AN SQP TRUST REGION METHOD FOR SOLVING THE DISCRETE–TIME LINEAR QUADRATIC CONTROL PROBLEM

EL-SAYED M.E. MOSTAFA

Department of Mathematics, Faculty of Science
Alexandria University, Moharam Bey, 21511, Alexandria, Egypt
e-mail: emostafa@alex-sci.edu.eg

In this paper, a sequential quadratic programming method combined with a trust region globalization strategy is analyzed and studied for solving a certain nonlinear constrained optimization problem with matrix variables. The optimization problem is derived from the infinite-horizon linear quadratic control problem for discrete-time systems when a complete set of state variables is not available. Moreover, a parametrization approach is introduced that does not require starting a feasible solution to initiate the proposed SQP trust region method. To demonstrate the effectiveness of the method, some numerical results are presented in detail.

Keywords: output feedback control design, sequential quadratic programming, trust region method.

1. Introduction

The design problem of an optimal output feedback controller for linear (continuous or discrete-time) control systems has been attracting much attention during the last four decades. The interest in this type of problems is motivated by numerous applications found in the system and control literature that are solved through this formulation; see, for instance, the benchmark collection outlined by Leibfritz (2004).

The determination procedure of the output feedback matrix that asymptotically stabilizes linear control systems is actually an open problem. Several methods have been proposed in the literature based on Lyapunov and Riccati equations, linear matrix inequalities, or eigenstructure assignment; see, e.g., the surveys of Mäkilä and Toivonen (1987) and Syrmos *et al.* (1997) as well as later ones, e.g., those by Varga and Pieters (1998), Garcia *et al.* (2001), Sulikowski *et al.* (2004), Zhai *et al.* (2005), Lee and Khargonekar (2007), Mostafa (2008), and the references therein.

Various gradient-based methods have been considered for solving the discrete-time LQR problem, among them the descent Anderson-Moore method (Mäkilä and Toivonen, 1987). This method solves the LQR problem by successfully minimizing a particular quadratic approximation of the objective function combined with the step-size rule. Mäkilä and Toivonen (1987) solve the discrete-time LQR problem with Newton's method with line search globalization. Mostafa (2008) introduces an unconstrained trust region method for solving the discrete-time LQR problem. A nonlinear conjugate gradient has recently been proposed by Mostafa (2012) which belongs to low storage methods for solving this problem. All these methods are based on formulating the LQR problem as an unconstrained matrix optimization one.

On the other hand, the LQR problem for continuous-time control systems has been stated in different constrained optimization forms and solved by various constrained optimization techniques. Leibfritz and Mostafa (2002) formulate the LQR problem as a nonlinear semi-definite programming one and suggest as a solution an interior-point trust region method. Moreover, Leibfritz and Mostafa (2003) introduced unconstrained and constrained trust region approaches for solving two formulations of the LQR problem. Kočvara *et al.* (2005) introduce an augmented Lagrangian semi-definite programming method for solving this problem. Mostafa (2005a; 2005b) suggests a trust region method for solving the decentralized LQR problem and an augmented Lagrangian SQP method for solving a special class of nonlinear semi-definite programming problems related to the LQR problem, respectively.

To the best of the author's knowledge, the LQR problem for discrete-time systems has not been solved yet with any higher-order constrained optimization tech-

niques. This represents the motivation of this paper.

Consider a discrete-time linear time-invariant system described by the following discrete state space equation:

$$x_{k+1} = Ax_k + Bu_k, \quad y_k = Cx_k, \quad k \geq 0, \quad (1)$$

where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^p$ is the control input vector and $y_k \in \mathbb{R}^r$ is the measured output vector. Moreover, $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{r \times n}$, $(r \leq n)$ are given constant matrices. Models of this form usually appear in the discretization of linear continuous-time models of system and optimal control applications; see, e.g., the benchmark collection (Leibfritz, 2004).

We shall consider the following well-known cost function to be minimized:

$$J = \mathbb{E}\Big[\sum_{k=0}^{\infty} \langle x_k, Qx_k \rangle + \langle u_k, Ru_k \rangle \Big], \quad (2)$$

where $\mathbb{E}[\cdot]$ is the expected value, $\langle \cdot, \cdot \rangle$ is the inner product operator, $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{p \times p}$ are given symmetric positive semi-definite and positive definite matrices, respectively. The objective function consists of two parts; the first one $\langle x_k, Qx_k \rangle$ concerns minimization over the state variable while $\langle u_k, Ru_k \rangle$ regards minimization over the control input variable.

The following control law is often used to close the above linear control system:

$$u_k = Fy_k, \quad (3)$$

where $F \in \mathbb{R}^{p \times r}$ is the output feedback gain matrix. The main objective is to calculate an optimal F that minimizes the cost function (2) and at the same time stabilizes the control system (1).

In the full state feedback problem, all components of the state vector x_k can be observed or measured. In such a case $C = I_n$ and $y_k = x_k$. Observe that in the discrete-time LQR problem considered we assume that a complete set of state variables is not available, which is a more general setting than in the case of the full state problem.

The derivation of the optimization problem from the corresponding optimal control one is included for completeness. From (1) and (3) we have $u_k = Fy_k = FCx_k$. Hence, the control system (1) is rewritten as

$$x_k = (A + BFC)x_{k-1} = A(F)x_{k-1} = A(F)^k x_0, \quad (4)$$

where $A(F) = (A + BFC)$. In order to stabilize the control system, the matrix variable F is restricted to lie within the following set of stabilizing output feedback gains:

$$\mathcal{D}_s = \Big\{ F \in \mathbb{R}^{p \times r} \: : \: \rho(A(F)) < 1 \Big\}, \quad (5)$$

where $\rho(\cdot)$ is the spectral radius. By choosing $F \in \mathcal{D}_s$ the state variables vector x_k decays to the zero state as $k \to \infty$.

From (4) in (2) we obtain

$$J(F) = \mathbb{E}\Big[x_0^T \sum_{k=0}^{\infty} [(A(F)^k)^T Q(F) A(F)^k] x_0 \Big]$$
$$= \mathbb{E}\big[x_0^T K(F) x_0 \big]$$
$$= \mathrm{Tr}\,(K(F)V), \quad (6)$$

where $Q(F) = Q + C^T F^T R F C$, $\mathrm{Tr}(\cdot)$ is the trace operator, $V = \mathbb{E}[x_0 x_0^T]$ is the covariance matrix, and

$$K(F) = \sum_{k=0}^{\infty} (A(F)^k)^T Q(F) A(F)^k \quad (7)$$

solves the discrete Lyapunov equation

$$K(F) = A(F)^T K(F) A(F) + Q(F). \quad (8)$$

The objective function depends on $x_0 \in \mathbb{R}^n$. To remove this dependence it is often assumed that x_0 is a random variable uniformly distributed over a unit sphere with $\mathbb{E}[x_0] = 0$, see, e.g., (Mäkilä and Toivonen, 1987). Equations (6)–(8) yield the unconstrained minimization problem:

$$\min_{F \in \mathcal{D}_s} \tilde{J}(F) = \langle K(F), V \rangle, \quad (9)$$

where $K(F)$ solves (8).

Note that an optimal $F \in \mathcal{D}_s$ that solves this optimization problem equivalently solves the linear quadratic control one (1)–(2). We consider both variables F and K to be independent and avoid the above dependency of K on F. This leads to the following constrained and matrix optimization problem:

$$\min_{(K,F) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s} J(K) = \langle K, V \rangle, \quad (10)$$

subject to

$$\tilde{H}(K, F) = K - A(F)^T K A(F) - Q(F) = 0. \quad (11)$$

A dual form to this problem is the following constrained optimization problem:

$$\min_{(L,F) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s} J(L, F) = \langle L, Q(F) \rangle, \quad (12)$$

subject to

$$H(L, F) = L - A(F) L A(F)^T - V = 0. \quad (13)$$

One can show the relationship between both problems easily as follows: Let $F \in \mathcal{D}_s$, and rewrite (11) and (13) as

$$Q(F) = K - A(F)^T K A(F), \quad (14)$$

$$V = L - A(F) L A(F)^T, \quad (15)$$

where K and L solutions of the discrete Lyapunov equations (11) and (13) are symmetric. Pre-multiplying the

first equation by L and the second one by K and then applying the trace operator gives

$$\text{Tr}\,(L\,Q(F)) = \text{Tr}\,(LK) - Tr\,(LA(F)KA(F)^T)$$

$$= \text{Tr}\,(KV). \qquad (16)$$

Hence, the objective functions of the two problems are equal.

A classical approach developed in the past for solving the discrete-time LQR problem solves the unconstrained problem (9) combined with (8) or the following unconstrained minimization problem (see Mäkilä and Toivonen, 1987):

$$\min_{F \in \mathcal{D}_s} \hat{J}(F) = J(L(F), F) = \langle L(F), Q(F) \rangle, \quad (17)$$

where $Q(F)$ is defined as above and $L(F)$ solves the discrete Lyapunov equation

$$L(F) = A(F)L(F)A(F)^T + V. \qquad (18)$$

Note that the problem (17)–(18) is an unconstrained version of the constrained problem (12)–(13). In the remainder of the paper we consider the constrained problem (12)–(13). The following are important issues which should be taken into consideration when solving any of the two problems:

- The equality constraint is not treated as a Lyapunov equation and therefore it is nonlinear.

- The variable L in (12)–(13) (or K in (10)–(11)) is larger in size than the variable F.

- Both problems are nonlinear and nonconvex. The nonconvexity is simply due to the eigenvalue constraint $F \in \mathcal{D}_s$. Moreover, we have matrix variables and therefore directional derivatives are used.

SQP methods have been shown to be among the most effective techniques for solving nonlinear constrained optimization problems. These methods generate steps by solving a sequence of quadratic programming problems. Moreover, SQP methods like Newton's one for unconstrained optimization are locally convergent and must be combined with a globalization strategy such as line search or the trust region. In this paper we study the numerical solution of the optimization problem (12)–(13) with an SQP method combined with a trust region globalization strategy. The interested reader is referred, for instance, to the books by Conn *et al.* (2000) as well as Nocedal and Wright (1999) for surveys on this topic.

Moreover, the existence of a feasible point $F_0 \in \mathcal{D}_s$ is necessary to start the sequence of iterations to any of the methods that seek an optimal or an suboptimal solution of the above minimization problems. A parametrization

approach is applied to the optimization problem (12)–(13) which does not require the existence of $F_0 \in \mathcal{D}_s$ to start the proposed method.

This paper is organized as follows. In the next section an outline of the SQP method and the corresponding QP subproblem is given. In Section 3 an SQP trust region algorithm (denoted by DISTR) is presented for solving the minimization problem (12)–(13). In Section 4 the trust region method is applied to a parameterized formulation of the problem (12)–(13) that does not require initial stabilizing $F_0 \in \mathcal{D}_s$. In Section 5 the performance of the DISTR method is tested numerically through several test problems from the literature. Finally, we end by a conclusion.

Notation. Throughout the paper, $\| \cdot \|$ denotes the Frobenius norm and I_n denotes the identity matrix of order n. The symbol $\langle \cdot, \cdot \rangle$ is the inner product of two matrices defined by $\langle M_1, M_2 \rangle = \text{Tr}\,(M_1^T M_2)$ for $M_1, M_2 \in \mathbb{R}^{n \times n}$. We omit the argument when it is known from the context, e.g., we use J_k to denote $J(L_k, F_k)$.

2. Outline and the QP subproblem

The Lagrangian function associated with the problem (12)–(13) is defined as

$$\ell(L, F, K) = J(L, F) + \langle K, H(L, F) \rangle, \qquad (19)$$

where $K \in \mathbb{R}^{n \times n}$ is the Lagrange multiplier. First and second-order directional derivatives of the Lagrangian function as well as first-order directional derivatives of the equality constraint (13) are obtained from the following lemma.

Lemma 1. *Let $F \in \mathcal{D}_s$ and $L, K \in \mathbb{R}^{n \times n}$ be given. Then, the function J and the constraint function H are twice continuously differentiable on their domains. Furthermore, the first and second order directional derivatives of the Lagrangian function and first-order derivatives of the constraint function are given by*

$$\langle \nabla_L \ell, \Delta L \rangle = \langle A(F)^T K A(F) + Q(F) \\ - K, \Delta L \rangle, \qquad (20)$$

$$\langle \nabla_F \ell, \Delta F \rangle = \langle 2N(F, K)LC^T, \Delta F \rangle, \qquad (21)$$

$$\langle \Delta L, \nabla^2_{LF} \ell \Delta F \rangle = \langle \Delta F, 2N(F, K)\Delta L\, C^T \rangle, \quad (22)$$

$$\langle \Delta L, \nabla^2_{LL} \ell \Delta L \rangle = 0, \qquad (23)$$

$$\langle \Delta F, \nabla^2_{FF} \ell \Delta F \rangle \\ = \langle \Delta F, 2N(F, K)\Delta F CLC^T \rangle, \qquad (24)$$

$$\nabla_L H \Delta L = \Delta L - A(F)\Delta L A(F)^T, \qquad (25)$$

$$\nabla_F H \Delta F = B\Delta F CLA(F)^T \\ + A(F)LC^T \Delta F^T B^T, \qquad (26)$$

where

$$N(F, K) = (B^T K A(F) + RFC).$$

Proof. The Lagrangian function is twice continuously differentiable in its domain. First and second-order directional derivatives of ℓ with respect to L and F yield the above equations (see, e.g, the work of Leibfritz and Mostafa (2003) for a similar result but for the continuous-time case). ∎

By using the result of Lemma 1, the optimality conditions of (12)–(13) imply the following Karush–Kuhn–Tucker necessary optimality conditions:

$$-K + A(F)^T K A(F) + Q(F) = 0, \qquad (27)$$

$$2(B^T K A(F) + RFC)LC^T = 0, \qquad (28)$$

$$-L + A(F)LA(F)^T + V = 0. \qquad (29)$$

The system (27)–(29) is nonlinear. Equations (27) and (29) can be regarded as discrete Lyapunov equations. However, it is not practical to consider both of them as Lyapunov equations in a numerical algorithm. This is due to the fact that one loses the nonlinearity of these equations. However, Levine–Athans based methods (see Mäkilä and Toivonen, 1987), consider one of the two equations as a Lyapunov equation and use (28) to eliminate F from the second Lyapunov equation. This yields a nonlinear matrix equation that must be solved at each iteration during the solution of the system (27)–(29). Such an approach lacks a fast local rate of convergence and it does not take into consideration the stability constraint.

SQP methods are based on successfully minimizing a quadratic programming problem of the form

$$\min_{(\Delta L, \Delta F)} q(\Delta L, \Delta F)$$

subject to

$$h(\Delta L, \Delta F) = 0, \qquad (30)$$

where

$$\begin{aligned} q(\Delta L, \Delta F) &= \langle \nabla_L \ell, \Delta L \rangle + \langle \nabla_F \ell, \Delta F \rangle \\ &\quad + \frac{1}{2}(\langle \Delta L, \nabla^2_{LL}\ell \Delta L \rangle + 2\langle \Delta L, \nabla^2_{LF}\ell \Delta F \rangle \\ &\quad + \langle \Delta F, \nabla^2_{FF}\ell \Delta F \rangle), \end{aligned} \qquad (31)$$

$$\begin{aligned} h(\Delta L, \Delta F) &= \nabla_L H(\cdot)\Delta L + \nabla_F H(\cdot)\Delta F + H(\cdot). \end{aligned} \qquad (32)$$

Note that all derivatives in this QP are given in Lemma 1.

A typical SQP method generates a sequence of iterates of the form

$$\begin{aligned} (L_{k+1}, &F_{k+1}, K_{k+1}) \\ &= (L_k + \Delta L, F_k + \Delta F, K_k + \Delta K), \quad (33) \end{aligned}$$

where $(\Delta L, \Delta F, \Delta K)$ is obtained as a solution to the augmented linear system:

$$\begin{bmatrix} \nabla^2_{LL}\ell_k & \nabla^2_{LF}\ell_k & \nabla_L H_k^* \\ \nabla^2_{FL}\ell_k & \nabla^2_{FF}\ell_k & \nabla_F H_k^* \\ \nabla_L H_k & \nabla_F H_k & 0 \end{bmatrix} \begin{bmatrix} \Delta L \\ \Delta F \\ \Delta K \end{bmatrix} = - \begin{bmatrix} \nabla_L \ell_k \\ \nabla_F \ell_k \\ H_k \end{bmatrix}. \qquad (34)$$

The SQP method combined with the trust region constraint for solving (12)–(13) is the topic of the next section. Moreover, the following three assumptions are made.

Assumption 1. The set $\mathcal{D}_s$ is assumed to be non-empty.

Assumption 2. The Jacobian ∇H^T is surjective.

Assumption 3. The Hessian matrix of the Lagrangian function ℓ is positive definite on the null space of ∇H^T, i.e., there exists a constant $\kappa > 0$ such that

$$\begin{aligned} \langle \Delta L, \nabla^2_{LL}\ell \Delta L \rangle &+ 2\langle \Delta L, \nabla^2_{LF}\ell \Delta F \rangle \\ &+ \langle \Delta F, \nabla^2_{FF}\ell \Delta F \rangle \geq \kappa \left[\|\Delta L\|^2 + \|\Delta F\|^2 \right] \quad (35) \end{aligned}$$

for all $(\Delta L, \Delta F)$ in the null space of ∇H^T.

3. Trust-region method

The trust-region problem combines the QP subproblem (30) with the trust-region constraint

$$\|(\Delta L, \Delta F)\| \leq \delta_k, \qquad (36)$$

where $\delta_k > 0$ is the trust-region radius. The trust-region radius is updated depending on how the predicted decrease in the merit function compares to the actual decrease. If there is compatibility between their values, then the trust-region radius is unaltered or increased, otherwise the radius is decreased. The inclusion of the trust-region constraint makes the trust region problem ((30) and (36)) considerably more difficult to solve. Moreover, the new model may not always have a solution because the linearized equality constraint and the trust-region constraint can be inconsistent, (see, e.g., Conn *et al.*, 2000).

3.1. Reduced Hessian with step decomposition.
The difficulty of the inconsistent constraint has been overcome in the literature via different approaches, among them the reduced Hessian with the step decomposition approach. In such an approach the step is calculated in two levels. In the first level, one tries to satisfy the linearized constraint while staying well inside the trust region. More precisely, one solves the following *vertical subproblem*:

$$\text{(TR1)} \quad \min_{\|\Delta L^v\| \leq \tau \delta_k} \|H_L(F, L)\Delta L^v + H(F, L)\|, \quad (37)$$

where $\tau \in (0, 1)$ is a parameter. The solution ΔL^v of this subproblem is called the *vertical step*.

The second step component is called the *tangential step*. This step component lies in the null space of the Jacobian of the equality constraint and obviously requires evaluating that operator (see, e.g., Conn *et al.*, 2000). Because of using directional derivatives, the Jacobian matrix is not given explicitly. However, the following lemma provides this null space operator implicitly.

Lemma 2. *Let* $(L, F, K) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s \times \mathbb{R}^{n \times n}$ *and* $(\Delta L, \Delta F) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{p \times r}$ *be given. The range space of the operator* $\mathcal{Z}(L, F)$ *defined by*

$$\mathcal{Z}(L, F) = (-\nabla_L H^{-1}(L, F)\nabla_F H(L, F), \mathcal{I}) \quad (38)$$

coincides with the null space of the Jacobian of the equality constraint, where $\mathcal{I}$ *is the identity operator.*

Proof. The operator $\nabla_L H(L, F)$ is linear as well as bijective. Then it is invertible and consequently the linearized equality constraint in (30) implies

$$\begin{aligned} \Delta L = &-\nabla_L H^{-1}(L, F)\nabla_F H(L, F)\Delta F \\ &- \nabla_L H^{-1}(L, F)H(L, F). \end{aligned} \quad (39)$$

This leads to the following decomposition of the step:

$$(\Delta L, \Delta F) = \mathcal{Z}(L, F)\Delta F + (-\nabla_L H^{-1}(L, F)H, \mathcal{O}), \quad (40)$$

where $\mathcal{O}$ is the zero operator. The null space of the Jacobian ∇H^T is given by

$$\begin{aligned} &\left\{ (\Delta L, \Delta F) \in \mathbb{R}^{n \times n + p \times r} : \nabla_L H \Delta L + \nabla_F H \Delta F = 0 \right\} \\ &+ \left\{ \mathcal{Z}(L, F)\Delta F, \ \Delta F \in \mathbb{R}^{p \times r} \right\} \\ &= \mathcal{R}(\mathcal{Z}(L, F)), \end{aligned} \quad (41)$$

where $\mathcal{R}(\mathcal{Z})$ is the range space of $\mathcal{Z}$. ∎

The decomposition of the step using the null space $\mathcal{Z}$ can be expressed as

$$\begin{aligned} (\Delta L, \Delta F) &= \mathcal{Z}(L, F)\Delta F + (-\nabla_L H^{-1}(L, F)H, 0) \\ &= (\Delta L^t(\Delta F), \Delta F) + (\Delta L^v, 0). \end{aligned} \quad (42)$$

According to this step decomposition the linearized equality constraint in Problem (30) is decomposed into two discrete Lyapunov equations.

Lemma 3. *Let* $(L, F, K) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s \times \mathbb{R}^{n \times n}$ *be given and let* $(\Delta L, \Delta F) \in \mathbb{R}^{n \times n + p \times r}$ *be the solution of* (30). *The linearized equality constraint of the problem* (30) *is decomposed into the following discrete Lyapunov equations:*

$$\Delta L^v = A(F)\Delta L^v (AF)^T + H(L, F), \quad (43)$$

$$\begin{aligned} \Delta L^t = &A(F)\Delta L^t A(F)^T + B\Delta FCLA(F)^T \\ &+ A(F)LC^T\Delta F^T B^T. \end{aligned} \quad (44)$$

Proof. (See also the work of Leibfritz and Mostafa (2002, Lemma 2.2) for a similar result but for the continuous-time case.) From the step decomposition the linearized equality constraint of (30) can be rewritten as

$$\nabla_L H \Delta L^t + \nabla_F H \Delta F + \nabla_L H \Delta L^v + H = 0. \quad (45)$$

Since $(\Delta L^t(\Delta F), \Delta F) = \mathcal{Z}(L, F)\Delta F$ lies in the null space of the Jacobian $\nabla H^T(L, F)$, we get

$$\nabla_L H(L, F)\Delta L^t + \nabla_F H(L, F)\Delta F = 0, \quad (46)$$

which implies (44). Hence, the linearized equality constraint reduces to

$$\nabla_L H(L, F)\Delta L^v + H(L, F) = 0, \quad (47)$$

which gives (43). ∎

Now, regarding the solution of the problem TR1, a straightforward approach is to calculate ΔL^v by solving the discrete Lyapunov equation (43) and then scale that solution by τ so that it satisfies the trust region constraint $\|\Delta L^v\| \leq \tau\delta$.

Having evaluated ΔL^v, one solves the following trust region subproblem over a smaller trust-region radius to yield the tangential step $\mathcal{Z}\Delta F = (\Delta L^t(\Delta F), \Delta F)$:

$$(\text{TR2}) : \begin{cases} \min_{\Delta F} & q(\Delta F) = \langle \Delta L^t(\Delta F), \nabla_L \ell \rangle \\ & + \langle \Delta F, \nabla_F \ell \rangle + \langle \Delta L^v, \nabla_{LF}^2 \ell \Delta F \rangle \\ & + \langle \Delta L^t(\Delta F), \nabla_{LL}^2 \ell \Delta L^v \rangle \\ & + \frac{1}{2}\langle \Delta F, \mathcal{Z}^T \nabla^2 \ell \mathcal{Z} \Delta F \rangle \\ \text{subject to} & \\ & \|\mathcal{Z}\Delta F\| \leq \sqrt{\delta^2 - \|\Delta L^v\|^2}, \\ & F + \Delta F \in \mathcal{D}_s, \end{cases}$$

where

$$\begin{aligned} &\langle \Delta F, \mathcal{Z}^T \nabla^2 \ell \mathcal{Z} \Delta F \rangle \\ &= \langle \Delta L^t(\Delta F), \nabla_{LL}^2 \ell \, \Delta L^t(\Delta F) \rangle \\ &\quad + 2\nabla_{LF}^2 \ell \mathcal{Z}\Delta F + \langle \Delta F, \nabla_{FF}^2 \ell \, \Delta F \rangle, \quad (48) \end{aligned}$$

and $\Delta L^t(\Delta F)$ solves the discrete Lyapunov equation (44).

The optimality conditions applied to Problem TR2 yield a linear matrix equation in the unknown ΔF combined with three discrete Lyapunov equations. This can be seen in the following lemma.

Lemma 4. *Let* $(L, F, K) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s \times \mathbb{R}^{n \times n}$ *be given. The step* $\Delta F \in \mathbb{R}^{p \times r}$ *in the solution of Problem TR2 satisfies the following linear matrix equation:*

$$\begin{aligned} &(B^T K B + R)\Delta FCLC^T + N(K, F)\Delta L^t(\Delta F)C^T \\ &\quad + B^T \Delta K(\Delta F)A(F)LC^T \\ &= -B^T SA(F)LC^T - N(K, F)(L + \Delta L^v)C^T, \end{aligned} \quad (49)$$

where $\Delta L^t(\Delta F)$, $\Delta K(\Delta F)$ *and S solve the following discrete Lyapunov equations:*

$$
\begin{aligned}
&\Delta L^t(\Delta F) \\
&= A(F)\Delta L^t(\Delta F)A(F)^T + \big(B\Delta FCLA(F)^T \\
&\quad + A(F)LC^T\Delta F^T B^T\big), \quad\quad (50)
\end{aligned}
$$

$$
\begin{aligned}
&\Delta K(\Delta F) \\
&= A(F)^T\Delta K(\Delta F)A(F) + N(K,F)^T\Delta FC \\
&\quad + C^T\Delta F^T N(K,F), \quad\quad (51)
\end{aligned}
$$

$$
S = A(F)^T SA(F) + (M + M^T), \quad\quad (52)
$$

where $M = A(F)^T KA(F) + Q(F) - K$ and $N(K,F)$ is as defined above.

Proof. By forming the Lagrangian function associated with Problem TR2, differentiating with respect to ΔF and then applying the optimality conditions, we get (49), where $\Delta L^t(\Delta F)$, $\Delta K(\Delta F)$ and S solve the discrete Lyapunov equations (50), (51) and (52), respectively. From Lyapunov theory, the exact solution of the discrete Lyapunov equation (50) can be expressed as the infinite sum:

$$
\begin{aligned}
\Delta L^t(\Delta F) = \sum_{k=0}^{\infty} A(F)^k \Big[B\Delta FCLA(F)^T \\
+ (B\Delta FCLA(F)^T)^T \Big](A(F)^T)^k (53)
\end{aligned}
$$

As an example of evaluating this derivative, the second and third terms on the left-hand side of (49) are obtained as follows:

$$
\begin{aligned}
&\frac{\partial}{\partial\Delta F}\langle \Delta F, N(K,F)\Delta L^t(\Delta F)C^T\rangle \\
&= N(K,F)\Delta L^t(\Delta F)C^T \\
&\quad + B^T\Delta K(\Delta F)A(F)LC^T, \quad\quad (54)
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta K = \sum_k^{\infty} (A(F)^T)^k [(B^T KA(F) + RFC)^T\Delta FC \\
+ C^T\Delta F^T(B^T KA(F) + RFC)]A(F)^k \quad (55)
\end{aligned}
$$

solves the discrete Lyapunov equation (51).

Note that, when evaluating the derivative of the term $\Delta L^t(\Delta F)$, we use the exact form (53) and the following properties of the inner product operator:

$$
\frac{\partial}{\partial X}\langle M_1^T, XM_2\rangle = M_1^T M_2^T, \quad\quad (56)
$$

$$
\frac{\partial}{\partial X}\langle M_1^T, X^T M_2\rangle = M_2 M_1, \quad\quad (57)
$$

where M_1 and M_2 are given constant matrices of appropriate dimensions. ■

It is important to emphasize that the solution of the subproblem TR2 cannot be calculated by the dogleg method unless one assures that the reduced Hessian $\mathcal{Z}^T\nabla^2\ell\mathcal{Z}$ is positive definite (see Conn *et al.*, 2000). However, Leibfritz and Mostafa (2003) apply a conjugate gradient trust-region method to solve a similar subproblem in the continuous-time case, where the idea of negative curvature is utilized. The same approach is considered in the DISTR method to calculate the tangential step $Z\Delta F$ such that $F + \Delta F \in \mathcal{D}_s$.

Having calculated ΔL^v and $Z\Delta F$, the obtained trail step is $(\Delta L, \Delta F) = (\Delta L^v, 0) + Z\Delta F$ and the new iterate is

$$
(L_{k+1}, F_{k+1}) = (L_k, F_k) + (\Delta L, \Delta F), \quad\quad (58)
$$

where k is the iteration counter.

3.2. Merit function/penalty parameter update.

A merit function that measures the progress made by the calculated step towards optimality and feasibility and also fits well with this approach is Fletcher's augmented Lagrangian function:

$$
\psi^\mu(L, F, K) = \ell(L, F, K) + \frac{\mu}{2}\|H(L,F)\|^2. \quad (59)
$$

According to the trust region strategy, the two quantities $ared$ and $pred$ of the *actual* and *predicted* decreases in the merit function are applied to measure such progress at each iteration (see, e.g., Conn *et al.*, 2000):

$$
\begin{aligned}
ared_k = \psi^{\mu_k}(L_k, F_k, K_k) \\
- \psi^{\mu_k}(L_k + \Delta L, F_k + \Delta F, K_{k+1}), \quad (60)
\end{aligned}
$$

and

$$
\begin{aligned}
pred_k = q_k(0) - q_k(\Delta F) \\
+ \mu_k\big[\|H_k\|^2 - \|DH_k\|^2\big], \quad\quad (61)
\end{aligned}
$$

where

$$
\begin{aligned}
DH = \nabla_L H(L,F)\Delta L \\
+ \nabla_F H(L,F)\Delta F + H(L,F) \quad (62)
\end{aligned}
$$

and $q_k(\Delta F)$ is as defined in Problem TR2. The ratio of the actual to the predicted decrease $r_k^d = ared_k/pred_k$ is used to measure the progress towards optimality and feasibility. According to the value of this ratio, the calculated step is accepted or rejected and, consequently, δ_k is increased or decreased.

The update rule of the penalty parameter is as follows (see, e.g., Conn *et al.*, 2000): If

$$
pred_k \geq \frac{\mu_{k-1}}{2}\big[\|H_k\|^2 - \|DH_k\|^2\big], \quad\quad (63)
$$

then set $\mu_k = \mu_{k-1}$. Otherwise, set

$$
\mu_k = \frac{2(q_k(\Delta F_k) - q_k(0)) + \langle \Delta K_k, DH_k\rangle}{\|H_k\|^2 - \|DH_k\|^2} + \bar{\mu}, \quad (64)
$$

where $\bar{\mu} > 0$ is a given constant.

Stopping test. The DISTR method terminates if an iterate (L_k, F_k, K_k) satisfies the condition

$$\|\nabla_F \ell(L_k, F_k, K_k)\| + \|H_k\| < \epsilon, \qquad (65)$$

where $\epsilon \in (0, 1)$ is a tolerance. Observe here that $\nabla_L \ell(L, F, K) = 0$ yields the discrete Lyapunov equation (27), which is solved every iteration in the algorithm to update the multiplier K_k. Therefore, it is unnecessary to include the term $\|\nabla_L \ell(L_k, F_k, K_k)\|$ in the stopping criterion (65). Hence, the method DISTR terminates after achieving a stationary point that satisfies the first-order optimality conditions.

Algorithm 1 DISTR (trust region method)

Initialization: Let $\epsilon \in (0, 1)$ be the tolerance, $0 < c_1 < c_2 < 1$ and $\bar{\mu} > 0, \mu_0 \geq 1$ be given. Choose $F_0 \in \mathcal{D}_s$ and let L_0 and K_0 be solutions of (29) and (27), respectively. Set $k := 0$.

While the stopping test (65) is not satisfied, do

1. *Compute ΔL_k^v as solution to TR1.*

2. *Compute ΔF as a solution to TR2 such that $F_k + \Delta F \in \mathcal{D}_s$. Given ΔF, solve (44) for $\Delta L^t(\Delta F)$ and then set $\Delta L_k = \Delta L_k^v + \Delta L^t(\Delta F)$.*

3. *Compute a new multiplier $K_{k+1} := K(F_k + \Delta F)$ according to (27).*

4. *Update the penalty parameter according to (63) and (64).*

5. *Trust region radius and step update: Compute the ratio r_k^d.*

    If $r_k^d < c_1$,

    > *set $\delta_{k+1} := \delta_k/2$ and reject the step;*

    Otherwise, if $c_1 \leq r_k^d < c_2$,

    > *set $\delta_{k+1} := \delta_k$, $F_{k+1} := F_k + \Delta F$ and $L_{k+1} = L(F_{k+1})$ solution of (29);*

    Otherwise, if $r_k^d \geq c_2$,

    > *set $\delta_{k+1} := 2\,\delta_k$, $F_{k+1} := F_k + \Delta F$ and $L_{k+1} = L(F_{k+1})$ as a solution to (29).*

    End (If)

6. *Set $k \leftarrow k + 1$ and go to 1.*

End(do)

---

4. Initial feasible solution

An important issue when solving any of the above-mentioned minimization problems is the requirement of an initial feasible point $F_0 \in \mathcal{D}_s$ to start the iteration sequence. A parametrization approach that allows starting the trust region method with $F_0 = 0$ is as follows. One replaces the system matrix A by $A_t = (1 - t)A$, where $0 \leq t < 1$ is a parameter. Then, instead of solving (12)–(13), we solve the following parameterized problem:

$$\min_{(L,F) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s^t} J(L, F) = \langle L, Q(F) \rangle, \qquad (66)$$

subject to

$$H_t(L, F) = L - A_t(F)LA_t(F)^T - V = 0, \quad (67)$$

where $A_t(F) = A_t + BFC$, and $Q(F)$ and V are as defined above. The matrix variable F must lie in the parameterized set of stabilizing output feedback gains:

$$\mathcal{D}_s^t = \{F \in \mathbb{R}^{p \times r} : \rho(A_t(F)) < 1\}. \qquad (68)$$

Obviously, $\mathcal{D}_s^t$ approaches $\mathcal{D}_s$ as $t \to 0$.

The parameter t is updated in the trust region main loop in a decreasing sequence $\{t_k\}_{k \geq 0}$ of the form

$$t_{k+1} = c^k t_k, \qquad (69)$$

where $0 < c < 1$ is a constant. This sequence or other sequences for updating t_k must be decreased at least at a super-linear rate so that the fast local rate of the trust region method is not disrupted. If A is Schur stable, i.e., $\rho(A) < 1$, then we take $F = 0$ and skip parameterizing the system matrix A.

The initialization of the parameterized trust region method is as follows. We choose $t_0 \in [0, 1)$ such that $\rho(A_{t_0}) < 1$ and set $F_0 = 0$. Initial L_0 and K_0 can be obtained by solving the discrete Lyapunov equations:

$$L = A_{t_0}(F_0)LA_{t_0}(F_0)^T + V, \qquad (70)$$

$$K = A_{t_0}(F_0)^T K A_{t_0}(F_0) + Q(F_0). \qquad (71)$$

Then Algorithm 1 can be applied directly to solve the parameterized problem (66)–(67) but with $A(F)$ replaced by $A_t(F)$.

Note that one can look at the trust region method with the parametrization approach as a two-phase method. In the first phase, where t is not too small, the goal is to have a stabilizing $F \in \mathcal{D}_s^t$, while in the second phase the goal is to achieve a stationary point $(F, L, K) \in \mathcal{D}_s^t \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ with t very small (by very small we mean $t < 10^{-16}$). Obviously, as $t \to 0$, both A_t and $\mathcal{D}_s^t$ coincide with A and $\mathcal{D}_s$, respectively.

5. Numerical results

In this section an implementation for the DISTR method is described. This method is compared numerically with

Newton's method combined with the Armijo line search globalization rule applied to the corresponding unconstrained problem (17)–(18) and is denoted by disArmijo (see Mäkilä and Toivonen, 1987).

Whenever the uncontrolled system matrix A is Schur stable, one can start with $F_0 = 0$; otherwise initial $F_0 \in \mathcal{D}_s$ must be given or the parametrization approach of Section 4 is applied. In order to compare both methods under the same circumstances, the parametrization approach is also considered for the unconstrained problem (17)–(18) so that the disArmijo method can start with $F_0 = 0$.

An initial $t_0 \in [0, 1)$ is chosen as the first candidate that makes $\rho(A_{t_0}) < 1$. Choosing other $t \in [t_0, 1)$ to start the sequence (69) influences the performance of both methods with respect to the number of iterations and/or the local rate of convergence. Clearly, t_0 adds the weight $(1 - t_0)$ to the system matrix A_t and, as t_k tends to 0, we have that A_{t_k} approaches A.

Several discrete Lyapunov equations have to be solved every iteration. The MATLAB dlyap$(\cdot, \cdot)$ function is used for computing these solutions.

For test purposes, several discrete-time test problems are generated from the COMPlib (2004) benchmark collection of continuous-time models. For all test problems the sample time for converting the control system is 0.1 s. In each example we are given the constant data matrices A, B and C of the control system (1). For all test problems the constant weight matrices Q and R as well as the covariance matrix V are equal to the identity, i.e., $Q = V = I_n$ and $R = I_p$. These six given constant matrices arise in the objective function and the equality constraint of the corresponding optimization problem (12)–(13) and have to be given in advance to execute the DISTR method.

The following values have been assigned to the parameters of Algorithm 1 during the implementation: $c_1 = 0.1$, $c_2 = 0.7$, $\bar{\mu} = 1.0$, $\delta_0 = \|\nabla_F \ell(L_0, F_0, K_0)\|$ and $\mu_0 = 1.0$. The tolerance is chosen as $\epsilon = 1 \times 10^{-7}$.

Example 1. This test problem represents a *transport aircraft* model (Leibfritz, 2004, AC15). The constant data matrices of the discrete-time model are the following:

$$A = \begin{bmatrix} 0.9901 & 0.0012 & 0.0000 & -0.0994 \\ -0.0315 & 1.0000 & 0.0989 & 0.0014 \\ -0.6269 & -0.0004 & 0.9773 & 0.0376 \\ 0.1237 & 0.0001 & 0.0016 & 0.9892 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.0005 & 0.0004 \\ 0.0004 & 0.0040 \\ 0.0078 & 0.0794 \\ -0.0086 & -0.0066 \end{bmatrix},$$

$$C^T = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The uncontrolled system is discrete-time Schur stable, where $\rho(A) = 0.999 < 1$. By the uncontrolled system we mean the plant without feedback. Hence, it is possible to start the DISTR method with $F_0 = 0$, and L_0 is the corresponding solution of the discrete Lyapunov equation (29). The goal is to calculate a stationary point $(L, F) \in \mathbb{R}^{n \times n} \times \mathcal{D}_s$ of the optimization problem (12)–(13), which equivalently solves the discrete-time LQR problem considered.

Table 1. Performance of the DISTR method in Example 1.

k	J_k	$\|\nabla_F \ell_k\| + \|H_k\|$	δ_k	$\rho(A(F_k))$
0	2.101e+5	1.889e+07	8.49e+07	9.98e-1
1	1.426e+5	8.393e+06	1.69e+08	9.98e-1
2	9.757e+4	3.730e+06	3.40e+08	9.96e-1
3	6.759e+4	1.657e+06	6.80e+08	9.95e-1
4	4.763e+4	7.358e+05	1.36e+09	9.95e-1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
19	1.612e+3	1.258e−02	1.00e+10	9.65e-1
20	1.612e+3	8.858e−06	1.00e+10	9.65e-1
21	1.612e+3	2.399e−12	1.00e+10	9.65e-1

The DISTR method, as we have seen, is iterative. Table 1 shows the convergence behavior of this method until a stationary point is reached. The first to the fifth columns in this table are the iteration counter k, the objective function J_k, the convergence criterion $\|\nabla_F \ell_k\| + \|H_k\|$, the trust region radius δ_k, and the spectral radius of the closed-loop system matrix $A(F_k)$, respectively. From the last column one sees that the method is feasible with respect to the stability constraint throughout iterations. Moreover, starting from a remote point, the method converges to a stationary point in 21 iterations with a fast local rate of convergence. The achieved discrete-time feedback gain is

$$F_* = \begin{bmatrix} 0.3489 & 1.3969 & 7.4745 \\ -1.1617 & -3.3304 & -4.6462 \end{bmatrix}. \quad (72)$$

On the other hand, the disArmijo method could not reach the prescribed accuracy and the convergence criterion stacks at 1.7717×10^{-6}, where Newton's direction is becoming non-descent. ♦

Example 2. The second test problem describes a *decentralized interconnected system*, cf. Leibfritz (2004, DIS1). By applying the MATLAB, function c2d the data matrices A, B and C of the discrete-time linear control system are obtained, which are not listed because of their high dimensions.

The uncontrolled system is discrete-time Schur stable, where $\rho(A) = 0.9912 < 1$. Hence, it is possible to start with $F_0 = 0$. The two methods successfully converge to the same solution. The DISTR and disArmijo methods achieve the prescribed accuracy in 10 and 12 iterations,

Table 2. Performance of the DISTR method in Example 2.

k	J_k	$\|\nabla_F J_k\| + \|H_k\|$	δ_k	$\rho(A(F_k))$
0	3.831e+2	5.137e+2	2.07e+3	9.87e-1
1	2.775e+2	2.388e+2	4.15e+3	9.84e-1
2	2.111e+2	9.972e+1	8.29e+3	9.78e-1
3	1.785e+2	3.909e+1	1.66e+4	9.69e-1
⋮	⋮	⋮	⋮	⋮
8	1.589e+2	1.190e−2	5.31e+5	9.51e-1
9	1.589e+2	6.318e−5	1.06e+6	9.51e-1
10	1.589e+2	1.528e−9	2.12e+6	9.51e-1

respectively, and the corresponding feedback gain is

$$F_* = \begin{bmatrix} 0.5496 & -0.7421 & 0.0752 & 0.0921 \\ -0.2487 & -0.0111 & -0.4751 & -0.8625 \\ -0.7048 & -0.2793 & -0.1213 & -0.1438 \\ -0.0475 & 0.0034 & -0.0946 & -0.0334 \end{bmatrix}. \tag{73}$$

Table 2 shows the convergence behavior of the DISTR method in this example. ♦

Example 3. This test problem represents a *wind energy conversion system* (cf. Leibfritz, 2004, WEC2). The dynamical system has the dimensions $n = 10, p = 3$, and $r = 4$. The data matrices A, B and C are not listed because of their dimensions.

Table 3. Performance of the DISTR method in Example 3.

k	J_k	$\|\nabla_F \ell_k\| + \|H_k\|$	δ_k	$\rho(A(F_k))$
0	2.406e+4	2.308e+5	1.03e+06	9.91e-1
1	1.364e+4	1.011e+5	2.06e+06	9.87e-1
2	7.111e+3	3.811e+4	4.12e+06	9.82e-1
3	4.932e+3	8.645e+3	8.24e+06	9.77e-1
⋮	⋮	⋮	⋮	⋮
14	3.669e+3	3.925e−3	1.00e+10	8.24e-1
15	3.669e+3	2.370e−5	1.00e+10	8.24e-1
16	3.669e+3	8.805e−7	1.00e+10	8.24e-1
17	3.669e+3	5.327e−7	1.00e+10	8.24e-1
18	3.669e+3	7.106e−8	1.00e+10	8.24e-1

The uncontrolled system is discrete-time Schur stable, because $\rho(A) = 0.9943 < 1$. The DISTR method converges to a stationary point in 18 iterations, while the disArmijo method could not achieve the same point. The convergence criterion $\|\nabla J(F_k)\|$ stacks at the value of 2.108×10^{-2}. Table 3 shows the convergence behavior of the method DISTR in this test problem. Although the DISTR method achieved the prescribed accuracy, the fast local rate is lost.

The achieved stationary point is (L_*, F_*), where

$$F_* = \begin{bmatrix} -0.0013 & -0.0048 & 1.9943 & -0.0935 \\ 0.0691 & 0.1490 & 1.1688 & -0.0953 \\ -0.0520 & 0.0449 & -0.6569 & -0.0505 \end{bmatrix}, \tag{74}$$

and L_* is the corresponding solution of the discrete Lyapunov equation (13). ♦

Example 4. This example is a discrete version of a *coupled spring experiment* (cf. Leibfritz, 2004, CSE1). Using the function c2d, we have obtained the discrete-time model, which has dimensions $n = 20, p = 2, r = 10$. Because of the problem dimensionality, we skip listing the data matrices A, B and C.

Table 4. Convergence behavior of the disArmijo and DISTR methods in Example 4.

k	J_k	$\|\nabla J(F_k)\|$	t_k
0	1.0755e+004	5.5801e+005	1.00e-004
1	1.7742e+004	1.6063e+006	9.00e-006
2	1.3338e+004	8.7950e+005	2.43e-007
3	9.1749e+003	3.9244e+005	1.97e-009
⋮	⋮	⋮	⋮
17	7.7718e+002	3.5526e-002	1.29e-093
18	7.7718e+002	2.9302e-004	1.50e-103
19	7.7718e+002	6.0517e-008	5.23e-114

k	J_k	$\|\nabla_F \ell_k\| + \|H_k\|$	t_k
0	7.4289e+003	2.480e+005	1.0e-004
1	6.4958e+003	1.824e+005	4.6e-005
2	5.2405e+003	1.104e+005	9.7e-006
3	3.8639e+003	5.241e+004	9.5e-007
⋮	⋮	⋮	⋮
15	7.7718e+002	1.861e-002	3.4e-045
16	7.7718e+002	2.963e-004	1.4e-050
17	7.7718e+002	1.403e-008	2.5e-056

The uncontrolled system is not discrete-time Schur stable, where $\rho(A) = 1$. In order to start with $F_0 = 0$, we use the parametrization approach and choose $c = 0.3$ of the decreasing rate in (69) and $t_0 = 1 \times 10^{-4}$. Both methods converge to the same local minimizer, where the method DISTR achieves the prescribed accuracy in 17 iterations and the method disArmijo requires 19 iterations. The final output feedback gain is as follows (entries are rounded to two digits after the decimal point):

$$F_* = \begin{bmatrix} -0.72 & -0.44 & -0.32 & -0.23 & -0.17 \\ 0.02 & 0.04 & 0.06 & 0.09 & 0.13 \\ -0.13 & -0.09 & -0.06 & -0.04 & -0.02 \\ 0.17 & 0.23 & 0.32 & 0.44 & 0.72 \end{bmatrix}. \tag{75}$$

Table 4 shows the performance of the two methods in this example. The objective function, the convergence criterion and the stabilizing parameter t_k are

listed column-wise. The initial value of the stabilizing parameter is $t_0 = 1.0 \times 10^{-4}$, which puts the weight $(1 - t_0) = 0.9999$ on the system matrix A_t. This weight approaches 1 as t_k tends to zero. Observe also that other values of $t_0 \in [1.0 \times 10^{-4}, 1)$ are allowed. ♦

Example 5. This example is a discrete version of a *binary distillation tower with pressure variation* (cf. Leibfritz, 2004, BDT1). Using the function c2ds we have obtained the discrete-time model, which has dimensions $n = 11$, $p = 3$, $r = 3$. Because of the system dimensionality, we skip listing the data matrices A, B, and C.

Table 5. Convergence behavior of the disArmijo and DISTR methods in Example 5.

k	J_k	$\|\nabla J(F_k)\|$	t_k
0	9.2009e+003	7.9073e+003	1.00e-004
1	1.4151e+004	1.9047e+004	9.00e-006
2	1.2631e+004	9.7797e+003	2.43e-007
3	1.1235e+004	3.9708e+003	1.97e-009
⋮	⋮	⋮	⋮
10	8.8409e+003	2.8054e-003	1.03e-038
11	8.8409e+003	3.4165e-007	5.47e-045
12	8.8409e+003	4.4884e-009	8.73e-052

k	J_k	$\|\nabla_F \ell_k\| + \|H_k\|$	t_k
0	7.6093e+003	2.789e+003	1.0e-004
1	1.1005e+004	5.043e+003	3.0e-005
2	1.0065e+004	2.647e+003	2.7e-006
3	9.3600e+003	1.003e+003	7.3e-008
⋮	⋮	⋮	⋮
7	8.8409e+003	2.854e-002	2.3e-019
8	8.8409e+003	3.792e-004	1.5e-023
9	8.8409e+003	1.731e-008	3.0e-028

The uncontrolled system is not discrete-time Schur stable. Therefore, we have to apply the parametrization approach to start the methods, where $F_0 = 0$ and $t_0 = 1.0 \times 10^{-4}$ imply that $\rho(A_{t_0}) = 0.9999 < 1$. The decreasing rate c of (69) is chosen as $c = 0.3$. The methods DISTR and disArmijo converge to the same stationary point in 9 and 12 iterations, respectively. The final output feedback gain is

$$F_* = \begin{bmatrix} 0.0005 & -0.0135 & -0.0129 \\ -0.0155 & -0.0034 & -0.0123 \\ -1.5675 & -0.3966 & -0.0822 \end{bmatrix}. \quad (76)$$

Table 5 shows the performance of the two methods in this example.

Tables 6 and 7 show the effect of the initial t_0 and the constant c on the two methods with respect to the number of iterations. The results show that the method DISTR is more flexible to the perturbation occurring in the given system matrix A_t. In all cases the method DISTR converges to the same local minimizer, while disArmijo either fails to reach the prescribed accuracy or requires more

iterations to converge to the same local minimizer. Such bad behavior of disArmijo occurred quite often in various unstable systems. The asterisks in Table 7 mean that the method fails to achieve the prescribed accuracy and stacks at the shown attached value.

Table 6. Effect of the changes in c and t_0 on the number of iterations of disArmijo and DISTR for the CSE1 test problem from the COMPlib benchmark.

	$t_0 = 1.0 \times 10^{-4}$				
c	0.15	0.30	0.45	0.75	0.95
disArmijo	21	20	20	22	27
DISTR	17	17	17	17	27
	$t_0 = 1.0 \times 10^{-2}$				
c	0.15	0.30	0.45	0.75	0.95
disArmijo	12	11	13	13	30
DISTR	10	9	10	13	13

Table 7. Effect of the changes in c and t_0 on the number of iterations of disArmijo and DISTR for the BDT1 test problem from the COMPlib benchmark.

	$t_0 = 1.0 \times 10^{-4}$				
c	0.15	0.30	0.45	0.75	0.95
disArmijo	$* 10^{-5}$	13	$* 10^{-4}$	$* 10^{-6}$	33
DISTR	9	9	10	14	31
	$t_0 = 1.0 \times 10^{-2}$				
c	0.15	0.30	0.45	0.75	0.95
disArmijo	12	$* 10^{-5}$	12	17	34
DISTR	12	12	12	15	34

The main conclusion that one can draw from the above results is that the DISTR method outperforms the disArmijo method on the analyzed set of test problems with respect to the number of iterations. There are several test problems as well, where both methods exhibit similar convergence behavior, which are not included.

The optimal choice for the starting $t_0 \in [0, 1)$ of the stabilizing parameter as well as the decreasing rate of the sequence $\{t_k\}$ are not yet determined and need further investigation. In particular, the choice of t_0 affects the convergence behavior of the method and its robustness.

6. Conclusion

In this paper, an SQP trust region method is analyzed and studied for solving the discrete-time linear quadratic control problem when a complete set of state variables is not available. The trust region method uses the reduced Hessian approach with step decomposition to handle the well-known difficulty of possible inconsistent constraints.

The trust region method requires an initial feasible point that satisfies an eigenvalue constraint. The method is parameterized so that a feasible starting point is not re-

quired. The proposed method performed quite well numerically on different test problems of engineering applications.

References

Conn, A.R., Gould, N.I.M. and Toint, Ph.L. (2000). *Trust-Region Methods*, SIAM, Philadelphia, PA.

Garcia, G., Pradin, B. and Zeng, F. (2001). Stabilization of discrete-time linear systems by static output feedback, *IEEE Transactions on Automatic Control* **46**(12): 1954–1958.

Kočvara M., Leibfritz, F., Stingl, M. and Henrion, D. (2005). A nonlinear SDP algorithm for static output feedback problems in COMPlib, *Proceedings of the 16th IFAC World Congress on Automatic Control, Prague, Czech Republic*, (on CD-ROM).

Lee, J.-W. and Khargonekar, P.P. (2007). Constrained infinite-horizon linear quadratic regulation of discrete-time systems, *IEEE Transactions on Automatic Control* **52**(10): 1951–1958.

Leibfritz, F. (2004). COMPlib: COnstraint Matrix-optimization Problem *lib*rary—A collection of test examples for nonlinear semi-definite programs, control system design and related problems, *Technical report*, http://www.complib.de/.

Leibfritz, F. and Mostafa, E.M.E. (2002). An interior point constrained trust region method for a special class of nonlinear semidefinite programming problems, *SIAM Journal on Optimization* **12**(4): 1048–1074.

Leibfritz, F. and Mostafa, E.M.E. (2003). Trust region methods for solving the optimal output feedback design problem, *International Journal of Control* **76**(5): 501–519.

Mäkilä, P.M. and Toivonen, H.T. (1987). Computational methods for parametric LQ problems—A survey, *IEEE Transactions on Automatic Control* **32**(8): 658–671.

Mostafa, E.M.E. (2005a). A trust region method for solving the decentralized static output feedback design problem, *Journal of Applied Mathematics & Computing* **18**(1–2): 1–23.

Mostafa, E.M.E. (2005b). An augmented Lagrangian SQP method for solving some special class of nonlinear semi-definite programming problems, *Computational and Applied Mathematics* **24**(3): 461–486.

Mostafa, E.M.E. (2008). Computational design of optimal discrete-time output feedback controllers, *Journal of the Operations Research Society of Japan* **51**(1): 15–28.

Mostafa, E.M.E. (2012). A conjugate gradient method for discrete-time output feedback control design, *Journal of Computational Mathematics* **30**(3): 279–297.

Nocedal J. and Wright, S.J. (1999). *Numerical Optimization*, Springer, New York, NY.

Peres, P.L.D. and Geromel, J.C. (1993). $\mathcal{H}_2$ control for discrete-time systems optimality and robustness, *Automatica* **29**(1): 225–228.

Sulikowski, B., Gałkowski, K., Rogers, E. and Owens, D.H. (2004). Output feedback control of discrete linear repetitive processes, *Automatica* **40**(12): 2167-2173.

Syrmos, V.L., Abdallah, C.T., Dorato, P. and Grigoriadis, K. (1997). Static output feedback—A survey, *Automatica* **33**(2): 125–137.

Varga, A. and Pieters, S. (1998). Gradient-based approach to solve optimal periodic output feedback control problems, *Automatica* **34**(4): 477–481.

Zhai, G., Matsumoto, Y., Chen, X., Imae, J. and Kobayashi, T. (2005). Hybrid stabilization of discrete-time LTI systems with two quantized signals, *International Journal of Applied Mathematics and Computer Science* **15**(4): 509–516.

El-Sayed M.E. Mostafa received the B.Sc. degree in mathematics from Alexandria University, Egypt, in 1989, the M.Sc. degree in industrial mathematics from the University of Kaisersluatern, Germany, in 1994, and the Ph.D. degree in numerical optimization from Alexandria University in 2000. Since 2008 he has been working as an associate professor at the Department of Mathematics, Faculty of Science, Alexandria University. His research interests include numerical optimization and optimal control with engineering applications.