Graphon Control of Large-scale Networks of Linear Systems

Shuang Gao, *Member, IEEE,* and Peter E. Caines, *Life Fellow, IEEE*

Abstract—To achieve control objectives for extremely large-scale complex networks using standard methods is essentially intractable. In this work a theory of the approximate control of complex network systems is proposed and developed by the use of graphon theory and the theory of infinite dimensional systems. First, graphon dynamical system models are formulated in an appropriate infinite dimensional space in order to represent arbitrary-size networks of linear dynamical systems, and to define the convergence of sequences of network systems with limits in the space. Exact controllability and approximate controllability of graphon dynamical systems are then investigated. Second, the minimum energy state-to-state control problem and the linear quadratic regulator problem for systems on complex networks are considered. The control problem for graphon limit systems is solved in each case and approximations are defined which yield control laws for the original control problems. Furthermore, convergence properties of the approximation schemes are established. A systematic control design methodology is developed within this framework. Finally, numerical examples of networks with randomly sampled weightings are presented to illustrate the effectiveness of the graphon control methodology.

Index Terms—Graphon control, large networks, complex networks, graphons, infinite dimensional systems

I. INTRODUCTION

Complex network systems such as the Internet of Things (IoT), electric, neuronal, food web, epidemic, stock market and social networks, are ubiquitous, and they have been the focus of much research over the past 20 years. In particular, researchers have been studying networks of interacting dynamical systems to learn which collective behaviours may emerge from system interactions over complex networks ([1]–[3]). Furthermore, in addition to the structural properties of networks, system theoretic notions such as controllability, observability, consensus dynamics and synchronization have been widely applied to systems on networks ([4]–[11]). However, to achieve general control objectives for extremely large scale networks with complex interconnections (henceforth, complex networks) using these standard methods is essentially an intractable task.

Graphon theory, introduced and developed in recent years by L. Lovász, B. Szegedy, C. Borgs, J. T. Chayes, V. T. Sós, and K. Vesztergombi among others (see [12]–[15]), provides a theoretical tool to characterize complex graphs and graph limits. This work draws on graph theory, measure theory, probability, and functional analysis, and has been applied in

different areas such as games [16], [17], signal processing [18], network centrality [19], and the heat equation [20].

We propose a graphon based control methodology for controlling complex network systems. The general graphon control strategy consists of the following steps:

1) Identify the graphon limit of the sequence $\tilde{S}$ of networks as the number of nodes goes to infinity.
2) Solve the corresponding control problem for the limit graphon dynamical system.
3) Approximate the control law for the limit system so as to generate control laws for the application to any given finite system along the sequence $\tilde{S}$ of finite network systems.

Specifically, in this paper, the minimum energy state-to-state control problem and the linear quadratic regulator problem are solved for complex network systems using this graphon control strategy.

The main contributions of this paper include:

1) the formulation of graphon differential equations and graphon dynamical control systems, which allows us to represent linear control systems on arbitrary size networks and compare systems of different sizes. This further permits us to design the graphon control methodology based on the network limit.
2) the graphon state-to-state control methodology to solve state-to-state control problem on complex networks.
3) the proposed graphon linear quadratic regulation methodology to solve linear quadratic regulator problems on large-scale networks.

This paper contains the complete proofs omitted in the previous articles [21]–[23] and the extension of previous results, as well as new numerical examples.

The paper is organized as follows: In Section II, the fundamentals of graphon theory are presented, followed by the development of the graphon unitary operator algebra and graphon differential equations. Section III introduces the network system model and its equivalent representation by the graphon dynamical system. Section IV presents the properties of graphon dynamical systems, including existence and uniqueness of the solution and controllability. In Section V and Section VI, the graphon control strategies for the state-to-state control problem and the linear quadratic regulator problem are presented respectively. For each problem, the approximation method is developed and the corresponding convergence properties are established. Section VII contains numerical examples to illustrate the graphon control methodology.

Notation: Bold face letters (e.g. $\mathbf{A}$, $\mathbf{B}$, $\mathbf{u}$) are used to represent graphons and functions. Blackboard bold letters (e.g.

$\mathbb{A}, \mathbb{B}$) are used to denote linear operators which are not necessarily compact. Let $\mathbb{I}$ denote the identity operator. Let $\langle \cdot, \cdot \rangle$ denote inner product for $L^2[0,1]$ and $\| \cdot \|$ represent norm. $\mathbb{1}_S(\cdot)$ denotes the indicator function for a set S, that is, $\mathbb{1}_S(x) = 1$ if $x \in S$, and 0 otherwise. $\mathbf{1}_S$ denotes the $L^2[0,1]$ function with 1 in $S \subset [0,1]$ and 0 in $[0,1]\backslash S$. The set of all real numbers and that of all natural numbers (excluding 0) are respectively denoted by $\mathbb{R}$ and $\mathbb{N}$.

II. PRELIMINARIES

A. Graphs, Adjacency Matrices and Pixel Pictures

The underlying structure of a network can be described by a graph $G = (V, E)$ specified by a node set V and an edge set E which represents the connections between nodes. An equivalent representation of a graph $G = (V, E)$ by a matrix called an *adjacency matrix* is defined to be the square $|V| \times |V|$ matrix A such that an element A_{ij} is one when there is an edge from node i to node j, and zero otherwise. If the graph is a weighted graph where edges are associated with weights, then the adjacency matrix has corresponding weighted elements.

Another representation of the adjacency matrix is given by a pixel diagram where the 0s are replaced by white squares and the 1s by black squares. The whole pixel diagram is presented in a unit square, so the square elements have sides of length $\frac{1}{n}$, where n is the number of nodes.

Fig. 1. Dürer Graph, adjacency matrix, pixel diagram

B. Graphons

A meaningful convergence with respect to the *cut metric* is defined for sequences of dense and finite graphs [15]. Graphons are then the limit objects of converging graph sequences. This concept is illustrated by a sequence of half graphs [15] represented by a sequence of pixel diagrams on the unit square converging to its limit in Fig. 2. Readers are referred to [15] for more examples of convergent graph sequences such as uniform attachment graphs, complete bipartite graphs, and Erdös-Rényi graphs. Exchangeable random graphs can also be modeled by graphons [24].
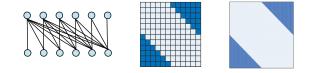

Fig. 2. Graph sequence converging to its limit

The set of finite graphs endowed with the cut metric gives rise to a metric space, and the completion of this space is the *space of graphons*. Graphons are represented by bounded symmetric Lebesgue measurable functions $\mathbf{W} : [0,1]^2 \to [0,1]$,

which can be interpreted as weighted graphs on the node set $[0,1]$. We note that in some papers, for instance [25], the word "graphon" refers to symmetric, integrable functions from $[0,1]^2$ to $\mathbb{R}$. In this paper, unless stated otherwise, the term "graphon" is used to refer to functions $\mathbf{W_1} : [0,1]^2 \to [-1,1]$ and $\mathcal{W}_1$ denotes the space of graphons. Let $\mathcal{W}_0$ represent the space of all graphons satisfying $\mathbf{W_0} : [0,1]^2 \to [0,1]$ and let $\mathcal{W}$ denote the space of all symmetric measurable functions $\mathbf{W} : [0,1]^2 \to \mathbb{R}$.

The cut norm of a graphon $\mathbf{W} \in \mathcal{W}_1$ is then defined as

$$\|\mathbf{W}\|_\square = \sup_{M,T \subset [0,1]} | \int_{M \times T} \mathbf{W}(x,y) dx dy | \qquad (1)$$

with the supremum taking over all measurable subsets M and T of $[0,1]$. Evidently, the following inequalities hold between norms on a graphon $\mathbf{W}$:

$$\|\mathbf{W}\|_\square \leq \|\mathbf{W}\|_1 \leq \|\mathbf{W}\|_2 \leq \|\mathbf{W}\|_\infty \leq 1, \qquad (2)$$

where the second to the forth norms are given by the corresponding L^p norms on $\mathcal{W}_1$. Denote the set of measure preserving bijections from $[0,1]$ to $[0,1]$ by $S_{[0,1]}$. The *cut metric* between two graphons $\mathbf{V}$ and $\mathbf{W}$ is then given by

$$\delta_\square(\mathbf{W}, \mathbf{V}) = \inf_{\phi \in S_{[0,1]}} \|\mathbf{W}^\phi - \mathbf{V}\|_\square, \qquad (3)$$

where $\mathbf{W}^\phi(x,y) = \mathbf{W}(\phi(x), \phi(y))$. We see that the cut metric $\delta_\square(\cdot, \cdot)$ is given by measuring the maximum discrepancy between the integrals of two graphons over measurable subsets of $[0,1]$, then minimizing the maximum discrepancy over all possible measure preserving bijections. Strictly speaking the cut metric is not a metric since the distance between two distinct graphons under the cut metric can be zero (see e.g. [13]). However, by identifying functions $\mathbf{V}$ and $\mathbf{W}$ for which $\delta_\square(\mathbf{V}, \mathbf{W}) = 0$, we can construct the metric space $\tilde{\mathcal{W}}_1$ which denotes the image of $\mathcal{W}_1$ under this identification. Similarly we can construct $\tilde{\mathcal{W}}_0$ from $\mathcal{W}_0$ and $\tilde{\mathcal{W}}$ from $\mathcal{W}$. See [15].

The L^2 *metric* for any graphons $\mathbf{W}$ and $\mathbf{V}$ is defined as

$$d_{L^2}(\mathbf{W}, \mathbf{V}) = \|\mathbf{W} - \mathbf{V}\|_2$$
$$= \left(\int_{[0,1]^2} |\mathbf{W}(x,y) - \mathbf{V}(x,y)|^2 dx dy \right)^{\frac{1}{2}} \quad (4)$$

and the δ_2 *metric* as

$$\delta_2(\mathbf{W}, \mathbf{V}) = \inf_{\phi \in S_{[0,1]}} d_{L^2}(\mathbf{W}^\phi, \mathbf{V}) = \inf_{\phi \in S_{[0,1]}} \|\mathbf{W}^\phi - \mathbf{V}\|_2; \qquad (5)$$

similarly, the L^1 *metric* and the δ_1 *metric* are defined respectively as

$$d_{L^1}(\mathbf{W}, \mathbf{V}) = \|\mathbf{W} - \mathbf{V}\|_1, \qquad (6)$$

$$\delta_1(\mathbf{W}, \mathbf{V}) = \inf_{\phi \in S_{[0,1]}} \|\mathbf{W}^\phi - \mathbf{V}\|_1. \qquad (7)$$

For any two graphons $\mathbf{W}$ and $\mathbf{V}$ the following inequalities hold immediately:

$$\delta_\square(\mathbf{W}, \mathbf{V}) \leq \delta_1(\mathbf{W}, \mathbf{V}) \leq \delta_2(\mathbf{W}, \mathbf{V}) \leq d_{L^2}(\mathbf{W}, \mathbf{V}). \quad (8)$$

The δ_2 (or δ_1) metric and $\delta_\square$ metric share the same equivalence classes under the measure preserving transformations [15, Corollary 8.14]. Clearly, the δ_2 (or δ_1) metric is also well defined on $\tilde{\mathcal{W}}_1$.

C. Compactness of the Graphon Space

Theorem 1 ([15]) *The space $(\tilde{\mathcal{W}}_0, \delta_\square)$ is compact.* □

This remains valid if $\tilde{\mathcal{W}}_0$ is replaced by any uniformly bounded subset of $\tilde{\mathcal{W}}$ closed in the cut metric [15].

Theorem 2 ([15]) *The space $(\tilde{\mathcal{W}}_1, \delta_\square)$ is compact.* □

Sets in $\tilde{\mathcal{W}}_1$ (or $\tilde{\mathcal{W}}_0$) compact with respect to the δ_2 metric are compact with respect to the cut metric. It follows immediately from (8) and Theorem 2 (or Theorem 1), if a graphon sequence is Cauchy in the δ_2 metric then it is also a Cauchy sequence in the cut metric and under both metrics, the limits are identical in $\tilde{\mathcal{W}}_1$ (or $\tilde{\mathcal{W}}_0$).

Define the L^p closed ball in $\tilde{\mathcal{W}}$ with radius $C > 0$ as $\mathcal{B}_{L^p}(C) := \{\mathbf{W} : \|\mathbf{W}\|_p \leq C, \mathbf{W} \in \tilde{\mathcal{W}}\}$.

Theorem 3 ([25]) *The space $(\mathcal{B}_{L^p}(C), \delta_\square)$ with $1 < p \leq \infty$ is compact.* □

By compactness, infinite sequences of graphons will necessarily possess one or more sub-sequential limits under the cut metric.

D. Step Function Graphons

A function $\mathbf{W} \in \mathcal{W}_1$ is called a *step function* if there is a partition $Q = \{Q_1, ..., Q_k\}$ of $[0,1]$ into measurable sets such that $\mathbf{W}$ is constant on every product set $Q_i \times Q_j$. The sets Q_i are the *steps* of $\mathbf{W}$.

Graphons generalize weighted graphs in the following sense. For every weighted graph G with N nodes, a step function $\mathbf{W}_G \in \mathcal{W}_1$ is given by partitioning $[0,1]$ into N measurable sets $Q_1, \cdots, Q_N$ of measure $\mu(Q_i) = \frac{\alpha_i}{\alpha(G)}$ and defining

$$\mathbf{W}_G(x,y) := \sum_{i=1}^{N}\sum_{j=1}^{N} \mathbb{1}_{Q_i}(x)\mathbb{1}_{Q_j}(y)\beta_{ij}(G), \quad (x,y) \in [0,1]^2, \tag{9}$$

where α_i denotes the node weight of ith node, $\alpha(G) = \sum_i \alpha_i$ and $\beta_{ij}(G)$ denotes the weight of the edge from node i to node j (i.e., $\beta_{ij}(G)$ is the ij^{th} entry in the adjacency matrix of G). Evidently the function $\mathbf{W}_G$ depends on the labeling of the nodes of G.

We define the *uniform partition* $P^N = \{P_1, P_2, ..., P_N\}$ of $[0,1]$ by setting $P_k = [\frac{k-1}{N}, \frac{k}{N}), k \in \{1, N-1\}$ and $P_N = [\frac{N-1}{N}, 1]$. Then $\mu(P_i) = \frac{1}{N}, i \in \{1, 2, ..., N\}$. Under the uniform partition, the step functions can be represented by the pixel diagram on the unit square. See [15].

E. Graphons as Operators

A graphon $\mathbf{W} \in \mathcal{W}_1$ can be interpreted as an operator $\mathbf{W} : L^2[0,1] \to L^2[0,1]$. The operation on $\mathbf{v} \in L^2[0,1]$ is defined as follows:

$$[\mathbf{W}\mathbf{v}](x) = \int_0^1 \mathbf{W}(x,\alpha)\mathbf{v}(\alpha)d\alpha. \tag{10}$$

The operator product is then defined by

$$[\mathbf{U}\mathbf{W}](x,y) = \int_0^1 \mathbf{U}(x,z)\mathbf{W}(z,y)dz, \tag{11}$$

where $\mathbf{U}, \mathbf{W} \in \mathcal{W}_1$. See [15] for more details. For simplicity of notation, $\mathbf{U}\mathbf{W}$ is used to denote the graphon given by the convolution in (11); similarly, $\mathbf{W}\mathbf{v}$ denotes the function defined by (10). Note that if $\mathbf{U} \in \mathcal{W}_1$ and $\mathbf{W} \in \mathcal{W}_1$, then $\mathbf{U}\mathbf{W} \in \mathcal{W}_1$, since for all $x, y \in [0,1]$

$$|[\mathbf{U}\mathbf{W}](x,y)| \leq \int_0^1 |\mathbf{U}(x,z)\mathbf{W}(z,y)|dz \leq 1. \tag{12}$$

Consequently, the power $\mathbf{W}^n$ of an operator $\mathbf{W} \in \mathcal{W}_1$ is defined as

$$\mathbf{W}^n(x,y) = \int_{[0,1]^n} \mathbf{W}(x,\alpha_1)\cdots\mathbf{W}(\alpha_{n-1},y)d\alpha_1\cdots d\alpha_{n-1}$$

with $\mathbf{W}^n \in \mathcal{W}_1$ ($n \geq 1$). $\mathbf{W}^0$ is formally defined as the identity operator $\mathbb{I}$ on functions in $L^2[0,1]$, but we note that $\mathbf{W}^0$ is not a graphon.

Any $\mathbf{A} \in \mathcal{W}_1$ gives a self-adjoint compact operator [26] and hence has a discrete spectral decomposition as follows:

$$\mathbf{A}(x,y) = \sum_{\ell=1}^{\infty} \lambda_\ell \mathbf{f}_\ell(x)\mathbf{f}_\ell(y), \quad (x,y) \in [0,1]^2, \tag{13}$$

where the convergence is in the $L^2[0,1]^2$ sense, $\{\lambda_1, \lambda_2, ...\}$ is the set of eigenvalues (which are not necessarily distinct) with decreasing absolute values, and $\{\mathbf{f}_1, \mathbf{f}_2, ...\}$ represents the set of the corresponding orthonormal eigenfunctions (i.e. $\|\mathbf{f}_\ell\|_2 = 1$, and $\langle \mathbf{f}_\ell, \mathbf{f}_k \rangle = 0$ if $l \neq k$). The only accumulating point of the eigenvalues is zero [15], that is, $\lim_{\ell\to\infty} \lambda_\ell = 0$. This implies that $\mathbf{A}$ can be approximated by a finite truncation of the spectral decomposition which preserves the most significant eigenvalues [26].

F. The Graphon Unitary Operator Algebra

It is evident that the operator composition defined in (11) above yields an operator algebra with a multiplicative binary operation possessing the associativity, left distributivity, right distributivity properties and compatibility with the scalar field $\mathbb{R}$, that is, for any $\mathbf{V}, \mathbf{W}, \mathbf{H}$ in the vector space $L^2[0,1]^2$ and $a, b \in \mathbb{R}$,

$$(\mathbf{V}\mathbf{W})\mathbf{H} = \mathbf{V}(\mathbf{W}\mathbf{H}),$$
$$\mathbf{V}(\mathbf{W} + \mathbf{H}) = \mathbf{V}\mathbf{W} + \mathbf{V}\mathbf{H},$$
$$(\mathbf{W} + \mathbf{H})\mathbf{V} = \mathbf{W}\mathbf{V} + \mathbf{H}\mathbf{V},$$
$$(a\mathbf{W})(b\mathbf{H}) = (ab)\mathbf{W}\mathbf{H}.$$

Thus we have an operator algebra $\mathcal{G}_\mathcal{A}$ over the field $\mathbb{R}$ acting on elements of $L^2[0,1]$ with operator multiplication as given in (10). By adjoining the identity element $\mathbb{I}$ to the algebra $\mathcal{G}_\mathcal{A}$ (see e.g. [27]) we obtain a unitary algebra $\mathcal{G}_{\mathcal{A}\mathcal{I}}$. The identity element $\mathbb{I}$ is defined as follows: for any $\mathbf{W} \in L^2[0,1]^2$

$$[\mathbb{I}\mathbf{W}](x,y) = \int_0^1 \mathbf{W}(z,y)\delta(x,z)dz = \mathbf{W}(x,y), \tag{14}$$

where $\delta(\cdot, z)dz$ is the measure satisfying $\int_0^1 u(z)\delta(x,z)dz = u(x)$ for all $u \in L^2[0,1]$, and in particular $\int_0^1 \delta(x,z)dz = 1$.

The *graphon unitary operator algebra* $\mathcal{G}_{\mathcal{A}\mathcal{I}}$ will be used in the definition of the graphon dynamical systems. More specifically, we use the subset $\mathcal{G}_{\mathcal{A}\mathcal{I}}^1 := \{(a\mathbb{I} + \mathbf{A}) : \mathbf{A} \in \mathcal{G}_\mathcal{A}^1, a \in \mathbb{R}\}$ where $\mathcal{G}_\mathcal{A}^1$ is the subset of $\mathcal{G}_\mathcal{A}$ that corresponds to $\mathcal{W}_1$.

G. Graphon Differential Equations

Let X be a Banach space. $\mathcal{L}(X)$ denotes the Banach algebra of all linear continuous mappings $\mathbb{T}: X \to X$, endowed with the norm $\|\mathbb{T}\|_{\text{op}} = \sup_{x \in X, \|x\|=1} \|\mathbb{T}x\|$. A mapping $\mathbb{S}: \mathbb{R} \to \mathcal{L}(X)$ is said to be a *strongly continuous semigroup* on X if the following properties hold:

1) $\mathbb{S}(0) = \mathbb{I}$, $\mathbb{S}(t + s) = \mathbb{S}(t)\mathbb{S}(s)$, $\quad \forall t, s \geq 0$
2) for all $x \in X$, $\mathbb{S}(\cdot)x$ is continuous on $\mathbb{R}$.

A *uniformly continuous semigroup* is a strongly continuous semigroup $\mathbb{S}$ such that $\lim_{t \to 0^+} \|\mathbb{S}(t) - \mathbb{I}\|_{\text{op}} = 0$, with $\| \cdot \|_{\text{op}}$ as the operator norm on a Banach space. The *infinitesimal generator* $\mathbb{A}$ of a strongly continuous semigroup $\mathbb{S}$ is the linear operator in X defined by $\mathbb{A}x = \lim_{t \to 0^+} \frac{1}{t}[\mathbb{S}(t)x - x]$, for all $x \in \mathcal{D}(\mathbb{A})$, where $\mathcal{D}(\mathbb{A}) = \{x \in X : \text{ s.t. } \lim_{t \to 0^+} \frac{1}{t}[\mathbb{S}(t)x - x] \text{ exists}\}$.

Let $\mathbf{A} : [0,1]^2 \to [-1,1]$ be a graphon and $\alpha \in \mathbb{R}$. Hence $(\alpha\mathbb{I} + \mathbf{A})$ is a bounded linear operator from $L^2[0,1]$ to $L^2[0,1]$. Following [28], $(\alpha\mathbb{I}+\mathbf{A})$ is the infinitesimal generator of the uniformly (and hence necessarily strongly) continuous semigroup

$$\mathbb{S}(t) := e^{(\alpha\mathbb{I} + \mathbf{A})t} = \sum_{k=0}^{\infty} \frac{t^k(\alpha\mathbb{I} + \mathbf{A})^k}{k!}.$$

Therefore, the initial value problem of the *graphon differential equation*

$$\dot{\mathbf{y}}_\mathbf{t} = (\alpha\mathbb{I} + \mathbf{A})\mathbf{y}_\mathbf{t}, \quad \mathbf{y_0} \in L^2[0,1] \tag{15}$$

has a solution given by $\mathbf{y}_\mathbf{t} = e^{(\alpha\mathbb{I}+\mathbf{A})t}\mathbf{y_0}$.

Lemma 1 *Let $\mathbf{A} : [0,1]^2 \to [-1,1]$ be a graphon and $\alpha \in \mathbb{R}$. Then $e^{(\alpha\mathbb{I}+\mathbf{A})t} = e^{\alpha t}e^{\mathbf{A}t}$ holds for all $t \geq 0$.* □

Readers can readily check this and hence we omit the proof.

Lemma 2 *Consider any $\mathbf{u} \in L^2[0,1]$, any $\mathbb{A} \in \mathcal{G}^1_{\mathcal{AI}}$, and any $t \in [0,T]$. Let $\mathbb{A} = \alpha\mathbb{I} + \mathbf{A}$, $\mathbf{A} \in \mathcal{W}_1$, then the following holds*

$$\left\|e^{\mathbb{A}t}\mathbf{u}\right\|_2 \leq e^{t(\alpha + \|\mathbf{A}\|_{\text{op}})}\|\mathbf{u}\|_2 \leq e^{t(\alpha + \|\mathbf{A}\|_2)}\|\mathbf{u}\|_2. \tag{16}$$
□

This result is immediate from Lemma 1 and Lemma 8.

Theorem 4 (Appendix B) *For any $\alpha_N, \alpha, t \in \mathbb{R}$, any $\mathbf{A_N}, \mathbf{A_*} \in \mathcal{W}_1$, and any $\mathbf{x} \in L^2[0,1]$, the following holds:*

$$\left\|e^{\mathbf{A_N}t}\mathbf{x} - e^{\mathbf{A_*}t}\mathbf{x}\right\|_2 \leq te^t\|\mathbf{A}_\Delta^\mathbf{N}\|_{\text{op}}\|\mathbf{x}\|_2, \tag{17}$$

$$\left\|e^{(\alpha_N\mathbb{I}+\mathbf{A_N})t}\mathbf{x} - e^{(\alpha\mathbb{I}+\mathbf{A_*})t}\mathbf{x}\right\|_2 \leq te^{(\alpha_N+1)t}\|\mathbf{A}_\Delta^\mathbf{N}\|_{\text{op}}\|\mathbf{x}\|_2$$
$$+ |\alpha - \alpha_N|te^{(L_\alpha + \|\mathbf{A_*}\|_{\text{op}})t}\|\mathbf{x}\|_2, \tag{18}$$

where $\mathbf{A}_\Delta^\mathbf{N} = \mathbf{A_N} - \mathbf{A_*}$ and $L_\alpha = \max\{|\alpha|, |\alpha_N|\}$. Furthermore if a sequence of graphons $\{\mathbf{A_N}\}_{N=1}^\infty$ and that of real numbers $\{\alpha_N\}_{N=1}^\infty$ converge as follows

$$\lim_{N \to \infty} \|\mathbf{A_A} - \mathbf{A_*}\|_{\text{op}} = 0 \quad and \quad \lim_{N \to \infty} |\alpha_N - \alpha| = 0, \tag{19}$$

then for any $\mathbf{x} \in L^2[0,1]$ and any $T > 0$,

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \left\|e^{\mathbf{A_N}t}\mathbf{x} - e^{\mathbf{A_*}t}\mathbf{x}\right\|_2 = 0, \tag{20}$$

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \left\|e^{(\alpha_N\mathbb{I}+\mathbf{A_N})t}\mathbf{x} - e^{(\alpha\mathbb{I}+\mathbf{A_*})t}\mathbf{x}\right\|_2 = 0. \tag{21}$$
□

The operator norm in (17), (18) and (19) can be replaced by the stronger $L^2[0,1]$ norm since $\|\mathbf{W}\|_{\text{op}} \leq \|\mathbf{W}\|_2$ for any $\mathbf{W} \in \mathcal{W}_1$ (see Lemma 7).

III. NETWORK SYSTEMS AND THEIR LIMIT SYSTEMS

A. Network System Models

Definition 1 (Network System) Consider an interlinked network S^N of linear (symmetric) dynamical subsystems $\{S_i^N; 1 \leq i \leq N\}$, each with an n dimensional state space. The subsystem S_i^N at the node V_i in the network $G_N(V,E)$ has interactions with $S_j^N, 1 \leq j \leq N$, specified as below:

$$S_i^N : \quad \dot{x}_t^i = \alpha_N x_t^i + \frac{1}{nN}\sum_{j=1}^N A_{ij}x_t^j + \beta_N u_t^i + \frac{1}{nN}\sum_{j=1}^N B_{ij}u_t^j,$$
$$x_t^i, u_t^i \in \mathbb{R}^n, i \in \{1, ..., N\}, \tag{22}$$

with $A_N = [A_{ij}], B_N = [B_{ij}] \in \mathbb{R}^{nN \times nN}$ as the (symmetric) block-wise adjacency matrices of $G_N(V,E)$ and of the input graph, where $A_{ij} = [0]$ if S_i^N has no connection to S_j^N and similarity for B_{ij}. We call S^N a *network system*. □

Then the (symmetric) linear dynamics for the network system $S^N(A_N, B_N, G_N)$ can be represented by

$$S^N : \quad \begin{aligned} \dot{x}_t &= \alpha_N x_t + A_N \circ x_t + \beta_N u_t + B_N \circ u_t, \\ x_t, u_t &\in \mathbb{R}^{nN}, A_N, B_N \in \mathbb{R}^{nN \times nN}, \end{aligned} \tag{23}$$

where $\circ$ denotes the so called averaging operator given by $A_N \circ x = \frac{1}{(nN)}A_N x$. Let $\mathcal{S} = \times_{N=1}^\infty \mathcal{S}^\mathcal{N}$ where $\mathcal{S}^\mathcal{N} = \cup_{A_N, B_N, G_N} S^N(A_N, B_N, G_N)$. For simplicity, we require the elements of A_N and B_N to be in $[-1,1]$ for each N (note that in general A_N and B_N have elements that are uniformly bounded real numbers for which case we would achieve similar results). In addition, we note that if we take the supremum norm on vectors in $\mathbb{R}^{nN}$, i.e. $\|x\|_\infty = \sup_i |x_i|$, and the corresponding $\circ$ operator norm of A, i.e. $\|A\|_{\text{op}\infty} = \sup_{\|x\|_\infty \neq 0} \frac{\|A \circ x\|_\infty}{\|x\|_\infty}$, then $\|A\|_{\text{op}\infty} \leq 1$.

B. Network Systems Represented by Step Functions

Let $\{(\alpha_N I + A_N; \beta_N I + B_N)\}_{N=1}^\infty \in \mathcal{S}$ be a sequence of systems with the node averaging dynamics each of which is described according to (23). Let $|A_{Nij}| \leq 1$ and $|B_{Nij}| \leq 1$ for all $i, j \in \{1, ..., nN\}$. Let $\mathbf{A^{[N]}}, \mathbf{B^{[N]}} \in \mathcal{W}_1$ be the step functions corresponding one-to-one to A_N and B_N; these are specified using the uniform partition P^{nN} of $[0,1]$ by the following *matrix to step function mapping M_G*:

$$\mathbf{A^{[N]}}(x,y) := \sum_{i=1}^{nN}\sum_{j=1}^{nN} \mathbb{1}_{P_i}(x)\mathbb{1}_{P_j}(y)A_{Nij}, \quad \forall(x,y) \in [0,1]^2 \tag{24}$$

and similar for $\mathbf{B^{[N]}}$. In fact, the step function $\mathbf{A^{[N]}}$ can represent a set of matrices of different sizes given by $\{A_N \otimes \text{one}_m, m \in \mathbb{N}\}$ where one_m is the $m \times m$ matrix of ones.

Define a *piece-wise constant (PWC) function* on $\mathbb{R}$ to be any function of the form $\sum_{k=1}^l \alpha_k \mathbb{1}_{I_k}(\cdot)$ where $\alpha_1, ..., \alpha_l$ are real numbers and each I_k is a bounded interval (open, closed, or half-open). Let $L^2_{pwc}[0,1]$ denote the space of piece-wise constant $L^2[0,1]$ functions under the uniform partition P^{nN}.

$\mathbf{u_t^{[N]}} \in L_{pwc}^2[0,1]$ corresponds one-to-one to $u_t \in \mathbb{R}^{nN}$ via the following *vector to PWC function mapping* also denoted by M_G:

$$\mathbf{u_t^{[N]}}(\alpha) := \sum_{i=1}^{nN} \mathbb{1}_{P_i}(\alpha) u_t(i), \quad \forall \alpha \in [0,1], \qquad (25)$$

and similarly $\mathbf{x_t^{[N]}} \in L_{pwc}^2[0,1]$ corresponds one-to-one to $x_t \in \mathbb{R}^{nN}$.

Lemma 3 (Appendix B) *The trajectories of the system in (23) correspond one-to-one under the mapping M_G to the trajectories of the system*

$$\dot{\mathbf{x}}_\mathbf{t}^{[N]} = (\alpha_N \mathbb{I} + \mathbf{A}^{[N]}) \mathbf{x_t^{[N]}} + (\beta_N \mathbb{I} + \mathbf{B}^{[N]}) \mathbf{u_t^{[N]}},$$
$$\mathbf{x_t^{[N]}}, \mathbf{u_t^{[N]}} \in L_{pwc}^2[0,1], \mathbf{A}^{[N]}, \mathbf{B}^{[N]} \in \mathcal{W}_1 \subset \mathcal{G}_{\mathcal{AI}}^1 \qquad (26)$$

with graphon operations defined according to (10). □

Since the system in (26) corresponds to a network system S^N in (23), we also refer to it as a *network system*. We use $(\mathbb{A}^{[N]}; \mathbb{B}^{[N]})$ to denote the network system in (26) where $\mathbb{A}^{[N]} = \alpha_N \mathbb{I} + \mathbf{A}^{[N]}$ and $\mathbb{B}^{[N]} = \beta_N \mathbb{I} + \mathbf{B}^{[N]}$.

C. Limits of Sequences of Network Systems

A sequence of network systems with node averaging dynamics in (23) can be represented by the sequence of systems $\{(\alpha_N \mathbb{I} + \mathbf{A}^{[N]}; \beta_N \mathbb{I} + \mathbf{B}^{[N]}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1\}_{N=1}^\infty$ in (26).

Definition 2 (System Sequence Convergence) A sequence of systems $\{(\alpha_N \mathbb{I} + \mathbf{A}^{[N]}; \beta_N \mathbb{I} + \mathbf{B}^{[N]}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1\}_{N=1}^\infty$ is convergent if the following two conditions hold
1) there exist $\alpha, \beta \in \mathbb{R}$ such that

$$\lim_{N \to \infty} \alpha_N = \alpha \quad \text{and} \quad \lim_{N \to \infty} \beta_N = \beta;$$

2) there exist $\mathbf{A}, \mathbf{B} \in \mathcal{W}_1$,

$$\lim_{N \to \infty} \|\mathbf{A} - \mathbf{A}^{[N]}\|_{op} = 0 \quad \text{and} \quad \lim_{N \to \infty} \|\mathbf{B} - \mathbf{B}^{[N]}\|_{op} = 0. \quad \square$$

The limit system is represented by $(\mathbb{A}; \mathbb{B})$ where $\mathbb{A} = \alpha \mathbb{I} + \mathbf{A}$ and $\mathbb{B} = \beta \mathbb{I} + \mathbf{B}$.

Since any $\mathbf{W} \in \mathcal{W}_1$ defines a self-adjoint and compact operator, the maximum absolute value of eigenvalues of $\mathbf{W}$ equals to the operator norm [29, Theorem 12.31], that is, $\|\mathbf{W}\|_{op} = \max_\ell |\lambda_\ell|$. Furthermore, the following inequalities between the cut norm and the operator norm hold [16], [30]:

$$\|\mathbf{W}\|_{\square} \leq \|\mathbf{W}\|_{op} \leq \sqrt{8\|\mathbf{W}\|_{\square}}. \qquad (27)$$

By Lemma 7, the following inequality also holds:

$$\|\mathbf{W}\|_{op} \leq \|\mathbf{W}\|_2. \qquad (28)$$

Based on (27) and (28), we obtain that the convergence of a sequence of graphons in $\|\cdot\|_2$ or $\|\cdot\|_{\square}$ implies its convergence in $\|\cdot\|_{op}$. Under certain extra conditions, the convergence of a sequence of graphons under the cut norm implies its convergence under the $L^2[0,1]$ norm [31, Corollary 1.1].

Let the graphon sequences $\{\mathbf{A}^{[N]}\}$ and $\{\mathbf{B}^{[N]}\}$ be Cauchy sequences of step functions in $L^2[0,1]^2$. Due to the completeness of $L^2[0,1]^2$, the respective graphon limits $\mathbf{A}$ and $\mathbf{B}$ exist and these will then necessarily be the limits in the cut metric (see Section II-C and [15]).

IV. THE LIMIT GRAPHON SYSTEM AND ITS PROPERTIES

We follow [32] and specialize both the Hilbert space of states H, and that of controls U appearing there, to the space $L^2[0,1]$. Let $L^2([0,T]; L^2[0,1])$ denotes the Hilbert space of equivalence classes of strongly measurable (in the Böchner sense [33, p.103]) mappings $[0,T] \to L^2[0,1]$ that are integrable with norm $\|\mathbf{f}\|_{L^2([0,T];L^2[0,1])} = \left[\int_0^T \|\mathbf{f}(s)\|_2^2 ds\right]^{\frac{1}{2}}$.

Definition 3 (Graphon Systems) We formulate an infinite dimensional linear system, which we call a *graphon system* $(\mathbb{A}; \mathbb{B})$, as follows:

$$\dot{\mathbf{x}}_\mathbf{t} = \mathbb{A}\mathbf{x_t} + \mathbb{B}\mathbf{u_t}, \quad \mathbf{x_0} \in L^2[0,1], \qquad (29)$$

where $\mathbb{A}, \mathbb{B} \in \mathcal{G}_{\mathcal{AI}}^1$, and are hence bounded operators on $L^2[0,1]$, $\mathbf{x_t} \in L^2[0,1]$ is the system state at time t and $\mathbf{u_t} \in L^2[0,1]$ is the control input at time t. □

Notice that the network system in (26) is a special case of the graphon system in (29).

A solution $\mathbf{x}_{(\cdot)} \in L^2([0,T]; L^2[0,1])$ is a *(mild) solution* of (29) if $\mathbf{x_t} = e^{(t-a)\mathbb{A}}\mathbf{x_a} + \int_a^t e^{(t-s)\mathbb{A}}\mathbb{B}\mathbf{u_s}ds$ for all a and t in $[0,T]$, taken to be $a \leq t$ (see [32]). Let $C([0,T]; L^2[0,1])$ denote the set of continuous mappings from $[0,T]$ to $L^2[0,1]$.

Proposition 1 *The graphon system in (29) has a unique solution $\mathbf{x} \in C([0,T]; L^2[0,1])$ for all $\mathbf{x_0} \in L^2[0,1]$ and all $\mathbf{u} \in L^2([0,T]; L^2[0,1])$.* □

PROOF $\mathbb{A}$ as a bounded linear operator generates a uniformly continuous semigroup and $\mathbb{B} \in \mathcal{G}_{\mathcal{AI}}^1$ as a bounded linear operator forms a continuous linear mapping from control space $L^2[0,1]$ to the state space $L^2[0,1]$. Hence following [32, p.385], the system (29) has a unique solution $\mathbf{x} \in C([0,T]; L^2[0,1])$ for all $\mathbf{x_0} \in L^2[0,1]$ and all $\mathbf{u} \in L^2([0,T]; L^2[0,1])$. ■

A system $(\mathbb{A}; \mathbb{B})$ is *exactly controllable* on $[0,T]$ if for any initial state $\mathbf{x_0} \in L^2[0,1]$ and any target state $\mathbf{x}_f \in L^2[0,1]$, there exists a control $\mathbf{u} \in L^2([0,T]; L^2[0,1])$ driving the system from $\mathbf{x_0}$ to $\mathbf{x}_f$, i.e. $\mathbf{x}_T = \mathbf{x}_f$ with $\mathbf{x}_T = e^{\mathbb{A}T}\mathbf{x_0} + \int_0^T e^{\mathbb{A}(T-t)}\mathbb{B}\mathbf{u_t}dt$. A system $(\mathbb{A}; \mathbb{B})$ is *approximately controllable* on $[0,T]$ if for any initial state $\mathbf{x_0} \in L^2[0,1]$, any target state $\mathbf{x}_f \in L^2[0,1]$ and any $\varepsilon > 0$, there exists a control $\mathbf{u} \in L^2([0,T]; L^2[0,1])$ which drives the system state from $\mathbf{x_0}$ into the $L^2[0,1]$ ε-neighborhood of $\mathbf{x}_f$, i.e., $\|\mathbf{x}_T - \mathbf{x}_f\|_2 \leq \varepsilon$.

The *controllability Gramian operator* $\mathbb{W}_T : L^2[0,1] \to L^2[0,1]$ is defined as

$$\mathbb{W}_T := \int_0^T e^{\mathbb{A}(t-s)}\mathbb{B}\mathbb{B}^\mathsf{T} e^{\mathbb{A}^\mathsf{T}(t-s)}ds, \quad T > 0. \qquad (30)$$

A necessary and sufficient condition for exact controllability on $[0,T]$ is the uniform positive definiteness of $\mathbb{W}_T$:

$$\langle \mathbb{W}_T h, h \rangle \geq c_T \|h\|^2 \qquad (31)$$

for all $h \in L^2[0,1]$, where $c_T > 0$ and $\|\cdot\|$ is the $L^2[0,1]$ norm. The positive definiteness of the controllability Gramian operator $\mathbb{W}_T$ as a kernel is equivalent to the approximate controllability of the corresponding system (see [32], [34]).

Define the *kernel space* (or *null space*) of a linear operator $\mathbb{T}$ on $L^2[0,1]$ as: $\ker(\mathbb{T}) := \{x \in L^2[0,1] : \mathbb{T}x = 0\}$. The *spectrum* $\sigma(\mathbb{T})$ of a bounded linear operator $\mathbb{T}$ on $L^2[0,1]$ is the set of all (complex or real) scalars λ such that $\mathbb{T} - \lambda\mathbb{I}$ is not invertible. Thus $\lambda \in \sigma(\mathbb{T})$ if and only if at least one of the following two statements is true:

(i) The range of $\mathbb{T} - \lambda\mathbb{I}$ is not all of $L^2[0,1]$, i.e., $\mathbb{T} - \lambda\mathbb{I}$ is not onto.

(ii) $\mathbb{T} - \lambda\mathbb{I}$ is not one-to-one.

If (ii) holds, λ is said to be an *eigenvalue* of $\mathbb{T}$; the corresponding eigenspace is $\ker(\mathbb{T} - \lambda\mathbb{I})$; each $x \in \ker(\mathbb{T} - \lambda\mathbb{I})$ (except $x = 0$) is an *eigenvector* of $\mathbb{T}$; it satisfies the equation $\mathbb{T}x = \lambda x$. See [29].

Theorem 5 (Appendix C) *Let $\mathbb{A}$ be an element in $\mathcal{G}_{\mathcal{AI}}^1$ and let $\mathbb{B}$ be a bounded linear operator on $L^2[0,1]$. The linear system $(\mathbb{A}; \mathbb{B})$ is exactly controllable on a finite time horizon $[0,T]$ if all the values in the spectrum of $\mathbb{B}\mathbb{B}^T$ are lower bounded by a strictly positive constant.* □

Proposition 2 (Appendix C) *Let $\mathbf{A}$ and $\mathbf{B}$ be graphons in $\mathcal{W}_1$. Then the linear system $(\mathbf{A}; \mathbf{B})$ is not exactly controllable on any finite time horizon $[0,T]$.* □

The results in Theorem 5 and Proposition 2 generalize to the case where the underlying graphons lie in any uniformly bounded subset of $\mathcal{W}$, i.e., any set of symmetric measurable functions $\mathbf{W} : [0,1]^2 \to I$ where I is a closed bounded interval in $\mathbb{R}$.

V. GRAPHON STATE-TO-STATE CONTROL OF NETWORK SYSTEMS

A. Approximation of $L^2[0,1]$ Functions

Theorem 6 ([35, Theorem 13.23]) *Let μ be any measure on $\mathbb{R}$ and $\mathcal{B}_\mu$ be the σ-algebra of μ-measurable sets, and let $1 \le p < \infty$. Then piece-wise constant functions on $\mathbb{R}$ form a dense subset of $L^p(\mathbb{R}, \mathcal{B}_\mu, \mu)$.* □

Proposition 3 *Let $\mathbf{v} \in L^2[0,1]$ be approximated by $\mathbf{v}^{[\mathbf{N}]} \in L^2_{pwc}[0,1]$ as follows: for all $x \in Q_i$ and for all $i \in \{1, \ldots, nN\}$,*

$$\mathbf{v}^{[\mathbf{N}]}(x) = \frac{1}{\mu(Q_i)} \int_{Q_i} \mathbf{v}(\alpha)d\alpha \tag{32}$$

with the partition $\{Q_1, Q_2, \ldots, Q_{nN}\}$ of $[0,1]$ where $\mu(Q_i)$ denotes the measure of Q_i. Then $\|\mathbf{v}^{[\mathbf{N}]}\|_2 \le \|\mathbf{v}\|_2$. □

PROOF Applying the Cauchy-Schwarz inequality yields

$$\|\mathbf{v}\|_2^2 = \int_0^1 \mathbf{v}^2(x)dx = \sum_{i=1}^{nN} \int_{Q_i} \mathbf{v}^2(x)dx$$

$$\ge \sum_{i=1}^{nN} \frac{1}{\mu(Q_i)} \left[\int_{Q_i} \mathbf{v}(x)dx\right]^2 \tag{33}$$

$$= \sum_{i=1}^{nN} \mu(Q_i) \left[\frac{1}{\mu(Q_i)} \int_{Q_i} \mathbf{v}(x)dx\right]^2 = \|\mathbf{v}^{[\mathbf{N}]}\|_2^2.$$

∎

In this paper we wish to approximate any control input $\mathbf{u_t} \in L^2[0,1], 0 \le t \le T$, through a piece-wise constant function

in $L^2[0,1]$ denoted by $\mathbf{u_t^{[N]}}$. Specifically, the approximation of an input $\mathbf{u_t}$ via the function $\mathbf{u_t^{[N]}}$ with the partition $Q = \{Q_1, Q_2, \cdots, Q_{nN}\}$ of $[0,1]$ will be specified as follows: for all $Q_i, i \in \{1, 2, \ldots, nN\}$,

$$\mathbf{u_t^N}(\alpha) = \frac{1}{\mu(Q_i)} \int_{Q_i} \mathbf{u_t}(\beta)d\beta, \quad \forall \alpha \in Q_i, \tag{34}$$

where $\mu(Q_i)$ denotes the measure of Q_i.

B. Limit Control for Network Systems

Theorem 7 (Appendix D) *Consider the problem of driving the systems $(\mathbb{A}^{[N]}; \mathbb{B}^{[N]})$ in (26) and $(\mathbb{A}; \mathbb{B})$ in (29) from the origin to some target state. Let $\mathbb{A}^{[N]} = \alpha_N\mathbb{I} + \mathbf{A}^{[N]}$, $\mathbb{B}^{[N]} = \beta_N\mathbb{I} + \mathbf{B}^{[N]}$, $\mathbb{A} = \alpha\mathbb{I} + \mathbf{A}$ and $\mathbb{B} = \beta\mathbb{I} + \mathbf{B}$. Let $\mathbf{x_T}(\mathbf{u})$ represent the terminal state of $(\mathbb{A}; \mathbb{B})$ under control $\mathbf{u}$ and $\mathbf{x_T^N}(\mathbf{u^{[N]}})$ represent the terminal state of $(\mathbb{A}^{[N]}; \mathbb{B}^{[N]})$ under control $\mathbf{u^{[N]}}$. Then for any $T > 0$, there exists a control $\mathbf{u^{[N]}}$ for $(\mathbb{A}^{[N]}; \mathbb{B}^{[N]})$ approximating the control $\mathbf{u}$ for $(\mathbb{A}; \mathbb{B})$ such that*

$$\|\mathbf{x_T}(\mathbf{u}) - \mathbf{x_T^N}(\mathbf{u^{[N]}})\|_2$$
$$\le |\alpha - \alpha_N| (|\beta| + \|\mathbf{B}\|_{\text{op}})$$
$$\int_0^T (T-t)e^{(L_\alpha + \|\mathbf{A}\|_{\text{op}})(T-t)} \|\mathbf{u_t}\|_2 dt$$
$$+ \|\mathbf{A_\Delta^N}\|_{\text{op}}(|\beta| + \|\mathbf{B}\|_{\text{op}}) \int_0^T e^{(\alpha_N+1)(T-t)}(T-t)\|\mathbf{u_t}\|_2 dt$$
$$+ (|\beta - \beta_N| + \|\mathbf{B_\Delta^N}\|_{\text{op}})$$
$$\int_0^T e^{\left(\alpha_N + \|\mathbf{A}^{[N]}\|_{\text{op}}\right)(T-t)} \|\mathbf{u_t}\|_2 dt$$
$$+ |\beta_N| \int_0^T e^{\left(\alpha_N + \|\mathbf{A}^{[N]}\|_{\text{op}}\right)(T-t)} \left\|\mathbf{u_t} - \mathbf{u_t^{[N]}}\right\|_2 dt, \tag{35}$$

where $L_\alpha = \max\{|\alpha_N|, |\alpha|\}$, $\mathbf{A_\Delta^N} = \mathbf{A} - \mathbf{A}^{[N]}$, $\mathbf{B_\Delta^N} = \mathbf{B} - \mathbf{B}^{[N]}$, and the control approximation is given in the following:

$$\mathbf{u_t^{[N]}}(\alpha) = nN \int_{P_i} \mathbf{u_t}(\beta)d\beta, \tag{36}$$

for all $\alpha \in P_i$, $t \in [0,T]$, with the uniform partition $P^{nN} = \{P_1, \cdots, P_{nN}\}$. Furthermore, if a sequence of network systems $\{(\mathbb{A}^{[N]}; \mathbb{B}^{[N]})\}_{N=1}^\infty$ converges to a graphon system $(\mathbb{A}; \mathbb{B})$ as in Definition 2, then for any $\varepsilon > 0$ there exists $N_\varepsilon > 0$ such that each $N \ge N_\varepsilon$,

$$\left\|\mathbf{x_T}(\mathbf{u}) - \mathbf{x_T^N}(\mathbf{u^{[N]}})\right\|_2 < \varepsilon. \tag{37}$$
□

The control law $u_{(\cdot)}^N$ for the finite network system $(\alpha_N I + A_N; \beta_N I + B_N)$ is given by

$$u_t^N(i) = \mathbf{u_t^{[N]}}(\alpha), \quad \forall i \in \{1, ..., nN\}, \forall \alpha \in P_i, t \in [0,T].$$

Note that u^N always exists by definition since the control approximation given in the definition (34) uses the same uniform partition as the step function approximation in the graphon space.

The operator norm in (35) can be replaced by the $L^2[0,1]^2$ norm since $\|\cdot\|_{\text{op}} \le \|\cdot\|_2$.

C. The Graphon State-to-state Control (GSSC) Strategy

Consider the control problem of steering the states of each member of $\{(\alpha_N I + A_N; \beta_N I + B_N)\}_{N=1}^{\infty} \in \mathcal{S}$ to each of a sequence of desired states $\{x_T^N \in \mathbb{R}^{nN}\}_{N=1}^{\infty}$.

The **Graphon State-to-state Control (GSSC) Strategy** consists of four steps:

S.1 *Let $\{(\alpha_N \mathbb{I} + \mathbf{A}^{[\mathbf{N}]}; \beta_N \mathbb{I} + \mathbf{B}^{[\mathbf{N}]}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1\}_{N=1}^{\infty}$ be the sequence of graphon dynamical systems equivalent to $\{(\alpha_N I + A_N; \beta_N I + B_N)\}_{N=1}^{\infty} \in \mathcal{S}$ under the mapping M_G, and assume it converges to the graphon system $(\alpha \mathbb{I} + \mathbf{A}; \beta \mathbb{I} + \mathbf{B}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1$ as in Definition 2. Let $\{\mathbf{x_T^N} \in L^2[0,1]\}_{N=1}^{\infty}$ be the image of $\{x_T^N \in \mathbb{R}^{nN}\}_{N=1}^{\infty}$ under M_G, which is assumed to converge to some $\mathbf{x_T^\infty} \in L^2[0,1]$ in the $L^2[0,1]$ norm.*

S.2 *Specify the corresponding state to state control problem CP^∞ for $(\alpha \mathbb{I} + \mathbf{A}; \beta \mathbb{I} + \mathbf{B}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1$ with $\mathbf{x_T^\infty}$ as the target terminal state and choose a tolerance $\varepsilon > 0$.*

S.3 *Find a control law $\mathbf{u}^\infty := \{\mathbf{u}_\tau \in L^2[0,1], \tau \in [0,T]\}$ solving CP^∞.*

S.4 *Then generate the control law $\{\mathbf{u}^{[\mathbf{N}]}\}$ according to Theorems 7 for which the convergence of $\{\mathbf{x_T^N}(\mathbf{u^N})\}$ to $\mathbf{x_T^\infty}$ is guaranteed. Together with the assumed convergence of $\{\mathbf{x_T^N} \in L^2[0,1]\}_{N=1}^{\infty}$ to $\mathbf{x_T^\infty}$, it yields N_ε such that $\mathbf{x_T^N}(\mathbf{u^N})$ is within ε of $\mathbf{x_T^N}$ for all $N \geq N_\varepsilon$ under the $L^2[0,1]$ norm.*

The notion of the effectiveness of the GSSC strategy for a sequence of network systems is defined to mean that (1) the terminal state is close to that achieved by the minimum energy control; (2) the computation for generating the control law is tractable.

The basic assumptions for the GSSC strategy are that (i) a sequence of finite network systems of interest converges to a limit graphon system (as in Definition 2) or a given instance of the network sequence can be closely approximated by a graphon system, and (ii) the corresponding state-to-state control problem for the (limit) graphon system is tractable.

These assumptions, together with the approximation theorem (i.e. Theorem 7), guarantee the effectiveness of the GSSC strategy for the finite network, that is to say, the GSSC strategy can achieve the target terminal state with a small error by means of a tractable computation.

D. Min-Energy State-to-state Control for Graphon Systems

A specific control law which may be used in **S.2** of the GSSC strategy is described in this section.

Define the energy cost by the control over the time horizon $[0,T]$ as $J(\mathbf{u}) = \int_0^T \|\mathbf{u}_\tau\|^2 d\tau, (0 < T < \infty)$. The objective is to drive the system from some initial state $\mathbf{x_0} \in L^2[0,1]$ to some target state $\mathbf{x_T} \in L^2[0,1]$ using minimum control energy. A function $\mathbf{u}^* \in L^2([0,T]; L^2[0,1])$ is called an optimal control if $J(\mathbf{u}^*) \leq J(\mathbf{u})$, for all $\mathbf{u} \in L^2([0,T]; L^2[0,1])$ which drive the system from $\mathbf{x_0}$ to $\mathbf{x_T}$.

Theorem 8 (Appendix E) *If the graphon system $(\mathbb{A}; \mathbb{B})$ in (29) with $\mathbb{W}_T$ as its graphon controllability Gramian operator is exactly controllable, then the inverse operator $\mathbb{W}_T^{-1}$ exists and is a bounded operator.* □

Assume the system $(\mathbb{A}; \mathbb{B})$ is exactly controllable, then $\mathbb{W}_T^{-1}$ exists and the optimal control law that achieves the minimum energy control is given by

$$\mathbf{u}_\tau^* = \mathbb{B}^\mathsf{T} e^{\mathbb{A}^\mathsf{T}(T-\tau)} \mathbb{W}_T^{-1}(\mathbf{x_T} - e^{\mathbb{A}(T)}\mathbf{x_0}), \quad \tau \in [0,T]. \quad (38)$$

The minimum energy for controlling the system in time horizon $[0,T]$ is

$$\|\mathbf{u}\|_{L^2([0,T];L^2[0,1])}^2 = [\mathbf{x_T} - e^{\mathbb{A}(T)}\mathbf{x_0}]^\mathsf{T} \mathbb{W}_T^{-1}[\mathbf{x_T} - e^{\mathbb{A}(T)}\mathbf{x_0}]. \quad (39)$$

Denote the spectral decomposition of $\mathbf{A}$ is as follows $\mathbf{A}(x,y) = \sum_{\ell \in I_\lambda} \lambda_\ell \mathbf{f}_\ell(x)\mathbf{f}_\ell(y)$, where $\mathbf{f}_\ell$ is the normalized eigenfunction corresponding to the eigenvalue λ_ℓ and I_λ is the index set for non-zero eigenvalues of $\mathbf{A}$, which contains a countable number of elements [15].

Proposition 4 (Appendix E) *Consider a graphon system $(\mathbf{A}; \mathbb{I})$. Then*

1) the controllability Gramian operator is given by

$$\mathbb{W}_T = T\mathbb{I} + \sum_{\ell \in I_\lambda} \left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T\right) \mathbf{f}_\ell \mathbf{f}_\ell^\mathsf{T}; \quad (40)$$

2) the inverse of the controllability Gramian operator for $(\mathbf{A}; \mathbb{I})$ is given by

$$\mathbb{W}_T^{-1} = \frac{1}{T}\mathbb{I} - \frac{1}{T} \sum_{\ell \in I_\lambda} \frac{\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T}{\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1]\right)} \mathbf{f}_\ell \mathbf{f}_\ell^\mathsf{T}. \quad (41)$$
□

Note that $\lim_{\lambda_\ell \to 0}\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T\right) = 0$. We further note that approximate controllability is not sufficient to achieve state-to-state control since in that case the inverse operator $\mathbb{W}_T^{-1}$ may not be bounded on certain subspaces in $L^2[0,1]$, and moreover the corresponding energy required would be unbounded.

VI. Graphon Linear Quadratic Regulation (LQR) of Network Systems

A. LQR Problems for Graphon Dynamical Systems

For finite $T > 0$, consider the problem of minimizing the cost given by

$$J(\mathbf{u}) = \int_0^T \left(\|\mathbb{C}\mathbf{x}_\tau\|_2^2 + \|\mathbf{u}_\tau\|_2^2\right) d\tau + \langle \mathbb{P}_0 \mathbf{x_T}, \mathbf{x_T} \rangle \quad (42)$$

over all controls $\mathbf{u} \in L^2([0,T]; L^2[0,1])$ subject to the system model constrains in (29).

Assumption 1 $\mathbb{P}_0 \in \mathcal{L}(L^2[0,1])$ *is Hermitian and non-negative;* $\mathbb{C} \in \mathcal{L}(L^2[0,1])$

Finding the feedback control via dynamic programming consists of the two following standard steps [32]:

1) Solving the Riccati equation

$$\dot{\mathbb{P}} = \mathbb{A}^\mathsf{T}\mathbb{P} + \mathbb{P}\mathbb{A} - \mathbb{P}\mathbb{B}\mathbb{B}^\mathsf{T}\mathbb{P} + \mathbb{C}^\mathsf{T}\mathbb{C}, \quad \mathbb{P}_0 \in \mathcal{L}(L^2[0,1]); \quad (43)$$

2) Given the solution $\mathbb{P}$ to the Riccati equation, the optimal control $\mathbf{u}^*$ is given by

$$\mathbf{u_t^*} = -\mathbb{B}^\mathsf{T}\mathbb{P}_{T-t}\mathbf{x_t^*}, \quad t \in [0,T] \quad (44)$$

and the optimal trajectory $\mathbf{x}^*$ is then the solution to the closed loop equation

$$\dot{\mathbf{x}}_{\mathbf{t}} = \mathbb{A}\mathbf{x}_{\mathbf{t}} - \mathbb{B}\mathbb{B}^{\mathsf{T}}\mathbb{P}_{T-t}\mathbf{x}_{\mathbf{t}}, \ t \in [0,T], \mathbf{x}_0 \in L^2[0,1]. \tag{45}$$

Let $\Sigma(L^2[0,1]) = \{\mathbb{T} \in \mathcal{L}(L^2[0,1]) : \mathbb{T} \text{ is Hermitian}\}$ and

$$\begin{aligned} &\Sigma^+(L^2[0,1]) \\ &= \{\mathbb{T} \in \Sigma(L^2[0,1]) : \langle \mathbb{T}\mathbf{v}, \mathbf{v}\rangle \geq 0, \ \forall \mathbf{v} \in L^2[0,1]\}. \end{aligned} \tag{46}$$

Denote the topological space of all strongly continuous mappings $\mathbb{F} : I \to \Sigma(L^2[0,1])$ endowed with strong convergence (see [32]) by $C_s(I; \Sigma(L^2[0,1]))$ where I denotes a compact interval, that is, the convergence of $\mathbb{F}_N$ to $\mathbb{F}$ in $C_s(I; \Sigma(L^2[0,1]))$ is defined by

$$\forall \mathbf{v} \in L^2[0,1], \quad \lim_{N \to \infty} \sup_{t \in I} \|\mathbb{F}_N(t)\mathbf{v} - \mathbb{F}(t)\mathbf{v}\|_2 = 0. \tag{47}$$

Proposition 5 ([32, Part IV-1]) *Under* Assumption 1*, there exist a unique solution to the Riccati equation (43) in $C_s([0,T]; \Sigma^+(L^2[0,1]))$ and an optimal solution pair $(\mathbf{u}^*, \mathbf{x}^*)$ to (44) and (45) where $\mathbf{x}^* \in C([0,T]; L^2[0,1])$ and $\mathbf{u}^* \in L^2([0,T]; L^2[0,1])$.* □

B. The Graphon-Network LQR (GLQR) Strategy

Consider the control problem of regulating the states of each member of $\{(\alpha_N I + A_N; \beta_N I + B_N)\}_{N=1}^\infty \in \mathcal{S}$.

The **Graphon-Network LQR (GLQR) Strategy** is as follows:

S.1 *Let $\{(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1\}_{N=1}^\infty$ be the sequence of equivalent representation of network systems $\{(\alpha_N I + A_N; \beta_N I + B_N)\}_{N=1}^\infty \in \mathcal{S}$ under the mapping M_G and assume that it converges to the graphon system $(\mathbb{A}; \mathbb{B}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1$ as in Definition 2.*

S.2 *Define the linear quadratic cost for $(\mathbb{A}; \mathbb{B})$ as*

$$J(\mathbf{u}) = \int_0^T \left[\|\mathbb{C}\mathbf{x}_\tau\|_2^2 + \|\mathbf{u}_\tau\|_2^2\right] d\tau + \langle \mathbb{P}_0\mathbf{x}_{\mathbf{T}}, \mathbf{x}_{\mathbf{T}}\rangle \tag{48}$$

and the linear quadratic cost for $(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]})$ as

$$\begin{aligned} J(\mathbf{u}^{[\mathbf{N}]}) = \int_0^T &\left[\|\mathbb{C}^{[\mathbb{N}]}\mathbf{x}_{\mathbf{t}}^{[\mathbf{N}]}\|_2^2 + \|\mathbf{u}_{\mathbf{t}}^{[\mathbf{N}]}\|_2^2\right] dt \\ &+ \langle \mathbb{P}_0^{[\mathbb{N}]}\mathbf{x}_{\mathbf{T}}^{[\mathbf{N}]}, \mathbf{x}_{\mathbf{T}}^{[\mathbf{N}]}\rangle \end{aligned} \tag{49}$$

where it is assumed that $\mathbb{C}^{[\mathbb{N}]} \to \mathbb{C}$ and $\mathbb{P}_0^{[\mathbb{N}]} \to \mathbb{P}_0$ in the strong operator sense. Solve the infinite dimensional Riccati equation for $(\mathbb{A}; \mathbb{B})$ to generate the solution $\mathbb{P}$.

S.3 *Approximate $\mathbb{P}$ to generate $\widetilde{\mathbb{P}}^{\mathbb{N}}$ and hence the control law $\mathbf{u}_{\mathbf{t}}^{[\mathbf{N}]} = -\mathbb{B}^{[\mathbb{N}]\mathsf{T}}\widetilde{\mathbb{P}}_{T-t}^{\mathbb{N}}\mathbf{x}_{\mathbf{t}}^{[\mathbf{N}]}$ for $(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]})$.*

Parallel to the state-to-state control problem, we take the notion of the effectiveness of the GLQR strategy for a sequence of network systems to be that (1) the regulation cost and the state trajectory are close to those achieved by the optimal LQR control; (2) the computation for generating the control law is tractable.

In analogy with the state-to-state control problem, the basic assumptions for the GLQR strategy are that the sequence of finite network systems converges to a limit graphon system (as in Definition 2) or that a given network system can be closely approximated by a graphon system, and that the corresponding LQR problem for the (limit) graphon system is tractable.

These assumptions, together with Theorem 10, guarantee the effectiveness of the GLQR strategy for the finite network systems that are sufficiently close to the limit graphon system for sufficiently large node cardinality.

C. Control Law Approximations

By approximating the Riccati equation solution $\mathbb{P}$ for $(\mathbb{A}; \mathbb{B})$ we can generate $\widetilde{\mathbb{P}}^{\mathbb{N}}$ that provides the control law for the finite dimensional network system:

$$\mathbf{u}_{\mathbf{t}}^{[\mathbf{N}]} = -\mathbb{B}^{[\mathbb{N}]\mathsf{T}}\widetilde{\mathbb{P}}_{(T-t)}^{\mathbb{N}}\mathbf{x}_{\mathbf{t}}^{[\mathbf{N}]}. \tag{50}$$

Consider the strongly continuous linear operator $\mathbb{P}$ in $C_s([0,T]; \Sigma^+(L^2[0,1]))$. Its approximation is given by

$$\widetilde{\mathbb{P}}_t^{\mathbb{N}}(x,y) = \frac{\langle \mathbf{1}_{Q_i}, \mathbb{P}_t \mathbf{1}_{Q_j}\rangle}{\mu(Q_i)\mu(Q_j)}, \quad \forall t \in [0,T], \quad \forall (x,y) \in Q_i \times Q_j, \tag{51}$$

where $\{Q_1, Q_2, \ldots, Q_{nN}\}$ forms a partition of $[0,1]$ and $\mu(Q_i)$ represents the length of the interval Q_i. In the case of uniform partition, $\mu(Q_i) = \frac{1}{nN}$.

Lemma 4 *Let $\widetilde{\mathbb{P}}^{\mathbb{N}}$ be generated by the step function approximation of $\mathbb{P}$ via the N uniform partition of $[0,1]$ according to (51). Then*

$$\lim_{N \to \infty} \widetilde{\mathbb{P}}^{\mathbb{N}} = \mathbb{P}, \quad \text{in } C_s([0,T]; \Sigma(L^2[0,1])). \tag{52}$$
□

PROOF Consider an arbitrary function $\mathbf{v} \in L^2[0,1]$. Based on the definition of the step function approximation in (51), for any $t \in [0,T]$, $\widetilde{\mathbb{P}}_t^{\mathbb{N}}\mathbf{v}$ is the piece-wise-constant function approximation of $\mathbb{P}_t\mathbf{v} \in L^2[0,1]$ as in (32) and by Proposition 3, $\|\widetilde{\mathbb{P}}_t^{\mathbb{N}}\mathbf{v}\|_2 \leq \|\mathbb{P}_t\mathbf{v}\|_2$. Since the Riccati equation (43) over the closed bounded interval $[0,T]$ has a solution $\mathbb{P}$ (see Proposition 5), for any $\mathbf{v} \in L^2[0,1]$ there exist $C_v > 0$ such that

$$\|\mathbb{P}_{t_0}\mathbf{v} - \mathbb{P}_t\mathbf{v}\|_2 \leq C_v|t_0 - t|, \quad \forall t_0, t \in [0,T]. \tag{53}$$

Note that $(\widetilde{\mathbb{P}}_{t_0}^{\mathbb{N}} - \widetilde{\mathbb{P}}_t^{\mathbb{N}})\mathbf{v} \in L^2[0,1]$ is an approximation of $(\mathbb{P}_{t_0} - \mathbb{P}_t)\mathbf{v} \in L^2[0,1]$ following (32). Therefore by the contraction property in Proposition 3, we obtain

$$\left\|\widetilde{\mathbb{P}}_{t_0}^{\mathbb{N}}\mathbf{v} - \widetilde{\mathbb{P}}_t^{\mathbb{N}}\mathbf{v}\right\|_2 \leq C_v|t_0 - t|, \quad \forall t_0, t \in [0,T], \tag{54}$$

and hence the sequence of functions $\{\widetilde{\mathbb{P}}_{(\cdot)}^{\mathbb{N}}\mathbf{v}\}_{N=1}^\infty$ is equicontinuous (see e.g. [29, p.43]). Furthermore, $\widetilde{\mathbb{P}}_{(\cdot)}^{\mathbb{N}}\mathbf{v}$ and $\mathbb{P}_{(\cdot)}\mathbf{v}$ are continuous functions defined over the closed bounded time interval $[0,T]$. Hence by the Arzelà-Ascoli Theorem,

$$\forall t \in [0,T], \quad \lim_{N \to \infty} \left\|\widetilde{\mathbb{P}}_t^{\mathbb{N}}\mathbf{v} - \mathbb{P}_t\mathbf{v}\right\|_2 = 0 \tag{55}$$

implies that, for any $\mathbf{v} \in L^2[0,1]$,

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \left\|\widetilde{\mathbb{P}}_t^{\mathbb{N}}\mathbf{v} - \mathbb{P}_t\mathbf{v}\right\|_2 = 0, \tag{56}$$

which gives the result in (52). ∎

Lemma 5 (Appendix F) *Let* $\widetilde{\mathbb{P}}^{\mathbb{N}}$ *be generated by step function approximation from* $\mathbb{P}$ *via the* N *uniform partition of* $[0,1]$ *according to* (51). *If* $\mathbb{P}^{[\mathbb{N}]}$ *converges strongly to the solution* $\mathbb{P}$ *in* $C_s([0,T]; \Sigma(L^2[0,1]))$, *then for any* $\mathbf{x} \in L^2[0,1]$,

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \left\| \widetilde{\mathbb{P}}_t^{\mathbb{N}} \mathbf{x} - \mathbb{P}_t^{[\mathbb{N}]} \mathbf{x} \right\|_2 = 0. \qquad \square$$

Let $Ricc(\mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{P}_0)$ denote the Riccati equation in (43) with initial condition $\mathbb{P}_0$.

Assumption 2

1) *For any* $N \geq 1$, $\mathbb{P}_0^{[\mathbb{N}]} \in \mathcal{L}(L^2[0,1])$ *is Hermitian and non-negative, and* $\mathbb{C}^{[\mathbb{N}]} \in \mathcal{L}(L^2[0,1])$.
2) *The system sequence* $\{(\mathbb{A}^{[\mathbb{N}]}, \mathbb{B}^{[\mathbb{N}]})\}$ *converges to* $(\mathbb{A}; \mathbb{B})$ *as in* Definition 2.
3) *The sequences* $\{\mathbb{C}^{[\mathbb{N}]}\}$ *and* $\{\mathbb{P}_0^{[\mathbb{N}]}\}$ *converge strongly to* $\mathbb{C}$ *and* $\mathbb{P}_0$, *respectively, as* $N \to \infty$.
4) $\mathbb{C}$ *and* $\mathbb{C}^{[\mathbb{N}]}$ *are self-adjoint linear operators.*

Theorem 9 *Consider a sequence of network systems* $\{(\alpha_N I + A_N; \beta_N I + B_N)\}_{N=1}^{\infty}$ *with* $\{(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]}) \in \mathcal{G}_{\mathcal{AI}}^1 \times \mathcal{G}_{\mathcal{AI}}^1\}_{N=1}^{\infty}$ *as the equivalent representation. Let* $\mathbb{P}$ *and* $\mathbb{P}^{[\mathbb{N}]}$ *be the solutions to* $Ricc(\mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{P}_0)$ *and* $Ricc(\mathbb{A}^{[\mathbb{N}]}, \mathbb{B}^{[\mathbb{N}]}, \mathbb{C}^{[\mathbb{N}]}, \mathbb{P}_0^{[\mathbb{N}]})$ *respectively. If* Assumption 2 *holds, then for any horizon* $[0,T]$, $T > 0$,

$$\lim_{N \to \infty} \mathbb{P}^{[\mathbb{N}]} = \mathbb{P} \quad \text{in } C_s([0,T]; \Sigma(L^2[0,1])). \qquad \square$$

PROOF From Theorem 4, we know for all $T > 0$ and all $\mathbf{x} \in L^2[0,1]$, $\lim_{N \to \infty} e^{t\mathbb{A}^{[\mathbb{N}]}} \mathbf{x} = e^{t\mathbb{A}} \mathbf{x}$ uniformly in $[0,T]$. Since the system sequence $\{(\mathbb{A}^{[\mathbb{N}]}, \mathbb{B}^{[\mathbb{N}]})\}$ converges to $(\mathbb{A}; \mathbb{B})$ as in Definition 2, $\{\mathbb{B}^{[\mathbb{N}]}\}$ converges to $\mathbb{B}$ in the strong operator sense. We can now apply [32, Theorem 2.2, Part IV], specialized to the Hilbert space $L^2[0,1]$. Since its hypotheses are then satisfied in the present case, the desired result follows. ∎

Let $\mathbb{P}^{[\mathbb{N}]}$ denote the solution to the Riccati equation for $(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]})$ that converges strongly to the solution $\mathbb{P}$ of the Riccati equation for $(\mathbb{A}; \mathbb{B})$. Let $\widetilde{\mathbb{P}}^{\mathbb{N}}$ be the step function approximation of $\mathbb{P}$ generated via the N uniform partition of $[0,1]$ according to (51).

Theorem 10 (Appendix F) *Consider the time horizon* $[0,T]$. *Assume the sequence of initial conditions* $\{\mathbf{x}_0^{[\mathbb{N}]} \in L^2[0,1]\}$ *is convergent and* Assumption 2 *holds. Let the optimal linear quadratic control law for* $(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]})$ *be generated by*

$$\mathbf{u}_t^{N*} = -\mathbb{B}^{[\mathbb{N}]^{\mathsf{T}}} \mathbb{P}_{(T-t)}^{[\mathbb{N}]} \mathbf{x}_t^{N*}, \qquad (57)$$

where the optimal state trajectory is given by $\mathbf{x}^{N*}$, *and let the graphon approximate control law for* $(\mathbb{A}^{[\mathbb{N}]}; \mathbb{B}^{[\mathbb{N}]})$ *be given by*

$$\mathbf{u}_t^{[\mathbb{N}]} = -\mathbb{B}^{[\mathbb{N}]^{\mathsf{T}}} \widetilde{\mathbb{P}}_{(T-t)}^{\mathbb{N}} \mathbf{x}_t^{[\mathbb{N}]}, \qquad (58)$$

where the corresponding state trajectory is given by $\mathbf{x}^{[\mathbb{N}]}$. *Then*

$$\forall t \in [0,T], \quad \lim_{N \to \infty} \left\| \mathbf{x}_t^{N*} - \mathbf{x}_t^{[\mathbb{N}]} \right\|_2 = 0,$$

and $\lim_{N \to \infty} \left| J(\mathbf{u}^{N*}) - J(\mathbf{u}^{[\mathbb{N}]}) \right| = 0$. $\qquad \square$

VII. NUMERICAL EXAMPLES

A. Convergent Network Sequences with Sampled Weightings

The generation of a randomly sampled network of size N from a graphon $\mathbf{A}$ is specified as follows:

1) Sample N points from a uniform distribution in $[0,1]$. Sort the sample points in the decreasing order of their values and label them from node 1 to node N. Denote the node set by V_N and the value of node $i \in V_N$ by v_i.
2) Connect the nodes $i, j \in V_N$ with edge weight $\mathbf{A}(v_i, v_j)$ to generate the network G_N. Then $A_{Nij} = \mathbf{A}(v_i, v_j)$ is the ij^{th} element of the adjacency matrix of G_N.

If $\mathbf{A}$ is almost everywhere continuous, then the step function $\mathbf{A}^{[\mathbb{N}]}$ of $A_N = [A_{Nij}]$ converges to $\mathbf{A}$ in the δ_1 metric as $N \to \infty$ (see e.g. [36]), that is, $\delta_1(\mathbf{A}^{[\mathbb{N}]}, \mathbf{A}) \to 0$, as $N \to \infty$. Furthermore, this implies $\delta_2(\mathbf{A}^{[\mathbb{N}]}, \mathbf{A}) \to 0$, as $N \to \infty$ since $\mathbf{A} \in \mathcal{W}_1$ is uniformly bounded. By the generation procedure, we obtain the labeling that approximates the minimum distance between the network and the limit, and hence the sequence of networks converge in the $L^2[0,1]^2$ metric to the limit.

As an example, we consider the following sinusoidal graphon $\mathbf{A}$: for all $x, y \in [0,1]$,

$$\mathbf{A}(x,y) = 0.5 \cos(2\pi(x-y)) + 0.25 \cos(4\pi(x-y)).$$

The normalized eigenfunctions are $\mathbf{f}_1 = \sqrt{2} \cos 2\pi(\cdot)$, $\mathbf{f}_2 = \sqrt{2} \cos 4\pi(\cdot)$, $\mathbf{f}_3 = \sqrt{2} \sin 2\pi(\cdot)$ and $\mathbf{f}_4 = \sqrt{2} \sin 4\pi(\cdot)$ with eigenvalues $\lambda_1 = \lambda_3 = \frac{1}{4}$ and $\lambda_2 = \lambda_4 = \frac{1}{8}$.

B. Minimum Energy Graphon State-to-state Control

Consider a network system evolving according to node averaging dynamics with G_N describing the dynamic interactions. Suppose each node has an independent input channel. Denote the system by $(A_N; I_N)$, where A_N is the adjacency matrix of G_N and I_N is the identity input mapping. The network system $(A_N; I_N)$ with node averaging dynamics is therefore described by

$$\dot{x}_t^i = \frac{1}{N} \sum_{j=1}^{N} A_{Nij} x_t^j + u_t^i, \quad x_t^i, u_t^i \in R, i \in \{1, ..., N\}, \quad (59)$$

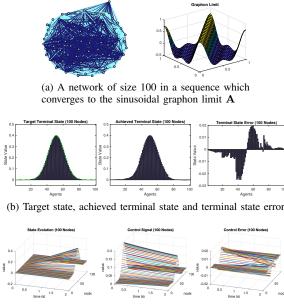where A_{Nij} is sampled from the sinusoidal graphon.

We solve the minimum energy control problem of driving the states of the network system $(A_N; I_N)$ to a terminal state x_T^N from the origin over the time horizon $[0,T]$ with $T = 2$. Here we consider the limit target terminal state $\mathbf{x_T} = \frac{1}{\sqrt{2\pi}} e^{-50(\alpha-0.5)^2}, \alpha \in [0,1]$.

Based on Proposition 4, the system $(\mathbf{A}; \mathbb{I})$ is exactly controllable and the inverse of the controllability Gramian operator is explicitly given by (41). The minimum control law based on (38) is explicitly given by

$$\mathbf{u_t} = \frac{1}{T} \mathbf{x_T}$$
$$+ \frac{1}{T} \sum_{\ell=1}^{4} \mathbf{f}_\ell \langle \mathbf{x_T}, \mathbf{f}_\ell \rangle \Big[- \frac{2\lambda_\ell \left(\frac{e^{2\lambda_\ell T} - 1}{2\lambda_\ell} - T \right) \left(e^{\lambda_\ell(T-t)} - 1 \right)}{e^{2\lambda_\ell T} - 1}$$
$$+ \left(e^{\lambda_\ell(T-t)} - 1 \right) - \frac{2\lambda_\ell \left(\frac{e^{2\lambda_\ell T} - 1}{2\lambda_\ell} - T \right)}{e^{2\lambda_\ell T} - 1} \Big], \quad t \in [0,T],$$

with $\langle \mathbf{x_T}, \mathbf{f}_1 \rangle = -0.116$, $\langle \mathbf{x_T}, \mathbf{f}_2 \rangle = 0.064$, $\langle \mathbf{x_T}, \mathbf{f}_3 \rangle = \langle \mathbf{x_T}, \mathbf{f}_4 \rangle = 0$. Then the control law $u_{(\cdot)}^N$ for a network system $(A_N; I_N)$ generated based on the approximation in (36). The error $\|\mathbf{x_T}(\mathbf{u}) - \mathbf{x_T^N}(\mathbf{u^{[N]}})\|_2$ is bounded as in Theorem 7 and converges to 0 as $N \to \infty$. The result of a simulation for a network system with 100 nodes using the GSSC strategy is shown in Figure 3.

(a) A network of size 100 in a sequence which converges to the sinusoidal graphon limit $\mathbf{A}$

(b) Target state, achieved terminal state and terminal state error

(c) State trajectory, control signal and its deviation from the optimal control signal

Fig. 3. Minimum energy graphon state-to-state control

C. Graphon-Network LQR

Consider a network system with dynamics given by

$$\dot{x}_t^i = \alpha x_t^i + \frac{1}{N} \sum_{j=1}^N A_{N_{ij}} x_t^j + \beta u_t^i, \quad x_t^i, u_t^i \in \mathbb{R}. \quad (60)$$

The objective is to design a control law that minimizes the following cost with network coupling:

$$J(u) = \frac{1}{N} \sum_{i=1}^N \Big[\int_0^T \Big(q \big(x_t^i - \frac{\eta}{N} \sum_{j=1}^N A_{N_{ij}} x_t^j \big)^2 + (u_t^i)^2 \Big) dt$$

$$+ q_T \big(x_T^i - \frac{\eta}{N} \sum_{j=1}^N A_{N_{ij}} x_T^j \big)^2 \Big], \quad (61)$$

where $q, q_T \geq 0$. That is we want to regulate the state of each subsystem to be close to the local weighted network average with small control effort. The equivalent formulation of this problem for the graphon system following (26) is given by

$$\dot{\mathbf{x_t^{[N]}}} = \alpha \mathbf{x_t^{[N]}} + \mathbf{A^{[N]}} \mathbf{x_t^{[N]}} + \beta \mathbf{u_t^{[N]}}$$

$$J(\mathbf{u^{[N]}}) = \int_0^T \Big[q \left\| (\mathbb{I} - \eta \mathbf{A^{[N]}}) \mathbf{x_t^{[N]}} \right\|_2^2 + \left\| \mathbf{u_t^{[N]}} \right\|_2^2 \Big] dt \quad (62)$$

$$+ q_T \left\| (\mathbb{I} - \eta \mathbf{A^{[N]}}) \mathbf{x_T^{[N]}} \right\|_2^2,$$

where $\mathbf{x_t^{[N]}}, \mathbf{u_t^{[N]}} \in L_{pwc}^2[0,1]$, $\mathbf{A^{[N]}} \in \mathcal{W}_1$. The limit problem (if exists) or the approximate problem is given by

$$\dot{\mathbf{x_t}} = \alpha \mathbf{x_t} + \mathbf{A} \mathbf{x_t} + \beta \mathbf{u_t}$$

$$J(\mathbf{u}) = \int_0^T \Big[q \left\| (\mathbb{I} - \eta \mathbf{A}) \mathbf{x_t} \right\|_2^2 + \left\| \mathbf{u_t} \right\|_2^2 \Big] dt \quad (63)$$

$$+ q_T \left\| (\mathbb{I} - \eta \mathbf{A}) \mathbf{x_T} \right\|_2^2,$$

where $\mathbf{A} \in \mathcal{W}_1$, $\mathbf{x_t}, \mathbf{u_t} \in L^2[0,1]$. Let us consider a special case where $\mathbf{A}$ permits an exact finite spectral decomposition as follows:

$$\mathbf{A}(x,y) = \sum_{\ell=1}^d \lambda_\ell \mathbf{f}_\ell(x) \mathbf{f}_\ell(y), \quad (x,y) \in [0,1]^2, \quad (64)$$

where $\lambda_1, \lambda_2, ..., \lambda_d$ are non-zero eigenvalues of $\mathbf{A}$ and $\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_d$ represent orthonormal eigenfunctions. Then the solution $\mathbb{P}$ to the Riccati equation

$$\dot{\mathbb{P}}_t = (\alpha \mathbb{I} + \mathbf{A})^\mathsf{T} \mathbb{P}_t + \mathbb{P}_t(\alpha \mathbb{I} + \mathbf{A}) - \beta^2 (\mathbb{P}_t)^2 + q(\mathbb{I} - \eta \mathbf{A})^2$$

$$\mathbb{P}_0 = q_T (\mathbb{I} - \eta \mathbf{A})^2$$

$$(65)$$

is given by $\mathbb{P}_t = \breve{\Pi}_t \mathbb{I} + \sum_{\ell=1}^d \left(\Pi_t^\ell - \breve{\Pi}_t \right) \mathbf{f}_\ell \mathbf{f}_\ell^\mathsf{T}$, where $\breve{\Pi}$ and Π^ℓ are the solutions to the following scalar Riccati equations

$$\dot{\breve{\Pi}}_t = 2\alpha \breve{\Pi}_t - \beta^2 (\breve{\Pi}_t)^2 + q,$$

$$\dot{\Pi}_t^\ell = 2(\alpha + \lambda_\ell) \Pi_t^\ell - \beta^2 (\Pi_t^\ell)^2 + q(1 - \eta \lambda_\ell)^2, \quad (66)$$

$$\breve{\Pi}_0 = q_T, \quad \Pi_0^\ell = q_T (1 - \eta \lambda_\ell)^2, \quad 1 \leq \ell \leq d.$$

The optimal control for the limit problem is then given by

$$\mathbf{u_t} = -\beta \breve{\Pi}_{(T-t)} \mathbf{x_t} - \beta \sum_{\ell=1}^d (\Pi_{(T-t)}^\ell - \breve{\Pi}_{(T-t)}) \langle \mathbf{x_t}, \mathbf{f}_\ell \rangle \mathbf{f}_\ell.$$

See [26], [37] for the details of the solution method, which provides solutions to a more general class of graphon control problems with network couplings in states, controls and costs.
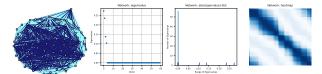
If $\mathbf{A^{[N]}} \to \mathbf{A}$ as $N \to \infty$ in $L^2[0,1]^2$, then all the conditions in Assumption 2 are satisfied. Hence one can generate approximate control for the original network system.

Consider the following parameters: $\alpha = 2$, $\beta = 1.5$, $q = 3$, $q_T = 7$, $\eta = 3$, $n = 1$ and $N = 100$. The numerical example is shown in Figure 4.

A direct solution to the original N-dimensional network LQR problem involves solving an $N \times N$ dimensional Riccati equation. However, the graphon approximate control method here involves only solving $d+1$ scalar Riccati equations, where d is the number of non-zero eigenvalues of the graphon limit $\mathbf{A}$. If the network is extremely large in size and the limit $\mathbf{A}$ permits simple spectral representations, then the graphon control method would significantly reduce the computation complexity.

VIII. DISCUSSION

The basic assumptions justifying the application of graphon control strategies are, first, that a given sequence of finite network systems converges to a unique limit graphon system (as in Definition 2) or that a given instance can be closely approximated by a graphon system, along with the measure

(a) This is a network of size 100 in a sequence that converges to the graphon **A**. This figure illustrates the structure, spectral properties and the step function representation.

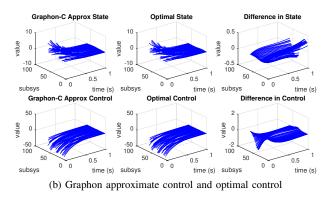(b) Graphon approximate control and optimal control

Fig. 4. Graphon approximate control and optimal control are applied to a network of size 100 in the sequence converging to the graphon limit. With the graph interpreted as an $L^2[0,1]^2$ function, the distance between the graph and the graphon limit is 0.061 in the operator norm and is 0.109 in the $L^2[0,1]^2$ norm. The maximum trajectory difference from the optimal control is less than 3.320% of the maximum of initial states. The graphon approximate control cost is 0.087% higher than the optimal LQR control cost.

preserving bijections that achieve the best fit, and second, that the corresponding control problem for the (limit) graphon system is tractable. Under these assumptions, Theorems 7 and 10 guarantee the effectiveness graphon control strategies for the finite large-scale complex network systems.

A plausible empirical approach to model the required infinite limit graphon G_∞ is to fit two dimensional Fourier series to the step function representation of the adjacency matrix. Such parametric modelling of empirical data could resemble parametric estimation in statistics and system identification. Moreover, due to the compactness of graphon operators, representations or approximations by simple spectral decomposition are possible [38] and will be analyzed in future work.

The generation of the graphon approximation models inevitably deals with relabellings. Although in the graphon control design methodology we do not restrict the labeling to be that of the best fit to the data, the control error still depends on the labeling of the nodes. Furthermore, the labeling of the nodes on the networks is necessary for control implementation. To find the best labellings for general graphs can be a complex combinatorial task. Consequently, we underline that it is assumed in this paper that the best labeling is known beforehand, either through a specific way of growing the networks with labels that ensure the best fit to the limit or through graphon estimation methods [39].

IX. CONCLUSION

We propose a method to approximately control networks of linear systems using the inherent limit described by graphons. Important aspects requiring further investigations include: (1) the application of the proposed limit graphon control strategy

to asymmetric network systems where the interactions of dynamics are described by directed networks; (2) the creation of an equivalent theory for sparse networks to the dense case developed here; (3) the generation of a methodology for systematically fitting bivariate analytic models to network data; (4) the application of graphon control to stochastic linear quadratic Gaussian problems; (5) the analysis of decentralized graphon control via Mean Field Game theory [17]; (6) the graphon control analysis to problems with non-symmetric local dynamic such as harmonic oscillator dynamics [40].

APPENDIX A
LEMMAS 6-8

Lemma 6 *Consider a step function $\mathbf{A}^{[\mathbf{N}]} \in \mathcal{W}_1$ defined via a partition $P = \{P_1, ..., P_{nN}\}$ and $\mathbf{u}_\tau^{[\mathbf{N}]} \in L^2_{pwc}[0,1]$ defined via the same partition P by*

$$\mathbf{u}_\tau^{[\mathbf{N}]}(\alpha) = nN \int_{P_i} \mathbf{u}_\tau(\beta)d\beta, \quad \forall \alpha \in P_i,$$

where $\mathbf{u}_\tau \in L^2[0,1]$. Then the following result holds:

$$(\mathbf{A}^{[\mathbf{N}]})^k \mathbf{u}_\tau^{[\mathbf{N}]} = (\mathbf{A}^{[\mathbf{N}]})^k \mathbf{u}_\tau, \quad k \geq 1. \qquad (67)$$

□

PROOF Let $A_{ij}^{[N]} = \mathbf{A}^{[\mathbf{N}]}(x,y)$, for all $(x,y) \in (P_i, P_j)$. Then for all $x \in P_i$,

$$\begin{aligned}
[\mathbf{A}^{[\mathbf{N}]}\mathbf{u}_\tau](x) &= \int_0^1 \mathbf{A}^{[\mathbf{N}]}(x,y)\mathbf{u}_\tau(y)dy \\
&= \sum_j \int_{P_j} A_{ij}^{[N]}\mathbf{u}_\tau(y)dy = \sum_j A_{ij}^{[N]} \int_{P_j} \mathbf{u}_\tau(y)dy \qquad (68) \\
&= \sum_j A_{ij}^{[N]} \cdot \mu(P_j) \cdot \mathbf{u}_\tau^{[\mathbf{N}]}(x),
\end{aligned}$$

$$\begin{aligned}
[\mathbf{A}^{[\mathbf{N}]}\mathbf{u}_\tau^{[\mathbf{N}]}](x) &= \sum_j \int_{P_j} A_{ij}^{[N]}\mathbf{u}_\tau^{[\mathbf{N}]}(y)dy, \\
&= \sum_j A_{ij}^{[N]} \cdot \mu(P_j) \cdot \mathbf{u}_\tau^{[\mathbf{N}]}(x).
\end{aligned} \qquad (69)$$

(68) and (69) give the equality $\mathbf{A}^{[\mathbf{N}]}\mathbf{u}_\tau^{[\mathbf{N}]} = \mathbf{A}^{[\mathbf{N}]}\mathbf{u}_\tau$, which immediately implies (67). ∎

Lemma 7 *For any graphon $\mathbf{W}$ or any function $\mathbf{W}$ in $L^2[0,1]^2$, $\|\mathbf{W}\|_{\text{op}} \leq \|\mathbf{W}\|_2$.*

□

PROOF

$$\|\mathbf{W}\|_{\text{op}} = \sup_{x \neq 0, x \in L^2[0,1]} \frac{\|\mathbf{W}x\|_2}{\|x\|_2}$$

$$= \sup_{x \neq 0, x \in L^2[0,1]} \frac{\sqrt{\int_0^1 \left[\int_0^1 \mathbf{W}(\alpha,\beta)x(\beta)d\beta\right]^2 d\alpha}}{\|x\|_2}$$

$$\leq \sup_{x \neq 0, x \in L^2[0,1]} \frac{\sqrt{\int_0^1 \left[\int_0^1 \mathbf{W}^2(\alpha,\beta)d\beta \int_0^1 x^2(\beta)d\beta\right] d\alpha}}{\|x\|_2}$$

$$= \sup_{x \neq 0, x \in L^2[0,1]} \frac{\|x\|_2\sqrt{\int_0^1 \int_0^1 \mathbf{W}^2(\alpha,\beta)d\beta d\alpha}}{\|x\|_2} = \|\mathbf{W}\|_2.$$

Lemma 8 *For any* $\mathbf{u} \in L^2[0,1]$, *any* $\mathbf{A} \in \mathcal{W}_1$, *and any* $t \in [0,T]$, *the following inequalities hold*

$$\left\| e^{\mathbf{A}t}\mathbf{u} \right\|_2 \le e^{t\|\mathbf{A}\|_{\mathrm{op}}}\|\mathbf{u}\|_2 \le e^{t\|\mathbf{A}\|_2}\|\mathbf{u}\|_2. \qquad \square$$

PROOF By recursively applying the definition of the operator norm, we obtain that $\|\mathbf{W}^k\|_{\mathrm{op}} \le \|\mathbf{W}\|_{\mathrm{op}}^k$, $k \in \mathbb{N} \cup \{0\}$, for any $\mathbf{W} \in \mathcal{W}_1$. Hence,

$$
\begin{aligned}
\left\| e^{\mathbf{A}t}\mathbf{u} \right\|_2 &\le \sum_{k=0}^{\infty} \frac{1}{k!} t^k \|\mathbf{A}^k \mathbf{u}\|_2 \le \sum_{k=0}^{\infty} \frac{1}{k!} t^k \|\mathbf{A}\|_{\mathrm{op}}^k \|\mathbf{u}\|_2 \\
&= e^{t\|\mathbf{A}\|_{\mathrm{op}}}\|\mathbf{u}\|_2 \le e^{t\|\mathbf{A}\|_2}\|\mathbf{u}\|_2. \quad \text{(by Lemma 7)}
\end{aligned}
$$
∎

Results in Lemma 6, Lemma 7 and Lemma 8 generalize to functions in any uniformly bounded subsets of $\mathcal{W}$, that is, any set of symmetric measurable functions $\mathbf{W} : [0,1]^2 \to I$ where I is a bounded interval in $\mathbb{R}$.

APPENDIX B
PROOFS OF GRAPHON SYSTEM PROPERTIES

A. Proof of Lemma 3

PROOF Let $P^{nN} = \{P_1, ..., P_{nN}\}$ be the uniform partition of $[0,1]$ with $P_i = [\frac{i-1}{nN}, \frac{i}{nN}), 1 \le i < nN$, and $P_{nN} = [\frac{nN-1}{nN}, 1]$. Consider any $\mathbf{x_s} \in L^2_{pwc}[0,1]$ and its corresponding vector $x \in \mathbb{R}^{nN}$ following the vetor-to-PWC-function mapping M_G. Since

$$[\mathbf{A}^{[\mathbf{N}]}\mathbf{x_s}](\alpha) = \int_0^1 \mathbf{A}^{[\mathbf{N}]}(\alpha, \beta)\mathbf{x_s}(\beta)d\beta, \quad \mathbf{x_s} \in L^2_{pwc}[0,1],$$

it follows that for all $\alpha \in P_i$,

$$
\begin{aligned}
[\mathbf{A}^{[\mathbf{N}]}\mathbf{x_s}](\alpha) &= \sum_{j=1}^{nN} \int_{P_j} \mathbf{A}^{[\mathbf{N}]}(\alpha, \beta)\mathbf{x_s}(\beta)d\beta \\
&= \sum_{j=1}^{nN} \int_{P_j} A_{Nij} x_j d\beta = \sum_{j=1}^{nN} \frac{1}{nN} A_{Nij} x_j \quad (70) \\
&= \frac{1}{nN}[A_N x]_i = [A_N \circ x]_i,
\end{aligned}
$$

where x_j denotes the jth element of $x \in R^{nN}$ and $[A_N x]_i$ denotes the ith element of $A_N x \in R^{nN}$. This implies that the step function $\mathbf{A}^{[\mathbf{N}]}$ in the graphon space, considered as an operator, represents a mapping in $L^2[0,1]$; this operator is equivalent to the matrix transformation A_N with $\circ$ operation in R^{nN} and the corresponding mapping M_G. A similar conclusion holds for $\mathbf{B}^{[\mathbf{N}]}$ and B_N. Furthermore, it is obvious that

$$\forall \gamma \in P_i, \quad \alpha_N \mathbb{I}\mathbf{x_s}(\gamma) = \alpha_N \mathbf{x_s}(\gamma) = \alpha_N x_i, \quad (71)$$

and a similar conclusion holds for $\beta_N \mathbb{I}$ and $\beta_N I$. Hence we conclude that the trajectory of the system $(\alpha_N I + A_N; \beta_N I + B_N)$ corresponds one-to-one to that of $(\alpha_N \mathbb{I} + \mathbf{A}^{[\mathbf{N}]}; \beta_N \mathbb{I} + \mathbf{B}^{[\mathbf{N}]})$ under the corresponding vetor-to-PWC-function mapping M_G. ∎

B. Proof of Theorem 4

PROOF Let us define $P_k(x, y) = \sum_{i=0}^{k} x^{k-i} y^i$. Then $x^k - y^k = (x - y)P_{k-1}(x, y)$. We obtain that for $k \ge 1$

$$\mathbf{A_N}^k - \mathbf{A_*}^k = P_{k-1}(\mathbf{A_N}, \mathbf{A_*})(\mathbf{A_N} - \mathbf{A_*}).$$

Since $\mathbf{A}^{(k-i-1)}\mathbf{A_*}^i \in \mathcal{W}_1$, for all $i \in \{0, 1, ...k-1\}$, we know that $\|\mathbf{A}^{(k-i-1)}\mathbf{A_*}^i\|_2 \le 1$. Hence, by Lemma 7,

$$\|P_{k-1}(\mathbf{A_N}, \mathbf{A_*})\|_{\mathrm{op}} \le \sum_{i=0}^{k-1} \|\mathbf{A}^{(k-i-1)}\mathbf{A_s}^i\|_2 \le k \cdot 1. \quad (72)$$

For an arbitrary $\mathbf{x} \in L^2[0,1]$ and finite t, $0 \le t < \infty$,

$$
\begin{aligned}
\left\| e^{\mathbf{A_N}t}\mathbf{x} - e^{\mathbf{A_*}t}\mathbf{x} \right\|_2 &\sum_{k=1}^{\infty} \frac{t^k}{k!}\|(\mathbf{A_N}^k - \mathbf{A_*}^k)\mathbf{x}\|_2 \\
&\le \sum_{k=1}^{\infty} \frac{t^k}{k!}\|P_{k-1}(\mathbf{A_N}, \mathbf{A_*})\|_{\mathrm{op}} \cdot \|\mathbf{A_\Delta^N}\|_{\mathrm{op}}\|\mathbf{x}\|_2 \quad (73) \\
&\le \sum_{k=1}^{\infty} \frac{t^k}{k!} \cdot k \cdot \|\mathbf{A_\Delta}\|_{\mathrm{op}}\|\mathbf{x}\|_2 = te^t \cdot \|\mathbf{A_\Delta^N}\|_{\mathrm{op}}\|\mathbf{x}\|_2.
\end{aligned}
$$

It follows that for $t \in [0,T]$

$$\left\| e^{\mathbf{A_N}t}\mathbf{x} - e^{\mathbf{A_*}t}\mathbf{x} \right\|_2 \le Te^T \cdot \|\mathbf{A_\Delta^N}\|_{\mathrm{op}}\|\mathbf{x}\|_2. \quad (74)$$

Hence the convergence is point-wise in time and uniform in t over $[0,T]$ and hence (20) holds. Furthermore,

$$
\begin{aligned}
\left\| e^{(\alpha_N \mathbb{I} + \mathbf{A_N})t}\mathbf{x} - e^{(\alpha \mathbb{I} + \mathbf{A_*})t}\mathbf{x} \right\|_2 &= \left\| e^{\alpha_N t}e^{\mathbf{A_N}t}\mathbf{x} - e^{\alpha t}e^{\mathbf{A_*}t}\mathbf{x} \right\|_2 \\
&\le \left\| e^{\alpha_N t}(e^{\mathbf{A_N}t} - e^{\mathbf{A_*}t})\mathbf{x} \right\|_2 + \left\| (e^{\alpha_N t} - e^{\alpha t})e^{\mathbf{A_*}t}\mathbf{x} \right\|_2 \\
&\le e^{\alpha_N t}te^t\|\mathbf{A_\Delta^N}\|_{\mathrm{op}}\|\mathbf{x}\|_2 + |\alpha - \alpha_N|te^{(L_\alpha + \|\mathbf{A_*}\|_{\mathrm{op}})t}\|\mathbf{x}\|_2,
\end{aligned}
$$
(75)

where $L_\alpha = \max\{|\alpha|, |\alpha_N|\}$. The last step of (75) is due to equation (73), Lemma 8 and the following

$$
\begin{aligned}
|e^{\alpha t} - e^{\alpha_N t}| &= \left| \sum_{k=0}^{\infty} \frac{1}{k!}[(\alpha)^k - (\alpha_N)^k]t^k \right| \\
&\le \sum_{k=1}^{\infty} \frac{1}{k!}|\alpha - \alpha_N| \cdot kL_\alpha^{(k-1)} t^k = |\alpha - \alpha_N|te^{tL_\alpha}
\end{aligned}
$$
(76)

where $t > 0$ and $L_\alpha = \max\{|\alpha|, |\alpha_N|\}$. An immediate implication of (75) is that for $t \in [0,T]$,

$$
\begin{aligned}
&\left\| e^{(\alpha_N \mathbb{I} + \mathbf{A_N})t}\mathbf{x} - e^{(\alpha \mathbb{I} + \mathbf{A_*})t}\mathbf{x} \right\|_2 \\
&\le e^{\alpha_N T}Te^T\|\mathbf{A_\Delta^N}\|_{\mathrm{op}}\|\mathbf{x}\|_2 + |\alpha - \alpha_N|Te^{(L_\alpha + \|\mathbf{A_*}\|_{\mathrm{op}})T}\|\mathbf{x}\|_2,
\end{aligned}
$$
(77)

By the convergence of $\{\alpha_N\}$, we know $\{\alpha_N\}$ and $\{L_\alpha\}$ are both uniformly bounded. This, together with (77), implies the convergence in (21) which is uniform in time over a closed time horizon $[0,T]$. ∎

APPENDIX C
PROOFS FOR EXACT CONTROLLABILITY

A. Proof of Theorem 5

PROOF Since any $\mathbf{A} \in \mathcal{W}_1$ defines a self-adjoint and compact operator, it has a discrete spectrum [15], and the maximum absolute value of eigenvalues of $\mathbf{A}$ equals to the operator norm

[29, Theorem 12.31], that is, $\|\mathbf{A}\|_{\text{op}} = \max_\ell |\lambda_\ell|$, where $\{\lambda_\ell\}$ denotes the set of eigenvalues of $\mathbf{A}$. For any $\mathbf{A} \in \mathcal{W}_1$, $\mathbf{A}$ is a bounded operator on $L^2[0,1]$, that is, there exists some finite $c_1 > 0$, such that $\|\mathbf{A}\|_{\text{op}} \leq c_1$. Therefore for $\ell \in \mathbb{N}$, $c_1 \geq \lambda_\ell \geq -c_1$ and hence for $t > 0$, $\|e^{\mathbf{A}t}\|_{\text{op}} \geq e^{\lambda_\ell t} \geq e^{-c_1 t} > 0$. Hence based on Lemma 1, for $t > 0$,

$$\|e^{\mathbb{A}t}\|_{\text{op}} = \|e^{(\alpha\mathbf{I}+\mathbf{A})t}\|_{\text{op}} = e^{\alpha t}\|e^{\mathbf{A}t}\|_{\text{op}} \geq e^{(\alpha-c_1)t} > 0. \quad (78)$$

This implies $e^{\mathbb{A}t}$ as an operator is uniformly positive definite.

Since all the values in the spectrum of $\mathbb{B}\mathbb{B}^\intercal$ as a self-joint operator are lower bounded by a positive constant, there exists $c > 0$ such that, for all $x \in L^2[0,1]$, $\langle \mathbb{B}\mathbb{B}^\intercal x, x \rangle \geq c\|x\|_2^2$. See e.g. [29, Theorem 12.12]. Consider the time horizon $[0, T]$. For any $h \in L^2[0,1]$,

$$\langle \mathbb{W}_T h, h \rangle = \int_0^T \langle \mathbb{B}\mathbb{B}^\intercal e^{\mathbb{A}^\intercal t}h, e^{\mathbb{A}^\intercal t}h \rangle dt \geq c\int_0^T \|e^{\mathbb{A}^\intercal t}h\|_2^2 dt$$
$$\geq cT(e^{(\alpha-c_1)T})^2\|h\|_2^2, \quad (79)$$

and hence the system $(\mathbb{A}; \mathbb{B})$ is exactly controllable. ∎

B. Proof of Proposition 2

PROOF By Lemma 7, since $\mathbf{A}$ and $\mathbf{B}$ are graphons in $\mathcal{W}_1$, there exists $c_1 \geq 0$ and $c_2 \geq 0$, such that

$$c_1 \geq \|\mathbf{A}\|_2 \geq \|\mathbf{A}\|_{\text{op}} \quad \text{and} \quad c_2 \geq \|\mathbf{B}\|_2 \geq \|\mathbf{B}\|_{\text{op}}. \quad (80)$$

Hence

$$\|e^{\mathbf{A}t}\|_{\text{op}} \leq \sup_{x \in L^2[0,1], \|x\|_2=1} \sum_{k=0}^\infty \frac{1}{k!}\|\mathbf{A}t\|_{\text{op}}^k \|x\|_2 \quad (81)$$
$$= e^{\|\mathbf{A}\|_{\text{op}}t} \leq e^{c_1 t}, \quad t \in [0, T].$$

Therefore

$$\|\mathbb{W}_T\|_2 \leq \int_0^T \left\|e^{\mathbf{A}t}\mathbf{B}\mathbf{B}^\intercal e^{\mathbf{A}^\intercal t}\right\|_2 dt \leq \int_0^T \left\|e^{\mathbf{A}t}\mathbf{B}\right\|_2^2 dt$$
$$\leq \int_0^T \left(\|e^{\mathbf{A}t}\|_{\text{op}}\|\mathbf{B}\|_2\right)^2 dt \leq T(e^{c_1 T}c_2)^2 < \infty, \quad (82)$$

which implies $\mathbb{W}_T \in L^2[0,1]^2$ and hence $\mathbb{W}_T$ is a compact (and self-joint) operator on $L^2[0,1]$ functions (see e.g. [41, Chapter 2, Proposition 4.7]). This means that $\mathbb{W}_T$ has a countable number of nonzero (real) eigenvalues $\{\lambda_1, \lambda_2, ...\}$ such that $\lambda_n \to 0$, and each eigenvalue has finite multiplicity (see e.g. [32]). Therefore $\mathbb{W}_T$ is not uniformly positive definite and hence the system $(\mathbf{A}; \mathbf{B})$ is not exactly controllable. ∎

APPENDIX D
PROOFS FOR STATE-TO-STATE GRAPHON CONTROL

Lemma 9 *Consider any* $\mathbf{u} \in L^2\left([0,T]; L^2[0,1]\right)$, $\mathbb{A}$, $\mathbb{A}^{[\mathbb{N}]} \in \mathcal{G}^1_{\mathcal{AT}}$, *and any* $t \in [0,T]$. *Let* $\mathbb{A} = (\alpha\mathbb{I} + \mathbf{A})$, $\mathbb{A}^{[\mathbb{N}]} = (\alpha_N\mathbb{I} + \mathbf{A}^{[\mathbb{N}]})$. *Then the following inequality holds*

$$\left\|\int_0^T \left(e^{\mathbb{A}(T-t)} - e^{\mathbb{A}^{[\mathbb{N}]}(T-t)}\right)\mathbf{u_t}dt\right\|_2$$
$$\leq |\alpha - \alpha_N|\int_0^T L_\alpha(T-t)e^{(L_\alpha+\|\mathbf{A}\|_{\text{op}})(T-t)}\|\mathbf{u_t}\|_2 dt \quad (83)$$
$$+ \|\mathbf{A}_\Delta^\mathbf{N}\|_{\text{op}}\int_0^T e^{(\alpha_N+1)(T-t)}(T-t)\|\mathbf{u_t}\|_2 dt$$

where $L_\alpha = \max\{|\alpha|, |\alpha_N|\}$ *and* $\mathbf{A}_\Delta^\mathbf{N} = \mathbf{A} - \mathbf{A}^{[\mathbb{N}]}$. □

PROOF An application of (18) in Theorem 4 leads to the above result. ∎

A. Proof of Theorem 7

PROOF

$$\|\mathbf{x_T}(\mathbf{u}) - \mathbf{x_T^N}(\mathbf{u}^{[\mathbb{N}]})\|_2$$
$$= \left\|\int_0^T e^{\mathbb{A}(T-t)}\mathbb{B}\mathbf{u_t}dt - \int_0^T e^{\mathbb{A}^{[\mathbb{N}]}(T-t)}\mathbb{B}^{[\mathbb{N}]}\mathbf{u_t^{[\mathbb{N}]}}dt\right\|_2$$
$$\leq \left\|\int_0^T [e^{\mathbb{A}(T-t)} - e^{\mathbb{A}^{[\mathbb{N}]}(T-t)}]\mathbb{B}\mathbf{u_t}dt\right\|_2$$
$$+ \left\|\int_0^T e^{\mathbb{A}^{[\mathbb{N}]}(T-t)}\left[\mathbb{B}\mathbf{u_t} - \mathbb{B}^{[\mathbb{N}]}\mathbf{u_t^{[\mathbb{N}]}}\right]dt\right\|_2$$
$$\leq |\alpha - \alpha_N|(|\beta| + \|\mathbf{B}\|_{\text{op}})$$
$$\int_0^T (T-t)e^{(L_\alpha+\|\mathbf{A}\|_{\text{op}})(T-t)}\|\mathbf{u_t}\|_2 dt$$
$$+ \|\mathbf{A}_\Delta^\mathbf{N}\|_{\text{op}}(|\beta| + \|\mathbf{B}\|_{\text{op}})\int_0^T e^{(\alpha_N+1)(T-t)}(T-t)\|\mathbf{u_t}\|_2 dt$$
$$+ \left\|\int_0^T e^{\mathbb{A}^{[\mathbb{N}]}(T-t)}\left[\mathbb{B}\mathbf{u_t} - \mathbb{B}^{[\mathbb{N}]}\mathbf{u_t^{[\mathbb{N}]}}\right]dt\right\|_2 \quad \text{(by Lemma 9)}$$
$$\leq |\alpha - \alpha_N|(|\beta| + \|\mathbf{B}\|_{\text{op}})$$
$$\int_0^T (T-t)e^{(L_\alpha+\|\mathbf{A}\|_{\text{op}})(T-t)}\|\mathbf{u_t}\|_2 dt$$
$$+ \|\mathbf{A}_\Delta^\mathbf{N}\|_{\text{op}}(|\beta| + \|\mathbf{B}\|_{\text{op}})\int_0^T e^{(\alpha_N+1)(T-t)}(T-t)\|\mathbf{u_t}\|_2 dt$$
$$+ (|\beta - \beta_N| + \|\mathbf{B} - \mathbf{B}^{[\mathbb{N}]}\|_{\text{op}})$$
$$\int_0^T e^{\left(\alpha_N+\|\mathbf{A}^{[\mathbb{N}]}\|_{\text{op}}\right)(T-t)}\|\mathbf{u_t}\|_2 dt$$
$$+ |\beta_N|\int_0^T e^{\left(\alpha_N+\|\mathbf{A}^{[\mathbb{N}]}\|_{\text{op}}\right)(T-t)}\left\|\mathbf{u_t} - \mathbf{u_t^{[\mathbb{N}]}}\right\|_2 dt. \quad (84)$$

The last step is due to Lemma 2 and the following

$$\left[\mathbb{B}\mathbf{u_t} - \mathbb{B}^{[\mathbb{N}]}\mathbf{u_t^{[\mathbb{N}]}}\right] = \left[(\beta\mathbb{I} + \mathbf{B})\mathbf{u_t} - (\beta_N\mathbb{I} + \mathbf{B}^{[\mathbb{N}]})\mathbf{u_t^{[\mathbb{N}]}}\right]$$
$$= \left[(\beta\mathbf{u} - \beta_N\mathbf{u_t^{[\mathbb{N}]}}) + \mathbf{B}\mathbf{u_t} - \mathbf{B}^{[\mathbb{N}]}\mathbf{u_t^{[\mathbb{N}]}}\right]$$
$$= \left[(\beta\mathbf{u} - \beta_N\mathbf{u_t^{[\mathbb{N}]}}) + \mathbf{B}\mathbf{u_t} - \mathbf{B}^{[\mathbb{N}]}\mathbf{u_t}\right] \quad \text{(by Lemma 6)}$$
$$= (\beta - \beta_N)\mathbf{u_t} + \beta_N(\mathbf{u_t} - \mathbf{u_t^{[\mathbb{N}]}}) + (\mathbf{B} - \mathbf{B}^{[\mathbb{N}]})\mathbf{u_t}. \quad (85)$$

By the convergence of $\{\alpha_N\}, \{\beta_N\}, \{\mathbf{A}^{[\mathbb{N}]}\}$ and $\{\mathbf{B}^{[\mathbb{N}]}\}$, we obtain that they are uniformly bounded. Based on Proposition 3, $\|\mathbf{u_t^{[\mathbb{N}]}}\|_2 \leq \|\mathbf{u_t}\|_2$, for any $t \in [0,T]$ and any $N \in \mathbb{N}$. Hence from (84), we obtain

$$\lim_{N\to\infty}\left\|\mathbf{x_T}(\mathbf{u}) - \mathbf{x_T^N}(\mathbf{u}^{[\mathbb{N}]})\right\|_2 = 0. \quad ∎$$

APPENDIX E
INVERSE OF THE CONTROLLABILITY GRAMIAN

Let $\mathbb{T}$ and $\mathbb{S}$ be linear bounded operators on a Hilbert space.

Proposition 6 ([32]) *Assume that $\mathbb{T}$ and $\mathbb{S}$ are symmetric and nonnegative. Then $\mathbb{I} + \mathbb{T}\mathbb{S}$ is one-to-one and onto; moreover $\|\mathbb{S}(\mathbb{I} + \mathbb{T}\mathbb{S})^{-1}\| \le \|\mathbb{S}\|$, and $\|(\mathbb{I} + \mathbb{T}\mathbb{S})^{-1}\| \le 1 + \|\mathbb{T}\|\|\mathbb{S}\|$.* □

Following this we prove the result on the existence of the inverse mapping of graphon controllability Gramian operator when the system is exactly controllable.

A. Proof of Theorem 8

PROOF If the graphon system $(\mathbb{A}; \mathbb{B})$ is exactly controllable, then

$$\forall h \in L^2[0,1], \quad \exists c_T > 0, \quad (\mathbb{W}_T h, h) \ge c_T \|h\|^2.$$

Let $\mathbb{I}$ denote the identity operator from $L^2[0,1]$ to $L^2[0,1]$. Let $\mathbb{M} = \mathbb{W}_T - \frac{1}{2}c_T\mathbb{I}$, then $\mathbb{W}_T = \frac{1}{2}c_T(\mathbb{I} + \frac{2}{c_T}\mathbb{M})$. By definition, the operator $\mathbb{M}$ is nonnegative and symmetric and hence $\frac{2}{c_T}\mathbb{M}$ is nonnegative and symmetric. By Proposition 6, $(\mathbb{I} + \frac{2}{c_T}\mathbb{M})$ is one-to-one and onto and the inverse operator is bounded. By a scaling factor $\frac{1}{2}c_T$, $\mathbb{W}_T = \frac{1}{2}c_T(\mathbb{I} + \frac{2}{c_T}\mathbb{M})$ is one-to-one and onto and hence the inverse operator $\mathbb{W}_T^{-1}$ exists. Since the scaling factor $\frac{1}{2}c_T$ is strictly positive and finite, the inverse operator $\mathbb{W}_T^{-1}$ is also bounded. ∎

B. Proof of Proposition 4

PROOF The controllability Gramian is given by

$$\mathbb{W}_T = \int_0^T e^{\mathbf{A}t}e^{\mathbf{A}^\intercal t}dt = \int_0^T \left(\mathbb{I} + \sum_{i=1}^\infty (2\mathbf{A}t)^i \frac{1}{i!}\right)dt$$
$$= T\mathbb{I} + \sum_{\ell \in I_\lambda}\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T\right)\mathbf{f}_\ell\mathbf{f}_\ell^\intercal. \quad (86)$$

Suppose $\mathbf{u} = \mathbb{W}_T\mathbf{x}$. We need to find the operator that maps $\mathbf{u}$ to $\mathbf{x}$. So we set

$$\mathbf{u} = \mathbb{W}_T\mathbf{x} = T\mathbf{x} + \sum_{\ell \in I_\lambda}\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T\right)\mathbf{f}_\ell\mathbf{f}_\ell^\intercal\mathbf{x}.$$

Therefore $\mathbf{x} = \frac{1}{T}\mathbf{u} - \frac{1}{T}\sum_{\ell \in I_\lambda}\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T\right)\mathbf{f}_\ell\mathbf{f}_\ell^\intercal\mathbf{x}$. From the definition of $\mathbf{u}$, we obtain $\mathbf{f}_\ell^\intercal\mathbf{u} = \left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1]\right)\mathbf{f}_\ell^\intercal\mathbf{x}$. Therefore

$$\mathbf{x} = \frac{1}{T}\mathbf{u} - \frac{1}{T}\sum_{\ell \in I_\lambda}\frac{\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1] - T\right)}{\left(\frac{1}{2\lambda_\ell}[e^{2\lambda_\ell T} - 1]\right)}\mathbf{f}_\ell\mathbf{f}_\ell^\intercal\mathbf{u}.$$

Equivalently, we obtain the result in (41). ∎

APPENDIX F
PROOFS FOR GRAPHON-LQR

A. Proof of Lemma 5

PROOF By Lemma 4 and the definition of the convergence in $C_s([0,T]; \Sigma(L^2[0,1]))$, we obtain that for any $\mathbf{x} \in L^2[0,1]$

$$\lim_{N\to\infty}\sup_{t\in[0,T]}\|\widetilde{\mathbb{P}}_t^\mathbb{N}\mathbf{x} - \mathbb{P}_t\mathbf{x}\|_2 = 0,$$
$$\lim_{N\to\infty}\sup_{t\in[0,T]}\|\mathbb{P}_t\mathbf{x} - \mathbb{P}_t^{[\mathbb{N}]}\mathbf{x}\|_2 = 0. \quad (87)$$

Since

$$\sup_{t\in[0,T]}\|\widetilde{\mathbb{P}}_t^\mathbb{N}\mathbf{x} - \mathbb{P}_t^{[\mathbb{N}]}\mathbf{x}\|_2$$
$$\le \sup_{t\in[0,T]}\|\widetilde{\mathbb{P}}_t^\mathbb{N}\mathbf{x} - \mathbb{P}_t\mathbf{x}\|_2 + \sup_{t\in[0,T]}\|\mathbb{P}_t\mathbf{x} - \mathbb{P}_t^{[\mathbb{N}]}\mathbf{x}\|_2, \quad (88)$$

we obtain $\lim_{N\to\infty}\sup_{t\in[0,T]}\|\widetilde{\mathbb{P}}_t^\mathbb{N}\mathbf{x} - \mathbb{P}_t^{[\mathbb{N}]}\mathbf{x}\|_2 = 0$. ∎

Lemma 10 *If a sequence $\{\mathbb{T}_N \in \mathcal{L}(L^2[0,1])\}$ of bounded linear operators for $L^2[0,1]$ functions converges strongly to $\mathbb{T} \in \mathcal{L}(L^2[0,1])$, that is,*

$$\forall \mathbf{x} \in L^2[0,1], \quad \lim_{N\to\infty}\|\mathbb{T}_N\mathbf{x} - \mathbb{T}\mathbf{x}\|_2 = 0, \quad (89)$$

then there exists $c > 0$ such that

$$\forall N \in \mathbb{N}, \quad \|\mathbb{T}_N\|_{\mathrm{op}} \le c. \quad \square$$

PROOF Consider any fixed $\mathbf{x} \in L^2[0,1]$ and an arbitrarily fixed $\varepsilon > 0$. The strong convergence of $\{\mathbb{T}_N\}$ implies there exist $N_0 > 0$ such that for $N > N_0$, $\|\mathbb{T}_N\mathbf{x} - \mathbb{T}\mathbf{x}\|_2 \le \varepsilon$, which implies $\|\mathbb{T}_N\mathbf{x}\|_2 \le \varepsilon + \|\mathbb{T}\mathbf{x}\|_2$ for $N > N_0$. Let $L = \max\{\|\mathbb{T}_N\mathbf{x}\|_2 : 1 \le N \le N_0\}$. Then $\|\mathbb{T}_N\mathbf{x}\|_2 \le \max\{L, \varepsilon + \|\mathbb{T}\mathbf{x}\|_2\}$. That is, for any fixed $\mathbf{x} \in L^2[0,1]$, $\|\mathbb{T}_N\mathbf{x}\|_2$ is uniformly bounded in N. Since $\mathbb{T}_N$ is a linear bounded operator from $L^2[0,1]$ to $L^2[0,1]$, the Uniform Boundedness Principle applies here and hence $\|\mathbb{T}_N\|_{\mathrm{op}}$ is uniformly bounded in N. ∎

B. Proof of Theorem 10

PROOF The closed loop system with the optimal control law is given by

$$\dot{\mathbf{x}}_{\mathbf{t}}^{N*} = \left(\mathbb{A}^{[\mathbb{N}]} - \mathbb{B}^{[\mathbb{N}]}\mathbb{B}^{[\mathbb{N}]\intercal}\mathbb{P}_{(T-t)}^{[\mathbb{N}]}\right)\mathbf{x}_{\mathbf{t}}^{N*}, \quad t \in [0,T]; \quad (90)$$

the closed loop system under the graphon approximate control law is given by

$$\dot{\mathbf{x}}_{\mathbf{t}}^{[\mathbf{N}]} = \left(\mathbb{A}^{[\mathbb{N}]} - \mathbb{B}^{[\mathbb{N}]}\mathbb{B}^{[\mathbb{N}]\intercal}\widetilde{\mathbb{P}}_{(T-t)}^\mathbb{N}\right)\mathbf{x}_{\mathbf{t}}^{[\mathbf{N}]}, \quad t \in [0,T]. \quad (91)$$

The initial conditions for (90) and (91) are given by $\mathbf{x}_0 \in L^2[0,1]$. Let $\mathbf{e}_{\mathbf{t}}^{\mathbf{N}} := \mathbf{x}_{\mathbf{t}}^{N*} - \mathbf{x}_{\mathbf{t}}^{[\mathbf{N}]}$. By (90) and (91), we obtain

$$\dot{\mathbf{e}}_{\mathbf{t}}^{\mathbf{N}} = \mathbb{F}_t^N\mathbf{e}_{\mathbf{t}}^{\mathbf{N}} + \mathbf{v}_{\mathbf{t}}^{\mathbf{N}},$$
$$\mathbb{F}_t^N := \left(\mathbb{A}^{[\mathbb{N}]} - \mathbb{B}^{[\mathbb{N}]}\mathbb{B}^{[\mathbb{N}]\intercal}\mathbb{P}_{(T-t)}^{[\mathbb{N}]}\right), \quad (92)$$
$$\mathbf{v}_{\mathbf{t}}^{\mathbf{N}} := \mathbb{B}^{[\mathbb{N}]}\mathbb{B}^{[\mathbb{N}]\intercal}\left(\mathbb{P}_{(T-t)}^{[\mathbb{N}]} - \widetilde{\mathbb{P}}_{(T-t)}^\mathbb{N}\right)\mathbf{x}_{\mathbf{t}}^{N*}.$$

Since $\mathbf{e}_0^\mathbf{N} = 0 \in L^2[0,1]$, the integral representation of (92) is given by $\mathbf{e}_{\mathbf{t}}^{\mathbf{N}} = \int_0^t \mathbf{v}_\tau^{\mathbf{N}}d\tau + \int_0^t \mathbb{F}_\tau^N\mathbf{e}_\tau^{\mathbf{N}}d\tau$. Hence we obtain

$$\|\mathbf{e}_{\mathbf{t}}^{\mathbf{N}}\|_2 \le \int_0^t \|\mathbf{v}_\tau^{\mathbf{N}}\|_2 d\tau + \int_0^t \|\mathbb{F}_\tau^N\|_{\mathrm{op}}\|\mathbf{e}_\tau^{\mathbf{N}}\|_2 d\tau. \quad (93)$$

Applying the Grönwall-Bellman inequality [42, p.7], we obtain

$$\|\mathbf{e}_{\mathbf{t}}^{\mathbf{N}}\|_2 \le \int_0^t e^{\int_s^t \|\mathbb{F}_\tau^N\|_{\mathrm{op}}d\tau}\|\mathbf{v}_s^{\mathbf{N}}\|_2 ds. \quad (94)$$

By the convergence of $\mathbb{A}^{[\mathbb{N}]}, \mathbb{B}^{[\mathbb{N}]}$ and $\mathbb{P}^{[\mathbb{N}]}$, we obtain that the limit $\mathbb{F} := \mathbb{A} - \mathbb{B}\mathbb{B}^\intercal\mathbb{P}_{(\cdot)}$ of the sequence $\{\mathbb{F}^N\}$ exists in $C_s([0,T]; \Sigma(L^2[0,1]))$, that is,

$$\mathbb{F}^N \to \mathbb{F} \text{ as } N \to \infty \text{ in } C_s([0,T]; \Sigma(L^2[0,1])); \quad (95)$$

furthermore, $\mathbb{F}$ is bounded in $C_s([0,T]; \Sigma(L^2[0,1]))$, that is,

$$\forall \mathbf{x} \in L^2[0,1], \quad \sup_{t \in [0,T]} \|\mathbb{F}_t \mathbf{x}\|_2 < \infty. \qquad (96)$$

By the convergence of $\{\mathbb{F}^N\}$ in $C_s([0,T]; \Sigma(L^2[0,1]))$, we obtain that for every $\mathbf{x} \in L^2[0,1]$ there exists $c > 0$ such that $\sup_{t \in [0,T]} \|\mathbb{F}_t^N \mathbf{x}\|_2 \leq c$ holds for all N. Therefore,

$$\forall \mathbf{x} \in L^2[0,1], \quad \sup_{\mathbb{T} \in \mathcal{F}} \|\mathbb{T} \mathbf{x}\|_2 < \infty, \qquad (97)$$

where $\mathcal{F} := \{\mathbb{F}_t^N : t \in [0,T], N \in \{1, \ldots, \infty\}\}$. By the Uniform Boundedness Principle [41, Chapter III, Theorem 14.1], there exists $c_F > 0$ such that

$$\sup_{\mathbb{T} \in \mathcal{F}} \|\mathbb{T}\|_{\mathrm{op}} \leq c_F. \qquad (98)$$

Furthermore, since $\{\mathbb{B}^{[\mathbb{N}]}\}$ converges strongly to $\mathbb{B}$, $\{\|\mathbb{B}^{[\mathbb{N}]}\|_{\mathrm{op}}\}$ is uniformly bounded in N by Lemma 10. Let c_B denote this uniform bound. This, together with (94) and (98), yields

$$
\begin{aligned}
\|\mathbf{e_t^N}\|_2 &\leq \int_0^t e^{(t-s)c_F} \|\mathbf{v}_s^{\mathbf{N}}\|_2 \, ds \\
&\leq c_B^2 e^{tc_F} \int_0^t \left\| \left(\mathbb{P}_{(T-s)}^{[\mathbb{N}]} - \widetilde{\mathbb{P}}_{(T-s)}^{\mathbb{N}}\right) \mathbf{x}_s^{N*} \right\|_2 \, ds \\
&\leq c_B^2 e^{tc_F} \int_0^t \left(\left\| \left(\mathbb{P}_{(T-s)}^{[\mathbb{N}]} - \widetilde{\mathbb{P}}_{(T-s)}^{\mathbb{N}}\right)\|_{\mathrm{op}} \|(\mathbf{x}_s^{N*} - \mathbf{x}_s^*)\right\|_2 \right. \\
&\qquad \left. + \left\| \left(\mathbb{P}_{(T-s)}^{[\mathbb{N}]} - \widetilde{\mathbb{P}}_{(T-s)}^{\mathbb{N}}\right)\mathbf{x}_s^* \right\|_2 \right) ds,
\end{aligned}
\qquad (99)
$$

where $\mathbf{x}^*$ denotes the trajectory of system $(\mathbf{A}^{[\mathbf{N}]}; \mathbf{B}^{[\mathbf{N}]})$ with the optimal control in (57) under the initial condition $\mathbf{x}_0^*$ which is the limit of the convergent sequence $\{\mathbf{x_0^{[N]}}\}$. Note that

$$\sup_{t \in [0,T]} \|\mathbf{x_t^{N*}}\|_2 = \sup_{t \in [0,T]} \|e^{\int_0^t \mathbb{F}_s^N ds} \mathbf{x}_0^{N*}\|_2 \leq e^{Tc_F} \|\mathbf{x}_0^{N*}\|_2, \qquad (100)$$

where the initial conditions $\{\mathbf{x}_0^{N*} \in L^2[0,1]\}$ are assumed to converge to $\mathbf{x}_0^*$. Hence

$$
\begin{aligned}
\sup_{t \in [0,T]} \|\mathbf{x_t^{N*}} - \mathbf{x_t^*}\|_2 &= \sup_{t \in [0,T]} \|e^{\int_0^t \mathbb{F}_s^N ds}(\mathbf{x}_0^{N*} - \mathbf{x}_0^*)\|_2 \\
&\leq e^{Tc_F} \|\mathbf{x}_0^{N*} - \mathbf{x}_0^*\|_2,
\end{aligned}
\qquad (101)
$$

that is, $\{\mathbf{x_t^{N*}}\}$ converges to $\mathbf{x_t^*}$ with respect to N and are uniformly bounded with respect to t over the horizon $[0,T]$.

Following a similar argument in (98), $\{\|\mathbb{P}_t^{[\mathbb{N}]} - \widetilde{\mathbb{P}}_t^{\mathbb{N}}\|_{\mathrm{op}}\}$ is uniformly bounded for all $t \in [0,T]$ and all $N \in \mathbb{N}$. This, together with the uniform convergence of $\{\mathbf{x_t^{N*}}\}$ to $\{\mathbf{x_t^*}\}$ in (101) and the result in Lemma 5, implies for any $t \in [0,T]$,

$$\lim_{N \to \infty} \|\mathbf{e_t^N}\|_2 = 0, \quad \text{i.e.} \quad \lim_{N \to \infty} \|\mathbf{x}_t^{N*} - \mathbf{x_t^{[N]}}\|_2 = 0. \qquad (102)$$

Next we prove the convergence of the cost function.

$$
\begin{aligned}
&\left| J(\mathbf{u}^{N*}) - J(\mathbf{u}^{[\mathbf{N}]}) \right| \\
&\leq \int_0^T \left(\|\mathbb{C}^{[\mathbb{N}]}(\mathbf{x}_t^{N*} + \mathbf{x_t^{[N]}})\|_2 \|\mathbf{x}_t^{N*} - \mathbf{x_t^{[N]}}\|_2 \right. \\
&\qquad + \|\mathbf{u}_t^{N*} + \mathbf{u_t^{[N]}}\|_2 \|\mathbf{u}_t^{N*} - \mathbf{u_t^{[N]}}\|_2 \Big) dt \\
&\qquad + \left\| \mathbb{P}_0^{[\mathbb{N}]}(\mathbf{x_T^{N*}} + \mathbf{x_T^{[N]}}) \right\|_2 \left\| \mathbf{x_T^{N*}} - \mathbf{x_T^{[N]}} \right\|_2
\end{aligned}
\qquad (103)
$$

Since $\mathbf{x}_0^{N*} = \mathbf{x_0^{[N]}}$, following a similar argument in (100) we obtain there exists $c_{F1} > 0$ such that

$$\sup_{t \in [0,T]} \|\mathbf{x_t^{[N]}}\|_2 \leq e^{Tc_{F1}} \|\mathbf{x}_0^{N*}\|_2. \qquad (104)$$

The convergence of $\{\mathbf{x}_0^{N*}\}$ implies that there exists a uniform bound $c_x > 0$ for $\mathbf{x}_t^{N*}$ and $\mathbf{x_t^{[N]}}$ for all $N \in \{1, 2, \ldots\}$ and all $t \in [0,T]$. The strong convergence of $\{\mathbb{C}^{[\mathbb{N}]}\}$ implies that $\{\|\mathbb{C}^{[\mathbb{N}]}\|_{\mathrm{op}}\}$ is uniformly bounded in N by some constant $c_C > 0$ (see Lemma 10). Therefore,

$$
\begin{aligned}
&\lim_{N \to \infty} \int_0^T \left(\|\mathbb{C}^{[\mathbb{N}]}(\mathbf{x}_t^{N*} + \mathbf{x_t^{[N]}})\|_2 \|\mathbf{x}_t^{N*} - \mathbf{x_t^{[N]}}\|_2 \right) dt \\
&\leq \lim_{N \to \infty} c_x c_C \int_0^T \|\mathbf{x}_t^{N*} - \mathbf{x_t^{[N]}}\|_2 dt \\
&\leq c_x c_C \int_0^T \lim_{N \to \infty} \|\mathbf{x}_t^{N*} - \mathbf{x_t^{[N]}}\|_2 dt = 0
\end{aligned}
\qquad (105)
$$

Recall

$$\mathbf{u}_t^{N*} = -\mathbb{B}^{[\mathbb{N}]} \mathbb{P}_{(T-t)}^{[\mathbb{N}]} \mathbf{x}_t^{N*}, \quad \mathbf{u_t^{[N]}} = -\mathbb{B}^{[\mathbb{N}]} \widetilde{\mathbb{P}}_{(T-t)}^{\mathbb{N}} \mathbf{x_t^{[N]}}. \qquad (106)$$

Note that $\{\|\mathbb{B}^{[\mathbb{N}]}\|_{\mathrm{op}}\}$ is uniformly bounded in N, and $\{\|\mathbb{P}_t^{[\mathbb{N}]}\|_{\mathrm{op}}\}$ and $\{\|\widetilde{\mathbb{P}}_t^{\mathbb{N}}\|_{\mathrm{op}}\}$ are uniformly bounded in N and in t. These, together with the fact that $\mathbf{x}_t^{N*}$ and $\mathbf{x_t^{[N]}}$ are uniformly bounded in N and t, imply that there exists a uniform bound $c_u > 0$ for $\{\mathbf{u}_t^{N*}\}$ and $\{\mathbf{u_t^{[N]}}\}$ in N and in t. Hence

$$
\begin{aligned}
&\int_0^T \left(\|\mathbf{u}_t^{N*} + \mathbf{u_t^{[N]}}\|_2 \|\mathbf{u}_t^{N*} - \mathbf{u_t^{[N]}}\|_2 \right) dt \\
&\leq c_u \int_0^T \|\mathbf{u}_t^{N*} - \mathbf{u_t^{[N]}}\|_2 dt \\
&\leq c_u c_B \int_0^T \left(\|\left(\mathbb{P}_{(T-t)}^{[\mathbb{N}]} - \widetilde{\mathbb{P}}_{(T-t)}^{\mathbb{N}}\right)\mathbf{x}_t^{N*}\|_2 \right. \\
&\qquad \left. + \|\widetilde{\mathbb{P}}_{(T-t)}^{\mathbb{N}}(\mathbf{x}_t^{N*} - \mathbf{x_t^{[N]}})\|_2 \right) dt.
\end{aligned}
\qquad (107)
$$

Hence,

$$\lim_{N \to \infty} \int_0^T \left(\|\mathbf{u}_t^{N*} + \mathbf{u_t^{[N]}}\|_2 \|\mathbf{u}_t^{N*} - \mathbf{u_t^{[N]}}\|_2 \right) dt = 0. \qquad (108)$$

A similar argument yields

$$\lim_{N \to \infty} \|\mathbb{P}_0^{[\mathbb{N}]}(\mathbf{x_T^{N*}} + \mathbf{x_T^{[N]}})\|_2 \|\mathbf{x_T^{N*}} - \mathbf{x_T^{[N]}}\|_2 = 0. \qquad (109)$$

Therefore, we have $\lim_{N \to \infty} |J(\mathbf{u}^{N*}) - J(\mathbf{u}^{[\mathbf{N}]})| = 0$. ∎

REFERENCES

[1] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.

[2] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective motion, sensor networks, and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.

[3] P. Ogren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1292–1302, 2004.

[4] N. J. Cowan, E. J. Chastain, D. A. Vilhena, J. S. Freudenberg, and C. T. Bergstrom, "Nodal dynamics, not degree distributions, determine the structural controllability of complex networks," *PLoS ONE*, vol. 7, no. 6, 06 2012.

[5] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[6] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, "Synchronization in complex networks," *Physics Reports*, vol. 469, no. 3, pp. 93–153, 2008.

[7] X. F. Wang and G. Chen, "Pinning control of scale-free dynamical networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 310, no. 3-4, pp. 521–531, 2002.

[8] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási, "Spectrum of controlling and observing complex networks," *Nature Physics*, vol. 11, no. 9, pp. 779–786, 2015.

[9] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, 2014.

[10] K. You and L. Xie, "Network topology and communication data rate for consensusability of discrete-time multi-agent systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2262–2275, 2011.

[11] X. F. Wang and G. Chen, "Synchronization in small-world dynamical networks," *International Journal of Bifurcation and Chaos*, vol. 12, no. 01, pp. 187–192, 2002.

[12] L. Lovász and B. Szegedy, "Limits of dense graph sequences," *Journal of Combinatorial Theory, Series B*, vol. 96, no. 6, pp. 933–957, 2006.

[13] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, "Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing," *Advances in Mathematics*, vol. 219, no. 6, pp. 1801–1851, 2008.

[14] ——, "Convergent sequences of dense graphs ii. multiway cuts and statistical physics," *Annals of Mathematics*, vol. 176, no. 1, pp. 151–219, 2012.

[15] L. Lovász, *Large Networks and Graph Limits*. American Mathematical Soc., 2012, vol. 60.

[16] F. Parise and A. Ozdaglar, "Graphon games," *arXiv preprint arXiv:1802.00080*, 2018.

[17] P. E. Caines and M. Huang, "Graphon mean field games and the GMFG equations," in *Proceedings of the 57th IEEE Conference on Decision and Control (CDC)*, December 2018, pp. 4129–4134.

[18] M. W. Morency and G. Leus, "Signal processing on kernel-based random graphs," in *Proceedings of the 25th European Signal Processing Conference (EUSIPCO)*, 2017, pp. 365–369.

[19] M. Avella-Medina, F. Parise, M. Schaub, and S. Segarra, "Centrality measures for graphons: Accounting for uncertainty in networks," *IEEE Transactions on Network Science and Engineering*, 2018.

[20] G. S. Medvedev, "The nonlinear heat equation on dense graphs and graph limits," *SIAM Journal on Mathematical Analysis*, vol. 46, no. 4, pp. 2743–2766, 2014.

[21] S. Gao and P. E. Caines, "The control of arbitrary size networks of linear systems via graphon limits: An initial investigation," in *Proceedings of the 56th IEEE Conference on Decision and Control (CDC)*, Melbourne, Australia, December 2017, pp. 1052–1057.

[22] ——, "Graphon-LQR control of arbitrary size networks of linear systems," in *Proceedings of the 23rd International Symposium on Mathematical Theory of Networks and Systems*, Hong Kong, China, July 2018, pp. 120–127.

[23] ——, "Graphon linear quadratic regulation of large-scale networks of linear systems," in *Proceedings of the 57th IEEE Conference on Decision and Control (CDC)*, Miami Beach, FL, USA, December 2018, pp. 5892–5897.

[24] P. Orbanz and D. M. Roy, "Bayesian models of graphs, arrays and other exchangeable random structures," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 2, pp. 437–461, 2014.

[25] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao, "An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions," *arXiv preprint arXiv:1401.2906*, 2014.

[26] S. Gao and P. E. Caines, "Optimal and approximate solutions to linear quadratic regulation of a class of graphon dynamical systems," Accepted by *the 58th IEEE Conference on Decision and Control (CDC)*, December 2019.

[27] R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras Volume I: Elementary Theory*. American Mathematical Society, 1997, vol. 15.

[28] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, ser. Applied Mathematical Sciences. New York: Springer, 1983.

[29] W. Rudin, *Functional Analysis*, ser. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991.

[30] S. Janson, "Graphons, cut norm and distance, couplings and rearrangements," *arXiv preprint arXiv: 1009.2376*, 2010.

[31] B. Szegedy, "Limits of kernel operators and the spectral regularity lemma," *European Journal of Combinatorics*, vol. 32, no. 7, pp. 1156–1167, 2011.

[32] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. Mitter, *Representation and Control of Infinite Dimensional Systems*, 2nd ed. Springer Science & Business Media, 2007.

[33] R. E. Showalter, *Monotone operators in Banach space and nonlinear partial differential equations*. American Mathematical Soc., 1997, vol. 49.

[34] R. F. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*. Springer Science & Business Media, 1995, vol. 21.

[35] E. Hewitt and K. Stromberg, *Real and Abstract Analysis: a Modern Treatment of the Theory of Functions of a Real Variable*. Springer-Verlag, 1965.

[36] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi, "Limits of randomly grown graph sequences," *European Journal of Combinatorics*, vol. 32, no. 7, pp. 985–999, 2011.

[37] S. Gao and A. Mahajan, "Networked control of coupled subsystems: Spectral decomposition and low-dimensional solutions," Accepted by *the 58th IEEE Conference on Decision and Control (CDC)*, December 2019.

[38] S. Gao and P. E. Caines, "Spectral representations of graphons in very large network systems control," Accepted by *the 58th IEEE Conference on Decision and Control (CDC)*, December 2019.

[39] E. M. Airoldi, T. B. Costa, and S. H. Chan, "Stochastic blockmodel approximation of a graphon: Theory and consistent estimation," in *Advances in Neural Information Processing Systems*, 2013, pp. 692–700.

[40] S. Gao and P. E. Caines, "Graphon control and graphon regulation of large-scale networks of linear systems with asymmetric local dynamics," May 2019, Centre for Intelligent Machines (CIM) and ECE Research Report.

[41] J. B. Conway, *A Course in Functional Analysis*, 2nd ed. Springer-Verlag New York, 1990, vol. 96.

[42] J. Bebernes, *Differential Equations Stability, Oscillations, Time Lags (A. Halanay)*. Society for Industrial and Applied Mathematics, 1968.

Shuang Gao (M'14) received the B.E. degree in automation and M.S. in control science and engineering, from Harbin Institute of Technology, Harbin, China, in 2011 and 2013. He received the Ph.D. degree in electrical engineering from McGill University, Montreal, QC, Canada, in February 2019, under the supervision of Prof. Peter. E. Caines. He is a member of McGill Centre for Intelligent Machines and Groupe d'Études et de Recherche en Analyse des Décisions. His research interest includes control of network systems, optimization on networks, decentralized control, network modeling, mean field games.

Peter E. Caines (LF'11) received the BA in mathematics from Oxford University in 1967 and the PhD in systems and control theory in 1970 from Imperial College, University of London, under the supervision of David Q. Mayne, FRS. After periods as a postdoctoral researcher and faculty member at UMIST, Stanford, UC Berkeley, Toronto and Harvard, he joined McGill University, Montreal, in 1980, where he is Distinguished James McGill Professor and Macdonald Chair in the Department of Electrical and Computer Engineering. In 2000 the adaptive control paper he coauthored with G. C. Goodwin and P. J. Ramadge (IEEE Transactions on Automatic Control, 1980) was recognized by the IEEE Control Systems Society as one of the 25 seminal control theory papers of the 20th century. He is a Life Fellow of the IEEE, and a Fellow of SIAM, the Institute of Mathematics and its Applications (UK) and the Canadian Institute for Advanced Research and is a member of Professional Engineers Ontario. He was elected to the Royal Society of Canada in 2003. In 2009 he received the IEEE Control Systems Society Bode Lecture Prize and in 2012 a Queen Elizabeth II Diamond Jubilee Medal. Peter Caines is the author of Linear Stochastic Systems, John Wiley, 1988, republished as a SIAM Classic in 2018, and is a Senior Editor of Nonlinear Analysis-Hybrid Systems; his research interests include stochastic, mean field game, decentralized and hybrid systems theory, together with their applications in a range of fields.