Personalized Feedback Control, Social Contracts, and Compliance Strategies for Ensembles

Pietro Ferraro, Lianna Zhao, Christopher King, and Robert Shorten

Abstract—This article describes the use of distributed ledger technologies as a means to create personalized social nudges and to influence the behavior of agents in a smart city environment. Specifically, we present a scheme to price personalized risk in sharing economy applications. We provide proofs for the convergence of the proposed stochastic system and we validate our approach through the use of extensive Monte Carlo simulations.

Index Terms—Control theory, cyber–physical systems, distributed ledger technologies (DLTs), smart cities, social impacts.

I. INTRODUCTION

Our objective in this article is to develop a framework to underpin the design of new forms of social contracts to guide the interaction of citizens, IoT devices, and IoT-enabled urban infrastructures. Specifically, we wish to develop control theoretic algorithms that can be deployed using distributed ledger technologies (DLTs) that nudge agents to respect social contracts in a range of Sharing Economy applications. Such technologies are of great relevance in a range of Sharing Economy applications that often require the good behavior and compliance of humans. For example, people should return shared assets in a timely manner as promised, and in good condition. Our basic proposal is to deploy a number of digital tokens as a bond or deposit to incentivize compliance: if the agent remains in compliance then these tokens are returned; otherwise, they, or some part of them, are lost [1]. As such, our algorithms can be viewed as mechanisms to realize personalized nudges that use financial rewards and penalties in order to direct how individual agents behave. Our contribution, in this specific context, is to develop DLT-enabled-feedback strategies that manage the level of compliance based on personalized (but anonymous) interventions.

To provide some context on when the need for such a system may arise, we describe briefly a vehicle-to-grid (V2G) battery-swapping architecture for shared e-bikes that we have designed in our lab. Roughly speaking, our smart battery system is a unit that both charges and aggregates the batteries from a number of e-bikes. For example, this could be part of a system in an apartment block that provides backup power in the building for essential services, as well as charging individual batteries. To see how the need for our compliance system arises in the context of this example, we first note that the batteries are financially valuable parts of the e-bike system. In our system, apartment owners would give residents access to an e-bike. Residents purchase tokens (whose value would exceed that of a battery) and use a digital deposit system as described above; namely, in order to release a battery, users would deposit a token into the ChargeWall, which would then be returned when the discharged battery is returned (the social contract). Please refer to Fig. 1 for a visual description of this procedure. In the remainder of this article, we describe methods to set the correct number of tokens to encourage bike users to comply with the social contract of returning a battery to the ChargeWall.

The issue of designing engineering systems to enforce compliance with social contracts is not new and has indeed become very topical recently in a number of fields. In particular, in the context of Covid 19 (social distancing and mask wearing), several systems based on machine learning have been proposed to encourage compliance with social contracts. Much of this recent work on this topic involves using AI techniques to identify noncompliant actors and to rely on some complementary strategy to enforce the social contract [2], [3], [4], [5]. Our approach is quite different and does not separate the identification of noncompliance from the enforcement of the social contract. Rather, as we have mentioned, we use a personalized pricing strategy to encourage agents to comply with social contracts using financial rewards and penalties. Thus, our work is related to the general area of Transactive Control; namely, the use of financial transactions as a feedback signal to improve the quality of service in various domains: some examples of work in this area can be found in [6], [7], [8], [9], [10], [11].
The basic idea presented in this article is to design algorithms which implement social policies, using distributed ledger technology (often referred to as DLTs; one example is Blockchain). We combine the notions of digital identity and smart contracts coming from DLTs with the rigorous design methods afforded by control theory, in order to design personalized interventions for ensembles of agents.

In our context, a Social Contract is a set of rules, or policies, designed to govern the interaction between humans, other humans, and societal infrastructures. For example, in the context of shared assets such as a pool of vehicles, a social contract might require that vehicles are returned to a specified location at a contracted time. Another example of a social contract is the requirement that plastic bottles are returned to point-of-sale after use.

The basic idea is to use digital tokens pegged at a stable fixed value, in the form of a cryptocurrency, to nudge users to comply with a social contract. To be more specific, let $E$ represent a general statement, such as *people should wear a face mask or a shared vehicle needs to be returned at a specified time and place*. The agent purchases some number of digital tokens in order to participate in the social contract $E$: if the agent behaves in compliance with the contract then tokens are returned in their entirety, whereas if the agent does not then some tokens are lost. Thus, the risk of losing tokens is the mechanism that encourages agents to comply with these social contracts.

A DLT is nothing more than a shared database. By its very nature, it is decentralized and therefore no central authority is required in order to achieve consensus amongst users. In DLTs, transactions are pseudo-anonymous, and their content can be encrypted. This allows every agent to own and manage access to the data present in their own transactions. In our setting, the only requirement is that the ownership of the tokens, used in the control, needs to remain visible to the compliance control algorithms, whereas other information (e.g., user quality of service and statistics on the usage of the system) can be encrypted. This allows each user to maintain ownership of their data and to use them as they please (e.g., to monetize them at a later stage). Token balances, and records of compliance on the ledger, associated with digital identities are the basis of the compliance system. Finally, in order to enable the compliance-based control, we will focus on the use of DLTs built around directed acyclic graphs (DAGs) such as IOTA [22], [23], [24]. These kinds of ledgers facilitate large transaction speeds and are fee-less. In contrast, many standard payment systems (e.g., VISA and Mastercard) and classical Blockchain architectures (e.g., Bitcoin and Ethereum) require users to pay a fee for each transaction. This makes such systems inadequate to serve as the backbone for the proposed compliance scheme. It is essential that the deposited token is returned in its entirety to the owner in the event of full compliance with rule $E$: the proposed social compliance mechanism would break down if the agent were required to pay a fee every time a token is deposited or returned, as this would effectively erode the value of her tokens over time.

The proposed architecture is shown in Fig. 2. The scheme is divided into three main components:

1) The Distributed Ledger discussed in this section, whose purpose is threefold: first, it acts as the communication backbone for the whole infrastructure, second, it enables the deposit mechanism for the digital bond, and third, it provides the controller with the current state of the network.

2) The Physical Layer, in which agents interact with their environment in the setting of the social contract $E$.

3) The Controller Layer, whose task is to regulate the price of the token bonds in order to achieve the desired level of compliance. Notice that the controller is not centralized as, due to the nature of the Distributed Ledger, every agent can implement the compliance scheme locally.

The latter two components of the architecture will be the focus of the next two sections. A complete discussion of DAG-based DLTs and their comparison to classical Blockchain is beyond the scope of this article; the interested reader can refer to [1], [22], [25], [26], and [27] for a thorough discussion of

1https://laurencetennant.com/papers/anonymity-iota.pdf
Fig. 2: Three components of the proposed compliance architecture are: the Distributed Ledger that acts as the communication backbone of the infrastructure, the compliance policy, and the feedback controller.

their properties. For the purpose of this article, all we need to assume is that there is a fast and secure way to execute the deposit and retrieval of these token bonds.

Remark: Before proceeding to the analysis and the modeling of the proposed framework, it is worth stressing that the issue of compliance is often not incorporated into algorithms which are designed to regulate, control, and optimize city infrastructures. Many studies addressing human behavior in this context assume full compliance with policies that have been engineered to optimally organize city infrastructures. As an example, consider traffic flow optimization: a crucial element that is often left out is that humans break rules, and the effect of this rule-breaking profoundly affects how cities operate and how well the engineered algorithms actually perform.

III. POLICY CHOICES FOR COMPLIANCE WITH SOCIAL CONTRACTS

As explained before we are interested in using DLT’s to create a type of digital bond which will encourage compliance with social contracts. Fig. 3 provides a visual representation of this basic idea. In order to engage with this social scheme each agent stakes an amount of tokens (to which some monetary cost is associated) that acts as a bond. This is shown in Fig. 3, by the arrow that goes from the agent to the transaction $T_C$ in the distributed ledger. The transaction $T_C$ represents the deposit of the tokens from the wallet of the agent to the wallet of a smart contract (these are distributed computer programs which execute as soon as certain conditions are met [28], [29]). Once the smart contract verifies that the agent complied with the rules of social contract $E$, it returns the funds to the agent’s wallet (through a subsequent transaction). Notice that due to the nature of smart contracts, the operations of deposit and return of the tokens are carried out automatically. All these operations are recorded on a DLT that is shared amongst all agents (anonymously).

A basic question that arises is how to price this bond: namely, how many tokens should be required as a bond in order to assure compliance with a social contract? Clearly, if this number is too low, one can expect low levels of compliance, and if it is too high, activity will cease and the social contract will be meaningless. In what follows, we shall develop a method for personalized pricing of the bond based on a feedback signal. The feedback signal will be designed so that aggregate levels of compliance satisfy some constraints. Before proceeding, we present two examples of social contracts and show how they lead to different policy choices. The first example is akin to the traffic signal situation mentioned in the previous section. In this kind of application, a desirable policy might be the following: if the agent behaves, their token is returned; otherwise, they lose their token. The second example concerns an agent who enters and moves within a public building (such as an airport or a train station) where the social contract might represent a rule such as keep your mask on. Clearly, in this type of contract, if the agent does not remove their mask then all tokens are returned. But what should happen if the agent does remove their mask? One policy might be to issue and redeem tokens at discrete intervals of time. Participants who break the contract multiple times would pay the bond repeatedly; this would also incentivize those individuals who remove their mask to wear it again so as to avoid further penalty. An alternative policy would be for an agent to lose their tokens, but for the pricing algorithm to operate at discrete intervals without further loss of tokens. In this case, the agent is incentivized to wear the mask again so as to keep their personalized price as low as possible. Below we summarize some policies that are of interest to us.

1) Fixed Penalty Policy: Before participating in the social scheme each agent deposits a certain amount of tokens, the amount being set by the controller. When the action is completed or when the agent exits the scheme, all
tokens are returned in the event that they complied with rule $E$; otherwise, no tokens are returned to the agent. In the latter case, the pricing algorithm continues to adjust the price based on both the agents’ level of compliance and that of the network.

2) *Adaptive Penalty Policies:* Initially, each agent deposits a certain amount of tokens, the amount being set by the controller. The contract is reissued at every time step. At each time step, compliant agents retrieve their tokens and stake new ones to continue the activity. Noncompliant agents lose all their tokens every time they do not comply. At all time steps, the pricing algorithm continues to adjust the price based on both the agents’ level of compliance and that of the network.

3) *Adaptive Penalty Policies With Return:* Initially, each agent deposits a certain amount of tokens, the amount being set by the controller. The contract is reissued at every time step. At each time step, compliant agents retrieve their tokens and stake new ones to continue the activity. Noncompliant agents lose all their tokens every time they do not comply. If an agent that previously lost a token starts complying again, they will retrieve a portion of the lost tokens. At all time steps, the pricing algorithm continues to adjust the price based on both the agents’ level of compliance and that of the network.

4) *Event Driven Policies:* Initially, each agent deposits a certain amount of tokens. Whenever the agent fails to comply with rule $E$ the tokens are lost; in order to keep participating in the scheme, the agent needs to deposit more tokens. In this version of the scheme, the amount of tokens that are required varies as a bond changes value over time (again a smart contract can easily take care of the update process).

Clearly, these are just four possible policies that might be adopted by the issuer of a social contract, and many others are possible. Our main contribution in this article is to develop a modeling and feedback control strategy to describe and enable a wide class of policies that include the four aforementioned ones.

**IV. Mathematical Framework**

As previously explained, we are interested in designing a feedback mechanism to avoid scenarios in which the value of the bond is either too low (leading to noncompliance) or too high (meaning that agents would not engage in the scheme for fear of losing their tokens). The issue of finding this value is the subject of this section. We will use typical elements of control theory in a stochastic environment where a large number of agents interact with one another and are subject to rule $E$.

Accordingly, we consider $n$ agents and, for each of them, we define dependent binary random variables $\{M_i(k) \in \{0, 1\}\}_{i=1}^{n}$, for discrete values of $k$, such that

$$\mathbb{P}(i \text{ complies with rule } E \text{ at time } k) = \mathbb{P}(M_i(k) = 1). \quad (1)$$

Moreover, we assume that the probability of these events is entirely dependent on a constant $q_i$, which represents the proclivity of each agent to comply with rules, and two control variables, $C(k)$, $c_i(k)$. The variable $C(k) + c_i(k)$ represents the value of the token bond staked by agent $i$ at time step $k$. The combination $q_i + C(k) + c_i(k)$ determines the likelihood that agent $i$ will comply with the rule at time step $k+1$. Then, (1) can be expressed as

$$\mathbb{P}(M_i(k+1) = 1) = p(q_i + C(k) + c_i(k)) \quad (2)$$

with $p : \mathbb{R} \rightarrow [0, 1]$ being a monotone increasing function (which is used to bind the probability between 0 and 1). $C(k)$ and $c_i(k)$ represent, respectively, a global and an individual feedback signal whose purpose is to regulate the behavior of each agent so as to achieve the desired level of compliance. However, due to the fact that the agents use a Distributed Ledger as a medium of communication, they can only access past levels of compliance $\{M_i(k-m)\}_{m=1}^{n}$, where $m$ is a delay in the measurements caused by POW, synchronization across ledgers and verification time [25]. Accordingly, we consider the following control laws $\forall k \in \mathbb{N}$ and $\forall i \in \{1, \ldots, n\}$:

$$C(k+1) = C(k) + \alpha \left( Q^* - n^{-1} \sum_{j=1}^{n} M_j(k-m) \right) \quad (3)$$

$$c_i(k+1) = c_i(k) + \beta \left( Q^* - M_i(k-m) \right)$$

with $\alpha > 0$ and $\beta > 0$ being two constants, $Q^* \in [0, 1]$ being the desired level of compliance and $\overline{M_i(k)}$ representing a windowed time average of the compliance of agent $i$, defined as

$$\overline{M_i(k)} = (1 - \gamma)^{-1} \sum_{j=1}^{k} \gamma^{k-j} M_i(j). \quad (4)$$

In this last expression, the factor $(1 - \gamma)^{-1}$ plays the role of the length of the window for the average, with $\gamma < 1$.

Notice that the proposed framework is very flexible and it would be possible to employ more sophisticated control laws. In this article, however, we limit ourselves to the study of a proportional action and the extension to more complex feedback loops will be the subject of a future work.

The reason to use both a global and an individual control signal, as opposed to just an individual or a global one, is that these two feedback signals achieve different complementary goals.

1) *Fairness:* Due to differences in individual behavior, some agents are going to comply with rules less than others. This means that if only a global shared signal was used to control the behavior of multiple agents, the signal would be driven up by the behavior of the least complying users, and this would result in an unfair price for the most *virtuous* agents. On the other hand, the introduction of a personalized cost ensures that individuals are going to be priced according to their own behavior (e.g., the less you comply the more you are going to pay, and vice versa).

2) *Distributed Trading of Compliance Levels:* While the presence of a global cost is not necessary to guarantee the desired level of average compliance in the presented framework (because if every agent’s compliance signal
were equal to the target value \( Q^* \) then the overall average compliance would be \( Q^* \), its absence would make the system vulnerable to the repeated misbehavior of malicious agents who purposely attempt to drive down the compliance level. The introduction of a global signal ensures that the system is able to achieve the desired level of compliance even in the presence of this kind of disturbance. In effect, the global cost allows compliant agents to compensate for noncompliant ones. This aspect is further explored in Section VI.

3) Pricing Attacks: In traditional pricing models, even one nefarious agent could, in principle, drive up the cost for all agents simply by misbehaving. In view of the previous comment, a natural concern is that similar effects might be possible in our schemes. Fortunately, in our scheme, such attacks are not possible. Even though noncompliant agents may drive up the control signal \( C(k) \) and hence drive up the cost of the bond, compliant agents will always recognize their full deposit after the expiration of a contract, leaving them unaffected by the increased price. On the other hand, noncompliant agents would continue to lose tokens while they drive \( C(k) \) to a high value. Furthermore, in the event that noncompliant agents prevent the desired levels of compliance being reached, \( c_i(k) \) may, in fact, tend to zero for compliant agents, further rewarding their behavior.

Remark: Before moving on, we want to point out that, since we are considering a DLT as the communication backbone of the whole architecture, the loss of a token is recorded on each agent’s copy of the ledger. This means that everyone is aware at all times (subject to delay) of the level of compliance of every other actor in the scheme (as the loss of the token by agent \( i \) at time \( k \) is recorded by the noncompliance value \( M_i(k) = 0 \)). Therefore, (3) can be validated by each user individually. Also, notice that (3) are consistent with the scenarios described in Section III.

V. THEORETICAL ANALYSIS

In this section, we provide a theoretical analysis of the convergence properties of the stochastic processes \( \{C(k), c_i(k), M_i(k)\} \). Recall that the cost functions evolve over one time step as

\[
C(k + 1) = C(k) + \alpha \left( Q^* - \frac{1}{n} \sum_{i=1}^{n} M_i(k-m) \right)
\]

\[
c_i(k + 1) = c_i(k) + \beta \left( Q^* - M_i(k-m) \right)
\]

and the time-averaged variable \( \overline{M_i(k)} \) satisfies the recursion formula

\[
\overline{M_i(k+1)} = \gamma \overline{M_i(k)} + (1 - \gamma) M_i(k+1).
\]

We will use the following ReLU-type choice for the agent probability function:

\[
p(x) = \begin{cases} 
  x, & \text{if } x \in (0,1) \\
  0, & \text{if } x \leq 0 \\
  1, & \text{if } x \geq 1.
\end{cases}
\]

Our main result is a convergence of probability for the time-averaged compliance variables \( \overline{M_i(k)} \) around the target value \( Q^* \) in the regime

\[
\alpha << \begin{pmatrix} \beta \\ 1 - \gamma \end{pmatrix} << 1.
\]

Accordingly, we introduce a small parameter \( \epsilon \) and a window size parameter \( w \), and we allow \( \alpha, \beta, \) and \( \gamma \) to scale with \( \epsilon, w \) as follows:

\[
\alpha = \epsilon^{3/2} \alpha_0, \quad \beta = \epsilon \beta_0 w, \quad 1 - \gamma = \epsilon w
\]

where \( \alpha_0 \) and \( \beta_0 \) are fixed constants. Our results will hold for \( \epsilon \) sufficiently small. We will also assume that the parameter \( \beta_0 \) satisfies the following condition:

\[
0 < \beta_0 \leq 1.
\]

Since the recursion equations (5) involve the time delay \( m \), it is necessary to include initial conditions

\[
\mathcal{C} = \left\{ c_i(k-m), C(k-m), M_i(k-m), \overline{M_i(k-m)} \right\}
\]

for all \( k = 0, \ldots, m \) and \( i = 1, \ldots, n \). We say \( \mathcal{C} \) is normalized if for all \( k = 0, \ldots, m \) and \( i = 1, \ldots, n \)

\[
q_i + C(k-m) + c_i(k-m) \in [0,1] \quad M_i(k-m) \in [0,1].
\]

Theorem 1: For any \( Q^* \in (0,1) \), suppose that the processes \( \{C(k), c_i(k), M_i(k)\} \) satisfy the recursion equations (5) and (6) and the parameters \( \alpha, \beta, \gamma \) satisfy the relations (7) and (8). Then, there are positive constants \( B, \epsilon_0 \) such that for all \( \epsilon \leq \epsilon_0 \), all normalized initial conditions \( \mathcal{C} \) satisfying (10), and all \( \delta > 0 \) and \( i \in \{1, \ldots, n\} \)

\[
\limsup_{k \to \infty} \mathbb{P}\left( |M_i(k) - Q^*| > \delta \right) \leq B \delta^{-2} \epsilon.
\]

Taking \( \delta = \epsilon^{1/3} \) we obtain that for all \( \epsilon \leq \epsilon_0 \)

\[
\limsup_{k \to \infty} \mathbb{P}\left( |M_i(k) - Q^*| > \epsilon^{1/3} \right) \leq B \epsilon^{1/3}.
\]

Remark: The result of Theorem 1 says that with high probability, the time average compliance level \( \overline{M_i(k)} \) will be found close to the target value \( Q^* \) for every agent \( i \), for all \( k \) sufficiently large. Note that the recursion equations (5) and (6) can be recast as stochastic approximation equations with constant stepsize \( \epsilon \). As such, our model does not fit into the Robbins–Monro setting [30] where the stepsize decreases to zero as \( k \to \infty \), and so we cannot expect to derive concentration-type results showing that the sequence \( \overline{M_i(k)} \) converges to a neighborhood of \( Q^* \) as \( k \to \infty \). Instead Theorem 1 is a statement about the long-term distribution of \( \overline{M_i(k)} \) and shows that the distribution is concentrated close to \( Q^* \) for all \( k \) sufficiently large. Note that our proof also shows that with high probability the cost function \( q_i + C(k) + c_i(k) \) will be found close to the target value \( Q^* \) for every agent \( i \), for large \( k \).

Remark: Although we chose a specific form for the agent probability function \( p \), our results can be extended to allow personalized probability functions \( p_i \) for each agent, where in each case, \( p_i : \mathbb{R} \to [0,1] \) is a nondecreasing uniformly Lipschitz function.
A. Proof of Theorem 1

In order to prove Theorem 1, we will first construct deterministic difference equations for each agent which approximate the stochastic recursion equations (5) and (6). We will then show that the solutions of the stochastic and deterministic equations remain close in $L^2$ norm as $k \to \infty$, and that the solution of the deterministic difference equations converges to a small neighborhood of the target value $Q^*$. The $L^2$ norm will be written as

$$\|X\|_{L^2} = \left(\mathbb{E}[\|X\|^2]\right)^{1/2}. \quad (13)$$

Let $\mathcal{F}(k)$ be the $\sigma$-algebra generated by the random variables $\{M_i(j) : 1 \leq i \leq n, 1 \leq j \leq k\}$. Then conditioned on $\mathcal{F}(k)$, the Bernoulli random variables $\{M_i(k+1) : i = 1, \ldots, n\}$ are independent with the distribution

$$\mathbb{P}(M_i(k+1) = 1 | \mathcal{F}(k)) = p(q_i + C(k) + c_i(k)). \quad (14)$$

We also define the variables $\{\xi_i(k)\}$ by

$$\xi_i(k) = M_i(k) - p(q_i + C(k-1) + c_i(k-1)).$$

It follows that:

$$\mathbb{E}[\xi_i(k+1) | \mathcal{F}(k)] = 0$$

and therefore $\{\xi_i(k)\}$ form a martingale difference sequence with respect to $\mathcal{F}(k)$. For future use, we also note the bound

$$\mathbb{E}[\xi_i(k+1)^2 | \mathcal{F}(k)] = \text{VAR}[M_i(k+1) | \mathcal{F}(k)] \leq \frac{1}{4}. \quad (15)$$

The recursion formula (6) can be written as

$$M_i(k+1) = Y \cdot M_i(k) + (1 - Y) \cdot M_i(k+1)$$

$$= M_i(k) + (1 - Y) \cdot \xi_i(k+1)$$

$$+ (1 - Y) \left[p(q_i + C(k) + c_i(k)) - M_i(k)\right]. \quad (16)$$

We now choose some index $i$ corresponding to a particular agent: this index will be fixed throughout the proof. We define for each $k$

$$Y_1(k) = \frac{M_i(k)}{M_i(k+1)} \quad (17)$$

$$Y_2(k) = q_i + C(k) + c_i(k) \quad (18)$$

and write $Y(k) = (Y_1(k), Y_2(k))^T \in \mathbb{R}^2$. We define the function $h = (h_1, h_2) : \mathbb{R}^2 \to \mathbb{R}^2$ as follows:

$$h_1(y) = w \cdot (p(y_2) - y_1) \quad (19)$$

$$h_2(y) = \beta_0 w \cdot (Q^* - y_1). \quad (20)$$

Using these definitions, we can rewrite the recursion equation (16) as follows:

$$Y_1(k+1) = Y_1(k) + (1 - Y) \cdot \xi_i(k+1)$$

$$+ (1 - Y) \left[p(q_i + C(k) + c_i(k)) - M_i(k)\right]$$

$$= Y_1(k) + w \cdot \xi_i(k+1) + \epsilon \cdot h_1(Y(k)) \quad (21)$$

and similarly from (5)

$$Y_2(k+1) = Y_2(k) + \epsilon \cdot h_2(Y(k-m)) + \epsilon^{3/2} \cdot G(k) \quad (22)$$

where

$$G(k) = a_0 \left(Q^* - \frac{1}{n} \sum_{i=1}^{n} M_i(k-m)\right). \quad (23)$$

The initial conditions $C$ as in (9) define initial conditions $\{Y(k-m), k = 0, \ldots, m\}$ for (21) and (22).

We next introduce deterministic difference equations by dropping terms from (21) and (22)

$$y_1(k+1) = y_1(k) + \epsilon \cdot h_1(y(k)) \quad (24a)$$

$$y_2(k+1) = y_2(k) + \epsilon \cdot h_2(y(k-m)) \quad (24b)$$

where $y(k) = (y_1(k), y_2(k)) \in \mathbb{R}^2$. It is evident that $(Q^*, Q^*)$ is the unique fixed point of this system. The initial conditions for (24) are $\{y(k-m), k = 0, \ldots, m\}$. Elementary estimates (using 0 $\leq p(y_2) \leq 1$) show that if $y_1(0) \in [0, 1]$, then $y_1(k)$ $\in [0, 1]$ for all $k$, and therefore $|h_1(y(k))| \leq w$ for all $k \geq 0$. Assuming also that $y_1(k-m) \in [0, 1]$ for all $k = 0, \ldots, m$, it follows that $|h_2(y(k))| \leq \beta_0 w \leq w$ for all $j \geq -m$, and therefore

$$\|y(j+1) - y(j)\| \leq 2w \epsilon \quad \text{for all } j \geq 0. \quad (25)$$

We will say that the initial condition $\mathcal{Y} = \{y(k-m)\}$ is bounded at level $M$ if

$$y_1(k-m) \in [0, 1], \quad y_2(k-m) \leq M \quad \forall k = 0, \ldots, m. \quad (26)$$

Note that (25) holds for any initial condition $\mathcal{Y}$ bounded at level $M$, for any $M \geq 0$.

The proof of Theorem 1 proceeds by first bounding the difference between the stochastic sequence $Y(k)$ and the deterministic sequence $y(k)$. We assume that $Y(k)$ and $y(k)$ share the same initial conditions $\mathcal{Y}$. Then

$$Y_1(k) - Y_1(0) = \epsilon \sum_{j=0}^{k-1} h_1(Y(j)) + \epsilon w \sum_{j=0}^{k-1} \xi_i(j+1) \quad (27a)$$

$$y_1(k) - y_1(0) = \epsilon \sum_{j=0}^{k-1} h_1(y(j)). \quad (27b)$$

Since $Y_1(0) = y_1(0)$ we get

$$Y_1(k) - y_1(k) = \epsilon \sum_{j=1}^{k-1} h_1(Y(j)) - h_1(y(j))$$

$$+ \epsilon w \sum_{j=0}^{k-1} \xi_i(j+1).$$

The function $h_1$ is uniformly Lipschitz on $\mathbb{R}^2$, and

$$|h_1(y) - h_1(z)| \leq 2w \|y - z\| \quad \text{for all } y, z \in \mathbb{R}^2.$$

Therefore

$$\|Y_1(k) - y_1(k)\|_{L^2} \leq 2 \epsilon w \sum_{j=1}^{k-1} \|Y(j) - y(j)\|_{L^2} + \epsilon \frac{1}{2} \|k^{1/2} \quad (28)$$

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where we used martingale orthogonality and (15) to deduce that
\[ \sum_{j=0}^{k-1} \xi(j+1)^2 \leq \sum_{j=0}^{k-1} \mathbb{E} [\xi(j+1)^2] \leq \frac{k}{4}. \]
Similarly
\[ Y_2(k) - y_2(k) = \varepsilon \sum_{j=0}^{k-1} h_2(Y(j-m)) - h_2(y(j-m)) + \varepsilon^{3/2} \sum_{j=0}^{k-1} G(j). \]
Noting that \( |G(k)| \leq \alpha_0 \) and \( |h_2(y) - h mass(2)| \leq w|y - z| \) we deduce
\[ \|Y_2(k) - y_2(k)\|_{L^2} \leq \varepsilon w \sum_{j=m+1}^{k-1} \|Y(j-m) - y(j-m)\|_{L^2} + \varepsilon^{3/2} \alpha_0 k \quad (29) \]
(where we used \( Y(j-m) = y(j-m) \) for all \( j = 0, \ldots, m \)). Combining (28) and (29) gives
\[ \|Y(k) - y(k)\|_{L^2} \leq \varepsilon w \sum_{j=1}^{k-1} \|Y(j) - y(j)\|_{L^2} + \varepsilon w \frac{1}{2} k^{1/2} + \varepsilon^{3/2} \alpha_0 k. \quad (30) \]
Therefore, by applying the discrete Grönwall inequality, we deduce that for all \( k \)
\[ \sup_{0 \leq j \leq k} \|Y(j) - y(j)\| \leq \varepsilon^{1/2} \left( \frac{(ek)^{1/2} w^{1/2} + \varepsilon k \alpha_0}{2} \right) e^{3wke}. \quad (31) \]
The bound (31) implies that \( Y(k) \) remains close to \( y(k) \) over time intervals \( k \sim \varepsilon^{-1} \), for any initial conditions shared by \( Y \) and \( y \). We next show that \( y(k) \) converges to a neighborhood of \( (Q^*, Q^*), \) using some ideas and techniques from [31, Ch. 9].

**Lemma 1:** We next show that \( y(k) \) converges to a neighborhood of \( (Q^*, Q^*), \) using some ideas and techniques from [31, Ch. 9]. Accordingly, consider the following lemma:

a) For any \( M \geq 0 \), there are positive constants \( B_1, \tau \) such that for all \( \varepsilon \) sufficiently small and all initial conditions \( Y \) bounded at level \( M \), the sequence \( y(k) \) defined by (24) satisfies
\[ \|Y(y) - (Q^*, Q^*)\| \leq \frac{1}{2} \|y(0) - (Q^*, Q^*)\| + B_1 \varepsilon. \quad (32) \]

b) There is \( M_2 > 2 \) such that for all \( \varepsilon \) sufficiently small and all initial conditions \( Y \) bounded at level 1, the sequence \( Y(k) \) defined by (21) and (22) satisfies
\[ 0 \leq Y_1(k) \leq 1, -M_2 \leq Y_2(k) \leq M_2 \quad \text{for all} \ k \geq 0. \quad (33) \]

Lemma 1 will be proved in the Appendix. We can now complete the proof of Theorem 1. As a convenient shorthand, we will write
\[ Y(k : k+l) = \{Y(k), Y(k+1), \ldots, Y(k+l)\} \quad (34) \]
to denote a sequence of successive states of the process. Also, we write \( \mathbb{E}_y \) to denote expected value with initial condition \( Y \) for the system (21), (22). We define \( J = [\varepsilon^{-1} \tau] \) where the number \( \tau \) is the value defined in Lemma 1-a) for \( M = M_2 \). Since the distribution of the increment of the system (21), (22) depends only on the previous \( m + 1 \) states, it follows that for any \( Y \) and integer \( k \geq 0 \), we have
\[ \mathbb{E}_y \left[ \|Y(k + J) - (Q^*, Q^*)\| \mid Y(k + J - m : k) = Y \right] \]
\[ = \mathbb{E}_y \left[ \|Y(J) - (Q^*, Q^*)\| \right]. \quad (35) \]
We define \( y(l) \) as the solution of (24) with initial conditions \( Y \), and define
\[ B_2 = \left( \frac{\tau^{1/2} w}{2} + \tau \alpha_0 \right) e^{3w\tau}. \quad (36) \]
The bound (31) implies that
\[ \mathbb{E}_y \left[ \|Y(J) - y(J)\| \right] \leq \varepsilon B_2^2. \quad (37) \]
Furthermore, from Lemma 1-b), we may assume that \( Y \) is normalized at level \( M_2 \), and so from Lemma 1-a), we deduce that
\[ \|y(J) - (Q^*, Q^*)\| \leq \frac{\sqrt{2}}{4} \|y(0) - (Q^*, Q^*)\|^2 + B_1^2 \varepsilon^2 \quad (38) \]
where \( B_1^2 = (2 + \sqrt{2}) B_1^2 \). Combining (37) and (38) gives
\[ \mathbb{E}_y \left[ \|Y(J) - (Q^*, Q^*)\|^2 \right] \]
\[ \leq \frac{\sqrt{2}}{2} \|y(J) - (Q^*, Q^*)\|^2 + (2 + \sqrt{2}) \mathbb{E}_y \left[ \|Y(J) - y(J)\|^2 \right] \]
\[ \leq \frac{\sqrt{2}}{2} \|y(0) - (Q^*, Q^*)\|^2 + \varepsilon B_2^2 \quad (39) \]
where (using \( \varepsilon < 1 \))
\[ B_2^2 = \left( 2 + \sqrt{2} \right) B_1^2 + \sqrt{2} B_1^2. \quad (40) \]
Substituting into (35) gives
\[ \mathbb{E}_y \left[ \|Y(k + J) - (Q^*, Q^*)\| \mid Y(k + J - m : k) \right] \]
\[ \leq \varepsilon B_2^2 + \frac{1}{2} \|Y(k) - (Q^*, Q^*)\|^2 \quad (41) \]
and therefore
\[ \mathbb{E}_y \left[ \|Y(k + J) - (Q^*, Q^*)\|^2 \right] \]
\[ \leq \varepsilon B_2^2 + \frac{1}{2} \mathbb{E}_y \left[ \|Y(k) - (Q^*, Q^*)\|^2 \right]. \quad (42) \]
For any \( k \geq 1 \) we write \( k = nJ + \hat{k} \) where \( J = [\varepsilon^{-1} \tau], n \geq 0 \) and \( 0 \leq \hat{k} \leq J - 1 \). Then, applying (42) recursively and using the bound (33) gives
\[ \mathbb{E}_y \left[ \|Y(k) - (Q^*, Q^*)\|^2 \right] \]
\[ \leq 2 \varepsilon B_2^2 + 2^n \mathbb{E}_y \left[ \|Y(J) - (Q^*, Q^*)\|^2 \right] \]
\[ \leq 2 \varepsilon B_2^2 + 2^n \left( 1 + (1 + M_2)^2 \right). \quad (43) \]
Since \( n \to \infty \) as \( k \to \infty \), it follows that:
\[ \limsup_{k \to \infty} \mathbb{E}_y \left[ \|Y(k) - (Q^*, Q^*)\|^2 \right] \leq 2 \varepsilon B_2^2. \quad (44) \]
Finally, using Markov’s inequality, we get
\[
\mathbb{P}(|\bar{M}_i(k) - Q^d| > \delta) = \mathbb{P}(|Y_1(k) - Q^d| > \delta) \\
\leq \mathbb{P}(\|Y(k) - (Q^d, Q^d)\| > \delta) \\
\leq \delta^{-2} \mathbb{E}[\|Y(k) - (Q^d, Q^d)\|^2] \tag{45}
\]
and (11) then follows by combining (44) and (45), with
\[
B = 2B_d^2. \tag{46}
\]

Remark: As a final remark, we comment on the relationship between the proposed control algorithms, and the class of algorithms that are commonly known as consensus algorithms. Consensus problems typically consider partial exchange of information between agents with asynchronous updates and look for local update rules so that all agents converge to a common value (or variations thereof). However, our problem is much simpler; agents have access to a global signal and full information from other agents via the distributed ledger. So even though the agents do agree asymptotically (if this is the pricing strategy), as in a consensus problem, there is no real explicit asynchronous exchange of partial information; rather all agents have full access to all agents’ information. Thus, the most natural formulation is the one we chose; that of a conventional feedback regulation problem with the main theoretical contribution being that of convergence of the feedback loop in probability.

VI. SIMULATIONS

In this section, we provide simulations for the proposed control scheme, to show the effectiveness of our approach and to highlight the need for both a global and individual signal. Specifically, we consider the following scenarios.

1. A scenario where only the global signal \( C(k) \) is used to regulate the behavior of each agent (i.e., \( c_i(k) = 0 \ \forall k \ \forall i \)).

2. A scenario where both global and individual signals are used to regulate the behavior of each agent.

3. A scenario where only the individual signal is used to regulate the behavior of each agent and a subset of individuals \( \{1, \ldots, n\} \) refuses to comply with rule \( E \) (i.e., \( \mathbb{P}(M_i(k) = 1) = 0 \ \forall k \ \forall i \in D \)).

4. A scenario where both the global and the individual signals are used to regulate the behavior of each agent and a subset of individuals \( \{1, \ldots, n\} \) refuses to comply with rule \( E \).

5. The same as scenario II but with increasing values of the delay \( m \).

VI. To analyze the robustness of the system in a more realistic scenario, where each agent is connected to a wireless network, we consider random delays (i.e., we allow the values of \( m \) to be drawn according to a distribution) and the random event that at each time step an agent might experience disconnection or package drops.

In all scenarios, we set \( n = 1000, \alpha = 0.025, \beta = 0.1, Q^d = 0.85, \gamma = 0.95, C(0) = 0, c_i(0) = 0 \ \forall i \in \{1, \ldots, n\} \) and for scenarios I–IV, we consider \( m = 3 \). Base compliance levels \( q_i \) are sampled from the uniform distribution in [0.1, 0.35].

Moreover, due to the stochastic nature of the system, for each scenario, we perform 150 Monte Carlo simulations and we average over the obtained realizations in order to obtain statistically meaningful results. Simulations are performed on MATLAB 2020a and Python 3.9.9.

A. Scenarios I and II

Figs. 4 and 5 show the results of the simulations in scenarios I and II. Even by visual inspection, it is clear that the lack of an individual signal to control the agents’ behavior leads to unfair results: while the overall compliance converges to \( Q^d \), this is achieved at the expense of the users that would behave better under normal circumstances (i.e., the users with larger \( q_i \)), that are forced to comply with a higher probability than \( Q^d \) in order to compensate for the behavior of the less compliant agents. Of course, this is undesirable and the use of the personalized cost, as shown in Fig. 5, tackles this problem by adjusting the individual price depending on the past behavior of each agent.

B. Scenarios III and IV

While in scenarios I and II, we explored how the lack of an individual cost leads to unfair results, it is less clear why a global cost is needed at all. In fact, (3) shows that when \( C(k) \) is set to zero, the personalized control signals would be sufficient to drive the average behavior to the desired level of compliance. Nevertheless, without the global cost, the system might fail to achieve the desired target for compliance in scenarios

Fig. 4. Compliance control using only the global signal. Compliance is enforced but at the expense of some agents that have to comply more than others.

Fig. 5. Compliance control using both the global and the individual signals. Compliance is achieved and the use of personalized signals makes it so that every agent contributes in a fair way.
when for some reason a certain number of agents fail to comply repeatedly with rule \( E \). This could be due to malfunctions or malicious behavior. This is highlighted in Fig. 6, related to scenario III, where 10\% of the agents, for \( k \leq 100 \), does not comply with rule \( E \) and the system is not able to achieve the desired level of compliance \( Q^* \). On the other hand, in scenario IV, shown in Fig. 7, it is possible to see that the presence of the global signal corrects this disturbance, thus making the system more robust to malfunctions and malicious behavior (of course, the drawback is that the honest agents will have to comply more in order to compensate for the misbehavior of the noncompliant users).

\section*{C. Scenario V}

As per the last set of simulations, we show the behavior of the system for increasing values of \( m \). Figs. 8–11 show that while for values \( m = 10 \) and \( m = 15 \), the system maintains stability, and the distributions of the average compliance and of each individual compliance accumulate around \( Q^* \), for larger values, such as \( m = 25 \) and \( m = 50 \), the system ends up oscillating without ever reaching an equilibrium. This shows that the delay introduced by the DLT represents a crucial design parameter and that, for fast-paced applications, the choice of an architecture that allows quick approvals is of paramount importance. Of course, the results of this last scenario are not meant to represent how a group of agents would behave, in the aforementioned circumstances. This scenario is merely showing a situation in which the control signal ends up failing due to the presence of the delay.

\section*{D. Scenario VI}

In our final set of experiments, we are interested in the behavior of the system when agents are connected to a network subject to heterogenous delays, such as a wireless or 5G network. In this scenario, the system will experience random delays due to latency and potential disconnections or packet loss. While simulations based on the use of tools, such as ns3 [32], would be a more accurate representation of reality in such situations, as we are interested in the impact of delay on the feedback control algorithms, we restrict ourselves to Monte Carlo-based Python simulations to evaluate the performance of our algorithms. Integration of our work into an ns3 environment will be the subject of future work.

Specifically, in this scenario, we allow for \( m \) to be drawn from a Gaussian distribution with mean \( \mu \) and standard deviation \( \sigma \). This simulates the delay experienced by an agent connected to a wireless network (e.g., 3G or 4G). Moreover, at time step \( k \), the level of compliance of agent \( i \) will remain the same as its level of compliance at time step \( k - 1 \), with probability \( \eta \). This simulates the possibility that either the agent
disconnects at time $k$ or the information never makes it to the controller.

Accordingly, Figs. 12–16 show the behavior of the system for different values of $\mu$, $\sigma$, and $\eta$. More specifically, we set $\sigma = \mu/2$ and $\eta = 0.025$ and we allow $\mu$ to vary between 5 and 45 with a step size of 10. Notice that the amount of time elapsed between time step $k$ and time step $k+1$ is application dependent and therefore it would not be meaningful to provide $\mu$ expressed in physical time units.

As in scenario V, the system maintains stability and the distributions of the average compliance and of each individual compliance accumulate around $Q^*$, for small values of $\mu$, whereas for values of $\mu > 50$, the system becomes unstable. The probability $\eta$ of an agent disconnecting or the information not reaching the controller does not affect the overall stability of the system. Notice that in scenario V, the system became unstable for values of $m > 25$. Interestingly, drawing $m$ from a Gaussian distribution, rather than being a constant value seems to increase the stability of the system. This aspect will be investigated in a future work.

This shows that, similarly to the role of the delay introduced by a DLT, in scenario V, the delay introduced by wireless networks represents a crucial design parameter and that, for fast-paced applications, the choice of an architecture that allows quick approvals is of paramount importance.

VII. CONCLUSION

In this article, we explored the use of a feedback control system to regulate the behavior of stochastic agents and to enforce the desired level of compliance, both globally and individually. The use of personalized feedback signals takes into account the behavior of each agent and leads to fair regulation, with respect to each individual’s base compliance $q_i$, whereas the global signal increases the robustness of the control system to malfunctions and malicious behavior. We proved a theorem that establishes that the averaged compliance of each agent, under the proposed regulation scheme, will accumulate around the target compliance $Q^*$, and finally we validated our results through extensive Monte Carlo simulations. As per future lines
of research, we intend to provide theoretical results for the robustness of the proposed compliance control against malicious actors, and explore different formulations of fairness to include, as an example, the economic status of each agent. Furthermore, we intend to extend our framework by using elements of game theory to take into account more complex scenarios. As a further strand of future work, we also wish to integrate our work into network simulators, such as ns3, perhaps in combination with mobility simulators, to provide more detailed experimental validation of the proposed techniques. Finally, we recognize that the compliance work presented here involves the co-design of technology and behaviors and that enforcing compliance involves exploring and managing the appetite for risk in agents (including human decision makers). This suggests a strong connection to control strategies that involve simultaneous exploration and policy enforcement (such as reinforcement learning). We have already commenced work in this direction and future publications will report on this work.

APPENDIX

1) Proof of Lemma 1: Both parts a) and b) of Lemma 1 will follow by showing that \( y(k) \) is well approximated by the solution of the following system of differential equations:

\[
\frac{dz_1}{dt} = h_1(z(t)) = w(p(z_2(t)) - z_1(t)) \tag{47}
\]
\[
\frac{dz_2}{dt} = h_2(z(t)) = \beta_0 w (Q^* - z_1(t)). \tag{48}
\]

Defining \( z(t) = (z_1(t), z_2(t)) \), we write this system as

\[
\frac{dz}{dt} = h(z(t)). \tag{49}
\]

It is easy to check that if \( z_1(0) \in [0, 1] \) then \( z_1(t) \in [0, 1] \) for all \( t \geq 0 \), and hence (since \( \beta_0 \leq 1 \))

\[
\| \frac{dz}{dt} \| = \| h(z(t)) \| \leq 2w. \tag{50}
\]

In order to make the connection between \( y(k) \) and \( z(t) \), we introduce

\[
\hat{y}(t_k) = y(k) \quad \text{where} \quad t_k = \epsilon k, \ k \geq -m \tag{51}
\]

and then define \( \hat{y}(t) \) for all \( t \geq -me \) using linear interpolation of the values at \( \{t_k\} \). We consider the solution of (24) with some initial condition satisfying (26)

\[
\hat{y}_1(t_k) = y_1(0) + \epsilon \sum_{j=0}^{k-1} h_1(y(j))
\]
Therefore, we have the bound
\[ FERRARO et al.: PERSONALIZED FEEDBACK CONTROL, SOCIAL CONTRACTS, AND COMPLIANCE STRATEGIES 3953 \]
\[ \begin{align*}
\|z(t) - (Q^*, Q^*)\| & \leq \frac{1}{2} \|z(0) - (Q^*, Q^*)\| \\
& \quad \text{for all } t \geq \tau_2(M). \\
\end{align*} \]

Before proving Lemma 2, we use it to prove Lemma 1. Given \( M \geq 0 \), define
\[ B_5 = \left( 6 w^2 \tau + B_4 \right) e^{3 w \tau} \]
where \( \tau = \tau_2(M) \) is the number defined in Lemma 2-b). For any \( \mathcal{Y} \) which is bounded at level \( M \), let \( y(k) \) be the solution of (49) with initial condition \( \gamma^\prime \), and let \( z(t) \) be the solution of (49) with initial value \( z(0) = y(0) \). Then, (57) implies that
\[ \|z(t) - \gamma(t)\| \leq B_5 \epsilon. \]

Therefore
\[ \|y\left(\left\lceil t \epsilon^{-1}\right\rceil\right) - (Q^*, Q^*)\| \leq \|z(t) - (Q^*, Q^*)\| + \|z(0) - (Q^*, Q^*)\| \]
and
\[ \frac{1}{2} \|z(0) - (Q^*, Q^*)\| + B_5 \epsilon + 2 w \epsilon \]
(62)

where \( B_1 = B_5 + 2 w \), and this establishes Lemma 1-a). For Lemma 1-b), note first that the bound \( Y_1(k) \in [0, 1] \) follows directly from (17). Also, the bound \( |Y_2(k)| \leq M_2 \) will be implied by a bound on the duration of any excursion outside the region \( |Y_2| \leq 2 \). Indeed consider integers \( u, v \) such that \( |Y_2(u - 1)| \leq 2 \) and
\[ |Y_2(u + k)| > 2 \] for \( k = 0, \ldots, v. \)

Let \( y(k) \) be the solution of (24) with initial conditions \( Y(u - m : u) \), and let \( z(t) \) be the solution of (49) with \( z(0) = y(0) \). It follows that \( \xi_1(u + 1) = 0 \), and so \( y_1(u + 1) = Y_1(u + 1) \) and \( Y_2(u + 1) - y_2(u + 1) \leq e^{3/2} \alpha_0 \). Also \( \xi_1(u + k) = 0 \) for \( k = 1, \ldots, v \), and so by similar reasoning it follows that:
\[ y_1(u + k) = Y_1(u + k) \]
\[ \|y_2(u + k) - Y_2(u + k)\| \leq e^{3/2} \alpha_0 k \]
(64)

for all \( k \leq v \) such that \( e^{3/2} \alpha_0 k \leq 1 \). Furthermore, (58) and (57) imply that
\[ \left| Y_2(u + k) \right| \leq 1 + \epsilon B_6 \]
(65)
where \( B_6 = (6 w^2 \tau_1 + B_4) e^{3 w \tau_1} \) and \( \tau_1 \) is defined in Lemma 2-a). For \( \epsilon \) sufficiently small, we have \( e^{3/2} \alpha_0 \tau_1 \leq 1 \). If we suppose that \( v > \frac{1}{2} \epsilon^{-1} \), then (64) and (65) would imply that for \( \epsilon \) sufficiently small
\[ \left| Y_2(u + k) \right| \leq 1 + e^{3/2} \alpha_0 \tau_1 + \epsilon B_6 < 2. \]
(66)

This contradicts (63), so we conclude that \( v \leq \tau_1 \epsilon^{-1} \). Therefore
\[ \sup_{0 \leq k \leq v} |Y_2(u + k)| \leq 2 + \tau_1 \epsilon^{-1} \left( e w + e^{3/2} \alpha_0 \right). \]
This holds for any excursion outside the region $|y_2| \leq 2$, and so we conclude that
\[ \sup_{k \geq 0} |y_2(k)| \leq 2 + \tau_1 \left( w + \epsilon^{1/2} \alpha_0 \right) \] (68)
which proves Lemma 1-b) with $M_2 = 2 + \tau_1 (w + \alpha_0)$.

2) Proof of Lemma 2: Lemma 2-a) is trivial for $M \leq 1$, so suppose $M > 1$. In the region $|z_2| \geq 1$, the solution of (49) has the form
\[ z_2(t) = z_2(0) - \beta_0 w (1 - Q^*) + \beta_0 (1 - z_1(0)) (1 - e^{-w t}). \] (69)
Taking $z_2(0) = M$, (69) shows that $z_2(t_1) = 1$ for some $t_1 \leq M (\beta_0 w (1 - Q^*))^{-1}$, and $z_2(t) \leq M + \beta_0$ for all $0 \leq t \leq t_1$. Similar reasoning for the region $\{z_2 \leq -M\}$ with $z_2(0) = -M$ shows that $z_2(t_2) = 0$ for some $t_2 \leq (M + 1) (\beta_0 w Q^*)^{-1}$, and $z_2(t) \geq -M - \beta_0$ for all $0 \leq t \leq t_2$. This establishes (58) with $\tau_1(M) = \max(\tau_1, t_2)$, and also shows that
\[ \sup_{t \geq 0} |z_2(t)| \leq M + \beta_0. \] (70)

To prove Lemma 2-b), we consider separately the cases $Q^* \geq 1/2$ and $Q^* < 1/2$. Suppose first that $Q^* \geq 1/2$, and consider the function
\[ V(z) = \beta_0 \left( z_1 - Q^* \right)^2 + (z_2 - Q^*)^2. \] (71)
Note that $\dot{V} = -2 \beta_0 w (z_1 - Q^*)^2$ in the square $[0, 1]^2$, so $V$ is a Lyapunov function in the square $[0, 1]^2$. We define
\[ E = \left\{ z : V(z) \leq \beta_0 \left( 1 - Q^* \right)^2 \right\}. \] (72)
The condition $\beta_0 \leq 1$ implies that $E \subset [0, 1]^2$. Therefore, if the solution $z(t)$ enters $E$ then it will remain thereafter inside the square $[0, 1]^2$, and thus its future evolution is determined by the linear system
\[ \dot{z}_1 = w (z_2 - z_1), \quad \dot{z}_2 = \beta_0 w \left( Q^* - z_1 \right). \] (73)
This linear system can be written in matrix form as follows:
\[ \frac{dz}{dt} = w A z + b, \quad A = \begin{pmatrix} -1 & 1 \\ -\beta_0 & 0 \end{pmatrix}, \quad b = w \beta_0 Q^* \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \] (74)
It is easy to see that the matrix $A$ is stable with eigenvalues
\[ \lambda_{\pm} = -\frac{1}{2} \pm \frac{1}{2} \sqrt{1 - 4 \beta_0^2}. \] (75)
Therefore, the condition $\beta_0 > 0$ implies that $\lambda_{\pm}$ have negative real parts and, therefore, the solution of (47) converges exponentially to the fixed point $(Q^*, Q^*)$. It remains to show that the solution $z(t)$ enters the set $E$ within a bounded time. Accordingly, we define three closed subsets of the square $[0, 1]^2$ as follows:
\[ S_1 = \left\{ z : 0 \leq z_1 \leq Q^*, 0 \leq z_2 \leq 1 \right\}, \]
\[ S_2 = \left\{ z : Q^* \leq z_1 \leq 1, Q^* \leq z_2 \leq 1 \right\}, \]
\[ S_3 = \left\{ z : Q^* \leq z_1 \leq 1, 0 \leq z_2 \leq Q^* \right\}. \]
It follows from the definition of $E$ that:
\[ S_2 \cap S_3 = \left\{ z : Q^* \leq z_1 \leq 1, z_2 = Q^* \right\} \subset E. \]

Inspection of the system (49) shows that the solution follows a trajectory that spirals clockwise around the fixed point $(Q^*, Q^*)$. If $z(0) \in S_1$ then $z(t)$ must eventually reach $S_2$, either directly from $S_1$ or after an excursion into the region $z_2 > 1$. We define $\tau_{12}$ to be the supremum over all starting points $z(0) \in S_1$ of the time until first entering the set $S_2$. These times depend continuously on $z(0)$ and $S_1$ is compact, therefore, $\tau_{12} < \infty$. Similarly, $\tau_{23} < \infty$ is the maximum time to reach $S_3$ starting from $S_2$, and $\tau_{31} < \infty$ is the maximum time to reach $S_1$ starting from $S_3$. Therefore, there is $\tau_{12} \leq \tau_{12} + \tau_{23} + \tau_{31}$ such that starting from any point $z(0)$ in $[0, 1]^2$, the solution $z(t)$ will reach the interval $S_2 \cap S_3$ at some time $t \leq \tau_{12}^*$. If the trajectory starts at a point $z(0)$ satisfying $1 \leq |z_2(0)| \leq M$ then (58) implies it must enter the square $[0, 1]^2$ before time $t_1$, and so must reach the interval $S_2 \cap S_3$ before time $t_1^* + t_1$.

Since $S_2 \cap S_3 \subset E$, this shows that the solution $z(t)$ enters the set $E$ within time $t^*_1 + t_1$, starting from any point in the region $|z_2| \leq M$. As noted before, the system is a contraction in $E$, so for any $r > 0$, there is some $\tau(r) < \infty$ such that for all $z(t) \in E$
\[ \|z(t + s) - (Q^*, Q^*)\| \leq r \|z(t) - (Q^*, Q^*)\| \text{ for all } s \geq \tau(r). \]
If $z(0) \in E$ this establishes (59) with $\tau_2(M) = \tau(1/2)$. So assume that $z(0) \notin E$. Let $R_1 = \{ z : 0 \leq z_1 \leq 1, |z_2| \leq M \}$ and define
\[ \rho = \sup_{z \in E, w \in E \setminus E} \frac{\|z - y^*\|}{\|w - y^*\|} \] (76)
Then for all $z(0) \in R_1 \setminus E$, we know that $z(t)$ reaches $E$ by latest time $t^*_2 + t_1$, and then is contracted after time $\tau(r)$. So for all $t \geq \tau(r) + t^*_2 + t_1$, we have
\[ \|z(t) - (Q^*, Q^*)\| \leq r \rho \|z(0) - (Q^*, Q^*)\|. \] (77)
Finally, we choose $r \leq (2\rho)^{-1}$ and $\tau_2(M) = \tau(r) + t^*_2 + t_1$ and conclude that for all $z(0) \in R_1$
\[ \|z(t) - (Q^*, Q^*)\| \leq \frac{1}{2} \|z(0) - (Q^*, Q^*)\| \text{ for all } t \geq \tau_2(M). \]
A similar argument applies when $Q^* \leq 1/2$, and this establishes Lemma 2-b).

REFERENCES

[1] P. Ferraro, C. King, and R. Shorten, “Distributed ledger technology for smart cities, the sharing economy and social compliance,” IEEE Access, vol. 6, pp. 62728–62746, 2018.

[2] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, “A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the Covid-19 pandemic,” Measurement, vol. 167, Jan. 2021, Art. no. 108288.

[3] E. Mbugue, S. Simelane, S. G. Fashoto, B. Akinnuwesi, and A. S. Meftula, “Application of deep learning and machine learning models to detect Covid-19 face masks—A review,” Sustain. Oper. Comput., vol. 2, Sep. 2021, Art. no. 235245.

[4] X. Kong et al., “Real-time mask identification for Covid-19: An edge-computing-based deep learning framework,” IEEE Internet Things J., vol. 8, no. 21, pp. 15929–15938, Nov. 2021.

[5] A. Sanal and G. Udupa, “Machine learning based human body temperature measurement and mask detection by thermal imaging,” in Proc. 3rd Int. Conf. Intell. Comput. Instrum. Control Technol. (ICICICT), vol. 10, 2022, pp. 1339–1343.
[6] T. Phan, A. M. Annaswamy, D. Yanakiev, and E. Tseung, “A model-based dynamic toll pricing strategy for controlling highway traffic,” in Proc. IEEE Amer. Control Conf. (ACC), 2016, pp. 6245–6252.

[7] B. Park, K. Amasyali, Y. Chen, and M. Olama, “Hierarchical transactive control of flexible building loads under distribution LMP,” in Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT), 2022, pp. 1–5.

[8] T. R. Nudell et al., “Distributed control for polygeneration microgrids: A Dynamic market mechanism approach,” Control Eng. Pract., vol. 121, Apr. 2022, Art. no. 105052.

[9] F. Soylemezgiller, M. Kuscu, and D. Kilinc, “A traffic congestion avoidance algorithm with dynamic road pricing for smart cities,” in Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC), 2013, pp. 2571–2575.

[10] K. T. Bui, V. A. Huynh, and E. Frazzoli, “Dynamic traffic congestion pricing mechanism with user-centric considerations,” in Proc. 15th Int. IEEE Conf. Intell. Transp. Syst., 2012, pp. 147–154.

[11] A. M. Annaswamy, Y. Guan, H. E. Tseung, H. Zhou, T. Phan, and D. Yanakiev, “Transactive control in smart cities,” Proc. IEEE, vol. 106, no. 4, pp. 518–537, Apr. 2018.

[12] S. Widergren, J. Fuller, C. Marinovici, and A. Somani, “Residential transactive control demonstration,” in Proc. IEEE PES Innov. Smart Grid Technol. Conf. (ISGT), 2014, pp. 1–5.

[13] P. Huang et al., “Analytics and transactive control design for the pacific northwest smart grid demonstration project,” in Proc. 1st IEEE Int. Conf. Smart Grid Commun. (SmartGridComm), 2010, pp. 449–454.

[14] A. O. Kotb, Y.-C. Shen, X. Zhu, and Y. Huang, “Parker—A new smart car-parking system based on dynamic resource allocation and pricing,” IEEE Trans. Intell. Transp. Syst., vol. 17, no. 9, pp. 2637–2647, Sep. 2016.

[15] Y. Yao and P. Zhang, “Transactive control of air conditioning loads for mitigating microgrid tie-line power fluctuations,” in Proc. IEEE PES Gen. Meeting, 2016, pp. 1–5.

[16] S. Katiyarpanala, “Smart buildings can help smart grid: Transactive controls,” in Proc. IEEE PES Innov. Smart Grid Technol. (ISGT), 2012, p. 1.

[17] J. Hu, G. Yang, K. Kok, Y. Xue, and H. W. Bindner, “Transactive control: A framework for operating power systems characterized by high penetration of distributed energy resources,” J. Mod. Power Syst. Clean Energy, vol. 5, no. 3, pp. 451–464, 2017.

[18] J. Li, X. Lin, S. Nazarian, and M. Pedram, “CTSM2: Concurrent task scheduling and storage management for residential energy consumers under dynamic energy pricing,” IET Cyber-Phys. Syst. Theory Appl., vol. 2, no. 3, pp. 111–117, 2017.

[19] D. A. Chekhirev, L. Khoulghi, and H. T. Mouftah, “Decentralized cloud-SDN architecture in smart grid: A dynamic pricing model,” IEEE Trans. Ind. Informat., vol. 14, no. 3, pp. 1220–1231, Mar. 2018.

[20] A. K. Bejestani, A. Annaswamy, and T. Samad, “A hierarchical transactive control architecture for renewables integration in smart grids: Analytical modeling and stability,” IEEE Trans. Smart. Grid, vol. 5, no. 4, pp. 2054–2065, Jul. 2014.

[21] H. Hao, C. D. Corbin, K. Kalsi, and R. G. Pratt, “Transactive control of commercial buildings for demand response,” IEEE Trans. Power Syst., vol. 32, no. 1, pp. 774–783, Jan. 2017.

[22] S. Popov, “The tangle-version 1.4.3,” Apr. 2018. [Online]. Available: https://iota.org/IOTA_Whitepaper.pdf

[23] S. Muller, A. Penzkofer, N. Polyanskiy, J. Theis, W. Sanders, and H. Moog, “Tangle 2.0 leaderless Nakamoto consensus on the heaviest DAG,” IEEE Access, vol. 10, pp. 105807–105842, 2022.

[24] S. Popov, O. Saa, and P. Finaridi, “Equilibria in the tangle,” 2017, arXiv:1712.05385.

[25] A. Cullen, P. Ferraro, C. King, and R. Shorten, “On the resilience of DAG-based distributed ledgers in IoT applications,” IEEE Internet Things J., vol. 7, no. 8, pp. 7112–7122, Aug. 2020.

[26] M. Moschella, P. Ferraro, E. Crisostomi, and R. Shorten, “Decentralized assignment of electric vehicles at charging stations based on personalized cost functions and distributed ledger technologies,” IEEE Internet Things J., vol. 8, no. 14, pp. 11112–11122, Jul. 2021.

[27] P. Ferraro, C. King, and R. Shorten, “On the stability of unverified transactions in a DAG-based Distributed Ledger,” IEEE Trans. Auton. Control, vol. 65, no. 5, pp. 3772–3783, Sep. 2020.

[28] V. Butler, “A next-generation smart contract and decentralized application platform,” Ethereum, Zug, Switzerland, White Paper, 2014.

[29] Z. Wang, H. Jin, W. Dai, K.-K. R. Choo, and D. Zou, “Ethereum smart contract security research: survey and future research opportunities,” Front. Comput. Sci., vol. 15, no. 2, 2021, Art. no. 152802.

[30] H. Robbins and S. Monro, “A stochastic approximation method,” Ann. Math. Stat., vol. 22, no. 3, pp. 400–407, 1951.

[31] V. S. Borkar, Stochastic Approximation a Dynamical Systems Viewpoint. New Delhi, India: Hindustan Book Agency, 2008.

[32] G. F. Riley and T. R. Henderson, “The NS-3 network simulator,” in Modeling and Tools for Network Simulation, K. Wehrle, and M. Güne¸s, and J. Gross, Eds. Heidelberg, Germany: Springer, 2010, pp. 15–34.

Pietro Ferrari received the Ph.D. degree in control and electrical engineering from the University of Pisa, Pisa, Italy, in 2018.

He is currently a Lecturer with the Dyson School of Design Engineering and Imperial-X, Imperial College London, London, U.K. His research interests include control theory applied to the sharing economy domain and distributed ledger technology and machine learning.

Lianna Zhao received the M.E. degree from Northeastern University, Shenyang, China, in 2020. She is currently pursuing the Ph.D. degree with the Dyson School of Design Engineering, Imperial College London, London, U.K.

She is sponsored by the IOTA Foundation, Berlin, Germany, and also works with IOTA Foundations Networking Research Team. Her research is about distributed ledger technology and its applications, such as Internet of Things system, smart cities, and social compliance.

Christopher King received the Ph.D. degree in physics from Harvard University, Cambridge, MA, USA, in 1984.

He was a Professor with the Mathematics Department, Northeastern University, Boston, MA, USA. His research interests include dynamical systems, quantum information theory, and mathematical physics.

Robert Shorten received the Ph.D. degree from University College Dublin, Dublin, Ireland, in 1996.

He is currently a Professor of Cyber–Physical Systems with the Dyson School of Design Engineering, Imperial College London, London, U.K. He is also with University College Dublin. He was a Co-Founder of the Hamilton Institute, Maynooth University, Maynooth, Ireland, and led the Optimization and Control Team, IBM Research Smart Cities Research Lab, Dublin. He was a Visiting Professor with the Berlin Institute of Technology, Berlin, Germany, and a Research Visitor with Yale University, New Haven, CT, USA, and Technion, Haifa, Israel. He is the coauthor of the recently published books AIMD Dynamics and Distributed Resource Allocation (SIAM, 2016) and Electric and Plug-In Vehicle Networks: Optimisation and Control (CRC Press, Taylor & Francis Group, 2017).

Prof. Shorten is the Irish Member of the European Union Control Systems Association, a member of the IEEE Control Systems Society Technical Committee on Smart Cities, and a member of the IFAC Technical Committees for Automotive Control and for Discrete Event and Hybrid Systems.