Distributed Control with Low-Rank Coordination

Madjidian, Daria; Mirkin, Leonid

2014

Distributed Control with Low-Rank Coordination

Daria Madjidian and Leonid Mirkin

***Abstract*— A common approach to distributed control design is to impose sparsity constraints on the controller structure. Such constraints, however, may greatly complicate the control design procedure. This paper promotes an alternative structure, which is not sparse, yet might nevertheless be well suited for distributed control purposes. The structure appears as the optimal solution to a class of coordination problems arising in multi-agent applications. The controller comprises a diagonal (decentralized) part, complemented by a rank-one coordination term. Although this term relies on information about all subsystems, its implementation only requires a simple averaging operation.**

I. Introduction

The ability to cope with complexity is one of the contemporary challenges of control engineering. Already an established research area by the late 1970s [1]–[3], control of complex systems reinvigorated during the last decade, impelled by recent technological progress, networking and integration trends, efficiency demands, etc.

Complexity may be manifested through different attributes, one of which is the presence of a very large number of sensors and actuators. In such situations, fully centralized, structureless, information processing becomes infeasible. This motivates the quest for distributed control methods, with various constraints on information exchange between subsystems and information processing in the controller. Such structural constraints are conventionally expressed in terms of sparsity patterns [3]–[5], with nonzero elements corresponding to permitted coordination between subsystems. Sometimes, delay constraints on the communication between subsystems are considered [5].

Although a sparse structure can effectively limit the amount of information processing in the controller, it substantially complicates the analysis and design of control systems. Many well understood problems might turn acutely opaque when sparsity constraints on the controller are added [6]. Analysis is simplified if the plant happens to possess a compatible sparsity pattern (the quadratic invariance condition [5], [7]) or if additional constraints are imposed on the closed-loop behavior (like positivity [8], [9]). But even then, the computational burden grows rapidly with the problem dimension and, more importantly, structural properties of the resulting controller are rarely transparent. Revealing such properties proved to be a challenge even in seemingly simple problem formulations, see [10], [11] and the references therein.

This paper puts forward an alternative structure, composed of a block-diagonal term perturbed by a block-rank-one component. Although not sparse, this structure might be feasible in various distributed control applications. Indeed, the only non-sparse, centralized task that has to be performed by the controller is an averaging operation. This is a relatively simple numerical operation, which might be robust to sensor imperfections for large groups. It can be performed either locally, by each agent, or globally, by a coordinator.

This type of control structure has previously been suggested in [12]. There, a diagonal (decentralized) controller is complemented by a low-rank component in order to improve robustness with respect to a certain type of model uncertainties. The controller is obtained by an LMI-design procedure, which relies on a special parametrization to ensure that the controller is of the desired form. However, in order to guarantee a convex synthesis procedure, only some degrees of freedom in the parametrization are exploited. The remaining variables are treated as fixed parameters, meaning that the approach might not fully utilize the proposed control structure. Also, in the design procedure it is not evident to what extent the addition of the low-rank component improves performance for a given plant, or for which class of plants it is most beneficial.

Here, we arrive at the diagonal plus rank-one controller from a different direction. We identify a class of large scale coordination problems that happen to admit a solution of this form. Specifically, we consider a group of autonomous agents having identical dynamics and local criteria, with coordination requirements introduced through a global constraint. This formulation is motivated by certain control tasks arising in wind farms. We show that, without any structural constraints, the problem admits a solution of the proposed form. The solution has additional appealing properties, one of which is that the computational effort required to obtain the optimal control law is independent of the number of agents. Another is that the structure of the controller has a clear interpretation. The diagonal part merely comprises the local, uncoordinated control laws. The rank-one part is then responsible for coordination via fine-tuning the local controllers on the basis of measurements of an "average" agent.

The averaging of measured variables of individual subsystems may be viewed as a spatial counterpart of the generalized sampling operation [13]. This is in contrast to the decentralized structure, which may be thought of as a form of the ideal sampling, that ignores the intersample information.

This research was supported by the Swedish Research Council through the LCCC Linnaeus Center and by the European commission through the project AEOLUS.

D. Madjidian is with the Department of Automatic Control, Lund University, Box 118, SE–221 00 Lund, Sweden. E-mail: *daria@control.lth.se.*

L. Mirkin is with the Faculty of Mechanical Eng., Technion—IIT, Haifa 32000, Israel. E-mail: *mirkin@technion.ac.il.*

With this analogy in mind, we expect the proposed controller structure to be advantageous even in problems outside the class studied in this paper.

The paper is organized as follows. In Section II, we consider a coordination problem arising in wind farms. This problem serves a motivation for the theoretical developments in Sections III where we study LQR-problems with hard coordination constraints. This section also contains an extensive discussion on properties of the resulting controller configuration and the structure of the optimal cost. The developments are illustrated by numerical studies of the same wind farm coordination problem. Concluding remarks are provided in Section IV.

Notation: The transpose of a matrix M is denoted as M'. By e_i we understand the ith standard basis of an Euclidean space and by I_n—the $n \times n$ identity matrix (we drop the dimension subscript when the context is clear). The Kronecker product of two matrices is defined as

$$A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{p1}B & \cdots & a_{pm}B \end{bmatrix}.$$

The L^2 norm of a signal ξ is denoted as $\|\xi\|_2$.

II. Motivating Example: Coordination in Wind Farms

Wind energy is an increasingly active application area for control, see [14] and the references therein. Lately, the focus is shifting from control of a stand-alone wind turbine (WT) to coordinated control of networks of WTs, commonly known as wind power plants (WPP) or wind farms. In this section we consider a coordination problem arising in large-scale WPPs, which is used to motivate the problem studied in this paper.

A. Problem description

We consider the problem discussed in [15], [16], where a WPP is required to meet a certain power demand. To achieve this, the WTs need to coordinate their power production. Since there are multiple WTs in the farm, certain freedom exists in distributing the power demand among them. This freedom can be used to address local objectives of individual turbines, such as regulating rotor speed, reducing fatigue loads, preventing excessive pitch action, etc. Thus, instead of following a fixed portion of the power demand, a WT can be allowed to continuously adjust its power production in response to local wind speed fluctuations. Since wind speed fluctuations are not the same across the WPP, changes in power production that benefit one WT can be compensated for by changes at WTs with opposite needs.

For control design purposes, it is common practice to model a WT as a linear system around an operating point. It may also be natural to make two additional simplifying assumptions.

1) WTs in a WPP are often identical in their design. By assuming that they operate around the same mean wind speed and mean power production, the WTs may be considered to have equal dynamics.
2) Due to a large distance between individual turbines in WPPs, it is common to assume [17], [18] that wind speed variations experienced by them are uncorrelated.

With these observations in mind, below we address a coordination problem with a stripped-down[1] version of individual WTs models and performance indices studied in [18]. We consider a WPP comprising ν turbines, each operating around a mean wind speed of 10 m/s and a nominal power production of 2 MW. Each turbine is modeled as

$$\dot{x}_i = Ax_i + B_w w_i + B_u u_i, \quad i = 1, \dots, \nu$$

where $\left[\begin{array}{ccc} A & B_w & B_u \end{array}\right]$ take the following numerical values:

$$\left[\begin{array}{ccccc:c:c} 0 & 120 & -0.92 & 0 & 0 & 0 & 0 \\ 0.0084 & -0.032 & 0 & 0 & 0 & 0.12 & -0.021 \\ 0 & 150 & -1.6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0.021 & 0.054 & 0 & -4 & -0.32 & 0.2 & 0 \end{array}\right].$$

Here the state vector spells out as

$$x_i = \begin{bmatrix} \text{pitch angle} \\ \text{rotor speed} \\ \text{internal controller state} \\ \text{nacelle fore-aft position} \\ \text{nacelle fore-aft speed} \end{bmatrix}$$

and the exogenous disturbance w_i is the deviation in wind speed from its nominal value, modeled as a white noise process with unit intensity. In the model each WT is assumed to be equipped with an internal controller that receives a power reference and manipulates the pitch angle and generator torque. The control signal u_i is the deviation in the power reference from its nominal value. The model neglects generator dynamics, which makes u_i equal to the actual deviation in the power production of the WT.

Following [18], we assume that each turbine aims at achieving a trade-off between regulating the rotor speed, reducing fatigue loads on the tower, and preventing excessive pitch activity and power deviations. The performance of the ith turbine is quantified as the variance of the regulated variable $z_i = C_z x_i + D_{zu} u_i$, where

$$\left[\begin{array}{c:c} C_z & D_{zu} \end{array}\right] = \left[\begin{array}{c:c} \operatorname{diag}\{\sqrt{0.1}, 100, 0, 100, 0\} & \begin{matrix} 0 \\ 1 \end{matrix} \end{array}\right].$$

In other words, for each turbine we consider the state-feedback H^2 problem for the closed-loop system from w_i to z_i.

The combined power production of the WTs must satisfy a power demand to the WPP, which is assumed to be the sum of nominal WT power productions. Since u_i is the deviation

[1] In particular, we measure the input in MW, use neither a dynamic model of the effective wind speed (its DC gain is absorbed into the model) nor dynamic weights on regulated signals (we use approximate static weights instead).

from nominal WT power production, this requirement can be imposed as the constraint

$$\sum_{i=1}^{\nu} u_i = 0, \tag{1}$$

which introduces coordination between individual WTs.

The resulting constrained H^2 problem can be converted to a standard unconstrained one by resolving (1) for any i, say as $u_1 = -(u_2 + \cdots + u_\nu)$. This results in an H^2 problem with ν subsystems and $\nu - 1$ control signals. Yet the dynamics of subsystems and the cost function in the latter problem are coupled. This might, especially if the number of turbines in the WPP is very large, considerably complicate both the solution procedure (the curse of dimensionality) and the implementation of the resulting controllers. Therefore a *scalable* solution is of interest.

B. Towards a scalable solution

As discussed in the Introduction, the conventional approach in the field is to impose some kind of sparsity constraints on the controller and seek a scalable optimization procedure to solve it. By limiting the information exchange between subsystems, a sparse structure can ensure that the information processing at each subsystem remains viable as the number of subsystems grows. This property is important, so it frequently preponderates over inevitable losses of performance. The problem is that imposing sparsity constraints might significantly complicate the design. Once the constraint (1) is resolved, our problem only satisfies the quadratic invariance condition of [7] for a handful of structural constraint options (e.g., block triangular). Another choice discussed in the Introduction, imposing positivity constraints on the closed-loop dynamics [8], is not suitable for our problem because we work in deviations from nominal values. We thus may consider resorting to nonconvex optimization procedures, relying upon a proper choice of initial parameter guess.

To provide a flavor of such an approach, we confine our attention to static state-feedback controllers, $u = Fx$, and add the constraint $F \in \mathcal{H}_\eta$, where for a given $\eta \in \mathbb{N}$

$$\mathcal{H}_\eta := \{F : F_{ij} = 0 \text{ whenever } |i-j| > \eta\}$$

and the addition in the spatial variable is performed modulo-ν (e.g., $\nu+1 = 1$). The parameter η determines the number of "neighbors" of each subsystem and can be interpreted as the degree of sparsity. The problem is indeed not quadratically invariant under $F \in \mathcal{H}_\eta$. To end up with a (potentially sub-optimal) controller having the desired structure, we use an approach similar to that proposed in [16]:

1) Resolve constraint (1) by the change of variables $u_i = \hat{u}_i - (\hat{u}_{i-1} + \hat{u}_{i+1})/2$, which couples the subsystems.
2) Apply the distributed gradient method in [19] to find a feedback law, $\hat{u} = \hat{F}x$, with $\hat{F} \in \mathcal{H}_{\eta-1}$.
3) Transform back, $\hat{u} \to u$, to obtain F. Note that, $\hat{F} \in \mathcal{H}_{\eta-1}$ is sufficient but not necessary to ensure $F \in \mathcal{H}_\eta$.

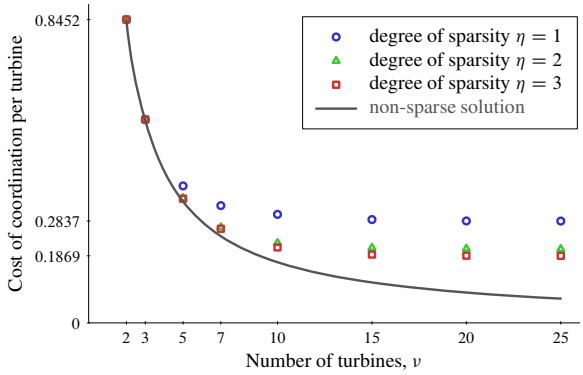

Fig. 1. Cost of coordination per turbine under different degrees of sparsity.

Fig. 1 shows the normalized difference between the H^2 performances attained with and without the coordination constraint (1) (the normalized cost of coordination) as a function of ν for different degrees of sparsity η. We can see that as the sparsity restriction is relaxed, i.e., as η increases, the performance improves. Besides, the performance of sparse controllers improves as ν increases. We can also see that the improvement is not as fast as in the non-sparse (shown by the solid line) case. This, however, was rather expected and is not the main focus of this example.

Rather, we would like to emphasize difficulties encountered in designing the sparse controllers. Although not visible on the plot, these difficulties are readily appreciable. The fact that the problem is not convex renders the whole procedure sensitive to the choice of initial values for the feedback gain. We did experience convergence to local minima, so the solutions presented in Fig. 1 are the result of multiple runs of the algorithm. And we still cannot guarantee that the results are globally optimal[2]. Moreover, potential conservatism might have been introduced by considering the more restrictive class $\mathcal{H}_{\nu-1}$ in the second step of the design and by fixing the feedback gain to be static (not a clear choice in the of structural constraints [11]). In addition, the optimization procedure itself is quite demanding, its computational complexity grows with the increase of ν. Finally, the results of the numerical procedure are not quite transparent, with no indication of what effect small changes of the system parameters might have on it.

To conclude, from the distributed control viewpoint the problem appears to be a challenge. Nonetheless, in the next section we show that its non-sparse version has a closed-form scalable solution, with additional appealing properties.

III. LQR WITH COORDINATION CONSTRAINTS

Motivated by the problem considered in Section II, in this section we study an optimization problem for non-interacting subsystems, having identical dynamics, with coordination constraints. To simplify the presentation, we consider an LQR version of the problem, although the extension to the H^2 formalism (external disturbances) is straightforward.

[2]In fact, they are not, as attested by the sub-optimality of the resulting cost in the case of $\eta = 3$ and $\nu = 7$, for instance.

A. Problem statement

Consider ν independent systems

$$\Sigma_i : \dot{x}_i(t) = Ax_i(t) + Bu_i(t), \quad x_i(0) = x_{i0} \tag{2}$$

where $x_i(t) \in \mathbb{R}^n$ can be measured, $u_i(t) \in \mathbb{R}^m$, and the pair (A, B) is stabilizable. Associate with each of these systems the performance index

$$\mathcal{J}_i = \int_0^\infty \big(x_i'(t)Q_\alpha x_i(t) + u_i'(t)u_i(t)\big)\mathrm{d}t \tag{3}$$

for some $Q_\alpha \geq 0$ and such that the pair (Q_α, A) has no unobservable modes on the jω-axis. Minimizing $\mathcal{J}_i$ for Σ_i would be a set of ν standard uncoupled LQR problems. We couple these problems by constraining the behavior of the *center of mass* of Σ_i, understood as the system

$$\bar{\Sigma} : \dot{\bar{x}}(t) = A\bar{x}(t) + B\bar{u}(t), \quad \bar{x}(0) = \bar{x}_0 \tag{4}$$

connecting the signals

$$\bar{u}(t) := \sum_{i=1}^{\nu} \mu_i u_i(t) \quad \text{and} \quad \bar{x}(t) := \sum_{i=1}^{\nu} \mu_i x_i(t), \tag{5}$$

where the weights $\mu_i \neq 0$ may be thought of as the masses of each subsystem. Coordination is understood as imposing a desired behavior on $\bar{u}$. Our formulation is as follows:

$$\begin{aligned} \text{minimize} \quad & \mathcal{J} := \sum_{i=1}^{\nu} \mathcal{J}_i && \text{(6a)} \\ \text{subject to} \quad & \Sigma_i, \quad i = 1, \ldots, \nu && \text{(6b)} \\ & \bar{u} - \bar{F}\bar{x} = 0 && \text{(6c)} \end{aligned}$$

for some "gain" $\bar{F}$ (the problem studied in Section II corresponds to $\bar{F} = 0$). We implicitly restrict our attention to stabilizing controllers only. Without loss of generality, we also assume that the weights are normalized as $\sum_i \mu_i^2 = 1$.

Remark 3.1 (minimizing weighted sum of $\mathcal{J}_i$): The weights μ_i may be manipulated to assign importance to each subsystem. This can also be attained via replacing $\mathcal{J}$ in (6a) with the weighted sum $\mathcal{J} = \sum_i \lambda_i \mathcal{J}_i$ for some $\lambda_i > 0$. The addition of λ_i, however, does not enrich the design. It is only a matter of scaling each x_i and u_i by $\sqrt{\lambda_i}$ and then replacing μ_i with $\mu_i/\sqrt{\lambda_i}$ (with the normalization assumption $\sum_i \mu_i^2/\lambda_i = 1$). In the choice between 'μ' and 'λ' scalings we picked the former because it allows negative weights. ▽

B. Problem solution

We start with rewriting (6) in an aggregate form using the Kronecker product notation. Introduce the unit vector

$$\mu := \begin{bmatrix} \mu_1 & \cdots & \mu_\nu \end{bmatrix}'$$

and the aggregate state and control signals $x := \sum_i e_i \otimes x_i$ and $u := \sum_i e_i \otimes u_i$, respectively. In this notation, the dynamics of the aggregate systems are

$$\dot{x}(t) = (I_\nu \otimes A)x(t) + (I_\nu \otimes B)u(t), \tag{7}$$

the cost function in (6a) is

$$\mathcal{J} = \int_0^\infty \big(x'(t)(I_\nu \otimes Q_\alpha)x(t) + u'(t)u(t)\big)\mathrm{d}t, \tag{8}$$

and the constraint (6c) reads

$$(\mu' \otimes I_m)u - (\mu' \otimes \bar{F})x = 0,$$

The key idea behind our solution is to apply coordinate transformations to the state and input signals that decouple constraint (6c) on the one hand, while preserving the uncoupled structure of the system and cost on the other. This can be achieved by the coordinate transformations

$$\tilde{x} := (U \otimes I_n)x \quad \text{and} \quad \tilde{u} := (U \otimes I_m)u \tag{9}$$

for some *unitary* matrix $U \in \mathbb{R}^{\nu \times \nu}$. Indeed, using the relation $(M_1 \otimes N_1)(M_2 \otimes N_2) = (M_1 M_2) \otimes (N_1 N_2)$, it is readily seen that both (7) and (8) remain the same, modulo the replacement of x and u with $\tilde{x}$ and $\tilde{u}$, respectively, while the coordination constraint changes and becomes

$$(\mu' U' \otimes I_m)\tilde{u} - (\mu' U' \otimes \bar{F})\tilde{x} = 0.$$

To achieve decoupling, we may consider the following requirements on U:

$$U\mu = e_1. \tag{10}$$

Because μ is assumed to be a unit vector, there is always a U satisfying (10). A possible choice is the matrix of transpose left singular vectors of μ.

Thus, when expressed in terms of $\tilde{x}$ and $\tilde{u}$ with U satisfying (10), problem (6) still has an uncoupled cost function and uncoupled dynamics. But now the constraint, which reads $(e_1' \otimes I_m)\tilde{u} - (e_1' \otimes \bar{F})\tilde{x} = 0$, is imposed only upon the first elements of $\tilde{u}$ and $\tilde{x}$, i.e., it reduces to

$$\tilde{u}_1 - \bar{F}\tilde{x}_1 = 0. \tag{11}$$

Hence, (6) splits into ν independent problems, with the ith problem depending only on the variables $\tilde{x}_i$ and $\tilde{u}_i$.

For $i = 2, \ldots, \nu$, we have identical unconstrained LQR problems with dynamics of the form (2) and cost functions of the form (3). Each one of these problems is then solved by the (stabilizing) control laws $\tilde{u}_i(t) = F_\alpha \tilde{x}_i(t)$, where $F_\alpha := -B'X_\alpha$ and $X_\alpha \geq 0$ is the stabilizing solution of the algebraic Riccati equation (ARE)

$$A'X_\alpha + X_\alpha A + Q_\alpha - X_\alpha BB'X_\alpha = 0. \tag{12}$$

These control laws achieve the optimal performance $\tilde{x}_{i0}' X_\alpha \tilde{x}_{i0}$.

For $i = 1$, constraint (11) completely determines $\tilde{u}_1$, rendering the optimization irrelevant. The plant dynamics then become

$$\dot{\tilde{x}}_1(t) = (A + B\bar{F})\tilde{x}_1(t), \quad \tilde{x}_1(0) = \tilde{x}_{10}$$

and the cost function

$$\int_0^\infty \tilde{x}_1'(t)(Q_\alpha + \bar{F}'\bar{F})\tilde{x}_1(t)\mathrm{d}t$$

(independent of $\tilde{u}_1$). The dynamics of $\tilde{x}$ are stable iff $A + B\bar{F}$ is Hurwitz and in this case the value of the cost function

is finite and equals $\tilde{x}_{10}'\bar{X}\tilde{x}_{10}$, where $\bar{X} \geq 0$ verifies the Lyapunov equation

$$(A + B\bar{F})'\bar{X} + \bar{X}(A + B\bar{F}) + Q_\alpha + \bar{F}'\bar{F} = 0. \quad (13)$$

The arguments above solve (6) in terms of the transformed variables in (9). What is left is to transform this solution back to x and u. This is done in the following theorem, which is the main technical result of this section:

Theorem 3.1: Let $A + B\bar{F}$ be Hurwitz and the pair (Q_α, A) have no unobservable pure imaginary modes. Then the ARE (12) and the Lyapunov equation (13) are solvable, with $\bar{X} \geq X_\alpha$, and the unique solution of (6) is

$$u_i(t) = F_\alpha x_i(t) + \mu_i(\bar{F} - F_\alpha)\bar{x}(t), \quad (14)$$

where $F_\alpha = -B'X_\alpha$ is the LQR gain, associated with the uncoordinated version of the problem, without (6c), and $\bar{x}$ is the state vector of the center of mass $\bar{\Sigma}$ defined by (5). The optimal performance attainable by this controller is

$$\mathcal{J}_{\text{opt}} = \sum_{i=1}^{\nu} \mathcal{J}_{i,\text{opt}} + \bar{x}_0'(\bar{X} - X_\alpha)\bar{x}_0, \quad (15)$$

where $\mathcal{J}_{i,\text{opt}} = x_{i0}'X_\alpha x_{i0}$ is the optimal uncoordinated costs of Σ_i and $\bar{x}_0$ is the initial condition of the center of mass.

Proof: The solvability of the Riccati equations under the conditions of the theorem is a standard result [20, Thm. 13.7]. The inequality $\bar{X} \geq X_\alpha$ follows by the fact that if $u_i = \bar{F}x_i$, then $\mathcal{J}_i = x_{i0}'\bar{X}x_{i0} \geq x_{i0}'X_\alpha x_{i0} = \mathcal{J}_{i,\text{opt}}$ for any x_{i0}. Now, the developments preceding the formulation of the theorem imply that $\tilde{u} = \tilde{F}\tilde{x}$, where

$$\tilde{F} = (I_\nu - e_1e_1') \otimes F_\alpha + (e_1e_1') \otimes \bar{F}.$$

Then (9) implies that $u = Fx = (U' \otimes I_m)\tilde{F}(U \otimes I_n)x$, so, with the help of (10), we obtain

$$F = I_\nu \otimes F_\alpha + (\mu\mu') \otimes (\bar{F} - F_\alpha), \quad (16)$$

which yields (14). Finally,

$$\mathcal{J}_{\text{opt}} = \tilde{x}_0'((I_\nu - e_1e_1') \otimes X_\alpha + (e_1e_1' \otimes \bar{X}))\tilde{x}_0 \quad (17a)$$
$$= x_0'(I_\nu \otimes X_\alpha + (\mu\mu') \otimes (\bar{X} - X_\alpha))x_0, \quad (17b)$$

from which (15) follows immediately. ■

Remark 3.2 (constraining a part of $\bar{u}$): If $\bar{F} = F_\alpha$, then the Lyapunov equation (13) is solved by $\bar{X} = X_\alpha$ and (14) reduces to the decentralized control law solving the uncoordinated version of (6). In other words, the coordination constraint becomes void if it attempts to mimic the optimal unconstrained dynamics. Likewise, we can constrain only a part of $\bar{u}$ by mimicking the optimal, with respect to (3), control trajectory of the partially constrained problem by its other part. Namely, let E be a tall matrix such that $E'E = I$. It can be shown that the optimization of (6), with (6c) replaced by the partial constraint $E'\bar{u} - \bar{F}_1\bar{x} = 0$, corresponds to the original formulation with

$$\bar{F} = E\bar{F}_1 - (I - EE')B'X_2$$

where $X_2 \geq X_\alpha$ is the stabilizing solution of the ARE

$$(A + BE\bar{F}_1)'X_2 + X_2(A + BE\bar{F}_1) + (Q_\alpha + \bar{F}_1'\bar{F}_1) - X_2B(I - EE')B'X_2 = 0$$

and the stabilizability of the pair $(A + BE\bar{F}_1, B(I - EE'))$ is required. Equation (13) is solved then by $\bar{X} = X_2$. ▽

C. Discussion

The remainder of this section is devoted to properties of the solution presented in Theorem 3.1. In particular, we discuss the structure of the optimal controller and its suitability for distributed control applications (§III-C.1), interpret the LQR problems in terms of the transformed variables (9) arising in the derivation as a technical step (§III-C.2), quantify the effect of the coordination constraint (6c) on the performance of each subsystem (§III-C.3), and explore the possibility of adding tracking requirements to the behavior of the center of mass (§III-C.4).

1) Control law: computation and structure: An important property of the solution of Theorem 3.1 is its computational scalability. To calculate the optimal controller, we only need to solve ARE (12), which is the Riccati equation associated with the local, unconstrained, LQR. The computational effort to obtain the solution is thus independent of the number of subsystems ν, which is an attractive property in the context of distributed control.

The low computational burden is not the only property of controller (14) that is appealing in distributed control applications. Its structure is even more intriguing. The optimal control law is a superposition of a local term, $F_\alpha x_i(t)$, and a (scaled) coordination term,

$$u_{\text{coord}}(t) := (\bar{F} - F_\alpha)\bar{x}(t). \quad (18)$$

The former is the optimal uncoordinated control law for Σ_i and is fully decentralized. Coordination then adds a "correction" of the form $\mu_i u_{\text{coord}}$ to this local controller. This term destroys the (sparse) decentralized structure as none of the elements of the overall feedback gain (16) is zero in general. Nonetheless, the resulting configuration might suit large-scale applications well.

The non-sparse coordination term, which may be thought of as a (block) rank-one correction to the (block) diagonal local controller, depends only on the behavior of the center of mass. Thus, although this term hinges upon information about all subsystems, the only operation required in its construction is averaging. This information clustering may be thought of as a form of *spatial generalized sampling* where the information required to form the correction component, u_{coord}, is obtained by aggregating distributed information in a weighted average. In this regard, the ideal sampling [21], where the intersample information is ignored, would correspond to the decentralized structure. The coordination term of the control law of Theorem 3.1 could then correspond to a generalized sampling [13], where the intersample information is aggregated via weighted averaging.

The information aggregation via $\bar{x}$ is clearly less demanding, from both computation and communication viewpoints, than an individual processing of each x_i. Hence, the control law (14), although centralized, may be feasible for distributed control. Measurements of the center of mass could, in principle, be done either globally, by a coordinator, or even locally, by each subsystem.

Remark 3.3 (interpretation of the coordination policy): Constraint (6c) can be satisfied without information exchange if each subsystem applies $u_i = \bar{F}x_i$. The term $(\bar{F} - F_\alpha)x_i$ can then be interpreted as a desired violation of this strategy in order to improve the performance with respect to $\mathcal{J}_i$. By rewriting the coordination term (18) as

$$u_{\text{coord}}(t) = \sum_{i=1}^{\nu} \mu_i(\bar{F} - F_\alpha)x_i(t),$$

we see that the benefit of exchanging information (coordination) is that subsystems can compensate each other's violations. ▽

2) LQR problems in terms of $\tilde{x}_i$ and $\tilde{u}_i$: The transformation of state and input coordinates defined by (9) and (10) serves the purpose of decomposing the problem into one problem with a prespecified control law and and $\nu - 1$ unconstrained LQRs. These problems have meaningful interpretations.

First, a comparison of (11) and (6c) suggests that

$$\tilde{x}_1 = \bar{x} \quad \text{and} \quad \tilde{u}_1 = \bar{u}.$$

This is indeed true, as can be seen through $\tilde{x}_1 = (e_1' \otimes I_n)\tilde{x} = ((e_1'T^{-\prime}) \otimes I_n)x = (\mu' \otimes I_n)x = \bar{x}$, for instance. Thus, the constrained problem is concerned with the center of mass (4) and its solution results in the dynamics

$$\dot{\bar{x}}(t) = (A + B\bar{F})\,\bar{x}(t), \tag{19}$$

as expected.

The other components of $\tilde{x}$ and $\tilde{u}$ do not possess such interpretations per se, they are not even unique. Nevertheless, the unconstrained LQR cost built on them,

$$\tilde{\mathcal{J}} := \sum_{i=2}^{\nu} \int_0^\infty \big(\tilde{x}_i'(t)Q_\alpha\tilde{x}_i(t) + \tilde{u}_i'(t)\tilde{u}_i(t)\big)\mathrm{d}t$$

(this is what the control law (14) actually minimizes), can be interpreted. To this end, rewrite

$$\sum_{i=2}^{\nu} \tilde{u}_i'\tilde{u}_i = \tilde{u}'((I - e_1e_1') \otimes I_m)\tilde{u} = u'((I_\nu - \mu\mu') \otimes I_m)u$$

(the last equality is obtained by (9) and (10)) and, likewise, $\sum_{i=2}^{\nu} \tilde{x}_i'Q_\alpha\tilde{x}_i = x'((I_\nu - \mu\mu') \otimes Q_\alpha)x$. It can be shown, by routine regrouping, that

$$I_\nu - \mu\mu' = \sum_{i=1}^{\nu}(e_i - \mu_i\mu)(e_i - \mu_i\mu)' \tag{20a}$$

$$= \sum_{i=1}^{\nu-1}\sum_{j=i+1}^{\nu}(\mu_je_i - \mu_ie_j)(\mu_je_i - \mu_ie_j)', \tag{20b}$$

Form (20a),

$$\tilde{\mathcal{J}} = \sum_{i=1}^{\nu} \int_0^\infty \big((x_i - \mu_i\bar{x})'Q_\alpha(x_i - \mu_i\bar{x}) + (u_i - \mu_i\bar{u})'(u_i - \mu_i\bar{u})\big)\mathrm{d}t.$$

In other words, $\tilde{\mathcal{J}}$ may be thought of as the cost of deviating from the normalized center of mass. The normalization becomes particularly transparent if all systems have equal masses, i.e., if $\mu_i = 1/\sqrt{\nu}$. In this case $\mu_i\bar{x} = \frac{1}{\nu}\sum_i x_i$ and $\mu_i\bar{u} = \frac{1}{\nu}\sum_i u_i$ are merely the average state and input signals and $\tilde{\mathcal{J}}$ quantifies the cumulative deviation from the average. In the same vein, (20b) leads to

$$\tilde{\mathcal{J}} = \sum_{i=1}^{\nu-1}\sum_{j=i+1}^{\nu} \int_0^\infty \big((\mu_jx_i - \mu_ix_j)'Q_\alpha(\mu_jx_i - \mu_ix_j) + (\mu_ju_i - \mu_iu_j)'(\mu_ju_i - \mu_iu_j)\big)\mathrm{d}t,$$

which penalizes mutual deviations of each subsystem from the others (the scaling factors μ_i and μ_j just align the subsystems to render the comparison meaningful), thus encouraging the achievement of an optimal *consensus*.

Summarizing, by solving (6) we effectively reach two goals: impose a required behavior on the center of mass and minimize discrepancy between subsystems. The optimal $\tilde{\mathcal{J}}$ can then be viewed as a measure of "gregariousness" or, perhaps, as a "herd instinct index" in the aggregate system (7). It follows from the proof of Theorem 3.1 (cf. (17a)) that

$$\tilde{\mathcal{J}}_{\text{opt}} = \tilde{x}_0'\big((I_\nu - e_1e_1') \otimes X_\alpha\big)\tilde{x}_0 = x_0'\big((I_\nu - \mu\mu') \otimes X_\alpha\big)x_0 = \sum_{i=1}^{\nu}\mathcal{J}_{i,\text{opt}} - \bar{x}_0'X_\alpha\bar{x}_0. \tag{21}$$

Thus, the attainable local uncoordinated costs $\mathcal{J}_{i,\text{opt}}$ also determine the cumulative closeness of systems Σ_i to each other. It is worth emphasizing that $\tilde{\mathcal{J}}_{\text{opt}}$ does not depend on the constraint imposed on the behavior of the center of mass. This separation is an intriguing property of the solution of (6).

3) Cost of coordination per subsystem: The last term in the right-hand side of (15) quantifies the deterioration of the (aggregate) performance $\mathcal{J}$ due to the coordination constraint (6c). Below, we look into the effect of coordination on the performance of individual subsystems.

We begin with the following result:

Proposition 3.2: The value of the ith performance index $\mathcal{J}_i$ under the control law (14) is

$$\mathcal{J}_i = \mathcal{J}_{i,\text{opt}} + \mu_i^2\,\bar{x}_0'(\bar{X} - X_\alpha)\bar{x}_0, \tag{22}$$

where $\bar{x}_0$ is the initial condition of the center of mass.

Proof: The control law (14) is a superposition of the locally optimal control law and the signal $v_i = \mu_i(\bar{F} - F_\alpha)\bar{x}$. It is known (see the proof of [20, Thm. 14.2]) that for any v_i,

$$\mathcal{J}_i = \mathcal{J}_{i,\text{opt}} + \|v_i\|_2^2.$$

As follows from (19), the last term in the right-hand side above equals $\mu_i^2 \bar{x}_0' X_v \bar{x}_0$, where $X_v \geq 0$ solves the Lyapunov equation

$$(A + B\bar{F})' X_v + X_v (A + B\bar{F}) + (\bar{F} - F_\alpha)'(\bar{F} - F_\alpha) = 0.$$

(22) then follows by the fact that $X_v = \bar{X} - X_\alpha$, which can be verified by straightforward algebra. ■

The second term in the right-hand side of (22) is exactly the cost of coordination for the ith subsystem. It is a function of the other subsystems through the vector $\bar{x}_0$. The dependence of $\bar{x}_0$ on an unspecified relation between the initial states of all subsystems complicates the analysis of the cost of coordination. If, for instance, $\bar{x}_0 = 0$, then $\mathcal{J}_i = \mathcal{J}_{i,\text{opt}}$ and the coordination in that case comes at no cost. But if every $x_{i0} = \mu_i \chi$ for some $\chi \in \mathbb{R}^n$, then $\bar{x}_0 = \chi$ and we end up with $\mathcal{J}_i = x_{i0}' \bar{X} x_{i0}$. This is what we would have if the control laws $u_i = \bar{F} x_i$ were applied to each subsystem, which would correspond to an attempt to enforce (6c) without communication between subsystems. To avoid the dependence on $\bar{x}_0$, we assume hereafter that $\bar{x}_0$ is bounded as a function of the number of subsystems ν. In this case the term $\bar{x}_0'(\bar{X} - X_\alpha)\bar{x}_0$ is bounded as well and the cost of coordination becomes quadratically proportional to the corresponding "mass" μ_i.

Consider now what happens with the cost of coordination per subsystem when the number of subsystems $\nu \to \infty$. It follows from the normalization assumption $\sum_{i=1}^{\nu} \mu_i^2 = 1$ that at most a finite number subsystems may have $\mu_i \not\to 0$ in this case. If such subsystems do exist, they dominate (5) and we then effectively have coordination between a finite number of subsystems. It is then natural that the cost of coordination for those subsystems does not vanish as ν grows. If, however, all $\mu_i \to 0$ as $\nu \to \infty$, the situation is different. In this case the coordination constraint (6c) is, in a sense, spread among all subsystems and the cost of coordination per subsystem vanishes with the increase of ν. For example, if we assign equal weights to each subsystems, i.e., if every $\mu_i = 1/\sqrt{\nu}$, then the coordination toll per subsystem decreases inversely proportional to the number of subsystems. The decrease of the coordination cost is intuitive, as the addition of more subsystems brings more opportunities for coordination.

4) Tracking: Constraint (6c) can be modified to incorporate tracking requirements on the center of mass (4). For example, we may consider the constraint

$$\bar{u} = \bar{F}\bar{x} + r$$

for a reference signal r. This would yield the control law

$$u_i(t) = F_\alpha x_i(t) + \mu_i(\bar{F} - F_\alpha)\bar{x}(t) + \mu_i r(t),$$

instead of (14), and the following dynamics of the center of mass:

$$\dot{\bar{x}}(t) = (A + B\bar{F})\,\bar{x}(t) + Br(t)$$

(in lieu of (19)). The cost function $\mathcal{J}$ in this case is no longer relevant per se, it might even be unbounded. Still, the "measure of gregariousness" interpretation of the unconstrained part of the optimization, as discussed in §III-C.2, remains valid. Moreover, the value of the cost function in (21) is finite and independent of r, so it can be used to quantify group tracking properties of the system.

D. Wind farm example (cont'd)

We are now in the position to return to the example studied in Section II. To render the current LQR problem formulations compatible with that in §II-A, we assume that $x_{i0} = B_w v_i$, where v_i are mutually independent random variables of unit variance. This yields $\bar{x}_0 = B_w \bar{v}$, where $\bar{v} := \sum_i \mu_i v_i$ is of unit variance as well. We then end up with (6) with $B = B_u$, $Q_\alpha = C_z' C_z$, $\bar{F} = 0$, and $\mu_i = 1/\sqrt{\nu}$ for all i.

By Theorem 3.1, the optimal control law is given by

$$u_i = F_\alpha x_i - F_\alpha x_\text{a},$$

where $x_\text{a} := \frac{1}{\nu}\sum_i x_i$ is the average state of wind turbines and the gain F_α is obtained by solving ARE (12). To calculate the cost of coordination depicted in Fig. 1 by the solid line, we use Proposition 3.2 to end up with the formula

$$\mathcal{J}_i - \mathcal{J}_{i,\text{opt}} = \tfrac{1}{\nu} B_w'(\bar{X} - X_\alpha) B_w,$$

where $\bar{X}$ is the observability Gramian of (C_z, A). This cost tends to zeros as $\nu \to \infty$.

With its structural properties revealed, the non-sparse solution to (6) compares favorably with the sparsity-based one considered in §II-B. Our calculations are scalable, in fact, they are independent of the number of turbines. The result is always globally optimal. The effect of the coordination constraint on the local performance of each turbine is transparent and easy to calculate as well. The price we pay is that the resulting controller is centralized. This might not be feasible in some situations where communication constraints are restrictive. Still, the only centralized information processing that is required to execute the control law is the averaging operation to calculate $F_\alpha x_\text{a}$. This does not require an individual processing of the global state of the whole farm by each turbine. It thus could be feasible even for a large farm.

IV. Concluding Remarks

In this paper we have studied a class of LQR problems, where autonomous agents with identical dynamics seek to reduce their own costs while coordinating their center of mass (average behavior). We have shown that the solution to these problems has two important *scalable* properties. First, the problem decomposes into two independent LQR problems: one for a single uncoordinated agent and one for the center of mass, whose dynamics has the same dimension as those of individual agents. Hence, the computational effort required to obtain the solution is independent of the number of agents. Second, the structure of the resulting controller is transparent, comprising a (block) diagonal decentralized part and a (block) rank-one coordination term. The coordination term relies on information about all subsystems, but only

requires a simple averaging operation. This renders the structure well suited for implementation in distributed control applications.

We have also revealed several other properties of the optimal solution. In particular, the cost of coordination incurred by each subsystem has been quantified and shown to vanish as the number of subsystems grows; the coordination problem has been interpreted in terms of a consensus-like cost function; the cost of the cumulative deviation of subsystems from the center of mass has shown to be independent of the behavior of the center of mass itself.

Although we have studied only the specific LQR problem, the diagonal-plus-low-rank structure may show up in a wider spectrum of applications. Relatively straightforward extensions include replacing the coordination constraint with a coordination penalty (results in a diagonal-plus-rank-one solution), problems with r coordination constraints (would result in a diagonal-plus-rank-r configuration), and output-feedback H^2 formulations (adding local estimators). Other directions may be less trivial. For instance, it may be important to account for additional constraints on the information exchange between agents, like delays or a sampled-data structure. Another possible direction, which would require a substantial alternation of the solution procedure, is to consider coordination among heterogeneous agents. Furthermore, it is interesting to investigate the possibility of reducing information processing / complexity by imposing the diagonal-plus-low-rank structure in problems, where it does not arise as an outcome of the unconstrained optimization procedue.

Last but not least, up to this point we managed to discuss distributed control without mentioning the word "graph."

Acknowledgments

The authors owe a debt of gratitude to Bo Bernhardsson for posing thought-provoking questions and to Anders Rantzer for several helpful discussions.

References

[1] M. D. Mesarović, D. Macko, and Y. Takahara, *Theory of Hierarchical, Multilevel Systems*. NY: Academic Press, 1970.

[2] M. G. Singh and A. Titli, Eds., *Systems: Decomposition, Optimization and Control*. Oxford, UK: Pergamon Press, 1978.

[3] D. D. Šiljak, *Large-Scale Dynamic Systems: Stability and Structure*. NY: North-Holland, 1978.

[4] A. I. Zečević and D. D. Šiljak, *Control of Complex Systems: Structural Constraints and Uncertainty*. NY: Springer-Verlag, 2010.

[5] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, "Information structures in optimal decentralized control," in *Proc. 51st IEEE Conf. Decision and Control*, Maui, HW, 2012, pp. 1291–1306.

[6] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control*, vol. 6, no. 1, pp. 131–147, 1968.

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

[8] T. Tanaka and C. Langbort, "The bounded real lemma for internally positive systems and H-infinity structured static state feedback," *IEEE Trans. Automat. Control*, vol. 56, no. 9, pp. 2218–2223, 2011.

[9] A. Rantzer, "Distributed control of positive systems," in *Proc. 50th IEEE Conf. Decision and Control*, Orlando, FL, 2011, pp. 6608–6611.

[10] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," in *Proc. 2012 American Control Conf.*, Montréal, Canada, 2012, pp. 6314–6321.

[11] P. Shah and P. A. Parrilo, "$\mathcal{H}_2$-optimal decentralized control over posets: A state-space solution for state-feedback," *IEEE Trans. Automat. Control*, vol. 59, 2014, to appear.

[12] A. I. Zečević and D. D. Šiljak, "Global low-rank enhancement of decentralized control for large scale systems," *IEEE Trans. Automat. Control*, vol. 50, no. 5, pp. 740–744, 2005.

[13] M. Araki, "Recent developments in digital control theory," in *Proceedings of 12th IFAC World Congress*, vol. 9, Sidney, Australia, 1993, pp. 251–260.

[14] L. Y. Pao and K. E. Johnson, "Control of wind turbines," *IEEE Control Syst. Mag.*, vol. 31, no. 2, pp. 44–62, 2011.

[15] V. Spudić, M. Jelavić, and M. Baotić, "Wind turbine power references in coordinated control of wind farms," *Automatika – J. Control, Measurements, Electronics, Computing and Commun.*, vol. 52, no. 2, pp. 82–94, 2011.

[16] B. Biegel, D. Madjidian, V. Spudić, A. Rantzer, and J. Stoustrup, "Distributed low-complexity controller for wind power plant in derated operation," in *Proc. 2013 IEEE Multi-Conf. Systems and Control*, Hyderabad, India, 2013.

[17] A. Vigueras-Rodríguez, P. Sørensen, A. Viedma, M. H. Donovan, and E. Gómez Lázaro, "Spectral coherence model for power fluctuations in a wind farm," *J. Wind Eng. Ind. Aerodyn.*, vol. 102, pp. 14–21, 2012.

[18] M. Kristalny, D. Madjidian, and T. Knudsen, "On using wind speed preview to reduce wind turbine tower oscillations," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1191–1198, 2013.

[19] K. Mårtensson and A. Rantzer, "A scalable method for continuous-time distributed control synthesis," in *Proc. 2012 American Control Conf.*, Montréal, Canada, 2012, pp. 6308–6313.

[20] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[21] K. J. Åström and B. Wittenmark, *Computer-Controlled Systems: Theory and Design*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1997.