a splitting of $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ correspondingly such that

$$T_1 = \begin{pmatrix} \nabla_u \mathcal{J}_1 \\ -\nabla_v \mathcal{J}_1 \end{pmatrix} \qquad T_2 = \begin{pmatrix} \nabla_u \mathcal{J}_2 + \mathcal{N}_U \\ -\nabla_v \mathcal{J}_2 + \mathcal{N}_V \end{pmatrix} \tag{8.2.8}$$

and $-l(u,v)$ is part of $\mathcal{J}_1(u,v)$.

If $(\mathcal{D})$ is the special linear-quadratic saddle point problem corresponding to Problem $(\mathcal{O}')$, then, the T in (8.2.8) is

$$T(u,v)=\begin{bmatrix} P_1u_1+p_1+G_1^Tv_{21}-B_1^Ty_2+\mathcal{N}_{U_1}(u_1)\\ \vdots \\ P_Nu_N+p_N+G_N^Tv_{2N}-B_N^Ty_{N+1}+\mathcal{N}_{U_N}(u_N)\\ Q_1v_1+D_1u_1+C_1x_1-q_1+\mathcal{N}_{V_1}(v_1)\\ \vdots \\ Q_Nv_N+D_Nu_N+C_Nx_N-q_N+\mathcal{N}_{V_N}(v_N)\\ Q_{N+1}v_{N+1}+C_{N+1}x_{N+1}-q_{N+1}+\mathcal{N}_{V_{N+1}}(v_{N+1})\end{bmatrix}, \tag{8.2.9}$$

where Q_i, D_i, C_i, v_i, q_i are given in Theorem 8.2.3.

Before we consider using the EFBS algorithms in the chapter 7, we need to consider the strong monotonicity of the T in (8.2.7) since the convergence of the algorithms requires the strong monotonicity of T. This will be given in next section.

8.3. Methods to Obtain the Strong Monotonicity of Operators

When Problem $(\mathcal{O})$ or $(\mathcal{O}')$ has mixed inequality constraints, the Q_i for $i=1,\ldots,N+1$ in the corresponding problem $(\mathcal{D})$ will not be all positive definite, that is, the corresponding T^{-1} and T can not generally be strongly monotone. Therefore, the convergence conditions for the EFBS algorithms in Chapter 7 can not be satisfied. Hence the convergence of the EFBS algorithms can not be guaranteed. Similarly, when P_i for $i=1,\ldots,N$ are not all positive definite in Problem $(\mathcal{O}')$, we have the same situation.

In order to guarantee the convergence of the FBS algorithms, we use proximal point methods or *penalty operator* techniques. Here we introduce proximal point methods from Rockafellar first.

As discussed in Section 2.5, if T is given by (8.2.7), then

$$z^{k+1}=(I+\lambda T)^{-1}(z^k) \tag{8.3.1}$$

is equivalent to

$$z^{k+1} = (u^{k+1}, v^{k+1}) = \underset{u \in U \; v \in V}{\text{argminimax}} \quad \Lambda_k(u, v) \tag{8.3.2}$$

where

$$\Lambda_k(u, v) = \mathcal{J}(u, v) + \frac{1}{2\lambda}\|u - u^k\|^2 - \frac{1}{2\lambda}\|v - v^k\|^2. \tag{8.3.3}$$

Equality (8.3.3) tells us that the Λ_k is a strongly convex-concave function. Therefore, we can apply the EFBS algorithms to it to obtain (u^{k+1}, v^{k+1}) for each k. The convergence review of the proximal point method can be also found in Section 2.5. In an Euclidean space, $\{z^k\}$ generated from (8.3.1) will converge to a zero point of T if $T^{-1}(0) \neq \emptyset$ and $\lambda > 0$. However, (8.3.1) gives

$$z^{k+1} + \lambda T(z^{k+1}) \ni z^k \qquad \text{or} \qquad \frac{z^k - z^{k+1}}{\lambda} \in T(z^{k+1}). \tag{8.3.4}$$

Hence, when λ is very large, z^1 will be an approximate solution of $0 \in T(z)$ for a given z^0. If we only need an approximate solution, we may just solve one or two proximal point iterations for a given large λ.

Now let us generalize the penalty function and regularization techniques in nonlinear programming to *operator regularization or penalty operator* techniques here.

Let T be a general maximal monotone operator and T_0 be a given single-valued, continuous, monotone operator such that $T + \gamma T_0$ is strongly monotone for any $\gamma > 0$. Hence, if T is not strongly monotone, then the following operator will be strongly monotone:

$$T_\gamma = T + \gamma T_0, \qquad \text{the scalar} \quad \gamma > 0. \tag{8.3.5}$$

Here we call T_0 *penalty operator.*

If z_γ is the unique zero point of T_γ, then

$$0 \in T_\gamma(z_\gamma) = T(z_\gamma) + \gamma T_0(z_\gamma), \qquad \text{or} \qquad -\gamma T_0(z_\gamma) \in T(z_\gamma). \tag{8.3.6}$$

Proposition 8.3.1 (Boundedness of z_γ). *Let T, T_γ, T_0 and z_γ be defined as the above. If $T^{-1}(0) \neq \emptyset$, then the set $S_\mu := \{z_\gamma \,|\, \mu \geq \gamma > 0\}$ is bounded for*

$\infty > \mu > 0$. *Furthermore, when T_0 is also strongly monotone, the set $S_\infty :=$ $\{z_\gamma \,|\, \gamma > 0\}$ is bounded as well.*

Proof. Let α be the strongly monotone modulus of T_γ and $\bar{z}$ be a zero point of T. Then, from (8.3.6) and $0 \in T(\bar{z})$,

$$\begin{aligned} \alpha\|z_\gamma - \bar{z}\|^2 &\le \langle z_\gamma - \bar{z}, 0 - (0 + \gamma T_0(\bar{z})\rangle \\ &\le \|z_\gamma - \bar{z}\|\|\gamma T_0(\bar{z})\|. \end{aligned} \tag{8.3.7}$$

Hence,

$$\|z_\gamma - \bar{z}\| \le \frac{\gamma}{\alpha}\|T_0(\bar{z})\|. \tag{8.3.8}$$

Therefore, S_μ is bounded for every $\mu < \infty$.

When T_0 is strongly monotone with modulus $\alpha_1 > 0$, from (8.3.6) and the monotonicity of T,

$$\begin{aligned} 0 &\le \langle z_\gamma - \bar{z}, -\gamma T_0(z_\gamma) - 0\rangle \\ &= \langle z_\gamma - \bar{z}, \gamma[T_0(\bar{z}) - T_0(z_\gamma)]\rangle - \langle z_\gamma - \bar{z}, \gamma T_0(\bar{z})\rangle. \end{aligned}$$

This and the strong monotonicity of T_0 imply

$$\begin{aligned} -\gamma\langle z_\gamma - \bar{z}, T_0(\bar{z})\rangle &\ge \gamma\langle z_\gamma - \bar{z}, T_0(z_\gamma) - T_0(\bar{z})\rangle \\ &\ge \alpha_1\gamma\|z_\gamma - \bar{z}\|^2. \end{aligned} \tag{8.3.9}$$

Therefore, for any $\gamma > 0$,

$$\|z_\gamma - \bar{z}\| \le \frac{1}{\alpha_1}\|T_0(\bar{z})\|. \tag{8.3.10}$$

This implies that S_∞ is also bounded. $\square$

Proposition 8.3.2 (Properties of z_γ). *Under the above definitions of z_γ, T_γ, T and T_0 in an Euclidean space, we have*

(a) any cluster point of any set $W = \{z_{\gamma_n} \,|\, n = 1, 2, \ldots; \gamma_n \to 0^+\}$ is a zero point of T;

(b) $T^{-1}(0) = \emptyset$ if and only if $\|z_\gamma\| \to \infty$ as $\gamma \to 0^+$;

(c) $T^{-1}(0) \ne \emptyset$ if and only if $S_\mu = \{z_\gamma \,|\, \mu \ge \gamma > 0\}$ is bounded for every μ with $0 < \mu < \infty$.

Proof of (a). If z^* is a cluster point of W, then there is a subsequence $\{z_{\gamma_i}\} \subset W$ such that $z_{\gamma_i} \to z^*$ as $i \to \infty$.

Now from (8.3.6), for any z and $y \in T(z)$,

$$0 \leq \langle z_{\gamma_i} - z, -\gamma_i T_0(z_{\gamma_i}) - y \rangle \tag{8.3.11}$$

As $i \to \infty$, we have $\gamma_i \to 0^+$ and obtain from the continuity of T_0 and convergence of $\{z_{\gamma_i}\}$ that

$$0 \leq \langle z^* - z, 0 - y \rangle, \qquad \text{for any} \quad z \quad \text{and} \quad y \in T(z). \tag{8.3.12}$$

The maximal monotonicity of T implies that $0 \in T(z^*)$.

Proof of (b). If $T^{-1}(0) = \emptyset$, then Part (a) implies

$$\lim_{\gamma \to 0+} \|z_\gamma\| = \infty. \tag{8.3.13}$$

On the other hand, if (8.3.13) is true, Proposition 8.3.1 guarantees $T^{-1}(0) = \emptyset$.

Proof of (c). If $T^{-1}(0) \neq \emptyset$, Proposition 8.3.1 gives that S is bounded. Conversely, if S is bounded, Part (b) implies that $T^{-1}(0) \neq \emptyset$. □

Proposition 8.3.3 (Continuity of z_γ). *Under the assumptions on z_γ, T_γ, T and T_0, z_γ is continuous in γ for $\gamma > 0$.*

Proof. Let γ and β be any two positive scalars. Then $T_\gamma(z_\beta) = T(z_\beta) + \gamma T_0(z_\beta)$ and $-\beta T_0(z_\beta) \in T(z_\beta)$ from (8.3.6). Hence, if α is the strong monotone modulus of T_γ,

$$\begin{aligned} \alpha \|z_\gamma - z_\beta\|^2 &\leq \langle z_\gamma - z_\beta, 0 - [-\beta T_0(z_\beta) + \gamma T_0(z_\beta)] \rangle \\ &\leq |\gamma - \beta| \|z_\gamma - z_\beta\| \|T_0(z_\beta)\|. \end{aligned} \tag{8.3.14}$$

Thus, we have

$$\|z_\gamma - z_\beta\| \leq \frac{|\beta - \gamma|}{\alpha} \|T_0(z_\beta)\|. \tag{8.3.15}$$

Therefore,

$$\|z_\gamma - z_\beta\| \to 0 \qquad \text{as} \qquad \gamma \to \beta. \tag{8.3.16}$$

This proves the continuity of z_γ in $\gamma > 0$. □

Theorem 8.3.4 (Convergence of z_γ). *Under the assumptions on T, T_0 and z_γ, we have*
(a) either $\|z_\gamma\| \to \infty$ as $\gamma \to 0^+$ if $T^{-1}(0) = \emptyset$;
(b) or $z_\gamma \to z^*$ as $\gamma \to 0^+$ with $z^* \in T^{-1}(0)$.

Proof. Part (a) is direct from Proposition 8.3.2. As for Part (b), let us now suppose the conclusion of Part (b) is not true. This means from Proposition 8.3.2 that there are two sequences $\{z_{\gamma_i}\}$, $\{z_{\gamma_j}\}$ and two zero points z^*, $\bar{z}$ of T such that $\gamma_i \to 0$, $\gamma_j \to 0$, $z_{\gamma_i} \to z^*$ and $z_{\gamma_j} \to \bar{z}$ with $z^* \neq \bar{z}$ when $i \to \infty$ and $j \to \infty$. Let α_0 be the scalar such that

$$\langle z^* - \bar{z}, T_0(z^*) - T_0(\bar{z})\rangle = \alpha_0 \|z^* - \bar{z}\|^2. \tag{8.3.17}$$

Then, from the strong monotonicity of T_γ for $\gamma > 0$,

$$0 < \langle z^* - \bar{z}, 0 + \gamma T_0(z^*) - 0 - \gamma T_0(\bar{z})\rangle.$$

Therefore, $\alpha_0 > 0$ in (8.3.17). For any $\varepsilon > 0$, there is a N such that

$$\begin{aligned} &\langle T_0(z_{\gamma_j}) - T_0(z^*), z_{\gamma_j} - z^*\rangle > (\alpha_0 - \varepsilon)\|z_{\gamma_j} - z^*\|^2, \\ &\langle T_0(z_{\gamma_i}) - T_0(\bar{z}), z_{\gamma_i} - \bar{z}\rangle > (\alpha_0 - \varepsilon)\|z_{\gamma_i} - \bar{z}\|^2 \end{aligned} \tag{8.3.18}$$

when $j > N$ and $i > N$ from the continuity of T_0 and the convergence of $\{z_{\gamma_i}\}$ and $\{z_{\gamma_j}\}$. However, similar to the proof of (8.3.9), we have

$$\begin{aligned} &\langle z_{\gamma_j} - z^*, T_0(z_{\gamma_j}) - T_0(z^*)\rangle \le \langle z_{\gamma_j} - z^*, -T_0(z^*)\rangle, \\ &\langle z_{\gamma_i} - \bar{z}, T_0(z_{\gamma_i}) - T_0(\bar{z})\rangle \le \langle z_{\gamma_i} - \bar{z}, -T_0(\bar{z})\rangle. \end{aligned} \tag{8.3.19}$$

Combine (8.3.19) with (8.3.18) and let j and i go to infinite. Then,

$$\begin{aligned} &\langle \bar{z} - z^*, -T_0(z^*)\rangle \ge (\alpha_0 - \varepsilon)\|\bar{z} - z^*\|^2, \\ &\langle z^* - \bar{z}, -T_0(\bar{z})\rangle \ge (\alpha_0 - \varepsilon\|z^* - \bar{z}\|^2 \end{aligned} \tag{8.3.20}$$

for any $\varepsilon > 0$. Adding the both sides of the two inequalities in (8.3.20) gives

$$\langle z^* - \bar{z}, T_0(z^*) - T_0(\bar{z})\rangle \ge 2(\alpha_0 - \varepsilon)\|z^* - \bar{z}\|^2 \tag{8.3.21}$$

for any $\varepsilon > 0$. Letting $\varepsilon \to 0^+$ implies

$$\langle z^* - \bar{z}, T_0(z^*) - T_0(\bar{z})\rangle \geq 2\alpha_0\|z^* - \bar{z}\|^2. \tag{8.3.22}$$

(8.3.22) contradicts (8.3.17). This contradiction proves the conclusion of Part (b). □

Theorem 8.3.4 states that z_γ is an approximate solution of $0 \in T(z)$ if γ is small. Therefore, when T is not strongly monotone, we may use this kind of *penalty operator technique* for solving $0 \in T_\gamma(z_\gamma)$ instead of $0 \in T(z^*)$ to approach the solution.

Let us consider several specific examples of T_0 here.

Example 8.3.1. T_0 is linear, that is,

$$T_0(z) = Hz + h, \qquad H \text{ is semi-positive definite }, \tag{8.3.23}$$

where h is a constant vector.

When $H = I$ and $h = z_k$ in (8.3.23), $0 \in (T + \gamma T_0)(z_\gamma)$ in (8.3.6) with T_0 given by (8.3.23) is just a proximal point step $z_\gamma = (I + \gamma^{-1}T)^{-1}(z^k)$.

Example 8.3.2. If $T = \partial f$ and $T_0(z) = z - a$, then (8.3.6) is equivalent to penalty function methods for minimizing points of f if f is convex.

Example 8.3.3. If T is given by (8.2.9), then $T_0 = \mathrm{diag}\{H_1, \ldots, H_{2N+1}\}$ can be a choice with given positive definite matrices H_i, $i = 1, \cdots, 2N + 1$. Thus, $T_\gamma = T + \gamma T_0$ is strongly monotone ($\gamma > 0$).

8.4. Implementation Detail and EFBS Algorithms

From the discussion in previous sections, we see more issues that need to be shown for the implementation of FBS algorithm. Hence, we shall consider the choice of splitting strategies $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$, the determination of the metric matrix H and the "preconditioner" matrix Q_{1i} to make FBS algorithms converge faster.

For the choices of step sizes $\{\lambda_k\}$, we have discussed them theoretically in Chapter 6. In practice, they are much more complicated. We shall leave them to a special separated paper.

From Chapter 5, we know that if we can choose $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$, H and Q_{1i} such that the linear convergence rate $\Theta(\lambda_k)$ is as small as possible, then we can obtain better Q-linear convergence rate. However, those choices for the metric H and splitting strategies have been given in Chapter 6. Here, we just apply them to our special case.

Theorem 8.4.1 (splitting of T). *Let T be given by (8.2.9) and $\mathcal{J}$ be defined in Section 8.2. If we restrict the across terms between u and v into $\mathcal{J}_1$, then the best splitting of $T = T_1 + T_2$ under the meaning of (6.5.4) with T_1, T_2 from (8.2.8) is*

$$\mathcal{J}_1(u,v) = -l(u,v) - \sum_{i=1}^{N} v_i^T D_i u_i, \quad \mathcal{J}_2(u,v) = \mathcal{J}(u,v) - \mathcal{J}_1(u,v). \tag{8.4.1}$$

Proof. This is a direct corollary of Theorem 6.5.4. □

Theorem 8.4.2 (Choice of H). *When $T = T_1 + T_2$ is split by (8.2.8) and (8.4.1), an approximate best choice of H in (8.1.3) under the meaning of (6.6.7) (minimal linear convergence rate) with $\mathcal{S}$ given by (6.6.8) is*

$$H = \mathrm{diag}\{P_1, \cdots, P_N, Q_{11}, \gamma I, Q_{12}, \gamma I, \cdots, Q_{1(N+1)}, \gamma I\} \tag{8.4.2}$$

where I is identity matrix, γ is a small positive scalar and Q_{1i} ($i = 1, \cdots, N+1$) are defined in Theorem 8.2.3.

Proof. Let the penalty operator $T_0 = \mathrm{diag}\{0, \cdots, 0, 0, I, 0, I, \cdots, 0, I\}$ in penalty operator techniques, where 0 is a zero matrix of proper dimensionality. We solve $0 \in T_\gamma(z) = T(z) + \gamma T_0(z)$ by applying the EFBS method from (8.1.3) to it. Then, the zero point of T_γ is a approximate solution of finding a zero point of T from Section 8.3 for small $\gamma > 0$. However, applying Theorem 6.6.4 to T_γ gives the best choice of H in (8.4.2). This gives our conclusion. □

Remark: In the above proof, when there are no inequality constraints in Problem $(\mathcal{O}')$, T_0 becomes a zero operator and $D_i \equiv 0$, $q_i \equiv 0$, $Q_i \equiv Q_{1i}$, $C_i \equiv C_{1i}$, $v_i \equiv v_{1i}$ in (8.2.6). In this case, $T_\gamma = T$ is strongly monotone.

Now let us give the EFBS algorithms which apply here. By applying the results of Section 7.5 and Section 7.8 to our optimal control problems through the penalty operator techniques described in the proof of Theorem 8.4.2, we obtain the following detailed two algorithms under the choices of Theorem 8.4.1 and Theorem 8.4.2.

ALGORITHM 8.1 (without line search):

0). **Initial input**: Input an initial point (u^0, v^0), initial $\lambda_0 > 0$, an error bound $\varepsilon > 0$ and the matrix H giving by (8.4.2). Let $k := 0$.

1). **EFBS iteration:** Use the EFBS scheme (8.1.3) and Theorem 7.8.1 to get (u^{k+1}, v^{k+1}) from (u^k, v^k). That is,

$$\begin{aligned} u_i^{k+1} &= \operatorname*{argmin}_{u \in U_i} \left[p_i^T u + \frac{1}{2} u^T P_i u + \frac{1}{2\lambda_k} \| u - P_i^{-1} \hat{u}_i^k \|^2 \right], \\ \hat{u}_i^k &= P_i u_i^k + \lambda_k [B_i^T y_{i+1}^k - G_i^T v_{2i}^k], \qquad i = 1, 2, \cdots, N; \end{aligned} \tag{8.4.3}$$

and

$$\begin{aligned} v_i^{k+1} &= \operatorname*{argmin}_{v \in V_i} \left[\frac{1}{2} v^T \bar{Q}_i v - q_i^T v + \frac{1}{2\lambda_k} \| v - \bar{Q}_i^{-1} \hat{v}_i^k \|^2 \right], \\ \hat{v}_i^k &= \bar{Q}_i v_i^k - \lambda_k [C_i x_i^k - G_i u_i^k], \qquad i = 1, \cdots, N+1. \end{aligned} \tag{8.4.4}$$

Where $G_{N+1} = 0$, $\bar{Q}_i = \operatorname{diag}\{Q_{1i}, \gamma I\}$ and x_i^k, y_i^k are the solutions of the dynamical systems (8.2.1) and (8.2.2) associated with given u_i^k, v_i^k.

2). **Primal-dual gap check:** Compute $f_k = \mathcal{F}(u^{k+1})$ and $g_k = \mathcal{G}(v^{k+1})$, namely,

$$f_k = \mathcal{J}(u^{k+1}, \tilde{v}), \qquad g_k = \mathcal{J}(\tilde{u}, v^{k+1}) \tag{8.4.5}$$

where

$$\begin{aligned} \tilde{v}_i &= \bar{P}_{V_i} \big[\bar{Q}_i^{-1} \big(q_i + G_i u_i^{k+1} - C_i x_i^{k+1} \big) \big], \qquad i = 1, \cdots, N, N+1; \\ \tilde{u}_i &= \bar{P}_{U_i} \big[P_i^{-1} \big(B_i^T y_{i+1}^{k+1} - G_i^T v_{2i}^{k+1} - p_i \big) \big], \qquad i = 1, 2, \cdots, N \end{aligned} \tag{8.4.6}$$

and $\bar{P}_{U_i}$ is the projection operator from $\mathbb{R}^{n_c}$ to U_i. $\bar{P}_{V_i}$ has similar meaning. If

$$f_k - g_k \leq \varepsilon, \tag{8.4.7}$$

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

3). . **Distance check:** If

$$\|u^{k+1} - u^k\|_P + \|v^{k+1} - v^k\|_{\bar{Q}} \leq \varepsilon, \tag{8.4.8}$$

or the iteration number arrives at a maximal number requirement, stop iteration and output the ε-optimal or *max*-number solution $(u^*, v^*) \approx (u^{k+1}, v^{k+1})$. Otherwise, go to the next step. Here $P = \text{diag}\{P_1, \cdots, P_N\}$ and $\bar{Q} = \text{diag}\{\bar{Q}_1, \cdots, \bar{Q}_{N+1}\}$ in (8.4.8).

4). **Constant modification:** Choose suitable $\lambda_{k+1} > 0$ and let $k := k + 1$. Go back to Step 1).

ALGORITHM 8.2 (with line search):

0. **Initial input**: Input an initial point $(u^0, v^0) = (\bar{u}^0, \bar{v}^0)$, initial $\lambda_0 > 0$, an error bound $varepsilon > 0$ and the matrix H giving by (8.4.2). Let $k := 0$.
1. **EFBS iteration:** Same as Algorithm 1.
2. **Primal-dual gap check:** Same as Algorithm 1.
3. **Distance check:** Same as Algorithm 1.
4. **Line search:** Minimize $\mathcal{F}(u)$ over the line segment $[\bar{u}^k, u^{k+1}]$ to get $\bar{u}^{k+1}$ and maximize $\mathcal{G}(v)$ over the line segment $[\bar{v}^k, v^{k+1}]$ to get $\bar{v}^{k+1}$.
5. **Another primal-dual gap check:** Do the same as Step 2 but replace (u^{k+1}, v^{k+1}) by $(\bar{u}^{k+1}, \bar{v}^{k+1})$.
6. **Constant modification:** Choose suitable $\lambda_{k+1} > 0$ and let k be increased by 1. Return then to Step 1.

In the above two algorithms, the choice of $\{\lambda_k\}$ may be made from the results of Theorem 7.8.2. Equalities (8.4.3), (8.4.4), then (8.4.6) will become a group of one-dimensional subproblems when P, $\bar{Q}$ are diagonal and U, V are boxes. At the box-diagonal case, Step 4 of Algorithm 8.2 may be solved from (7.1.7) and (7.1.8).

8.5. Numerical Examples and Computing Results

A. Example 8.1: (Bertsekas 1982)

$$\min \int_0^1 [6u(t)^2 + 2x_1(t)^2 + x_2(t)^2]\, dt, \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 + u \end{pmatrix}, \tag{8.5.1}$$

$$|u(t)| \le B, \qquad x_1(0) = 15, \qquad x_2(0) = 5, \tag{8.5.2}$$

where B is a positive constant to be specified.

Suppose $x \in \mathbb{R}^{n_s}$, $u \in \mathbb{R}^{n_c}$ and the problem has n_g constraints at each stage. Then, let $\varepsilon = 10^{-10}\sqrt{N(n_c + 2n_s + n_g)}$. Apply Algorithm 8.1 to this problem. The N is the discretizing stage number of the problem into the form of Problem $(\mathcal{O}')$ through the technique in Section 7.8.

(1). Computational Results of Example 8.1:

Table 8.1. Comparison of EFBS and Interior Point Methods for Ex. 8.1

B	N	IPM1 IP/time	IPM2 IP/time	IPM objective	FBS1 iter/time	FBS2 iter/time	EFBS objective
1.00	10^2	12/0.48	20/1.61	473.5738	8/0.047	8/0.047	473.5733
	10^3	16/6.33	38/30.73	471.108521	8/0.43	8/0.41	471.108516
	10^4	18/71.16	67/558.25	470.86080522	7/3.77	7/3.80	470.86080519
	10^5				7/38.18	7/36.99	470.8360
1.25	10^2	16/0.62	23/2.29	473.4119	9/0.051	9/0.051	473.4113
	10^3	16/6.46	36/28.65	470.936984	8/0.42	8/0.41	470.936978
	10^4	19/75.71	66/545.20	470.68831616	8/4.28	8/4.20	470.68831615
	10^5				7/37.73	7/37.09	470.6634
1.40	10^2	13/0.51	20/1.59	473.3622	9/0.051	9/0.051	473.3614
	10^3	17/6.70	37/29.86	470.883134	8/0.42	8/0.41	470.883127
	10^4	19/75.31	65/539.69	470.63405897	8/4.27	8/4.44	470.63405899
	10^5				8/43.11	8/41.97	470.6091
2.00	10^2	9/0.36	13/1.06	473.3264	9/0.051	9/0.051	473.3254
	10^3	10/4.02	24/19.69	470.841834	9/0.46	9/0.47	470.841824
	10^4	10/41.02	41/340.39	470.59221705	8/4.42	9/4.77	470.59221695
	10^5				8/43.16	8/42.06	470.5672

We compare extended forward-backward splitting methods with the two interior point methods discussed in the paper of Steve Wright (1993). The program codes and input data for the interior point methods come from Dr. Wright directly. The different EFBS methods FBS1 and FBS2 represent the algorithm 8.1 and EFBS method with projection splitting without line search. We choose $\lambda_k = 100$ for the Algorithm 8.1 and $\lambda_k = 1.0$ for the FBS2 in Table 8.1. When $N = 100000$, our computer workstations do not have enough spaces for the interior point method but do have enough spaces for the EFBS method. All results in this section are computed on the DEC 5000/240 workstation in the computer laboratory of applied mathematics at the University of Washington.

(2). Results of the Method FBS1 for Different λ_k

We use EFBS method FBS1 (Algorithm 8.1 in Section 8.4) as example to see the influence of step sizes. $\lambda_k = 100$ and $\lambda_k = 1.0$ have been chosen to compute Example 1. The results are as follows.

Table 8.2. Influence of Step Sizes for EFBS Methods

N	B	$\lambda_k = 100$ iter/time	$\lambda_k = 1.0$ iter/time	B	$\lambda_k = 100$ iter/time	$\lambda_k = 1.0$ iter/time
1000	1.0	8/0.43	18/0.91	1.4	8/0.42	18/0.91
	1.25	8/0.42	18/0.91	2.0	9/0.46	18/0.91
10000	1.0	7/3.77	18/9.34	1.4	8/4.27	18/9.35
	1.25	8/4.28	18/9.34	2.0	8/4.42	18/10.02
100000	1.0	7/38.18	17/91.40	1.4	8/43.11	17/89.35
	1.25	7/37.73	17/88.63	2.0	8/43.16	17/91.43

(3). Results of the Method FBS1 for Different ε

We use EFBS method FBS1 (Algorithm 8.1 in Section 8.4) as an example to see the influence of the stopping criterion ε. We fix $\lambda_k = 100$ and $N = 200,000$ (hence the total variable number of u, v, x, and y is **1.4 millions!!**).

N=200000 and $\lambda_k = 100$

Table 8.3. Influence of Stopping Bounds for EFBS Methods

B	ε	iter/time		ε	iter/time
1.0	10^{-10}	10/130.6		N^{-1}	6/88.4
1.25		10/144.6		(5×10^{-6})	6/88.4
1.4		10/144.5			6/92.6

B. Example 8.2: (Wright 1992)

$$\min \int_0^1 [2x_1(t)^2 + x_2(t)^2]\, dt, \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 + u \end{pmatrix}, \tag{8.5.3}$$

$$|u(t)| \le B, \qquad x_1(0) = 15, \qquad x_2(0) = 5, \tag{8.5.4}$$

where B is a positive constant to be specified.

This problem has a bang-bang solution. Let t_s be the switching time such that $u(t)$ switchs from $u(t) = -B$ to $u(t) = B$ at t_s. I have used penalty operator methods to compute it for satisfying convergence conditions. Hence, I have added a term $10^{-4}u(t)$ into the integrand.

We compare the extended forward-backward splitting method Algorithm 8.1 (FBS1) of Section 8.4 with the two interior point methods in the paper of Steve Wright (1992). We choose $\lambda_k = 100$ for the FBS1 and the same stopping criterion as Example 1 in the following.

(4). Computational Results for Example 8.2

Table 8.4. Comparison of EFBS and Interior Point Methods for Ex. 8.2

B	N	IPM1 IP/time	IPM2 IP/time	IPM t_s	IPM objective	FBS1 it/time	EFBS t_s	EFBS objective
1.0	10^3	15/5.94	37/29.82	0.512	468.084757	9/0.46	0.510	468.084750
	10^4	16/66.2	64/535.1	0.512	467.834964	8/4.25	0.5110	467.834964
	10^5					8/43.0	0.5111	467.809974
2.0	10^3	13/5.23	36/29.63	0.500	462.563387	20/1.00	0.499	462.563385
	10^4	16/65.5	64/538.9	0.500	462.304680	13/7.54	0.4991	462.304680
	10^5					12/63.3	0.4992	462.278798

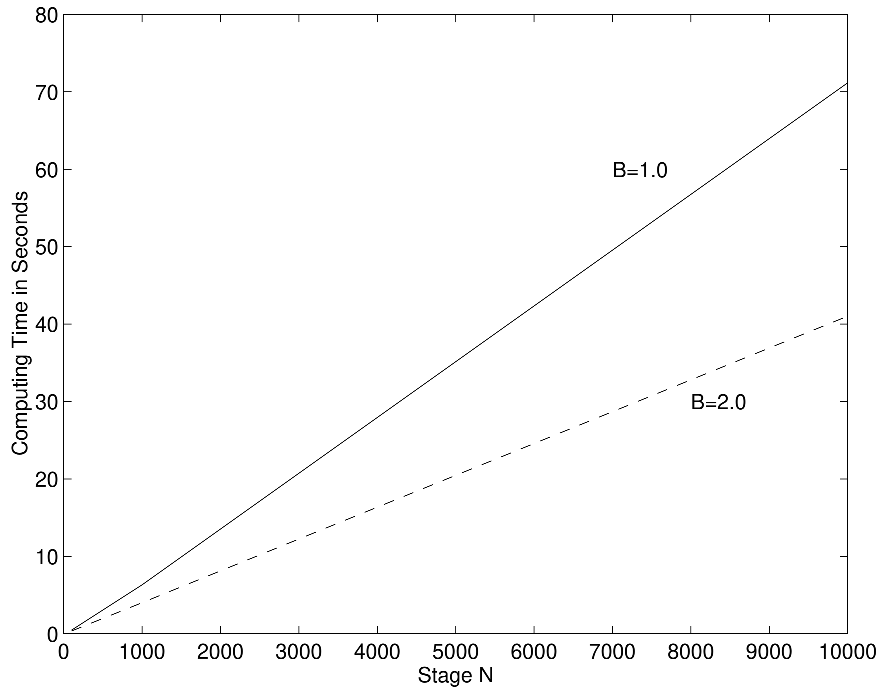

Figure 8.1: Computing Time of Example 8.1 by IPM1

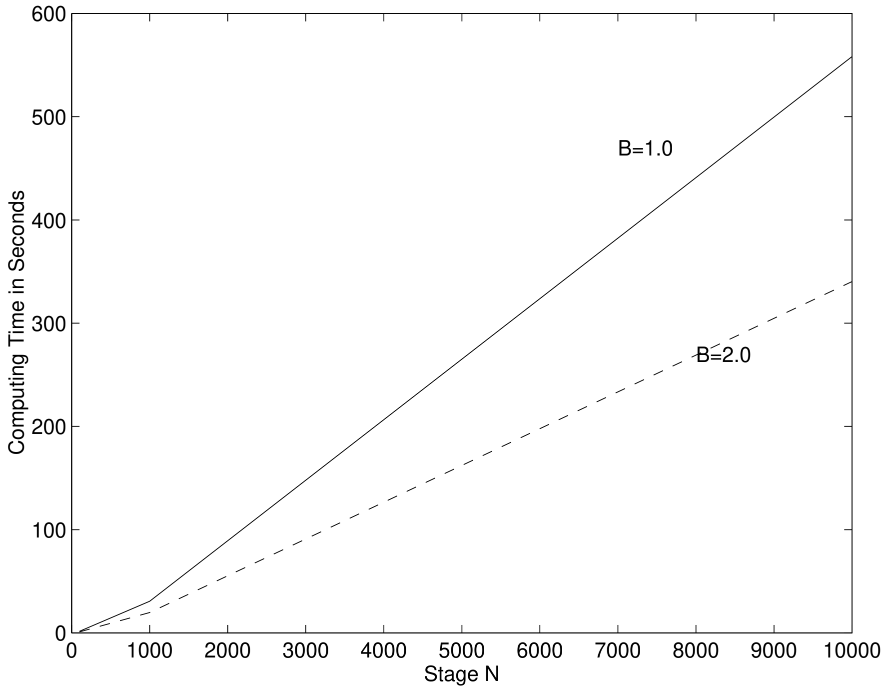

Figure 8.2: Computing Time of Example 8.1 by IPM2

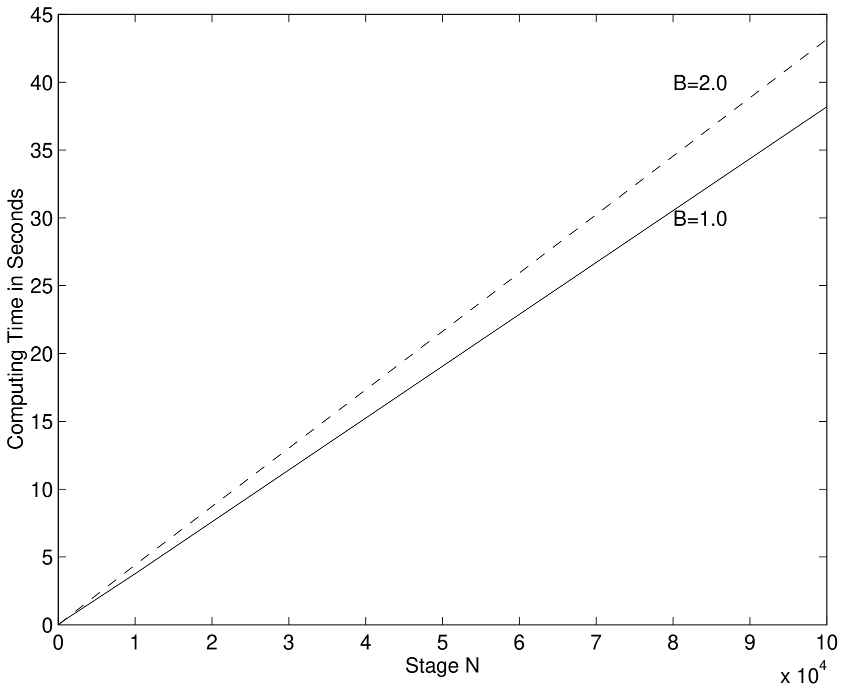

Figure 8.3: Computing Time of Example 8.1 by FBS1

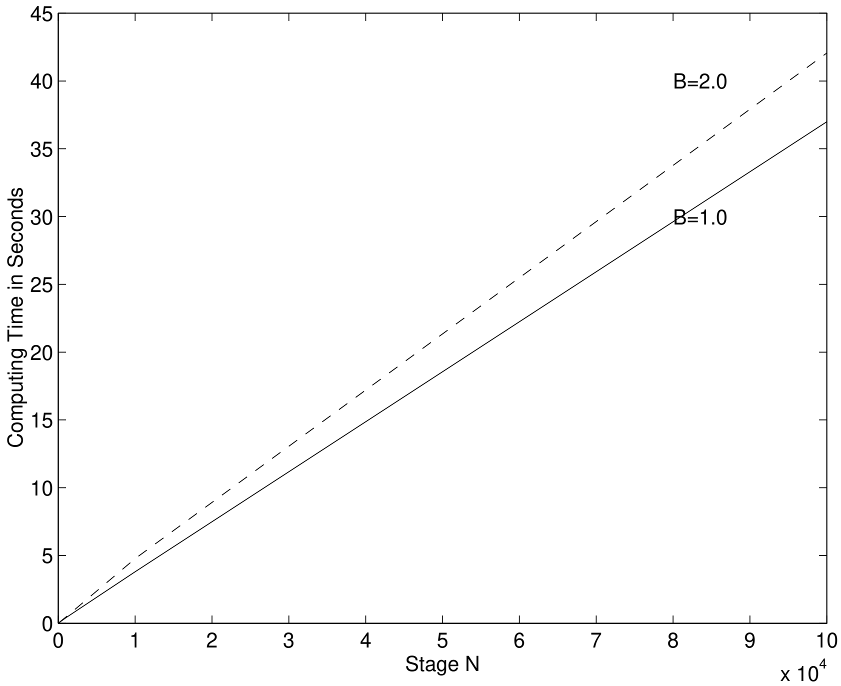

Figure 8.4: Computing Time of Example 8.1 by FBS2

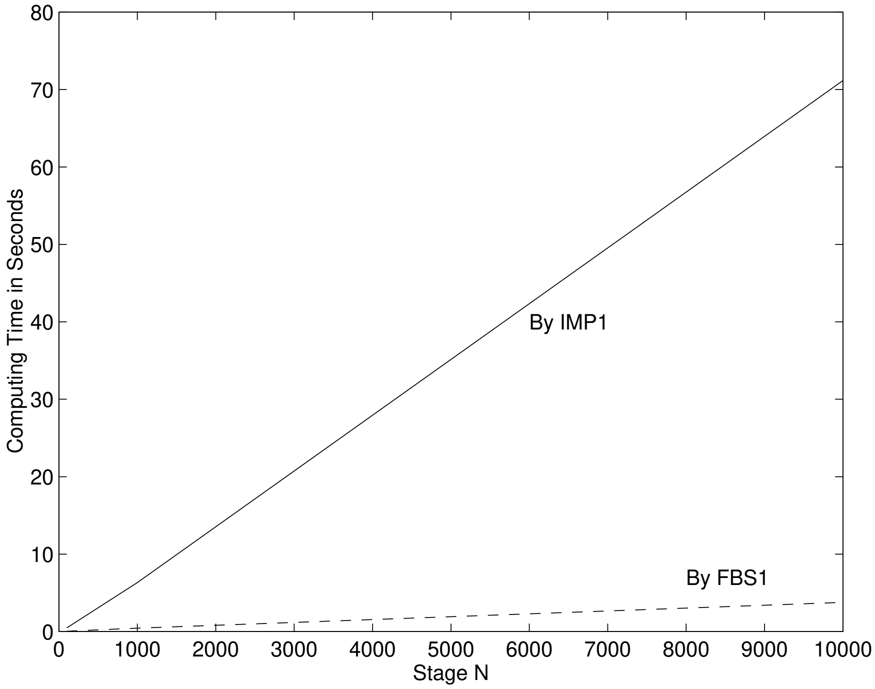

Figure 8.5: Comparison of IPM1 and FBS1 for B=1.0

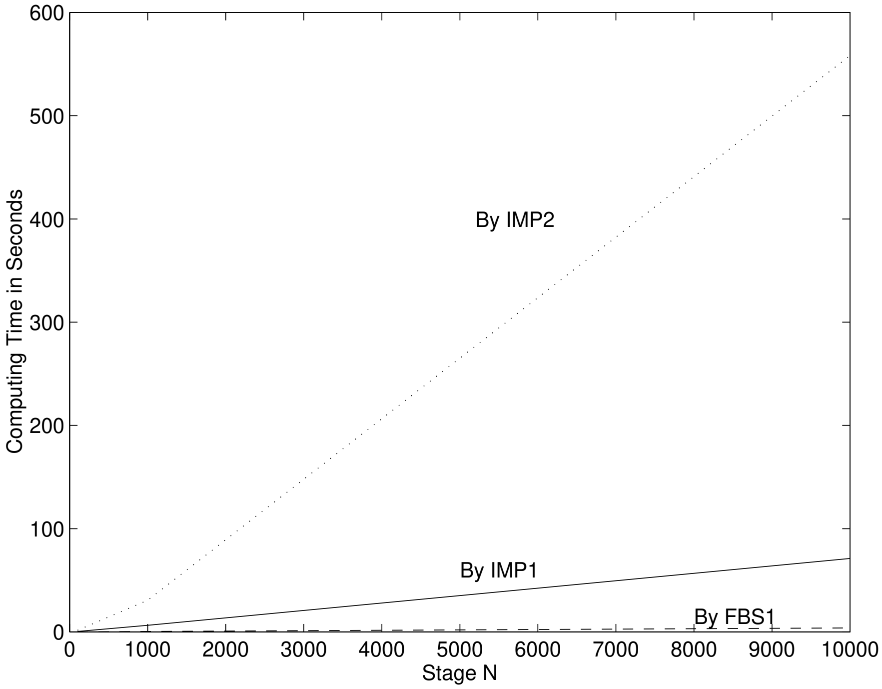

Figure 8.6: Comparison of IPM1, IPM2, & FBS1 for B=1.0