Distributed Multi-Parametric Quadratic Programming

Nader Motee, *Member, IEEE*, and Ali Jadbabaie, *Senior Member, IEEE*

Abstract—One of the fundamental problems in the area of large-scale optimization is to study locality features of spatially distributed optimization problems in which the variables are coupled in the cost function as well as constraints. Such problems can motivate the development of fast and well-conditioned distributed algorithms. In this paper, we study spatial locality features of large-scale multi-parametric quadratic programming (MPQP) problems with linear inequality constraints. Our main application focus is receding horizon control of spatially distributed linear systems with input and state constraints. We propose a new approach for analysis of large-scale MPQP problems by blending tools from duality theory with operator theory. The class of spatially decaying matrices is introduced to capture couplings between optimization variables in the cost function and the constraints. We show that the optimal solution of a convex MPQP is piecewise affine- represented as convolution sums. More importantly, we prove that the kernel of each convolution sum decays in the spatial domain at a rate proportional to the inverse of the corresponding coupling function of the optimization problem.

Index Terms—Multi-parametric quadratic programming (MPQP), receding horizon control, spacially decaying (SD) matrix.

I. INTRODUCTION

THE problem of performing distributed computations over a network to implicitly solve a global optimization problem has been an active area of research over the past few years [1]. There are many important problems that have been cast in the form of a large-scale finite-dimensional or an infinite-dimensional constraint optimization problem. Substantial progress has been made in understanding of the fundamental issues regarding this class of problems (cf. [1]–[3] and references therein). One of the fundamental problems in this area is to study locality features of spatially distributed optimization problems. Such features can be exploited in the development of fast and well-conditioned parallel and distributed algorithms that are scalable to large interconnected systems.

In parallel to this progress, there has been a rapidly growing interest in systems and control community in the study of coordination and control algorithms for networked dynamic systems. From consensus and agreement problems to formation control, sensing, and coverage, researchers have been interested in algorithms that are spatially distributed and would achieve a global objective using local interactions [4]–[7]. Despite the resurgence of interest in this area, there are few results on optimality of the proposed distributed control schemes [8]–[10].

Among such results, only a handful of proposed methods can handle state and control constraints. With advances in real-time optimization, this has started to change. There have been several recent attempts to develop distributed control algorithms that are based on receding horizon control, where the control action is obtained by solving an online, finite horizon, open-loop optimal control problem, and applying the first step of the optimal control trajectory. Repeating this process and moving the horizon gives rise to a time-invariant feedback policy [11].

While receding horizon control has been around for more than three decades, emergence of its distributed implementations is more recent. Examples of recent application of distributed receding horizon control range from formation control of multi-vehicle systems [12]–[14] to applications in manufacturing and process industry where multiple units cooperatively produce a product [15], [16], and large scale power systems [17]–[22].

In [23], the authors proposed a distributed receding horizon control algorithm for systems which consist of subsystems whose dynamics and constraints are uncoupled, and couplings are imposed through a single performance cost function. Stability analysis is based on the fact that the optimal state trajectory of each subsystem satisfies a compatibility constraints condition, and that the receding horizon updates happen sufficiently fast. In [13], a decentralized receding horizon control scheme for systems whose coupling is through cost function and constraints, is proposed. Each subsystems uses only local information of itself and every neighbor to compute the optimal trajectory. Stability and feasibility issues regarding this distributed algorithm is also discussed and compared to those of others being proposed earlier in the literature. Another related work on this subject was reported in [24] where the authors solve a min-max problem for each subsystem. In this work, coupling comes from dynamics and the stability of the proposed algorithm is ensured by imposing a contractive constraint, called stability constraint. In [18], [22], [25], unconstrained coupled subsystems are addressed with a separable quadratic cost function. The primary objective of these papers is to develop decomposition algorithms, with stability and feasibility guarantees, to solve the centralized receding horizon control problem in a distributed fashion.

Despite the above recent advances, connections between centralized and decentralized implementations of receding horizon control problems on spatially distributed systems is not clear. Such systems consist of large, possibly infinite, number of subsystems coupled either through their dynamics or through a

N. Motee is with the Control and Dynamical Systems Department, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: motee@cds.caltech.edu).

A. Jadbabaie is with the Department of Electrical and Systems Engineering and GRASP Laboratory, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: jadbabai@seas.upenn.edu).

single cost function, which represents some common goal or objective. Our main objective is to explore and quantify the locality features of solutions of receding horizon control problems for spatially distributed systems in which coupling between subsystems decay as function of distance between them. In this paper, we focus our attention on analyzing properties of Multi-Parametric Quadratic Programming (MPQP) problems. We refer to [26], [27] for finite-dimensional analysis of MPQP problems.

The key idea is to employ the operator theoretic tools developed in [28] to study spatially distributed systems. We introduce the Banach algebra of infinite-dimensional spatially decaying (SD) matrices. Furthermore, under some mild conditions we show that the algebra of SD matrices is also closed under inversion operation [29]. Using duality theory and complimentary slackness, we show that the optimal solution of an infinite-dimensional MPQP problem is a piecewise affine map of parameters, and can be represented as convolution sums. Most importantly, we prove that the kernel of each convolution sum decays in the spatial domain at a rate proportional to the inverse of the corresponding coupling function of the optimization problem. In other words, a change of parameters in one node mainly affects the optimal solution of its nearest neighbors. It is important to stress that this locality feature is inherent and, therefore, verify the feasibility of spatial truncation without too much loss of performance. Indeed, in a MPQP problem we can quantify the relations between couplings in the cost function and constraints and locality features of the optimal solution. The importance of this result is that a significant drop-off in complexity can be achieved by localization.

This paper is organized as follows. We introduce the notation and the basic concepts used throughout the paper in Section II. The MPQP problem is presented in Section III. In Section IV, it is shown that receding horizon control of spatially distributed linear systems can be formulated as MPQP. The class of SD matrices is introduced in Section V. This sections studies various properties of SD matrices. In Section VI, we study the locality features of large-scale MPQP problems. Simulations results are discussed in Section VII.

II. PRELIMINARIES

The notation used in this paper is fairly standard. Symbol $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}^+$ the set of nonnegative real numbers, $\mathbb{Z}$ the set of integer numbers, $\mathbb{C}$ the set of complex numbers, and $\mathbb{S}^1$ the unit circle in $\mathbb{C}$. We refer to $\mathbb{G}$ as the spatial domain. Examples of typical spatial domains include $\mathbb{Z}^n$ for some positive integer n. The Banach space $\ell_p(\mathbb{G})$ for $1 \leq p < \infty$ is defined to be the set of all sequences $x = (x_i)_{i \in \mathbb{G}}$ in which $x_i \in \mathbb{R}$ satisfying

$$\sum_{i \in \mathbb{G}} |x_i|^p < \infty$$

endowed with the norm

$$\|x\|_p \triangleq \left(\sum_{i \in \mathbb{G}} |x_i|^p \right)^{1/p}.$$

The Banach space $\ell_\infty(\mathbb{G})$ denotes the set of all bounded sequences endowed with the norm

$$\|x\|_\infty \triangleq \sup_{i \in \mathbb{G}} |x_i|.$$

Throughout the paper, we will use the shorthand notation ℓ_p for $\ell_p(\mathbb{G})$. A linear functional F on the space ℓ_p is a linear mapping from ℓ_p to $\mathbb{R}$. We will use the notation $\langle x, F \rangle$ to denote $F(x)$. An operator $Q : \ell_p \to \ell_q$ for $1 \leq p, q < \infty$ is bounded if it has a finite induced norm, i.e., the following quantity:

$$\|A\|_{p,q} \triangleq \sup_{\|x\|_p = 1} \|Ax\|_q \qquad (1)$$

is bounded. The same definition holds when $p = q = \infty$. The identity operator is denoted by I. The set of all bounded linear operators of ℓ_p into ℓ_p for some $1 \leq p < \infty$ (also for $p = \infty$) is denoted by $\mathscr{B}(\ell_p)$. The space $\mathscr{B}(\ell_p)$ equipped with norm (1) is a Banach space (cf. [30]). The dual space of a Banach space X, denoted by X^*, is the space of all bounded linear functionals on X. Since we are interested in Banach space ℓ_p in this paper, we have $(\ell_p)^* = \ell_q$ where $1/p + 1/q = 1$.

An operator $A \in \mathscr{B}(\ell_p)$ has an *algebraic* inverse [30] if it has an inverse A^{-1} in $\mathscr{B}(\ell_p)$.

Definition 1: Let $A : \ell_p \to \ell_q$ be a bounded linear operator. The adjoint operator $A^* : (\ell_q)^* \to (\ell_p)^*$ is defined by the following equation:

$$\langle x, A^* y^* \rangle = \langle Ax, y^* \rangle$$

for all $x \in \ell_p$ and $y^* \in (\ell_q)^*$.

The space ℓ_2 is a Hilbert space with inner product

$$\langle x, y \rangle \triangleq \sum_{i \in \mathbb{G}} x_i y_i$$

for all $x, y \in \ell_2$. An operator $A \in \mathscr{B}(\ell_2)$ is self-adjoint if $A = A^*$.

Definition 2: An operator $A \in \mathscr{B}(\ell_2)$ is *positive definite*, shown as $A \succ 0$, if there exists a number $\alpha > 0$ such that

$$\langle x, Ax \rangle > \alpha \|x\|_2^2$$

for all nonzero $x \in \ell_2$.

Definition 3: A subset $\mathscr{P}$ of a linear vector space is called a cone if for every $x \in \mathscr{P}$ and $\theta \in \mathbb{R}^+$ it satisfies $\theta x \in \mathscr{P}$. A subset $\mathscr{P}$ of a real vector space is a convex cone if it is convex and a cone, which means that for any $x_1, x_2 \in \mathscr{P}$ and $\theta_1, \theta_2 \in \mathbb{R}^+$, it satisfies

$$\theta_1 x_1 + \theta_2 x_2 \in \mathscr{P}.$$

Definition 4: Let $\mathscr{P}$ be a convex cone in a vector space V. For $x, y \in V$, we write $x \succeq y$ (with respect to $\mathscr{P}$) if $x - y \in \mathscr{P}$. The cone $\mathscr{P}$ defining this relation is called the positive cone in V.

In most situations, the choice of $\mathscr{P}$ will arise naturally. For example, the set of all continuous functions from $D \subseteq \mathbb{R}$ into $\mathbb{R}$ is a vector space $\mathscr{F}$ over $\mathbb{R}$. The positive cone in $\mathscr{F}$ is the set of all continuous functions in the space that are nonnegative
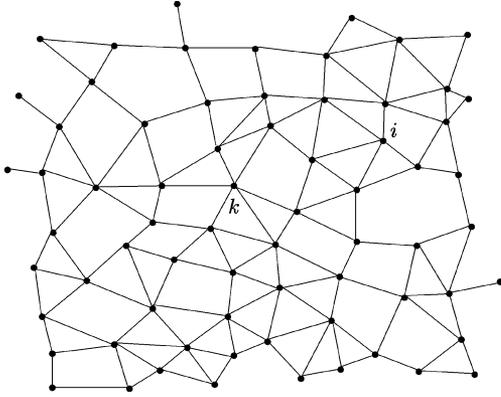
Fig. 1. A spatially distributed system is represented as a connected graph. Coupling between two agents is shown by an undirected edge between them.

everywhere on D. Therefore, for $f, g \in \mathscr{F}$, the notation $f \preceq g$ means the pointwise inequality $f(t) \leq g(t)$ for all $t \in D$.

For a positive cone $\mathscr{P}$ in Banach space X, the corresponding positive convex cone $\mathscr{P}^{\oplus}$ in the dual space X^* is defined by

$$\mathscr{P}^{\oplus} = \{ x^* \in X^* : \langle x, x^* \rangle \geq 0 \text{ for all } x \in \mathscr{P} \}.$$

In this paper we are interested in linear operators which have matrix representations

$$A = \begin{bmatrix} \ddots & & \\ & A_{ki} & \\ & & \ddots \end{bmatrix}$$

where $A_{ki} \in \mathbb{R}$.

Given linear operators $X, Y, Z, W : \ell_p \to \ell_p$, the row and column block composition of these operators are defined entry-wise as follows:

$$\begin{bmatrix} X \\ Y \end{bmatrix}_{ki} = \begin{bmatrix} X_{ki} \\ Y_{ki} \end{bmatrix}, \quad [Z \quad W]_{ki} = [Z_{ki} \quad W_{ki}].$$

More complex block compositions can be defined in terms of these elementary operations. Similarly, we can define the row and column block compositions of elements in ℓ_p.

III. MULTI-PARAMETRIC QUADRATIC PROGRAMMING (MPQP)

In this paper, we consider large-scale MPQP problems defined by

$$\inf_{x \in \Omega} \frac{1}{2} \langle x, Qx \rangle + \langle c, x \rangle$$
$$\text{subject to} : Gx \preceq b \qquad (2)$$

where $Q, G \in \mathscr{B}(\ell_2)$, Q is self-adjoint and positive definite, $c, b \in \ell_2$, and Ω is a convex subset of ℓ_2. We can associated an undirected weighted graph to problem (2) as follows (see Fig. 1). Let denote $\mathbb{G}$ as the index set of nodes of the graph. For each $x = (x_k)_{k \in \mathbb{G}}$ in ℓ_2, element x_k represents the variable corresponding to node k. For a given pair of nodes (k, i), the block elements Q_{ki} and G_{ki} can be thought of as coupling between nodes k and i in the cost function and the constraints. In distributed control applications, very often the underlying system

is spatially distributed over an arbitrary graph. Each node corresponds to an individual dynamical subsystem which is coupled to the other subsystems in the network through their dynamics and collective performance objective function. There are numerous problems concerning this class of systems that can be posed as problem (2). For example, finite-horizon control (also known as model predictive control) of spatially distributed systems with discrete-time linear models (see Section IV for details) and least square problems corresponding to parameter estimations in sensor networks can be cast as problem (2). We refer to [31] and [32] for more discussion on many applications that can be cast as problem (2).

Problem (2) is a large-scale convex MPQP problem. In general, it is a tedious task to find numerical (or explicit) solutions for this class of problems. Therefore, we focus our study on the structural properties of this class of problems. Specifically, we show that the optimal solution of (2) has, to some extend, spatial locality features. This result can be useful in developing efficient approximation methods to solve problem (2) in a distributed fashion.

One of the fundamental tools in deriving methods to compute exact or approximate solutions to the infinite-dimensional problems is duality theory [2], [31]. In spite of the finite-dimensional case, the relationship between primal and dual problems in the infinite-dimensional problems may not be simple. Problem (2) is a convex optimization problem and, therefore under some assumptions (see Theorem 3), strong duality holds. The strong duality (no duality gap) involves cases where the optimal solutions of both primal and dual problems exist and both have the same optimal cost. The strong duality relationship in the primal-dual pair provides a clear insight into the structure of the problem and its optimal solution.

In Section IV, we will show that receding horizon control problem of spatially distributed systems can be formulated as (2).

IV. FORMULATION OF RECEDING HORIZON CONTROL PROBLEM AS QUADRATIC PROGRAMMING PROBLEM

We consider the class of spatially distributed systems which can be described by a discrete-time linear time-invariant model

$$\begin{cases} \psi(t+1) = (A\psi)(t) + (Bu)(t) \\ y(t) = (C\psi)(t) + (Du)(t) \end{cases} \qquad (3)$$

subject to constraints

$$Gu(t) + Hy(t) + F \preceq 0 \qquad (4)$$

for all $t \geq 0$ and with the initial condition $\psi(0) = \psi_0$. All signals are assumed to be in $\ell_2(\mathbb{Z}^+, \ell_2)$ space: $\sum_{t \in \mathbb{Z}^+} \|\psi(t)\|_2^2 < \infty$, similarly summation of $u(t)$ and $y(t)$ over $t \in \mathbb{Z}^+$ are bounded. The state-space matrices A, B, C, D, G, H assumed to be time-independent, $F \in \ell_2$, and the pair (A, B) stabilizable. Note that the ordering in inequality (4) is defined with respect to the positive cone in ℓ_2. The following assumption guarantees existence and uniqueness of classical solutions of the system given by (3) (cf. [33, Chapter 3]). We assume that the semigroup generated by A is *strongly continuous* on ℓ_2.

The control objective is to regulate the state of system (3) to zero while satisfying constraints (4). In the sequel, we will

explain how to achieve this objective by employing receding horizon control techniques [11].

An equivalent representation of system (3) can be obtained by using block-composition operation as follows:

$$\begin{bmatrix} \psi(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \psi(t) \\ u(t) \end{bmatrix}. \tag{5}$$

Furthermore, we assume that operator A is the infinitesimal generator of an exponentially stable C_0-semigroup on ℓ_2. We refer to [27] and references in there for further discussions on this assumption.

The receding horizon control problem for system (3) subject to constraint (4) can be formulated as follows:

$$\inf_{\mathbf{u} \triangleq [u(0)^T, \dots, u(T_u-1)^T]^T}$$

$$\left\{ \mathfrak{J}(\psi_0, \mathbf{u}) \triangleq \langle \psi(T_y), P\psi(T_y) \rangle \right.$$

$$\left. + \sum_{k=1}^{T_y-1} \langle \psi(k), Q\psi(k) \rangle + \sum_{k=0}^{T_u-1} \langle u(k), Ru(k) \rangle \right\}$$

subject to

$$\begin{aligned} Gu(k) + Hy(k) + F &\preceq 0, \quad k = 0, 1, \dots, T_c \\ \psi(0) &= \psi(t) \\ \psi(k+1) &= (A\psi)(k) + (Bu)(k), \ k = 0, 1, \dots, T_y \\ u(k) &= 0, \quad T_u \le k \le T_y - 1. \end{aligned} \tag{6}$$

The nonnegative integer number T_y is the state prediction horizon, T_u the control prediction horizon, and T_c the constraint horizon. Furthermore, we assume that $T_u \le T_y - 1$ and $T_c \le T_y - 1$ (cf. [27]). For simplicity, we will assume that $T_u = T_c = T_y - 1$. For $T_u < T_y - 1$ and $T_c < T_y - 1$, optimization is performed only over T_u control variables, and for the rest of the horizon we may use zero control inputs, which in turn will reduce the complexity of the problem.

The functional $\mathfrak{J}(\psi_0, \mathbf{u})$ can be interpreted as the collective performance objective of the entire system. Assume that matrices $Q \ge 0, R > 0$ are self-adjoint, and $(Q^{1/2}, A)$ detectable.

In the sequel, we show that similar to the finite-dimensional case [27], the receding horizon control problem (6) can be represented in the compact form of (2). Moreover, the terminal weighting cost P can be determined by solving the corresponding Lyapunov equation (cf. [34])

$$\langle A\phi, PA\phi \rangle - \langle \phi, P\phi \rangle + \langle \phi, Q\phi \rangle = 0 \tag{7}$$

for all $\phi \in D(\mathcal{A})$ (domain of the operator). A similar proof as shown in for the continuous-time case can be outlined to show that if A, Q are SD, then P the unique solution of the Lyapunov (7) is also SD.

The prediction model for system (5) is given by

$$\begin{bmatrix} \Psi \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \psi_0 \\ \mathbf{u} \end{bmatrix} \tag{8}$$

where

$$\Psi = \begin{bmatrix} \psi(1) \\ \vdots \\ \psi(T_y) \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y(0) \\ \vdots \\ y(T_y-1) \end{bmatrix}$$

and matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are completely determined from A, B, C, and D by using block-composition operations as follows:

$$\mathbf{A} = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^{T_y} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} B & 0 & \dots \\ AB & B & \dots \\ \vdots & & \ddots \\ A^{T_y}B & \dots & AB & B \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{T_y-1} \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} D & 0 & & \dots \\ CB & D & & \dots \\ CAB & CB & D & \dots \\ \vdots & & & \ddots \\ CA^{T_y-2}B & \dots & & CB & D \end{bmatrix}.$$

Therefore, the state prediction model for system (2) is given by

$$\Psi = \mathbf{A}\psi_0 + \mathbf{B}\mathbf{u}. \tag{9}$$

The quadratic cost in (6) can be rewritten in the following form:

$$\mathfrak{J}(\psi_0, \mathbf{u}) = \langle \Psi, \mathbf{Q}\Psi \rangle + \langle \mathbf{u}, \mathbf{R}\mathbf{u} \rangle \tag{10}$$

in which

$$\mathbf{Q} = \begin{bmatrix} Q & & & \\ & \ddots & & \\ & & Q & \\ & & & P \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} R & & \\ & \ddots & \\ & & R \end{bmatrix}$$

and the off-diagonal blocks are equal to zero. Substituting (9) into (10) gives us

$$\mathfrak{J}(\psi_0, \mathbf{u}) = \langle \psi_0, (Q + \mathbf{A}^*\mathbf{Q}\mathbf{A}) \psi_0 \rangle$$
$$+ 2 \langle \psi_0, \mathbf{A}^*\mathbf{Q}\mathbf{B}\mathbf{u} \rangle + \langle \mathbf{u}, (\mathbf{R} + \mathbf{B}^*\mathbf{Q}\mathbf{B}) \mathbf{u} \rangle.$$

The input and output constraints in (6) can be rewritten as

$$\mathbf{G}\mathbf{u} + \mathbf{H}\mathbf{y} + \mathbf{F} \preceq 0 \tag{11}$$

where

$$\mathbf{G} = \begin{bmatrix} G & & \\ & \ddots & \\ & & G \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} H & & \\ & \ddots & \\ & & H \end{bmatrix},$$

$$\mathbf{F} = \begin{bmatrix} F \\ \vdots \\ F \end{bmatrix}$$

and the off-diagonal blocks are equal to zero. In (11), the inequality is interpreted elementwise. From (8), we have

$$\mathbf{y} = \mathbf{C}\psi_0 + \mathbf{Du}. \tag{12}$$

Substituting (12) into (11), it follows that:

$$(\mathbf{G} + \mathbf{HD})\,\mathbf{u} \preceq -\mathbf{HC}\psi_0 - \mathbf{F}.$$

Therefore, problem (6) is equivalent to the following infinite-dimensional QP problem:

$$\inf_{\mathbf{u}} \frac{1}{2}\,\langle \mathbf{u}, \mathbf{L}_0\mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{L}_1\psi_0\rangle$$
$$\text{subject to}: \mathbf{Mu} \preceq \mathbf{N} + \mathbf{E}\psi_0 \tag{13}$$

where

$$\mathbf{L}_0 \triangleq \mathbf{R} + \mathbf{B}^*\mathbf{QB}, \quad \mathbf{L}_1 \triangleq \mathbf{B}^*\mathbf{QA} \tag{14}$$
$$\mathbf{M} \triangleq \mathbf{G} + \mathbf{HD}, \quad \mathbf{N} \triangleq -\mathbf{F}, \quad \mathbf{E} \triangleq -\mathbf{HC}. \tag{15}$$

In finite-dimensional case, for a given horizon length T_y there is a polyhedral set of initial conditions for which feasible trajectories exists, over which the receding horizon controller is stabilizing (cf. [11], [27], [35]–[37] and the references therein). For our formulation, we assume that the set of all initial conditions for which an optimal solution of (6) exist is a set $\mathcal{X}_0$ which can be characterized as follows (with respect to the positive cone in ℓ_2):

$$\mathcal{X}_0 = \{x \in \ell_2 : A_0 x \preceq b_0\}$$

where $A_0 \in \mathscr{B}\,(\ell_2)$. Formulation in (13) gives a clear picture of the relationship between the control input variables and initial condition ψ_0. Problem (13) is a multi-parametric optimization problem on $\mathbb{G}$ (the spatial domain) in which ψ_0 is treated as vector of parameters.

V. Spatially Decaying Matrices

Spatially-decaying (SD) matrices have been recently developed to study the structural properties of infinite-horizon optimal control of spatially distributed systems [28]. In general, the algebra of SD matrices furnishes a suitable framework to study infinite-dimensional systems. In the following, we will introduce the class of SD matrices. We will also show that the set of such matrices is closed under inversion (when the inverses exist). Specifically, it is shown that similar to translation invariant operators, if a spatially decaying matrix has an algebraic inverse in $\mathscr{B}\,(\ell_2)$ and some additional conditions, then the inverse matrix is spatially decaying as well.

A. Definitions

In the sequel, by a distance function on $\mathbb{G}$ we mean a single-valued function $\mathrm{dis} : \mathbb{G} \times \mathbb{G} \to \mathbb{R}^+$ which has the following properties:

1) $\mathrm{dis}(k,i) = 0$ iff $k = i$;
2) $\mathrm{dis}(k,i) = \mathrm{dis}(i,k)$;
3) $\mathrm{dis}(k,i) \leq \mathrm{dis}(k,j) + \mathrm{dis}(j,i)$;

for all $k,i,j \in \mathbb{G}$. As discussed in Section III, a graph can be associated to problem (2). In analysis of problem (2), notation $\mathrm{dis}(k,i)$ represents the spatial distance between nodes k and i and it is invariant under node relabeling.

Definition 5: Let $\rho : [0,\infty) \to [0,\infty)$ be a continuous concave function such that $\rho(0) = 0$. A coupling characteristic function χ is of the following form:

$$\chi(x) = e^{\rho(x)}$$

and it satisfies $\chi(0) = 1$ and $\chi(x+y) \leq \chi(x)\chi(y)$. A coupling characteristic function $\chi(x)$ is said to satisfy the GRS-condition (Gelfand–Raikov–Shilov condition) if

$$\lim_{n\to\infty} \chi(nx)^{\frac{1}{n}} = 1 \qquad \text{for all} \;\; x \in \mathbb{R}^+$$

or equivalently, we have

$$\lim_{x\to\infty} \frac{\rho(x)}{x} = 0.$$

We refer to [28] for a discussion on coupling characteristic functions. Typical examples of such coupling characteristic functions are polynomial functions $\chi_\tau(x) = (1 + |x|)^\tau$, sub-exponential functions of the form $\chi_\tau(x) = e^{\tau|x|^\beta}$, and logarithmic functions of the form $\chi_\tau(x) = \left(\log(e + |x|)\right)^\tau$ for all $\tau > 0$ and $0 < \beta < 1$ (see [29] for more details). The definition of a coupling characteristic function enables us to measure the decay rate of the coupling strength between nodes in a coupled network of subsystems as distance increases.

For a given spatial domain $\mathbb{G}$ and a parameterized coupling characteristic function, we define the class of infinite-dimensional matrices defined by the off-diagonal decay of their entries. The decay can be quantified by a coupling characteristic function.

Definition 6: We define the set of spatially-decaying (SD) matrices with decay margin $\tau > 0$ to be the subspace $\mathcal{S}_\tau^\infty$ of all matrices such that $|||A|||_\tau < \infty$ where

$$|||A|||_\tau := \max\left\{\sup_{k\in\mathbb{G}}\sum_{i\in\mathbb{G}} |A_{ki}|\chi_\tau(\mathrm{dis}(k,i)),\right.$$
$$\left. \sup_{i\in\mathbb{G}}\sum_{k\in\mathbb{G}} |A_{ki}|\chi_\tau(\mathrm{dis}(k,i))\right\}.$$

Intuitively, a matrix is SD, if the size of each block decays faster than inverse of a coupling characteristic function. It can be shown that $\mathcal{S}_\tau^\infty$ is a Banach algebra of bounded operators acting on ℓ_2 [29].

In Section VI, the closure properties of algebra $\mathcal{S}_\tau^\infty$ is utilized to show the structural properties of problem (2). In the sequel, it is shown that under a reasonable assumption, namely invertibility on $B(\ell_2)$ and a weak growth condition, the Banach algebra $\mathcal{S}_\tau^\infty$ is closed under inversion as well. As we will see in Section VI, the closure under inversion property of $\mathcal{S}_\tau^\infty$ plays a central role in proving the spatial locality features of the optimal solution of problem (2).

Remark 1: In this paper, we present our results for a countable index set $\mathbb{G}$ and an arbitrary distance function $\mathrm{dis}(.,.)$ on $\mathbb{G}$. The proof of Theorem 2 is mainly based on [29, Theor. 6]. The result of this theorem is expressed for index set $\mathbb{Z}^d$ and distance function $\mathrm{dis}(i,j) = |i - j|$. In the following discussion which

is adopted from [29, Remark 2], we briefly show that the proof of [6, Theor. 6] carries over to arbitrary countable index sets $\mathbb{G}$ endowed with a non-trivial metric $\mathrm{dis}(.,.)$. The key modification in the proof of Theorem 6 is when we apply Barnes's Lemma [29, Lemma 5]. In the proof of Theorem 6, we have to assume that the volume of balls $B(i, r) = \{j \in \mathbb{G} | \mathrm{dis}(i, j) \le r\}$ grows polynomially in the radius r and independently of $i \in \mathbb{G}$. Therefore, in the rest of the paper we assume that the distance function $\mathrm{dis}(.,.)$ satisfies the above condition. For more details, we refer to [29, Remark 2].

B. Closure Under Inversion

The study of optimal solutions of problem (2) involves solving linear equations of the following form:

$$Ax = y \tag{16}$$

where A is an invertible SD matrix, $x \in \ell_2$ is the unknown variable, and $y \in \ell_2$ is given.

In the following, it is shown that under some mild assumptions if $A^{-1} \in \mathscr{B}(\ell_2)$, then $A^{-1} \in \mathcal{S}_\tau^\infty$. Before stating the main results of this section, we recall a motivating result about translation invariant operators [38], [39]. As shown in [28], every translation invariant operator on $\ell_2(\mathbb{Z})$, which its Fourier transform has analytic continuation to an annulus with nonzero radius around the unit circle, belongs to $\mathcal{S}_\tau^\infty$ for some $\tau > 0$ that relates to the radius of the annulus. We begin by introducing the unit translation operator to the right with respect to the group operation "$+$" as follows:

$$\mathbf{T}u = \mathbf{T}(\ldots, |u_i, u_{i+1}, \ldots) = (\ldots, |u_{i-1}, u_i, \ldots).$$

Note that for this case $\mathbb{G} = \mathbb{Z}$. For a translation invariant operator $A \in \mathscr{B}(\ell_2(\mathbb{Z}))$ which is defined as

$$A = \sum_{k \in \mathbb{Z}} A_k \mathbf{T}^k$$

the discrete Fourier transform is defined by

$$\hat{A}(z) = \sum_{k \in \mathbb{Z}} A_k z^{-k}.$$

Theorem 1: Let A be a traslation invariant operator with discrete Fourier transform $\hat{A}(z)$. If condition $\det(\hat{A}(z)) \ne 0$ holds for all $z \in \mathbb{S}^1$, then A is invertible and the inverse operator can be represented as

$$A^{-1} = \sum_{k \in \mathbb{Z}} [A^{-1}]_k \mathbf{T}^k. \tag{17}$$

Furthermore, if $\hat{A}(z)^{-1}$ has analytic continuation to some annulus

$$\Omega = \left\{ z \in \mathbb{C} : e^{-\tau} < z < e^\tau, \tau > 0 \right\}$$

then the norm of matrix coefficients of operator A^{-1} decay exponentially in the spatial domain, that is, for all $k \in \mathbb{Z}$

$$\|[A^{-1}]_k\| \le C e^{-\alpha|k|} \tag{18}$$

for some $C > 0$ and $0 < \alpha < \tau$.

Proof: The proof follows from [28, Theorem 3]. ∎

The following theorem extends the above result to SD matrices and gives a similar decay result for the inverse matrix. The following result is a consequence of [29, Theorem 6].

Theorem 2: Suppose that $A \in \mathcal{S}_\tau^\infty$ where the coupling characteristic function satisfies the weak growth condition

$$\chi_\tau(x) \ge C(1 + |x|)^\delta$$

for some $\delta > 0$. If A has an algebraic inverse on $\mathscr{B}(\ell_2)$ then $A^{-1} \in \mathcal{S}_\tau^\infty$. Moreover, If A is a positive definite matrix on ℓ_2, then also the matrices corresponding to A^α for $\alpha \in \mathbb{R}$ are in $\mathcal{S}_\tau^\infty$.

Proof: It suffices to prove the theorem for symmetric matrix A. The reason is that for an invertible matrix the following relation holds:

$$A^{-1} = A^*(AA^*)^{-1}.$$

Therefore, without loss of generality we may assume that A is symmetric. According to [28, Theorem 6], the weak growth condition implies that

$$\sigma_{\mathcal{S}_\tau^\infty}(A) = \sigma_{\ell_2}(A)$$

where $\sigma_{\mathcal{S}_\tau^\infty}$ and σ_{ℓ_2} denote the spectrum of an operator acting on $\mathcal{S}_\tau^\infty$ and ℓ_2, respectively. From this, it follows that $0 \notin \sigma_{\mathcal{S}_\tau^\infty}(A)$. Thus, $A^{-1} \in \mathcal{S}_\tau^\infty$. The result of the second part of the theorem follows from Wiener-Levy theorem (see [29] and [40] for more details). ∎

Remark 2: It is important to emphasize that Theorem 2 is false when the coupling characteristic function does not satisfy the GRS-condition. For example, if the coupling characteristic function is an exponential function, i.e.

$$\chi_\tau(x) = e^{\tau|x|}$$

for some $\tau > 0$, then Theorem 2 is not valid anymore. In the following, we provide an explicit example. Consider a translation invariant operator A whose discrete Fourier transform is given by:

$$\hat{A}(\omega) = 3 - e^{i\omega} - e^{-i\omega}.$$

The matrix of A has entries $A_{ki} = 3$ if $k = i$, $A_{ki} = -1$ if $|k - i| = 1$, and $A_{ki} = 0$ otherwise. Therefore, A is a banded matrix and

$$\||A|\|_\tau < \infty$$

for all $\tau > 0$. It is easy to see that $\hat{A}(\omega) \ne 0$ for all ω. Thus, A is invertible on $\ell_2(\mathbb{Z})$. The Fourier transform of the inverse operator can be explicitly computed as follows:

$$\widehat{A^{-1}}(\omega) = \sum_{k=-\infty}^\infty \frac{1}{\sqrt{5}} e^{-\delta|k|} e^{-ik\omega}$$

where $\delta = -\ln(3 - \sqrt{5}/2)$. Now, if the coupling characteristic function is $\chi_\tau(x) = e^{\tau|x|}$, by choosing $\delta < \tau$ we have that

$$\||A^{-1}|\|_\tau = \frac{1}{\sqrt{5}} \sum_{k=-\infty}^\infty e^{-\delta|k|} e^{\tau|k|} = \infty.$$

This example shows that $A \in \mathcal{S}_\tau^\infty$ for all $\tau > 0$, but $A^{-1} \notin \mathcal{S}_\tau^\infty$ for all $\tau > 0$. Therefore, the GRS-condition is critical for Theorem 2 to be valid.

In the venue of the above results, from (16) it follows that:

$$x = A^{-1} y$$

and for any given $k \in \mathbb{G}$, we have

$$x_k = \sum_{i \in \mathbb{G}} [A^{-1}]_{ki} y_i.$$

According to Theorem 2, we get

$$\left\| [A^{-1}]_{ki} \right\| \le C \chi_\tau(\mathrm{dis}(k, i))^{-1}.$$

We will apply this result in Section VI to analyze the structural properties of the optimal solution of the infinite-dimensional quadratic programming (2).

Remark 3: Note that if $X, Y \in \mathcal{S}_\tau^\infty$, then

$$\left\| \left\| \left\| \begin{bmatrix} X \\ Y \end{bmatrix} \right\| \right\| \right\|_\tau = \max \left\{ \sup_{k \in \mathbb{G}} \sum_{i \in \mathbb{G}} \left\| \begin{bmatrix} X_{ki} \\ Y_{ki} \end{bmatrix} \right\| \chi_\tau(\mathrm{dis}(k, i)), \right.$$

$$\left. \sup_{i \in \mathbb{G}} \sum_{k \in \mathbb{G}} \left\| \begin{bmatrix} X_{ki} \\ Y_{ki} \end{bmatrix} \right\| \chi_\tau(\mathrm{dis}(k, i)) \right\} < \infty.$$

Similarly, one can show that

$$\left\| \left\| \left\| \begin{bmatrix} X & Y \end{bmatrix} \right\| \right\| \right\|_\tau < \infty.$$

Using this property, one can conclude that if $A, B, G, H \in \mathcal{S}_\tau^\infty$ in (6), then $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{G}, \mathbf{H} \in \mathcal{S}_\tau^\infty$. One can also show that if $A, Q \in \mathcal{S}_\tau^\infty$, then P the unique positive definite solution of (7) is in $\mathcal{S}_\tau^\infty$ (cf. , [41]). Therefore, it follows that $\mathbf{Q} \in \mathcal{S}_\tau^\infty$, and if $R \in \mathcal{S}_\tau^\infty$, also $\mathbf{R} \in \mathcal{S}_\tau^\infty$. Therefore, it follows that $\mathbf{L}_0$, $\mathbf{L}_1, \mathbf{M}, \mathbf{E} \in \mathcal{S}_\tau^\infty$. From positive definiteness of matrix R, it also follows that $\mathbf{L}_0 > 0$. Therefore, problem (13) is a convex quadratic programming problem, and that it can be written in the standard form (2).

VI. ANALYSIS OF INFINITE-DIMENSIONAL MPQP

In this section, we study the structural properties of the optimal solution of a large-scale MPQP problem

$$\inf_{x \in \Omega} \frac{1}{2} \langle x, Qx \rangle + \langle c, x \rangle \text{ subject to: } Gx \preceq b \quad (19)$$

where parameters are the components of vectors $b, c \in \ell_2$. We assume that Q is a self-adjoint and positive definite matrix, $Q, G \in \mathscr{B}(\ell_2)$, and Ω a convex subset of ℓ_2. Assume that there exists a set $\mathbb{P} \subset \ell_2$ of parameters for which for every $b \in \mathbb{P}$ an optimal solution of (19) exists. When problem (19) is considered on a finite-dimensional vector space, it can be shown that $\mathbb{P}$ can be partitioned into countably many partitions (cf. [27] and references in there) over each of which the optimal solution of (19) is an affine function of the parameters. In the following, we will show that a similar explicit representation exists for the optimal solution when problem (19) is treated in an infinite-dimensional vector space. In this scenario, however, the affine representation is in the form of a convolution sum with some matrix gains appearing as the kernel. We will prove that convolution

kernel corresponding to the optimal solution on each parameter set, have decay in the spatial domain at a rate proportional to the inverse of the corresponding coupling characteristic function. We recall the following theorem which is Theorem 1 in chapter 8 of [2].

Theorem 3: Let X be a vector space, Z a normed space, Ω a convex subset of X, and P the positive cone in Z. Assume that P contains an interior point. Let f be a real-valued convex functional on Ω and G a convex mapping from Ω into Z. Assume the existence of a point $x_1 \in \Omega$ for which $G(x_1) \prec 0$ (i.e., $G(x_1)$ is an interior point of $N = -P$). Let

$$\mu_0 = \inf_{x \in \Omega} f(x) \text{ subject to : } G(x) \preceq 0 \quad (20)$$

and assume μ_0 is finite. Then

$$\mu_0 = \max_{\lambda^* \ge 0} \inf_{x \in \Omega} \{ f(x) + \langle G(x), \lambda^* \rangle \} \quad (21)$$

and the maximum is achieved by an element $\lambda_0^* \ge 0$ in Z^*, where the inequality is defined with respect to the positive cone P^*. Furthermore, if the infimum is achieved in (20) by an $x_0 \in \Omega$, then

$$\langle G(x_0), \lambda_0^* \rangle = 0 \quad (22)$$

and x_0 minimizes $f(x) + \langle G(x), \lambda_0^* \rangle$ with $x \in \Omega$.

The above result provides a precise method to explain primal-dual relationship, and that formulate the necessary optimality conditions for a convex infinite-dimensional (clearly, finite-dimensional as well) optimization problem. The outcome of this theorem can be thought of as the generalized Karush–Kuhn–Tucker conditions in infinite-dimensions. In the following, as we will see the necessary optimality conditions for problem (19) are very similar to its finite-dimensional counterpart. In the following theorem, we assume that the existence assumptions of Theorem 3 hold for (19).

Theorem 4: In problem (19), we assume that $Q \in \mathcal{S}_\tau^\infty$ is a self-adjoint and positive definite, $G \in \mathcal{S}_\tau^\infty$, and $b, c \in \ell_2$. We assume that the assumptions of Theorem 2 hold and a strictly feasible point x_1 exists for problem (19). We suppose that at the optimal solution some combination of constraints in (19) are active and the corresponding rows to these active constraints from matrix G form an onto matrix $\tilde{G}$. Let $\mathbb{B} \subseteq \mathbb{P}$ be the set of all $b \in \mathbb{P}$ so that such combinations are active at the optimal solution. Then the optimal solution of (19), as well as the corresponding Lagrange multipliers to index $k \in \mathbb{G}$, are

(a) *affine* maps of b over $\mathbb{B}$, especially

$$\bar{x}_k = \sum_{i \in \mathbb{G}} K_{ki} b_i + \sum_{i \in \mathbb{G}} K'_{ki} c_i \quad (23)$$

for some linear bounded matrices K and K'.

(b) *spatially distributed*, in the sense that the coupling decays in the spatial domain at a rate proportional to the inverse of the corresponding coupling characteristic function, i.e.

$$|K_{ki}| \le \frac{\kappa}{\chi_\tau(\mathrm{dis}(k, i))} \text{ and } |K'_{ki}| \le \frac{\kappa'}{\chi_\tau(\mathrm{dis}(k, i))}$$

for some $\kappa, \kappa' > 0$.

Proof: We may assume that $\mathbb{P}$ is nonempty. Let fix $b \in \mathbb{B} \subseteq \mathbb{P}$. Since the optimal value is less than $f(x_1)$ with $f(x) = 1/2\langle x, Qx \rangle + \langle c, x \rangle$, it follows that $\mu_0 < \infty$ where:

$$\mu_0 = \inf_{x \in \Omega} f(x) \text{ subject to : } Gx \preceq b.$$

The minimum is achieved since one can reduce the problem to minimizing a continuous function over a compact set. Let denote the minimum by $\bar{x}$. The quadratic cost functional in (19) is Fréchet differentiable [2]. This allows us to define the gradient of the quadratic cost functional. Theorem 3 implies that the gradient of the Lagrangian calculated at the optimal λ must vanish. Condition (22) in Theorem 3 can be written as

$$\langle G(\bar{x}), \lambda \rangle = \sum_{i \in \mathbb{G}} \lambda_i \left(\sum_{k \in \mathbb{G}} G_{ik}\bar{x}_k - b_i \right) = 0. \quad (24)$$

Every term inside the summations in (24) is nonpositive. This follows from the fact that at optimum we have:

$$G(\bar{x}) = G\bar{x} - b \preceq 0 \text{ and } \lambda \succeq 0. \quad (25)$$

Therefore, each term has to be zero. Note that the positive cone in ℓ_2 is defined by

$$\mathscr{P} = \{x \in \ell_2 : x = (x_i)_{i \in \mathbb{G}}, x_i \geq 0 \text{ for all } i \in \mathbb{G}\}$$

and by definition $\mathscr{P}^{\oplus} = \mathscr{P}$. Therefore, by applying Theorem 3 to (19) we have the following conditions:

$$Q\bar{x} + c + G^*\lambda = 0 \quad (26)$$

$$\lambda_i \left(\sum_{k \in \mathbb{G}} G_{ik}\bar{x}_k - b_i \right) = 0 \quad (27)$$

$$\lambda_i \geq 0 \quad (28)$$

$$G\bar{x} \preceq b$$

$$i \in \mathbb{G} \quad (29)$$

where $\lambda = (\lambda_i)_{i \in \mathbb{G}} \in \ell_2$ is the corresponding Lagrange multipliers. Condition (27) is so called complementary slackness. We refer the reader to [31] (for example, see Chapter 7 and 8) for more details on more general cases of this problem.

Since Q is bounded and positive definite, it has an algebraic inverse on $\mathscr{B}(\ell_2)$. Equation (26) results in

$$\bar{x} = -Q^{-1}(G^*\lambda + c). \quad (30)$$

According to (27) and (28), all Lagrange multipliers λ_i corresponding to inactive constraints must be zero, and the Lagrange multipliers corresponding to active constraints are nonnegative numbers. We denote the index set corresponding to the active constraints by $\mathbb{A}$ and the corresponding Lagrange multipliers by $\tilde{\lambda} = (\lambda_i)_{i \in \mathbb{A}}$. Therefore, we can form linear matrix $\tilde{G} : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{A})$ whose entries $\tilde{G}_{ik}$ are obtained by deleting rows corresponding to the inactive constraints from G. We may equivalently represent this operation as $\tilde{G} = EG$ where E is a bounded matrix that is obtained by deleting rows from I (identity matrix) which correspond to the inactive constraints. One

can see that E is SD. This is because distance function $\text{dis}(i,j)$ corresponds to the actual spatial distance between nodes i and j on the underlying graph of the system and is invariant under relabeling the nodes. From (27) and our assumptions, for every $b \in \mathbb{B}$ we have the following equation:

$$\tilde{G}\bar{x} - \tilde{b} = 0 \quad (31)$$

in which $\tilde{b} = (\tilde{b}_i)_{i \in \mathbb{A}}$ or, equivalently, $\tilde{b} = Eb$. This equation allows us to solve it along with (30) for $\tilde{\lambda}$. Using (30), it follows that:

$$\bar{x} = -Q^{-1}\tilde{G}^*\tilde{\lambda} - Q^{-1}c. \quad (32)$$

Substituting (32) into (31), results in

$$\tilde{G}Q^{-1}\tilde{G}^*\tilde{\lambda} = -\tilde{b} - \tilde{G}Q^{-1}c.$$

Matrix $\tilde{G}$ is onto and $\tilde{G}^*$ is $(1-1)$. Thus, matrix $\tilde{G}Q^{-1}\tilde{G}^*$ is invertible and positive definite, and that it has an algebraic inverse on $\mathscr{B}(\ell_2(\mathbb{A}))$. Therefore, we have

$$\tilde{\lambda} = -\left(\tilde{G}Q^{-1}\tilde{G}^*\right)^{-1}\tilde{b} - \left(\tilde{G}Q^{-1}\tilde{G}^*\right)^{-1}\tilde{G}Q^{-1}c \quad (33)$$

and

$$\bar{x} = Kb + K'c \quad (34)$$

where

$$K \triangleq Q^{-1}\tilde{G}^*\left(\tilde{G}Q^{-1}\tilde{G}^*\right)^{-1}E$$

$$K' \triangleq Q^{-1}\tilde{G}^*\left(\tilde{G}Q^{-1}\tilde{G}^*\right)^{-1}\tilde{G}Q^{-1} - Q^{-1}.$$

Equation (34) can be written in the form of convolution sums as follows:

$$\bar{x}_k = \sum_{i \in \mathbb{G}} K_{ki}b_i + \sum_{i \in \mathbb{G}} K'_{ki}c_i$$

for all $k \in \mathbb{G}$. This proves part (a) of the theorem that the optimal solution $\bar{x}$ is an affine function of parameters b_i for all $i \in \mathbb{G}$.

Since $G \in \mathcal{S}_\tau^\infty$, $\tilde{G}$ is also SD. According to Theorem 2, $Q^{-1} \in \mathcal{S}_\tau^\infty$. It follows that $\tilde{G}Q^{-1}\tilde{G}^*$ is also SD with respect to coupling characteristic function χ_τ as an operator acting on $\ell_2(\mathbb{A})$. For simplicity of notation, we also use $\mathcal{S}_\tau^\infty$ to refer to the Banach algebra of SD operators acting on $\ell_2(\mathbb{A})$. Applying Theorem 2 one more time, it results in

$$\left(\tilde{G}Q^{-1}\tilde{G}^*\right)^{-1} \in \mathcal{S}_\tau^\infty.$$

Therefore, using the closedness under multiplication property of Banach algebra $\mathcal{S}_\tau^\infty$, it concludes that the gain matrix K is SD. Thus, it follows that:

$$|K_{ki}| \leq \frac{\kappa}{\chi_\alpha(\text{dis}(k,i))}$$

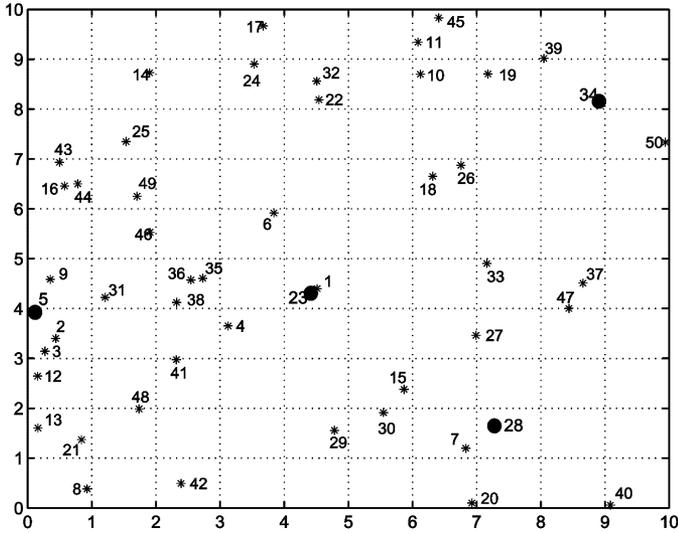for some $\kappa > 0$. It is straightforward to show that a similar result holds for matrix K', i.e.

Fig. 2. $N = 50$ nodes are randomly and uniformly distributed in a region of area $10 \times 10 \, (\text{units})^2$. A scalar variable x_k is associated to each node. The node variables are coupled through a central quadratic cost function and some linear constraints by using a given coupling characteristic function.

$$|K'_{ki}| \le \frac{\kappa'}{\chi_\alpha(\text{dis}(k,i))}$$

for some $\kappa' > 0$. ∎

The result of theorem 4 can be used to characterize the parameter set $\mathbb{B}$. Assume that $\mathbb{B}$ is nonempty. The optimal solution has to satisfy constraint (29) and by (28) the Lagrange multipliers (33) must be nonnegative. Therefore, the parameter set $\mathbb{B}$ can be represented as

$$\mathbb{B} = \{b \in \mathbb{P} : B_0 b + d_0 \preceq 0, B_1 b + d_1 \preceq 0\}$$

where

$$B_0 \triangleq GK - I, \quad d_0 \triangleq -GK'c$$
$$B_1 \triangleq \left(\tilde{G} Q^{-1} \tilde{G}^*\right)^{-1} E, \quad d_1 \triangleq \left(\tilde{G} Q^{-1} \tilde{G}^*\right)^{-1} \tilde{G} Q^{-1} c.$$

Note that matrices B_0 and B_1 are bounded.

The results of Theorem 4 are valid for infinite-dimensional MPQP problems. In Section VII, simulation results suggest that a finite-dimensional version of Theorem 4 may also hold. The finite-dimensional version of Theorem 4 is still an open problem.

VII. SIMULATION

In this section, we consider a large-scale quadratic programming problem as follows:

$$\min_x \quad \frac{1}{2} x^T Q x + c^T x$$
$$\text{subject to :} \quad G x \preceq b \quad (35)$$

where $Q, G \in \mathbb{R}^{N \times N}$ and $c, b \in \mathbb{R}^N$. We assume that $N = 50$ nodes are randomly and uniformly distributed in a region of area $10 \times 10 \, (\text{units})^2$ (see Fig. 2). A scalar variable x_k is associated to each node for all $k = 1, \ldots, N$. The node variables are coupled to each other through a common quadratic cost function and linear constraints by using a given coupling characteristic function χ_τ.

The weight matrices Q, G are defined as follows:

$$Q_{ki} = \frac{1}{\chi_\tau(\text{dis}(k,i))} \quad (36)$$

and

$$G_{ki} = \frac{(-1)^{|k-i|}}{\chi_\tau(\text{dis}(k,i))} \quad (37)$$

for all $k, i = 1, \ldots, N$ and $b = c = [1 \; 1 \; \ldots \; 1]^T$. The distance function is Euclidean. According to Theorem 4, the optimal solution corresponding to node k can be written as an affine function of parameters b and c

$$\bar{x}_k = \sum_{i=1}^N K_{ki} b_i + \sum_{i=1}^N K'_{ki} c_i.$$

Simulations are performed in MATLAB using the optimization toolbox. First, the optimal solution of problem (35) is computed numerically, and then using the active/inactive constraints technique (illustrated in the proof of Theorem 4) the gain matrices K and K' are derived.

In the sequel, two different scenarios are considered for the coupling characteristic function. In the first simulation, we assume that the coupling characteristic function is exponential function $\chi_\tau(x) = e^{\tau |x|^\beta}$ where $\tau = 1$ and $\beta = 0.7$. In the next simulation, we assume that the coupling characteristic function is algebraic $\chi_\tau(x) = (1 + \lambda x)^\tau$ where $\lambda = 1$ and $\tau = 2$. In both cases, computations show that $K = \mathbf{0}$.

In Figs. 3 and 4, the norm of coupling matrix K'_{ki} corresponding to agents $k = 5, 23, 28, 34$ (their locations are marked by bold stars in Fig. 2) is depicted versus the distance of node k from the other nodes. As seen from these simulations, for node k the norm of the coupling matrix K'_{ki} is approximately enveloped with function $\|K'_{kk}\|/\chi_\tau(\text{dis}(k,i))$. Therefore, the spatial decay rate of the optimal solution can be determined priory using the information of the coupling characteristic function $\chi_\tau(\text{dis}(k,i))$. As seen in Figs. 3 and 4, for each subsystem k, the corresponding optimal solution is effectively coupled only to those subsystems with distance closer than 3 (units). This suggests the possibility of formulating the optimal control problem in a distributed fashion, rather than solving a centralized high-dimension optimization problem such as (35).

VIII. CONCLUSION

We introduced the class of spatially decaying (SD) matrices and showed that the space of spatially decaying matrices forms a Banach algebra. It was shown that under the GRS-condition and a weak growth condition on the coupling characteristic function, the Banach algebra of SD matrices is also closed under inversion operation. Also, we formulated model predictive control problem for spatially distributed linear systems as quadratic programming. We employed the developed framework to analyze spatial structure of large-scale quadratic programming problems. By applying duality theory and complementary slackness conditions, we proved that the optimal solution of a multi-parametric quadratic programming is piecewise affine which can be represented as convolution sums. Furthermore, we proved that the kernel of each convolution sum decays in the spatial domain. It is important to stress that this locality feature is inherent, therefore, verifies the feasibility of spatial truncation in the optimal solution without too much loss of performance. The future research relies on developing computational tools and algorithms that exploit the spatial structure of SD matrices.
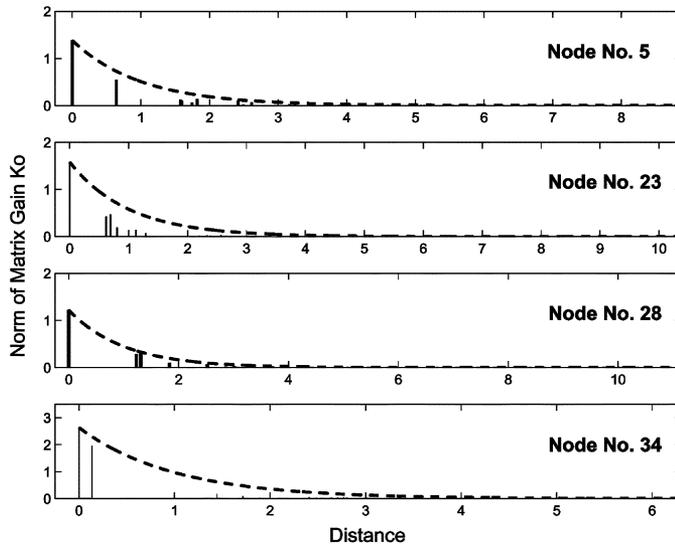
Fig. 3. Exponentially decaying coupling characteristic function: norm of matrix main $\|K'_{ki}\|$ (bar) and $\|K'_{kk}\|/\chi_\tau(\mathrm{dis}(k,i))$ when $\tau = 1$ and $\beta = 0.7$ (dashed) for nodes $k = 5,23,28,34$, respectively, from top to bottom.

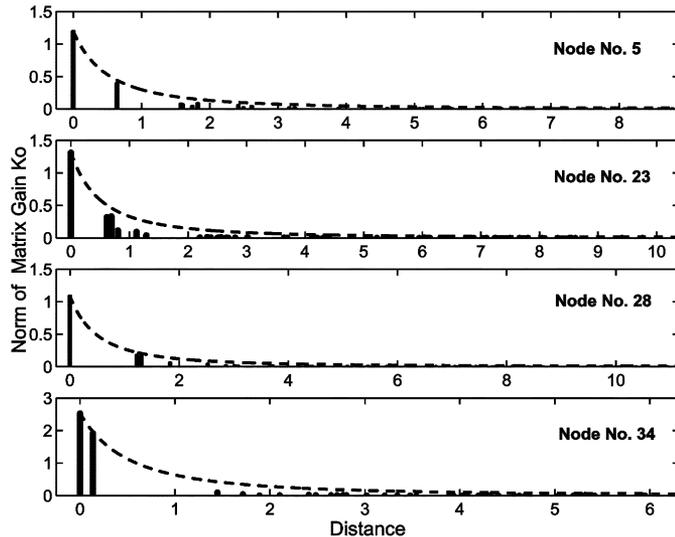Fig. 4. Algebraically decaying coupling characteristic function: norm of matrix main $\|K'_{ki}\|$ (bar) and $\|K'_{kk}\|/\chi_\tau(\mathrm{dis}(k,i))$ when $\tau = 1$(dashed) for nodes $k = 5,23,28,34$, respectively, from top to bottom.

REFERENCES

[1] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation*. Boston, MA: Athena Scientific, 1988.

[2] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[3] E. J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces*. New York: Wiley, 1991.

[4] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Automat. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[5] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Automat. Control*, vol. AC-31, no. 9, pp. 803–812, Sep. 1986.

[6] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Trans. Robot. Automat.*, vol. 20, no. 2, pp. 243–255, Feb. 2004.

[7] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[8] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Tran. Automat. Control*, vol. 51, no. 2, pp. 274–286, Feb. 2006.

[9] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed control of spatially invariant systems," *IEEE Trans. Automat. Control*, vol. 47, no. 7, pp. 1091–1107, Jul. 2002.

[10] G. Dullerud and R. D'Andrea, "Distributed control of heterogeneous systems," *IEEE Trans. Automat. Control*, vol. 49, no. 12, pp. 2113–2128, Dec. 2004.

[11] D. Q. Mayne, J. B. Rawlings, C. Rao, and P. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[12] W. B. Dunbar, "Distributed Receding Horizon Control With Application to Multi-Vehicle Formation Stabilization," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, 2004.

[13] T. Keviczky, F. Borrelli, and G. Balas, "Stability analysis of decentralized rhc for decoupled systems," in *Proc. 44th IEEE Conf. Decision Control, Eur. Control Conf. (CDC-ECC'05)*, 2005, pp. 1689–1694.

[14] A. Richards and J. How, "A decentralized algorithm for robust constrained model predictive control," in *Proc. Amer. Control Conf.*, Boston, MA, 2004, pp. 4261–4266.

[15] L. Acar, "Some examples for the decentralized receding horizon control," in *Proc. 31st IEEE Conf. Decision Control*, 1992, pp. 1356–1359.

[16] S. Oschs, S. Engell, and A. Draeger, "Decentralized vs. model predictive control of an industrial glass tube manufacturing process," in *Proc. IEEE Int. Conf. Control Appl.*, Trieste, Italy, Sep. 1998, pp. 16–20.

[17] Y. Guo, D. Hill, and Y. Wang, Nonlinear Decentralized Control of Large-Scale Power Systems Electrical and Information Engineering School, The University of Sydney, Sydney, Australia, Tech. Rep. EE-98020, 2006.

[18] E. Camponogara, D. Jia, B. Krogh, and S. Talukdar, "Distributed model predictive control," *IEEE Control Systems Magazine*, vol. 22, no. 1, pp. 44–52, Feb. 2002.

[19] P. Hines, D. Jia, and S. Talukdar, "Distributed model predictive control for electric grids," in *Proc. Carnegie Mellon Transmission Conf.*, Pittsburgh, PA, 2004, [CD ROM].

[20] C. Lin, "Towards Optimal Strain Actuated Aerolastic Control," Ph.D. dissertation, M.I.T, Cambridge, MA, 1996.

[21] M. Jamoom, E. Feron, and M. McConley, "Optimal distributed actuator control grouping schemes," in *Proc. 37th IEEE Conf. Decision Control*, 1998, pp. 1900–1905.

[22] N. Motee and B. Sayyar-Rodsari, "Optimal partitioning in distributed model predictive control," in *Proc. Amer. Control Conf.*, Denver, CO, 2003, pp. 5300–5305.

[23] W. B. Dunbar and R. M. Murray, "Distributed receding horizon control with application to multi-vehicle formation stabilization," *Automatica*, vol. 2, no. 4, pp. 549–558.

[24] D. Jia and B. H. Krogh, "Min-max feedback model predictive control for distributed control with communication," in *Proc. Amer. Control Conf.*, 2002, pp. 4507–4512.

[25] L. Acar, "Boundaries of the receding horizon control for interconnected systems," *J. Optim. Theory Appl.*, vol. 84, no. 2, pp. 251–271, 1995.

[26] A. Bemporad, F. Borrelli, and M. Morari, "Model predictive control based on linear programming—The explicit solution," *IEEE Trans. Automat. Control*, vol. 47, no. 12, pp. 1974–1985, Dec. 2002.

[27] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.

[28] N. Motee and A. Jadbabaie, "Optimal control of spatially distributed systems," *IEEE Trans. Automat. Control*, vol. 53, no. 7, pp. 1616–1629, Aug. 2008.

[29] K. Gröchenig and M. Leinert, "Symmetry and inverse-closedness of matrix algebras and symbolic calculus for infinite matrices," *Trans. Amer. Math. Soc.*, vol. 358, no. 6, pp. 2695–2711, 2006.

[30] P. Lax, *Functional Analysis*. New York: Wiley, 2002.

[31] K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications (Advances in Design and Control)*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2008.

[32] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization With PDE Constraints (Mathematical Modelling: Theory and Applications)*. New York: Springer, 2008.

[33] R. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*. New York: Springer-Verlag, 1995.

[34] J. B. Rawlings and K. R. Muske, "The stability of constrained receding horizon control," *IEEE Trans. Automat. Control*, vol. 38, no. 10, pp. 1512–1516, Oct. 1993.

[35] M. Seron, J. DeDona, and G. Goodwin, "Global analytic modelpredictive control with input constraints," in *Proc. IEEE Conf. Decision Control*, Sydney, Australia, 2000, pp. 154–159.

[36] D. Chmielewski and V. Manousiouthakis, "On constrained infinite-time linear quadratic optimal control," *Syst. Control Lett.*, vol. 29, pp. 121–129, 1996.

[37] P. Scokaert and J. Rawlings, "Constrained linear quadratic regulation," *IEEE Trans. Automat. Control*, vol. 43, no. 8, pp. 1163–1169, Aug. 1999.

[38] D. R. I. Gel'fand and G. Shilov, *Commutative Normed Rings*. New York: Chelsea Publishing Co, 1964.

[39] W. Rudin, *Fourier Analysis on Groups*. London, U.K.: Interscience, 1962.

[40] Q. Sun, "Wiener's lemma for infinite matrices," *Trans. Amer. Math. Soc.*, vol. 359, pp. 3099–3123, 2007.

[41] N. Motee and A. Jadbabaie, "Reply to 'Comments on "On optimal control of spatially distributed systems"'," *IEEE Trans. Automat. Control*, vol. 38, submitted for publication.

Nader Motee (S'99–M'08) received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2000, the M.Sc. degree in mechanical engineering from the Louisiana State University, Baton Rouge, in 2003, and the M.Sc. and Ph.D. degrees in electrical and systems engineering from the University of Pennsylvania, Philadelphia, in 2006 and 2007, respectively.

He is currently a Postdoctoral Scholar at the Control and Dynamical Systems Department, California Institute of Technology (Caltech), Pasadena. From January to November 2008, he was a Postdoctoral Researcher at the General Robotics, Automation, Sensing, and Perception (GRASP) Laboratory, University of Pennsylvania. His current research interests include the areas of network science, distributed control systems, distributed optimization, and nonlinear control systems, and their applications to robotics and biology.

Dr. Motee received the 2008 O. Hugo Schuck Award for Theory of the American Automatic Control Council, the Student Best Paper Award at the American Control Conference in 2007, the Joseph and Rosaline Wolf Award for Best Ph.D. Dissertation in 2008, and was a finalist for the Student Best Paper Award at the American Control Conference in 2006, and the IEEE Region 8 Student Paper Contest in 2000.

Ali Jadbabaie (S'99–M'03–SM'07) received the B.S. degree (with high honors) from Sharif University of Technology, Tehran, Iran, in 1995, the M.S. degree in electrical and computer engineering from the University of New Mexico, Albuquerque, in 1997, and the Ph.D. degree in control and dynamical systems from the California Institute of Technology (Caltech), Pasadena, in 2001.

From July 2001 to July 2002, he was a Postdoctoral Associate at the Department of Electrical Engineering, Yale University, New Haven, CT. Since July 2002, he has been with department of Electrical and Systems Engineering and GRASP Laboratory at the University of Pennsylvania, where he is currently the Skirkanich Associate Professor of Innovation. His research interests are on network science, motion coordination of multi-agent systems, optimization and optimal control, as well as applications of algebraic topology and computational geometry in deployment and design of sensor networks.

Dr. Jadbabie received the NSF Career Award, the ONR Young Investigator Award, the Student Best Paper Award (as advisor) in the American Control Conference, the O Hugo Schuck Award, and the George S. Axelby Outstanding Paper Award of the IEEE Control Systems Society.