2026 American Control Conference (ACC)
May 27-29, 2026. New Orleans, Louisiana, USA

Discrete-time linear quadratic stochastic control with equality-constrained inputs: Application to energy demand response

Léo Seugnet and Shuang Gao

***Abstract*— We investigate the discrete-time stochastic linear quadratic control problem for a population of cooperative agents under the hard equality constraint on total control inputs, motivated by demand response in renewable energy systems. We establish the optimal solution that respects hard equality constraints for systems with additive noise in the dynamics. The optimal control law is derived using dynamic programming and Karush-Kuhn-Tucker (KKT) conditions, and the resulting control solution depends on a discrete-time Riccati-like recursive equation. Application examples of coordinating the charging of a network of residential batteries to absorb excess solar power generation are demonstrated, and the proposed control is shown to achieve exact power tracking while considering individual State-of-Charge (SoC) objectives.**

I. Introduction

With the integration of renewable energy sources such as solar and wind power, modern energy grids are facing growing challenges due to the intermittent nature of these sources [1], [2]. The presence of energy storage units if used properly can help reduce uncertainties as well as accommodate temporary loss of generation capacity on the grids [1]. The scheduling of energy usage from the users' side (which is known as demand response) can also help reduce the fluctuations and peak demands (see e.g. [3]). The coordination of a group of cooperative energy users can be cast as a stochastic control problem where each user has a local dynamics representing the energy storage level and all users share a global objective of satisfying the demand response request to balance the load and supply. As an approximate model for demand response problems, we consider a group of users with linear stochastic dynamics and quadratic expected cost in discrete time to establish analytical solutions.

To reduce uncertainties in the mismatch between the supply and actual demand by collaboratively adjusting the schedules of electricity usages of a large group of users, a key aspect of our formulation is to introduce a hard equality constraint; more specifically, the total charging power of the group of collaborative users matches exactly a prescribed demand profile imposed by the virtual operator.

For deterministic linear quadratic (LQ) control problems in discrete time, inequality constraints have been treated in works such as [4]–[7], whereas equality constraints have been addressed in [8]–[11]. Constrained LQ problems can be solved using Quadratic Programming (see e.g. [4], [10]), however the computational complexity of directly applying Quadratic Programming to LQ problems scales cubically with respect to the time horizon (or the length of the trajectories) as discussed in [9], [10]. More efficient alternatives based on dynamic programming have been established to solve the LQ control problems subject to inequality constraints (e.g. [5], [6], [12]) and those subject to equality constraints (see [8]–[10]).

Stochastic control with constraints in continuous time or discrete time has been studied by many researchers (see e.g. [5], [8], [13]–[18]). In these previous works, the constraints are not the hard equality constraints on control inputs. For instance, when equality constraints are considered for discrete-time stochastic systems, they are imposed on the (expected) state [8], [13]–[15] or the terminal state [17]. The works [16], [18] focused on inequality constraints, which impose upper and lower bounds on the control inputs.

Contribution: This paper addresses the hard equality constraints on the sum of control inputs for stochastic linear quadratic control problems in discrete time. It establishes a Riccati-like difference equation for the optimal stochastic control solution satisfying the equality constraint, derived using dynamic programming together with the KKT conditions at each dynamic programming step. Such a solution method is extended to address intermittent equality constraints on the sum of the control inputs, where the hard equality constraint is active only during pre-specified time intervals. Additionally, to ensure smooth control transitions, a "switched" control strategy is established by solving the associated problems with both soft and hard constraints over different time intervals.

Notation: Let $\mathbb{R}$ and $\mathbb{N}$ denote the set of real and that of natural numbers, respectively. Let $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ denote the set of nonzero natural numbers. For any $a, b \in \mathbb{N}$ with $a \le b$, let $[\![a, b]\!] := \{a, a+1, \ldots, b\}$ denote the discrete interval. For a matrix A, $A^\intercal$ denotes its transpose. For a symmetric matrix Q and a vector z, let $|z|_Q^2 := z^\intercal Q z$. For any $n \in \mathbb{N}^*$, let $\mathbb{1}_n \in \mathbb{R}^n$ denote the column vector of all ones and $I_n \in \mathbb{R}^{n\times n}$ the identity matrix. Let $[N] := \{1, 2, \cdots, N\}$. Lastly, $\mathrm{diag}(M_1, \ldots, M_N)$ denotes the matrix with diagonal blocks $M_1, \cdots, M_N$ and zero elsewhere.

II. System model and problem formulation

A. System model

Consider $N \in \mathbb{N}^*$ collaborative agents with discrete time dynamics over a finite horizon $[\![0, T]\!]$. For an agent i, let $x_{i,t} \in \mathbb{R}^{d_x^i}$ (resp. $u_{i,t} \in \mathbb{R}^{d_u^i}$) denote its state (resp. control)

*This work is supported in part by NSERC grant RGPIN-2024-06612 and FRQNT grant 359699 (https://doi.org/10.69777/359699).

Léo Seugnet and Shuang Gao are with the Department of Electrical Engineering, Polytechnique Montréal, and Groupe d'études et de recherche en analyse des décisions (GERAD), Montréal, QC, Canada. Emails: {leo.seugnet,shuang.gao}@polymtl.ca

with d_x^i (resp. d_u^i) as the dimension. At time $t=0$, the system starts from an initial state $x_{i,0}$ and for $t \in [\![0, T-1]\!]$, the state of agent i evolves according to the linear dynamics

$$x_{i,t+1} = A_i x_{i,t} + B_i u_{i,t} + w_{i,t}, \tag{1}$$

where A_i and B_i are matrices of appropriate dimensions and $w_{i,t} \in \mathbb{R}^{d_x^i}$ is the process noise at time t.

Let $x_t = [x_{1,t}, \cdots, x_{N,t}]^\intercal$, $u_t = [u_{1,t}, \cdots, u_{N,t}]^\intercal$ and $w_t = [w_{1,t}, \cdots, w_{N,t}]^\intercal$. The compact representation of the N-agent system is then given by

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{2}$$

where $A = \text{diag}(A_1, \ldots, A_N)$, $B = \text{diag}(B_1, \ldots, B_N)$, and A_i and B_i are system parameters of agent i in (1). We note $n_{\text{tot}} := \sum_{i=1}^N d_x^i$ and $m_{\text{tot}} := \sum_{i=1}^N d_u^i$ the dimensions of the N-agent state x_t and control input u_t, respectively.

Assumption 1: $\{w_t\}_{t\geq 0}$ is an i.i.d. noise sequence with mean zero and finite covariance matrix $W \in \mathbb{R}^{n_{\text{tot}} \times n_{\text{tot}}}$.

B. System performance and control objective

At time $t \in [\![0, T-1]\!]$, agent i incurs an instantaneous cost

$$\ell_i(x_{i,t}, u_{i,t}) = |x_{i,t} - r_{i,t}|^2_{Q_i} + |u_{i,t}|^2_{R_i}, \tag{3}$$

and at the terminal time T, a terminal cost

$$\ell_{i,T}(x_{i,T}, u_{i,T}) = |x_{i,T} - r_{i,T}|^2_{Q_{i,T}}, \tag{4}$$

where Q_i, $Q_{i,T}$, and R_i are matrices of appropriate dimensions, and $r_{i,t}$ are a given deterministic reference trajectory.

The global system (2) incurs an instantaneous cost

$$\ell(x_t, u_t) = \sum_{i=1}^N \ell_i(x_{i,t}, u_{i,t}), \tag{5}$$

and at the terminal time T, a terminal cost

$$\ell_T(x_T) = \sum_{i=1}^N \ell_{i,T}(x_{i,T}). \tag{6}$$

Let $Q = \text{diag}(Q_1, \ldots, Q_N)$, $Q_T = \text{diag}(Q_{1,T}, \ldots, Q_{N,T})$, and $R = \text{diag}(R_1, \ldots, R_N)$.

Assumption 2: The matrices Q_i and $Q_{i,T}$ are symmetric and positive semi-definite and R_i is symmetric and positive definite.

Assumptions 1 and 2 follow standard assumptions in stochastic linear quadratic control problems.

Problem 1: *Choose a control trajectory $u : [\![0, T-1]\!] \to \mathbb{R}^{m_{tot}}$ to minimize*

$$J(u_t) = \mathbb{E} \sum_{t=0}^{T-1} \ell(x_t, u_t) + \mathbb{E}\ell_T(x_T) \tag{7}$$

subject to the dynamics (2) *and the equality constraint*

$$\mathbb{1}_{m_{\text{tot}}}^\intercal u_t = c_t, \qquad \forall t \in [\![0, T-1]\!] \tag{8}$$

where $c_t \in \mathbb{R}$ represents the total consumption requirement at time $t \in [\![0, T-1]\!]$. □

III. Main results on optimal control solutions

The optimal solution to Problem 1 is obtained using dynamic programming together with the KKT conditions at each dynamic programming step, and is given as follows.

Theorem 1: *Let Assumptions 1 and 2 hold. Let $P : [\![0, T]\!] \to \mathbb{R}^{n_{tot} \times n_{tot}}$ and $s : [\![0, T]\!] \to \mathbb{R}^{n_{tot}}$ be the solutions of the following backward recursions*

$$P_t = Q + A^\intercal P_{t+1} A - A^\intercal P_{t+1} B \Gamma_t B^\intercal P_{t+1} A, \tag{9}$$

$$s_t = \left[A^\intercal - A^\intercal P_{t+1} B \Gamma_t B^\intercal\right] s_{t+1} + A^\intercal P_{t+1} B \gamma_t - Q r_t, \tag{10}$$

with the final conditions $P_T = Q_T$ and $s_T = -Q_T r_T$, where

$$\Gamma_t = \Omega_t^{-1} - \frac{\Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}} \mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1}}{\mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}}}, \quad \gamma_t = \frac{\Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}} c_t}{\mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}}},$$
$$\Omega_t = R + B^\intercal P_{t+1} B.$$

Then the optimal control strategy for Problem 1 is given by

$$u_t = -\Gamma_t (B^\intercal P_{t+1} A x_t + B^\intercal s_{t+1}) + \gamma_t \tag{11}$$

for all $t \in [\![0, T-1]\!]$. □

PROOF: The cost of Problem 1 can be written as follows

$$J(u_t) = \mathbb{E} \sum_{t=0}^{T-1} \left[|x_t - r_t|^2_Q + |u_t|^2_R\right] + \mathbb{E}|x_T - r_T|^2_{Q_T}$$

Let $V_t(z)$ be the optimal cost-to-go (or the value function) at t starting from $x_t = z$, defined as

$$V_t(z) = \min_{u_t, \ldots, u_{T-1}} \mathbb{E}\left[\sum_{\tau=t}^{T-1} \ell(x_\tau, u_\tau) + \ell_T(x_T) \middle| x_t = z\right].$$

Based on the terminal cost, the value function at the terminal time T satisfies $V_T(z) = |z - r_T|^2_{Q_T} = z^\intercal Q_T z - 2r_T^\intercal Q_T z + r_T^\intercal Q_T r_T$. We proceed by backward induction, assuming the value function at time t is of form

$$V_t(z) = z^\intercal P_t z + 2 s_t^\intercal z + q_t \tag{12}$$

which holds for $t = T$ with $P_T = Q_T$ and $s_T = -Q_T r_T$. The Bellman recursion is as follows

$$V_t(z) = \min_u \left\{|z - r_t|^2_Q + |u|^2_R + \mathbb{E}\left[V_{t+1}(Az + Bu + w_t)\right]\right\}$$

subject to the constraint $\mathbb{1}_{m_{\text{tot}}}^\intercal u_t = c_t$. Assuming the previous form of the value function (12), the Bellman recursion can be written as

$$V_t(z) = |z - r_t|^2_Q + |z|^2_{A^\intercal P_{t+1} A} + 2 s_{t+1}^\intercal A z + \text{tr}(W P_{t+1}) + q_{t+1} + \min_u \left\{|u|^2_{\Omega_t} + 2u^\intercal f_t(z)\right\}$$

using $\mathbb{E}(w_{t+1}^\intercal P_{t+1} w_{t+1}) = \text{tr}(W P_{t+1})$, where $\Omega_t := (R + B^\intercal P_{t+1} B)$ and $f_t(z) = \left(B^\intercal P_{t+1} A z + B^\intercal s_{t+1}\right)$. To respect the constraint $\mathbb{1}_{m_{\text{tot}}}^\intercal u_t = c_t$, we introduce the following Lagrangian function

$$\mathcal{L}(u, \lambda) = u^\intercal \Omega_t u + 2u^\intercal f_t(z) + \lambda(\mathbb{1}_{m_{\text{tot}}}^\intercal u - c_t).$$

The KKT conditions [19] give $2\Omega_t u + 2 f_t(z) + \lambda \mathbb{1}_{m_{\text{tot}}} = 0$, which implies $u = -\Omega_t^{-1}(f_t(z) + \frac{\lambda}{2}\mathbb{1}_{m_{\text{tot}}})$. Substituting this into the constraint $\mathbb{1}_{m_{\text{tot}}}^{\intercal} u_t = c_t$ yields the Lagrange multiplier term

$$\frac{\lambda}{2} = -\frac{\mathbb{1}_{m_{\text{tot}}}^{\intercal} \Omega_t^{-1} f_t(z) + c_t}{\mathbb{1}_{m_{\text{tot}}}^{\intercal} \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}}}.$$

Plugging this back into the expression for u yields the optimal action $u_t^* = -\Gamma_t(B^\intercal P_{t+1} A x_t + B^\intercal s_{t+1}) + \gamma_t$ where

$$\Gamma_t = \Omega_t^{-1} - \frac{\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}\mathbb{1}_{m_{\text{tot}}}^{\intercal}\Omega_t^{-1}}{\mathbb{1}_{m_{\text{tot}}}^{\intercal}\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}}, \quad \gamma_t = \frac{\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}} c_t}{\mathbb{1}_{m_{\text{tot}}}^{\intercal}\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}}.$$

Since R is assumed to be positive definite and P_t is positive semidefinite (see Lemma 1 in the Appendix VII-B), we obtain $\Omega_t = R + B^\intercal P_{t+1} B > 0$. Hence that optimization problem at each dynamic programming step is strictly convex and the control identified using KKT at each time step is indeed the unique optimal control action.

Finally, by substituting the optimal control u_t^* back into the Bellman equation we get the following equality

$$\begin{aligned} V_t(z) &= z^\intercal \Big[Q + A^\intercal P_{t+1} A\Big] z + 2(s_{t+1}^\intercal A - r_t^\intercal Q) z + \theta_{t+1} \\ &+ \Big[-\Gamma_t f_t(z) + \gamma_t\Big]^\intercal \Omega_t \Big[-\Gamma_t f_t(z) + \gamma_t\Big] \\ &+ 2\Big[-\Gamma_t f_t(z) + \gamma_t\Big]^\intercal f_t(z), \end{aligned}$$

where $\theta_{t+1} := q_{t+1} + r_t^\intercal Q r_t + \mathrm{tr}(W P_{t+1})$. By substituting the expression for $f_t(z)$ and regrouping terms based on their dependency on z, we arrive at the following form

$$\begin{aligned} V_t(z) &= z^\intercal \Big[Q + A^\intercal P_{t+1} A - A^\intercal P_{t+1} B \Gamma_t B^\intercal P_{t+1} A\Big] z \\ &+ 2\Big[\left(A^\intercal - A^\intercal P_{t+1} B \Gamma_t B^\intercal\right) s_{t+1} + A^\intercal P_{t+1} B \gamma_t \\ &- Q r_t\Big]^\intercal z + q_{t+1} + r_t^\intercal Q r_t + \mathrm{Tr}(W P_{t+1}) \\ &+ \gamma^\intercal \Omega_t \gamma_t + 2 s_{t+1}^\intercal B \gamma_t - s_{t+1}^\intercal B \Gamma_t B^\intercal s_{t+1}. \end{aligned}$$

We recognize the form $V_t(z) = z^\intercal P_t z + 2 s_t^\intercal z + q_t$ which confirms our previous assumption on the form of V_t in (12). The identification of the terms P_t and s_t gives us the following recursion equations (9) and (10).

Remark 1: The structure of the optimal control law (11) admits an elegant geometric interpretation. The term $-\Omega_t^{-1}(B^\intercal P_{t+1} A x_t + B^\intercal s_{t+1})$ is the optimal unconstrained control law. The constrained problem requires this control action to be projected onto the affine subspace defined by the hard constraint (8). The operators Γ_t and γ_t perform this orthogonal projection in the Hilbert space of control actions endowed with this Ω_t^{-1}-weighted inner product, ensuring the solution lies on the required affine manifold to meet the equality constraint c_t.

The equation (9) is similar to but different from the generalized discrete-time Riccati equation [20], [21].

Remark 2: Although the problem is formulated with matrices A, B, W, Q, Q_T and R being block-diagonal, all the proof steps and hence the results in Theorem 1 apply to problems with general cases with non-block-diagonal matrices.

Remark 3: It is important to note that the feasibility of the constraint (8) is physically limited by the aggregate capacity of the N participating agents. For example, if agent $i \in [N]$ has minimum and maximum (scalar) control actions denoted by $u_{i,\min}$ and $u_{i,\max}$ respectively, then the constraint c_t must lie within the aggregate operational range of the system, i.e., $\sum_{i=1}^N u_{i,\min} \leq c_t \leq \sum_{i=1}^N u_{i,\max}$. Consequently, a larger number of users N may provide greater flexibility and enable the system to satisfy a wider range of constraints c_t.

Remark 4: Considering the case with $d_x^i = n$ and $d_u^i = m$ for all $i \in [N]$. The computation of the control in Theorem 1 involves a computation complexity $\mathcal{O}\left(TN^3\left(n^3 + m^3 + n^2 m + n m^2\right)\right)$ floating point operations (flops). The complexity scales linearly with the time horizon T. In contrast, applying generic Quadratic Programming to solve constrained LQ control problems incurs a computation complexity $\mathcal{O}(T^3 N^3 (2n+m)^3)$ flops scaling cubically with respect to the time horizon T as discussed in [22, p. 553] and [9], [10]. In demand response problems, such a complexity reduction with respect to the time horizon allows a longer planning horizon.

IV. Generalizations: Intermittent Constraints

The framework developed for hard equality constraints can be adapted to address scenarios with intermittent constraints. In this section we explore two such generalizations. First, we consider the case where the hard equality constraint is only active during specific, predetermined time intervals. Second, we consider the case that includes intermittent soft constraints where deviations from the target are penalized rather than strictly forbidden, along with intermittent hard equality constraints.

A. Intermittent hard constraint

Denote the set of time steps with active constraint by

$$\mathbb{S} := \{t \in [\![0, T-1]\!] \mid \sigma(t) = 1\}$$

where $\sigma : [\![0, T-1]\!] \to \{0, 1\}$ is a binary-valued signal satisfying

$$\sigma(t) = \begin{cases} 1 & \text{if the hard constraint is active at time } t; \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Problem 2: *Choose a control trajectory* $u : [\![0, T-1]\!] \to \mathbb{R}^{m_{tot}}$ *to minimize*

$$J(u_t) = \mathbb{E} \sum_{t=0}^{T-1} \ell(x_t, u_t) + \mathbb{E}\ell_T(x_T) \tag{14}$$

subject to the dynamics (2) *and the intermittent equality constraint*

$$\mathbb{1}_{m_{tot}}^\intercal u_t = c_t, \quad \forall t \in \mathbb{S} \subseteq [\![0, T-1]\!] \tag{15}$$

where $c_t \in \mathbb{R}$ *represents the intermittent total consumption requirement at time* $t \in \mathbb{S} \subseteq [\![0, T-1]\!]$. ☐

The solution to Problem 2 builds upon the solution in Theorem 1. Since the hard equality constraint is only active for time steps within the set $\mathbb{S}$, the optimal control law adopts a switched structure For any time $t \in \mathbb{S}$, the control is derived using the constrained formulation and is similar to the solution of Problem 1; for any time $t \notin \mathbb{S}$, the problem reduces to a standard unconstrained stochastic LQ problem. Proposition 1 below formalizes this mechanism by presenting a control law where matrices the Γ_t and γ_t switch between their constrained and unconstrained forms depending on the activation status of the constraint.

Proposition 1: *Let Assumptions 1 and 2 hold and let $P : [\![0,T]\!] \to \mathbb{R}^{n_{tot}\times n_{tot}}$ and $s : [\![0,T]\!] \to \mathbb{R}^{n_{tot}}$ be the solutions of the following backward recursions*

$$P_t = Q + A^\intercal P_{t+1}A - A^\intercal P_{t+1}B\Gamma_t B^\intercal P_{t+1}A, \tag{16}$$

$$s_t = \left[A^\intercal - A^\intercal P_{t+1}B\Gamma_t B^\intercal\right] s_{t+1} + A^\intercal P_{t+1}B\gamma_t - Qr_t, \tag{17}$$

with the final condition $P_T = Q_T$ and $s_T = -Q_T r_T$ where

$$\Gamma_t = \begin{cases} \Omega_t^{-1} - \dfrac{\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}\mathbb{1}_{m_{\text{tot}}}^\intercal\Omega_t^{-1}}{\mathbb{1}_{m_{\text{tot}}}^\intercal\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}}, & \text{if } \sigma(t) = 1 \\ \Omega_t^{-1}, & \text{if } \sigma(t) = 0 \end{cases}$$

$$\gamma_t = \begin{cases} \dfrac{\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}c_t}{\mathbb{1}_{m_{\text{tot}}}^\intercal\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}}, & \text{if } \sigma(t) = 1 \\ 0_{m_{\text{tot}}\times 1}, & \text{if } \sigma(t) = 0 \end{cases}$$

$$\Omega_t = R + B^\intercal P_{t+1}B.$$

Then the stochastic optimal control strategy for Problem 2 is given by

$$u_t = -\Gamma_t(B^\intercal P_{t+1}Ax_t + B^\intercal s_{t+1}) + \gamma_t, \tag{18}$$

for all $t \in [\![0, T-1]\!]$. □

PROOF: The optimal control law is derived by considering two cases based on the activity of the constraint.

For any time $t \in \mathbb{S}$, the equality constraint is active. The derivation of the optimal control law and the corresponding recursions for P_t and s_t follows directly from the proof steps of Theorem 1. For any time $t \notin \mathbb{S}$, the constraint is inactive. The problem reduces to a standard unconstrained stochastic LQ problem. The optimal control law for these time steps is given by $u_t^* = -\Omega_t^{-1}(B^T P_{t+1}Ax_t + B^T s_{t+1})$, which corresponds to setting $\Gamma_t = \Omega_t^{-1}$ and $\gamma_t = 0$. Substituting this unconstrained solution into the Bellman equation yields the standard Riccati recursions for P_t and s_t as stated in the proposition for $t \notin \mathbb{S}$.

The minimization at each dynamic programming step is a convex optimization problem with or without linear constraints, and hence the KKT conditions applied are sufficient for optimality. Combining these two cases above gives the desired forms for Γ_t and γ_t, and completes the proof. ■

B. Intermittent soft and hard constraints

Building on the intermittent control strategy presented previously, this section introduces a modified approach. Instead of having no constraint when the hard equality constraint is inactive, a soft constraint can be applied in the form of a quadratic penalty term in the cost function. Such formulation does not strictly guarantee the constraint is met at all times but penalizes deviations from the target, leading to smoother control actions.

More specifically, at time $t \in [\![0, T-1]\!]$, the global system (2) incurs an instantaneous cost

$$\ell^s(x_t, u_t) := \ell(x_t, u_t) + \eta(1-\sigma(t))(\mathbb{1}_{m_{\text{tot}}}^\intercal u_t - c_t)^2 \tag{19}$$

with $\sigma(t)$ defined in (13), and at the terminal time T, the system incurs a terminal cost

$$\ell_T^s(x_T) = \ell_T(x_T) \tag{20}$$

where $\eta \geq 0$ is a scalar penalty weight.

Problem 3: *Choose a control trajectory $u : [\![0, T-1]\!] \to \mathbb{R}^{m_{tot}}$ to minimize*

$$J(u_t) = \mathbb{E}\sum_{t=0}^{T-1}\ell^s(x_t, u_t) + \mathbb{E}\ell_T^s(x_T) \tag{21}$$

while respecting the following constraint

$$\mathbb{1}_{m_{\text{tot}}}^\intercal u_t = c_t, \quad \forall t \in \mathbb{S} \subseteq [\![0, T-1]\!] \tag{22}$$

where $c_t \in \mathbb{R}$ represents the intermittent total consumption requirement at time $t \in \mathbb{S} \subseteq [\![0, T-1]\!]$. □

Compared to the previous problems, Problem 3 incorporates an additional quadratic penalty term $\eta(1-\sigma(t))(\mathbb{1}_{m_{\text{tot}}}^\intercal u_t - c_t)^2$ directly into the cost function to reduce the mismatch between the sum of the control and the required consumption. Proposition 2 below provides the resulting stochastic optimal control solutions.

Proposition 2: *Let Assumptions 1 and 2 hold. Let $P : [\![0,T]\!] \to \mathbb{R}^{n_{tot}\times n_{tot}}$ and $s : [\![0,T]\!] \to \mathbb{R}^{n_{tot}}$ be the solutions of the following backward recursions*

$$P_t = Q + A^\intercal P_{t+1}A - A^\intercal P_{t+1}B\Gamma_t B^\intercal P_{t+1}A \tag{23}$$

$$s_t = \left[A^\intercal - A^\intercal P_{t+1}B\Gamma_t B^\intercal\right] s_{t+1} + A^\intercal P_{t+1}B\gamma_t - Qr_t \tag{24}$$

with the final condition $P_T = Q_T$ and $s_T = -Q_T r_T$ where

$$\Gamma_t = \begin{cases} \Omega_t^{-1} - \dfrac{\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}\mathbb{1}_{m_{\text{tot}}}^\intercal\Omega_t^{-1}}{\mathbb{1}_{m_{\text{tot}}}^\intercal\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}}, & \text{if } \sigma(t) = 1 \\ \Pi_t^{-1}, & \text{if } \sigma(t) = 0 \end{cases}$$

$$\gamma_t = \begin{cases} \dfrac{\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}c_t}{\mathbb{1}_{m_{\text{tot}}}^\intercal\Omega_t^{-1}\mathbb{1}_{m_{\text{tot}}}}, & \text{if } \sigma(t) = 1 \\ \eta c_t\Pi_t^{-1}\mathbb{1}_{m_{\text{tot}}}, & \text{if } \sigma(t) = 0 \end{cases}$$

$$\Omega_t = R + B^\intercal P_{t+1}B, \quad \Pi_t = R + B^\intercal P_{t+1}B + \eta\mathbb{1}_{m_{\text{tot}}}\mathbb{1}_{m_{\text{tot}}}^\intercal.$$

Then the optimal control strategy for Problem 3 is given by

$$u_t = -\Gamma_t(B^\intercal P_{t+1}Ax_t + B^\intercal s_{t+1}) + \gamma_t \tag{25}$$

for all $t \in [\![0, T-1]\!]$. □

For the detailed proof, please refer to the Appendix VII-A.

V. Numerical examples for demand response

Consider an illustrative example where the energy generated by a solar farm is about to reach its local storage capacity and excess solar power need be distributed to users in an energy network. We apply our control strategies to allocate solar power to a network of $N = 50$ residential batteries where each battery unit has a capacity limit of 80 kWh. The batteries are equally divided into two classes (α and β) with distinct SoC targets of 80% and 40% to model diverse household needs. The state $x_{i,t} \in \mathbb{R}$ (with $d_x^i = 1$) and control input $u_{i,t} \in \mathbb{R}$ (with $d_u^i = 1$) represent respectively the battery level in kWh and power consumption in kW of battery unit i. The power generation data is taken from the Canadian Weather Energy and Engineering Datasets (CWEEDS) from the climate station located in Montreal-East [23]. We consider a solar power plan with a panel area of 1000 m² and time span of $T = 24$ hours (a full day) with a period of 1 hour. The stochastic noise $\{w_{i,t}\}_{t\geq 0}$ is Gaussian process with mean zero and $\mathbb{E}[w_i^2] = 3$ (kWh)^2 and is assumed to be independent among agents.

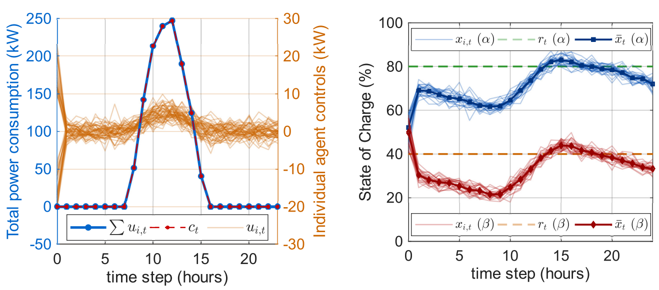
Consider a demand response call that imposes the equality constraint $\mathbb{1}_{m_{\text{tot}}}^\intercal u_t = c_t$, where c_t is the (predicted) excess power of the solar farm (in kW) and $\mathbb{1}_{m_{\text{tot}}}^\intercal u_t$ is the total power consumption of the N residential batteries (in kW). The parameters for user i are as follows:

$$A_i \in [0.96, 0.99], \; B_i = 1 \text{ kWh/kW},$$
$$Q_i = Q_{i,T} = I_{d_x^i}, \; R_i = 0.01 I_{d_u^i}.$$

We assume that A_i is uniformly distributed between 0.96 and 0.99 and that each user has an initial SoC uniformly distributed between 40% and 60% of its 80 kWh capacity.

A. Non-intermittent hard constraint

First, we solve Problem 1 using Theorem 1, where the constraint (8) is active the full duration of the simulation.



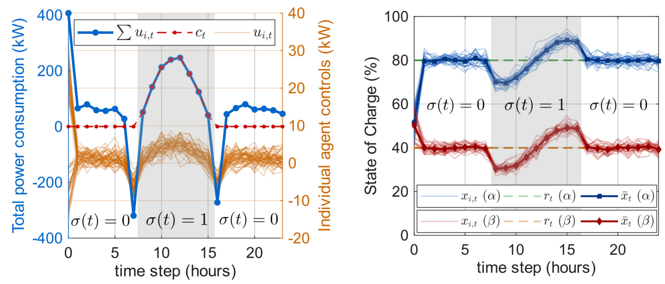
(a) Constraint verification (b) Reference tracking

Fig. 1: Hard non-intermittent constraint

The results in Fig. 1(a) demonstrate that the total power consumed by the network of residential batteries accurately matches the available solar generation at each time step, thereby satisfying the hard equality constraint. Furthermore, Fig. 1(b) shows that the SoCs of the two classes α and β, diverge after $t = 1$ to successfully track their respective nominal targets.

B. Intermittent hard constraint

Consider the case where the constraint (15) is respected intermittently when the excess power is generated by the solar power plant. Problem 2 is then solved using Proposition 1 where the hard constraint is active only for $t \in \mathbb{S} \subseteq [\![0, T-1]\!]$ where $\mathbb{S} = \{t \in [\![0, T-1]\!] \mid c_t \neq 0\}$.



(a) Constraint verification (b) Reference tracking

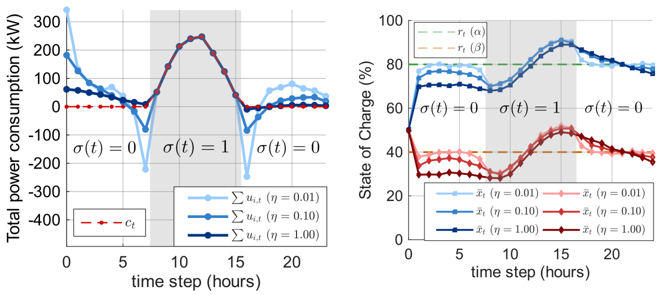
Fig. 2: Hard intermittent constraint

The results in Fig. 2 show that when $t \in \mathbb{S}$, the constraint is verified, while for $t \notin \mathbb{S}(\sigma(t) = 0)$, the individual SoC tracking for classes α and β is more accurate. The reason that the individual's tracking performance in Fig. 1 is less satisfactory compared in Fig. 2 in the unconstrained periods is due to the trade-off within the control problem formulation: the equality constraint forces the optimal control inputs to project onto a specific affine subspace (see Remark 1), removing operational degrees of freedom that would allow the control effort to better track the agents' individual SoC objectives. However, we observe that the transition between the unconstrained and constrained periods results in a sharp negative power peak, as seen in Fig. 2(a), which implies an undesirable discharge to other local storage devices or flexible loads. This negative peak is reflecting the controller's anticipatory behavior to make "room" in the batteries so that the impending solar charging doesn't cause the agents' SoC trajectories to overshoot their objectives.

C. Intermittent soft and hard constraints

To resolve the issue of the negative power peak (illustrated in Fig. 2(a)), we use a switched strategy; more specifically, we activate the hard constraint when $t \in \mathbb{S}$ and switch to a soft constraint when $t \notin \mathbb{S}$, as described in Proposition 2. In this formulation, the choice of the penalty weight η is critical for balancing the trade-off between enforcing the soft constraint and achieving individual agent tracking objectives. A small value of η would place a lower penalty on deviations from the target consumption c_t, allowing agents to prioritize tracking their individual SoC objectives more closely, but at the risk of insufficient smoothing during transition periods as seen in Fig. 2. Conversely, a very large η would cause the soft constraint to approximate a hard constraint.

As illustrated in Fig. 3, with the high value of $\eta = 1$, the combination of soft and hard intermittent constraints smooths the transition between the constrained and unconstrained periods. This "switched" approach eliminates the



(a) Constraint verification

(b) Reference tracking

Fig. 3: Soft and hard intermittent constraint

sharp negative power peaks observed in the case with purely intermittent hard constraint. Furthermore, a comparison of Fig. 3(b) with Fig. 2(b) and Fig. 1(b) reveals that during the soft constrained periods ($\sigma(t) = 0$), the agents' SoC trajectories track their reference paths more accurately as $\eta \to 0$. Specifically, decreasing the penalty parameter η shifts the optimization trade-off, causing the individual SoC tracking term to dominate the cost function during unconstrained periods.

VI. Conclusion

We solved the discrete-time linear quadratic stochastic control problem with hard equality constraints on the summation of control inputs. The results are generalized to allow for predetermined switches between hard and soft constraints in order to eliminate sharp power peaks that otherwise may occur with purely intermittent hard constraints. Numerical examples in the context of energy demand response demonstrated that the total consumptions of the network of batteries exactly matched the excess solar power generation, satisfying the hard constraint, while individual units tracked their distinct SoC objectives.

The collaborative agents face a trade-off between closely tracking their individual SoC objectives and satisfying the hard equality constraint. Depending on the constraints and individual objectives, a sufficiently large number of participating agents with flexibility in SoC may be needed, which will be investigated in the future. Other future work should investigate the solution to scenarios with constraints on low-dimensional subspace with network couplings and constraints with individual charging rates, formulations where agents have local objectives (similar to mean field games), and switching conditions to activate constraints to ensure desired system behavior.

VII. Acknowledgment

The authors gratefully acknowledge Roland Malhamé and Antoine Lesage-Landry for their insightful discussions, as well as the anonymous reviewers for their valuable feedback.

References

[1] T. Kataray, B. Nitesh, B. Yarram, S. Sinha, E. Cuce, S. Shaik, P. Vigneshwaran, and A. Roy, "Integration of smart grid with renewable energy sources: Opportunities and challenges – a comprehensive review," *Sustainable Energy Technologies and Assessments*, vol. 58, p. 103363, 2023.

[2] J. Pallage and A. Lesage-Landry, "(Trustworthy) AI for Québec's virtual power plant," Groupe d'études et de recherche en analyse des décisions, GERAD, Montréal QC H3T 2A7, Canada, Les Cahiers du GERAD G-2024-23, Mar. 2024.

[3] A. Lesage-Landry and J. Pallage, "Online dynamic submodular optimization," *Automatica*, vol. 167, p. 111758, 2024.

[4] P. Scokaert and J. Rawlings, "Constrained linear quadratic regulation," *IEEE Transactions on Automatic Control*, vol. 43, no. 8, pp. 1163–1169, 1998.

[5] D. Chmielewski and V. Manousiouthakis, "Constrained infinite-time quadratic optimal control: the linear stochastic and nonlinear deterministic cases," in *Proceedings of the American Control Conference (ACC)*, vol. 4, 1998, pp. 2093–2097 vol.4.

[6] J. B. Mare and J. A. De Doná, "Solution of the input-constrained LQR problem using dynamic programming," *Systems & Control Letters*, vol. 56, no. 5, pp. 342–348, 2007.

[7] I. Chang and J. Bentsman, "Constrained discrete-time state-dependent Riccati equation technique: A model predictive control approach," in *Proceedings of the Conference on Decision and Control (CDC)*, 2013, pp. 5125–5130.

[8] S. Ko and R. R. Bitmead, "Optimal control for linear systems with state equality constraints," *Automatica*, vol. 43, no. 9, pp. 1573–1582, 2007.

[9] A. Sideris and L. A. Rodriguez, "A Riccati approach to equality constrained linear quadratic optimal control," in *Proceedings of the American Control Conference (ACC)*, 2010, pp. 5167–5172.

[10] F. Laine and C. Tomlin, "Efficient computation of feedback control for equality-constrained LQR," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2019, pp. 6748–6754.

[11] A. Laurenzi, F. Ruscelli, and N. Tsagarakis, "A Constrained Iterative LQR Solver for the Trajectory Optimization Framework Horizon," Istituto Italiano di Tecnologia, Tech. Rep., Mar. 2025.

[12] R. Mitze and M. Mönnigmann, "A dynamic programming approach to solving constrained linear–quadratic optimal control problems," *Automatica*, vol. 120, p. 109132, 2020.

[13] D. Krokavec and A. Filasová, "Constrained control of discrete-time stochastic systems," in *Proceedings of the IFAC World Congress*, vol. 41, no. 2, 2008, pp. 15 315–15 320.

[14] D. Krokavec and A. Filasová, "Control of discrete-time stochastic systems with state equality constraints," in *Proceedings of the IFAC World Congress*, vol. 44, no. 1, 2011, pp. 3210–3215.

[15] D. Krokavec and A. Filasová, "H_∞ Control of Discrete-Time Stochastic State-multiplicative Systems Constrained in State by Equality Constraints," in *Proceedings of the IFAC World Congress*, vol. 47, no. 3, 2014, pp. 8699–8704.

[16] M. F. Hassan, M. T. Alrifai, H. M. Soliman, and M. A. Kourah, "Observer-based controller for constrained uncertain stochastic nonlinear discrete-time systems," *International Journal of Robust and Nonlinear Control*, vol. 26, no. 10, pp. 2090–2115, 2016.

[17] Y. Chen and Y. Zhu, "Indefinite LQ optimal control with equality constraint for discrete-time uncertain systems," *Japan Journal of Industrial and Applied Mathematics*, vol. 33, 05 2016.

[18] W. Wu, J. Gao, J.-G. Lu, and X. Li, "On continuous-time constrained stochastic linear–quadratic control," *Automatica*, vol. 114, p. 108809, 2020.

[19] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. Berkeley and Los Angeles: University of California Press, 1951, pp. 481–492.

[20] A. Ferrante and L. Ntogramatzidis, "The generalised discrete algebraic riccati equation in linear-quadratic optimal control," *Automatica*, vol. 49, no. 2, pp. 471–478, 2013.

[21] V. Ionescu and C. Oară, "Generalized discrete-time riccati theory," *SIAM Journal on Control and Optimization*, vol. 34, no. 2, pp. 601–619, 1996.

[22] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[23] Natural Resources Canada, "Canadian Weather Energy and Engineering Datasets (CWEEDS)," accessed: 2025-09-22. [Online]. Available: https://climate.weather.gc.ca/prods_servs/engineering_e.html

Appendix

A. Proof of Proposition 2

The proof proceeds by considering two cases based on the activation of constraints determined by the signal $\sigma(t)$.

1) Case 1: Active hard constraint ($\sigma(t) = 1$): For any time t such that $\sigma(t) = 1$, the penalty term in the cost function (19) is null since $(1 - \sigma(t)) = 0$. The problem thus reduces to minimizing the original cost function subject to the hard equality constraint $\mathbb{1}_{m_{\text{tot}}}^\top u_t = c_t$. This scenario is identical to Problem 1. Consequently, the derivation of the optimal control law and the corresponding backward recursions for P_t and s_t follow directly from the proof of Theorem 1. This yields the expressions for Γ_t and γ_t as defined in Proposition 2 for $\sigma(t) = 1$.

2) Case 2: Active soft constraint ($\sigma(t) = 0$): For any time t such that $\sigma(t) = 0$, the hard constraint is inactive, and the cost function includes the quadratic penalty term $\eta(\mathbb{1}_{m_{\text{tot}}}^\top u_t - c_t)^2$. This transforms the problem into an unconstrained stochastic LQ problem with a modified instantaneous cost.

The Bellman recursion at time t starting at state z satisfies

$$\begin{aligned}V_t(z) = \min_u \big\{&|z - r_t|_Q^2 + |u|_R^2 + \eta(\mathbb{1}_{m_{\text{tot}}}^\top u - c_t)^2 \\ &+ \mathbb{E}\big[V_{t+1}(Az + Bu + w_t)\big]\big\}\end{aligned}$$

Assuming the value function at time $t + 1$ is of the form $V_{t+1}(z) = z^\top P_{t+1} z + 2s_{t+1}^\top z + q_{t+1}$, and expanding the expectation, the Bellman recursion becomes

$$\begin{aligned}V_t(z) = &|z - r_t|_Q^2 + \operatorname{tr}(WP_{t+1}) + q_{t+1} + |Az|_{P_{t+1}}^2 + \eta c_t^2 \\ &+ 2s_{t+1}^\intercal Az + \min_u \Big\{|u|_{\Pi_t}^2 + 2u^\intercal g_t(z)\Big\}\end{aligned}$$

using $\mathbb{E}(w_{t+1}^\intercal P_{t+1} w_{t+1}) = \operatorname{tr}(WP_{t+1})$, where $\Pi_t := R + B^\intercal P_{t+1} B + \eta \mathbb{1}_{m_{\text{tot}}} \mathbb{1}_{m_{\text{tot}}}^\intercal$ and $g_t(z) = B^\intercal P_{t+1} Az + B^\intercal s_{t+1} - \eta c_t \mathbb{1}_{m_{\text{tot}}}$. The unconstrained minimization problem is $\min_u\{u^\intercal \Pi_t u + 2u^\intercal g_t(z)\}$. Setting the gradient with respect to u to zero yields $2\Pi_t u + 2g_t(z) = 0$, which gives the optimal control law:

$$u_t^* = -M_t(B^\top P_{t+1} A x_t + B^\top s_{t+1}) + \gamma_t$$

where $\Gamma_t = \Pi_t^{-1}$ and $\gamma_t = \eta c_t \Pi_t^{-1} \mathbb{1}_{m_{\text{tot}}}$. The resulting control at each dynamic programming step is optimal following the standard LQ control theory.

Substituting the optimal control law back into the Bellman recursion, we can identify the coefficients for the quadratic and linear terms in z. By matching the coefficients with the assumed form $V_t(z) = z^\intercal P_t z + 2s_t^\intercal z + q_t$, we obtain the following recursion equations for P_t and s_t:

$$\begin{aligned}P_t &= Q + A^\intercal P_{t+1} A - A^\intercal P_{t+1} B \Pi_t^{-1} B^\intercal P_{t+1} A \\ s_t &= \left[A^\intercal - A^\intercal P_{t+1} B \Pi_t^{-1} B^\intercal\right] s_{t+1} - Q r_t \\ &\quad + A^\intercal P_{t+1} B \Pi_t^{-1} \mathbb{1}_{m_{\text{tot}}} \eta c_t.\end{aligned}$$

This confirms our assumption on the quadratic form of the value function in both cases and completes the proof by induction. ■

B. Properties of the Riccati-like equation

Lemma 1: Let $R > 0$ and $Q \geq 0$ and $Q_T \geq 0$. Then solution P_t to the following recursive equation

$$P_t = Q + A^\intercal P_{t+1} A - A^\intercal P_{t+1} B \Gamma_t B^\intercal P_{t+1} A, \qquad (26)$$

with $P_T = Q_T$ and

$$\Gamma_t = \Omega_t^{-1}\left(I - \frac{\mathbb{1}_{m_{\text{tot}}} \mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1}}{\mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}}}\right), \quad \Omega_t = R + B^\intercal P_{t+1} B,$$

remains positive semidefinite for all $t \in [\![0, T]\!]$. □

PROOF: This can be proved by backward induction. At $t = T$, $P_T = Q_T \geq 0$ by assumption. Assume $P_{t+1} \geq 0$. Then we need to show that $P_t \geq 0$. To show this, we construct an auxiliary linear quadratic regulation problem from state $x \in \mathbb{R}^{n_{\text{tot}}}$ at time t, under dynamics

$$x_{t+1} = Ax_t + Bu_t$$

and input $u_t \in \mathbb{R}^{m_{\text{tot}}}$ satisfying $\mathbb{1}_{m_{\text{tot}}}^\intercal u_t = 0$. The cost is

$$J_t(x, u) := x^\intercal Q x + u^\intercal R u + (Ax + Bu)^\intercal P_{t+1}(Ax + Bu).$$

which can be represented as

$$J_t(x, u) = x^\intercal Q x + x^\intercal A^\intercal P_{t+1} A x + 2u^\intercal g + u^\intercal \Omega_t u,$$

with $g := B^\intercal P_{t+1} A x$. To solve the problem

$$\min_{u : \mathbb{1}_{m_{\text{tot}}}^\intercal u = 0} J_t(x, u)$$

subject to the linear dynamics above, we consider the following Lagrangian $\mathcal{L}(u, \lambda) = u^\intercal \Omega_t u + 2g^\intercal u + \lambda \mathbb{1}_{m_{\text{tot}}}^\intercal u$ and identify the first-order optimality conditions

$$2\Omega_t u + 2g + \lambda \mathbb{1}_{m_{\text{tot}}} = 0, \quad \mathbb{1}_{m_{\text{tot}}}^\top u = 0.$$

Since $R > 0$ and $P_{t+1} \geq 0$, we obtain that $\Omega_t = R + B^\intercal P_{t+1} B > 0$. Hence the optimal control action is given by $u^* = -\Omega_t^{-1}\left(g + \frac{\lambda}{2}\mathbb{1}_{m_{\text{tot}}}\right)$. Plugging into constraint

$$\mathbb{1}_{m_{\text{tot}}}^\intercal u^* = -\mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} g - \frac{\lambda}{2} \mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}} = 0.$$

and solving for λ yields $\frac{\lambda}{2} = -\frac{\mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} g}{\mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}}}$. Thus,

$$u^* = -\Gamma_t\, g, \quad \text{where} \quad \Gamma_t = \Omega_t^{-1} - \frac{\Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}} \mathbb{1}_{m_{\text{tot}}}^\intercal \Omega_t^{-1}}{\mathbb{1}_{m_{\text{tot}}}^\top \Omega_t^{-1} \mathbb{1}_{m_{\text{tot}}}}.$$

The minimum value of the cost is hence given by

$$\begin{aligned}&x^\intercal Q x + x^\intercal A^\intercal P_{t+1} A x - g^\intercal \Gamma_t g \\ &= x^\intercal\left(Q + A^\intercal P_{t+1} A - A^\intercal P_{t+1} B \Gamma_t B^\intercal P_{t+1} A\right) x.\end{aligned}$$

Therefore, we get the equation (26). From the cost minimization interpretation, we have

$$x^\intercal P_t x = \min_{u :\, \mathbb{1}_{m_{\text{tot}}}^\top u = 0} J_t(x, u)$$

and since all terms in $J_t(x, u)$ are nonnegative, the minimum must also be nonnegative. That is

$$x^\intercal P_t x = \min_{u :\, \mathbb{1}_{m_{\text{tot}}}^\top u = 0} J_t(x, u) \geq 0, \quad \forall x.$$

This implies $P_t \geq 0$. Thus, by backward induction, $P_t \geq 0$ for all t. ■