Near-Optimal Deviation-Proof Medium Access Control Designs in Wireless Networks

Khoa Tran Phan, Jaeok Park, and Mihaela van der Schaar, *Fellow, IEEE*

***Abstract*—Distributed medium access control (MAC) protocols are essential for the proliferation of low-cost, decentralized wireless local area networks (WLANs). Most MAC protocols are designed with the presumption that nodes comply with prescribed rules. However, selfish nodes have natural motives to manipulate protocols in order to improve their own performance. This often degrades the performance of other nodes as well as that of the overall system. In this paper, we propose a class of protocols that limit the performance gain from selfish manipulation while incurring only a small efficiency loss. The proposed protocols are based on the idea of a review strategy, with which nodes collect signals about the actions of other nodes over a period of time, use a statistical test to infer whether or not other nodes are following the prescribed behavior, and trigger a punishment if a deviation is inferred. We consider the cases of private and public signals and provide analytical and numerical results to demonstrate the properties of the proposed protocols.**

***Index Terms*—Deviation-proof protocols, game theory, medium access control (MAC) protocols, repeated games, review strategy.**

I. Introduction

In WIRELESS communication networks, multiple nodes often share a common channel and contend for access. To resolve contention among nodes, many different medium access control (MAC) protocols have been developed and are currently used in international standards (e.g., IEEE 802.11a/b/g protocols) [1]. When a MAC protocol is designed, two types of node behavior can be considered. One is *compliant* nodes that comply with prescribed protocols, and the other is *selfish* nodes that are willing to manipulate prescribed protocols to improve their own performance.[1] With compliant nodes, a MAC protocol can be designed to optimize the system performance without taking into account the possibility of selfish manipulation (see, for example, [3]–[6]). However, when such a protocol is used with selfish nodes, they may deviate from the protocol in pursuit of their self-interest [7], yielding a suboptimal outcome, different from the one desired by the protocol designer (see, for example, [8]–[10]). On the other hand, a MAC protocol can be designed assuming selfish nodes so that the protocol is *deviation-proof* in the sense that selfish nodes do not find it profitable to deviate from the protocol. However, incentive constraints imposed by selfish behavior in general restrict the system performance (see, for example, [6]). In this paper, we aim to resolve the tension between selfish manipulation and optimal performance by proposing a class of MAC protocols that limit the performance gain from selfish manipulation while incurring only a small efficiency loss compared to the optimal performance achievable with compliant nodes. Our design and analysis are conducted in the context of slotted multiaccess communication networks where nodes transmit data in a single hop interfering with each other.

Recently, selfish behavior in MAC protocols has also been analyzed using game theory. In [11], the authors establish the stability region for a slotted Aloha system with multipacket reception and selfish nodes. In [12], the authors study the existence of and convergence to Nash equilibrium in a slotted Aloha system where selfish nodes have quality-of-service requirements. It is often observed that selfish behavior leads to suboptimal outcomes. For example, a prisoners' dilemma phenomenon arises among selfish nodes adopting the generalized slotted Aloha protocols of [5]. A decrease in system throughput, especially when the workload increases due to the selfish behavior of nodes, is observed in [6]. In the 802.11 distributed MAC protocol, competition among selfish nodes results in an inefficient use of the shared channel in Nash equilibria [8].

In order to improve suboptimal outcomes with selfish nodes, various incentive schemes have been proposed in the literature. In [13], selfish nodes are induced to behave cooperatively in a slotted multiaccess network by introducing an intervening node that monitors the actions of nodes and decides its intervention level accordingly. Pricing has also been used as a method to incentivize selfish nodes. In [6], the degradation of system throughput due to selfish behavior is prevented by adding a cost of transmissions and retransmissions. In [14], the network charges nodes for each successfully transmitted packet in order to achieve a desired operating point. The above approaches, however, require a central entity, which may not be available in a distributed environment. In the case of an intervention scheme, an intervening node that is capable of monitoring and intervening should be present in the system. In the case of a pricing

Manuscript received June 16, 2010; revised June 05, 2011; accepted December 24, 2011; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor S. Shakkottai. Date of publication January 27, 2012; date of current version October 11, 2012.

K. T. Phan was with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA. He is now with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: khoa.phan@mail.mcgill.ca).

J. Park was with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA. He is now with the School of Economics, Yonsei University, Seoul 120-749, Korea (e-mail: jaeok.park@yonsei.ac.kr).

M. van der Schaar is with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail: mihaela@ee.ucla.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2011.2182359

[1]There can also be *malicious* nodes that aim to degrade the performance of other nodes. In this paper, we deal with selfish nodes exclusively. For a model with both selfish and malicious users, see [2].

scheme, a billing authority is needed to charge payments depending on the usage of the network. In this paper, we propose a decentralized approach where nodes carries out monitoring and punishment.

To this end, we rely on the theory of repeated games to sustain cooperation among selfish nodes. When the nodes in a system interact repeatedly, they can make their decisions dependent on their past observations. Thus, nodes can trigger a punishment when they observe a deviation from a predetermined operating point. If the loss in future payoff due to punishment outweighs the current gain from deviation, selfish nodes do not have an incentive to deviate from the predetermined operating point. The idea of using a repeated game strategy to build a deviation-proof protocol has recently been applied to several problems in communications and networking (see, for example, [15], [16]). However, most existing work assumes perfect monitoring, where players observe decisions that other players make. With perfect monitoring, deviations are always detected, and thus it is relatively straightforward to construct a deviation-proof protocol by having a sufficiently strong punishment triggered following a deviation.

In the scenario considered in this paper, the problem of constructing a deviation-proof protocol is complicated due to *imperfect* monitoring. The decisions of nodes are their transmission probabilities, but it is impossible for nodes to observe the transmission probabilities of others directly. Instead, they need to rely on their channel access outcomes or channel states in order to make inferences about the transmission probabilities. The deviation-proof protocol we construct is based on the idea of a review strategy [17], [18], with which nodes collect imperfect signals about the decisions of other nodes, perform a statistical test to determine whether or not a deviation has occurred, and trigger a punishment if they conclude so.

The main contributions of this paper can be summarized as follows.

- We model a slotted multiaccess communication scenario as a repeated game, which allows us to design a protocol based on a repeated game strategy, including a review strategy.
- We first consider the case where nodes observe private signals on the channel access outcomes. We design deviation-proof protocols assuming that a deviating node can employ only a constant transmission probability. We provide a necessary and sufficient condition for a given protocol to be deviation-proof. We show that the efficiency loss of a deviation-proof protocol can be made arbitrarily small if there is a statistical test that becomes perfect as more signals are accumulated.
- We also consider the case where nodes observe public signals on the channel access outcomes. We show that with public signals it is possible to design near-optimal deviation-proof protocols even when deviating nodes can use any deviation strategy.
- We discuss how our design methodology can be applied to CSMA/CA networks.
- We illustrate the properties of the proposed protocols with numerical results.

The proposed protocols are fully distributed in the sense that they require no central entity to coordinate the operation of nodes (with a possible exception of initial synchronization) and that nodes make decisions depending solely on their own local information without communicating with other nodes.

The rest of this paper is organized as follows. In Section II, we formulate a repeated game model for slotted multiaccess communications. In Section III, we propose and analyze deviation-proof protocols based on a review strategy when signals are private, with an example presented in Section IV. In Section V, we investigate deviation-proof protocols when signals are public, with an example presented in Section VI. In Section VII, we discuss a possible extension of the proposed protocols to a CSMA/CA network with selfish nodes. We conclude the paper in Section VIII.

II. Repeated Game Framework for Slotted Multiaccess Communications

A. Stage Game

We consider a single-hop wireless communication network where a set $\mathcal{N} = \{1, 2, \ldots, N\}$ of N nodes aims to transmit data to their receivers. Time is divided into slots of equal length, and in each slot, a node has a packet to transmit (i.e., saturated arrivals) and can attempt to send the packet or wait. Due to interference, a packet is transmitted successfully only if there is no other packet transmitted in the same slot. Otherwise, a collision occurs, and no packet is transmitted successfully. We model the interaction of nodes in a single slot as a noncooperative game in normal form, called the *random access game*.

The set of pure actions available to node $i \in \mathcal{N}$ in a slot is $A_i \triangleq \{T, W\}$, where T stands for "transmit" and W for "wait." We denote the pure action of node i by $a_i \in A_i$ and a pure action profile by $\mathbf{a} \triangleq (a_1, \ldots, a_N) \in \mathcal{A} \triangleq \prod_{i \in \mathcal{N}} A_i$. A mixed action for node i is a probability distribution on A_i. Since there are only two pure actions, a mixed action for node i can be represented by a transmission probability $p_i \in [0, 1]$, and the set of mixed actions for node i can be written as $P_i \triangleq [0, 1]$. A mixed action profile is denoted by $\mathbf{p} \triangleq (p_1, \ldots, p_N) \in \mathcal{P} \triangleq \prod_{i \in \mathcal{N}} P_i$. The payoff function of node i is defined by $u_i : \mathcal{A} \to \mathbb{R}$, where $u_i(\mathbf{a}) = 1$ if $a_i = T$ and $a_j = W$ for all $j \neq i$, and $u_i(\mathbf{a}) = 0$ otherwise. That is, a node receives payoff 1 if it has a successful transmission, and 0 otherwise. Then, the expected payoff of a node is given by the probability that it has a successful transmission, and with a slight abuse of notation, the payoff of node i when mixed action profile $\mathbf{p}$ is chosen can be written as

$$u_i(\mathbf{p}) = p_i \prod_{j \in \mathcal{N} \setminus \{i\}} (1 - p_j).$$

The random access game is defined by the tuple $\Gamma \triangleq \langle \mathcal{N}, (A_i)_{i \in \mathcal{N}}, (u_i)_{i \in \mathcal{N}} \rangle$. It is well known from the static analysis of the random access game that there is at least one node i choosing $p_i = 1$ at any Nash equilibrium (NE) [9], [13]. That is, when nodes myopically maximize their own payoffs, there is at least one node always transmitting its packets, and thus there can be at most one node obtaining a positive payoff. Moreover, in the unique symmetric NE, every node transmits with probability 1, which results in zero payoff for every node. On the other hand, the symmetric Pareto-optimal (PO) outcome

is achieved when each node chooses $p_c = 1/N$, which yields a positive payoff $u^{PO} = (1 - 1/N)^{N-1}/N$ for every node. We call p_c the *cooperation probability* and u^{PO} the optimal payoff.

B. Repeated Game

We now formulate the repeated random access game, where the actions of a node can depend on its past observations or information histories. Time slots are indexed by $t = 1, 2, \ldots$. At the end of each slot, nodes obtain signals on the pure action profile chosen in the slot. Let Z be the finite set of signals that each node can receive. Let Q be a mapping from $\mathcal{A}$ to Z^N, where $Q(\mathbf{a})$ represents the signals that nodes receive given pure action profile $\mathbf{a}$.[2] A *signal structure* is specified by the pair (Z, Q). In this paper, we restrict attention to symmetric signal structures where the signal that a node received is preserved under permutations of indices for nodes. We say that signals are *private* if there exist $\mathbf{z} \triangleq (z_1, \ldots, z_N) \in Z^N$ and $\mathbf{a} \in \mathcal{A}$ such that $z_i \neq z_j$ for some $i, j \in \mathcal{N}$ and $\mathbf{z} = Q(\mathbf{a})$. We say that signals are *public* if they are not private. That is, signals are private if it is possible for nodes to receive different signals, whereas signals are public if signal realization is the same for all nodes. We present two examples of signal structures that will be used later in this paper.

Example 1 (ACK Signals): In the slotted Aloha protocols in [5] and [19], a node receives an acknowledgement (ACK) signal if it transmits its packet successfully, and no signal otherwise. In the ACK signal structure, the signal space can be written as $Z = \{S, F\}$, where $z_i = S$ means that node i receives an ACK signal and F means that it does not. Provided that there is no error in the transmission and reception of ACK signals, the signal determination rule Q is such that the ith element of $Q(\mathbf{a})$ is S, while all the other elements are F if there exists i such that $a_i = T$ and $a_j = W$ for all $j \neq i$, and $Q(\mathbf{a}) = (F, \ldots, F)$ otherwise. ACK signals are private because when a success occurs, only one node receives signal S, while all the other nodes receive signal F.

Example 2 (Ternary Signals): In the ternary signal structure as in [20] and [21], we have $Z = \{0, 1, e\}$ and $Q(\mathbf{a}) = (0, \ldots, 0)$ if $\mathbf{a} = (W, \ldots, W)$, $Q(\mathbf{a}) = (1, \ldots, 1)$ if there exists i such that $a_i = T$ and $a_j = W$ for all $j \neq i$, and $Q(\mathbf{a}) = (e, \ldots, e)$ otherwise. That is, signals 0, 1, and e represent that the channel state is idle, success, and collision, respectively. Ternary signals are public because all the nodes obtain the same signal no matter what pure action profile is chosen.

The history of node i in slot t, denoted by h_i^t, contains the signals that node i has received by the end of slot $t-1$. That is, $h_i^t = (z_i^0, \ldots, z_i^{t-1})$, for $t = 1, 2, \ldots$, where z_i^t represents the signal that node i receives in slot t and z_i^0 is set as an arbitrary element z^0 of Z.[3] The set of all possible slot t histories of a node is given by $H^t \triangleq \{z^0\} \times Z^{t-1}$, and the set of all possible histories is $H \triangleq \cup_{t=1}^{\infty} H^t$. The (behavior) strategy for a node specifies a mixed action for it in the stage game conditional on a history it reaches. Thus, it can be represented by a mapping $\sigma : H \to [0, 1]$. We use Σ to denote the set of all strategies for a node. We define a *protocol* as a strategy profile $\underline{\sigma} \triangleq (\sigma_1, \ldots, \sigma_N) \in \Sigma^N$. To evaluate payoffs in the repeated game model, we use the limit of means criterion since the length of a slot is typically short.[4] A protocol $\underline{\sigma}$ induces a probability distribution on the sequences of mixed action profiles $\{\mathbf{p}^t\}_{t=1}^{\infty}$, where $\mathbf{p}^t$ is the mixed action profile in slot t. The payoff of node i under protocol $\underline{\sigma}$ can be expressed as

$$U_i(\underline{\sigma}) = \lim_{J \to \infty} E_{\underline{\sigma}} \left[\frac{1}{J} \sum_{t=1}^{J} u_i(\mathbf{p}^t) \right]$$

assuming that the limit exists, where $E_{\underline{\sigma}}$ denotes expectation with respect to the probability measure on $\mathcal{P}^{\infty}$ induced by $\underline{\sigma}$. If the limit does not exist, we replace the operator $\lim$ by $\liminf$.[5] We say that a protocol $\underline{\sigma}$ is symmetric if it prescribes the same strategy to every node, i.e., $\sigma_1 = \cdots = \sigma_N$. Note that a symmetric protocol can be represented by its common strategy, and thus we will sometimes refer to a strategy $\sigma \in \Sigma$ as a protocol in order to mean $\underline{\sigma} = (\sigma, \ldots, \sigma)$. In the following, we focus on symmetric protocols, which yield the same payoff to each node.

C. Deviation-Proof Protocols and the Efficiency Loss

The goal of this paper is to build a protocol that fulfills the following two requirements: 1) selfish nodes do not gain from manipulating the protocol; and 2) the protocol achieves an optimal outcome. We formalize the first requirement using the concept of deviation-proofness while evaluating the second requirement using the concept of efficiency loss. Since nodes are noncooperative (i.e., they are not able to form coalitions), we can focus on unilateral deviations. That is, a deviating node will compare its payoffs when it uses a deviation strategy and when it follows the prescribed strategy, provided that every other node follows the prescribed strategy. We use $U(\sigma'; \sigma)$ to denote the payoff that a node obtains when it uses strategy σ', while every other node follows strategy σ.

Definition 1: A protocol $\sigma \in \Sigma$ is *deviation-proof* (DP) against a strategy $\sigma' \in \Sigma$ if

$$U(\sigma; \sigma) \geq U(\sigma'; \sigma).$$

When σ is DP against σ', a node cannot gain by deviating to σ' while other nodes follow σ. Hence, if a deviating node has only one possible deviation strategy σ', a protocol σ that is DP against σ' can prevent selfish manipulation. However, in principle, a deviating node can choose any strategy in Σ, in which case we need a stronger concept than deviation-proofness. Let $\Sigma_c \subset \Sigma$ be the set of all constant strategies that prescribe a fixed transmission probability $p_d \in [0, 1]$, called the *deviation probability*, in every slot regardless of the history.

Definition 2: A protocol $\sigma \in \Sigma$ is *robust ϵ-deviation-proof* (robust ϵ-DP) if

$$U(\sigma; \sigma) + \epsilon \geq U(\sigma'; \sigma) \qquad \text{for all } \sigma' \in \Sigma_c.$$

[2]If signals are determined randomly given an action profile (e.g., due to errors), we can use $Q : \mathcal{A} \to \Delta(Z^N)$ instead, where $\Delta(Z^N)$ is the set of all possible probability distributions over Z^N.

[3]In slot $t \geq 2$, node i also knows its past mixed actions $(p_i^1, \ldots, p_i^{t-1})$ and their realizations $(a_i^1, \ldots, a_i^{t-1})$. However, since we focus on repeated game strategies using only past signals, we do not include them in our history specification.

[4]For example, the slot duration of the 802.11 DCF basic access method is 20 μs [1].

[5]Although we consider the limit of means criterion, the following results can be extended with a complication to the case of the discounting criterion as long as the discount factor is close to 1, as in [17].

TABLE I
SUMMARY OF MAIN RESULTS

Section	Signal	Test	Robustness to selfish manipulation	Optimality
III (Proposition 2)	Private (general)	Asymptotically perfect test	DP against a strategy using a constant transmission probability	δ-PO
IV (Theorem 2)	Private (ACK signal)	ACK ratio test	robust ϵ-DP	δ-PO
V (Proposition 5)	Public (general)	Asymptotically perfect test	DP against a strategy using a constant transmission probability in a review phase	δ-PO
VI (Theorem 4)	Public (ternary signal)	Idle slot ratio test	ϵ-NE	δ-PO

In other words, if a protocol σ is robust ϵ-DP, a node cannot gain more than ϵ by deviating to a constant strategy using a fixed deviation probability. If there is a fixed cost of manipulating a given protocol and a deviation strategy is constrained to constant strategies, then a robust ϵ-DP protocol can prevent a deviation by having ϵ smaller than the cost. When there is no restriction on possible deviation strategies, the following concept is relevant.

Definition 3: A protocol $\sigma \in \Sigma$ is an *ϵ-Nash equilibrium* (ϵ-NE) if

$$U(\sigma;\sigma) + \epsilon \geq U(\sigma';\sigma) \qquad \text{for all } \sigma' \in \Sigma.$$

We define the system payoff as the sum of the payoffs of all the nodes. Then, the system payoff when all nodes follow a protocol σ is given by $V(\sigma) \triangleq NU(\sigma;\sigma)$. Since Nu^{PO} is the maximum system payoff achievable with a symmetric mixed action profile, the *efficiency loss* of a protocol $\sigma \in \Sigma$ is measured by

$$C(\sigma) \triangleq Nu^{\mathrm{PO}} - V(\sigma). \tag{1}$$

Definition 4: A protocol $\sigma \in \Sigma$ is *δ-Pareto optimal* (δ-PO) if

$$C(\sigma) \leq \delta.$$

A δ-PO protocol is a protocol that yields an efficiency loss less than or equal to δ. Let σ^{c} be the strategy that prescribes the cooperation probability p_{c} in every slot regardless of the history.[6] Then, $U(\sigma^{\mathrm{c}};\sigma^{\mathrm{c}}) = u^{\mathrm{PO}}$, and thus σ^{c} achieves full efficiency (i.e., 0-PO). However, σ^{c} is not DP against a constant deviation strategy with $p_{\mathrm{d}} > p_{\mathrm{c}}$ as a deviating node can increase its payoff from $p_{\mathrm{c}}(1-p_{\mathrm{c}})^{N-1}$ to $p_{\mathrm{d}}(1-p_{\mathrm{c}})^{N-1}$. We construct DP protocols that achieve a near-optimal system payoff in the following sections, whose main results are summarized in Table I.

III. Deviation-Proof Protocols When Signals are Private

A. Description of Protocols With Private Signals

In this section, we consider private signals. As pointed out in [22], when signals are private, it is difficult, if not impossible, to construct an NE that has a simple structure and is easy to compute. Thus, we focus on a simpler problem of constructing a DP protocol against a constant deviation strategy $\sigma^{\mathrm{d}} \in \Sigma_{\mathrm{c}}$. Since a simple protocol such as σ^{c} is DP against σ^{d} with $p_{\mathrm{d}} \in [0, p_{\mathrm{c}}]$, we restrict our attention to deviation strategies with $p_{\mathrm{d}} \in (p_{\mathrm{c}}, 1]$.

[6]Note that σ^{c} corresponds to a slotted Aloha protocol that does not distinguish new and backlogged packets as in [12].

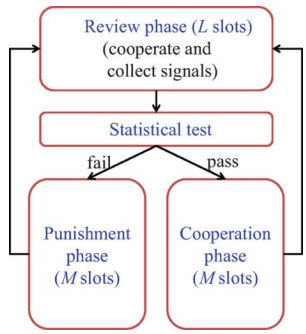

Fig. 1. Review strategy with private signals.

Note that the restriction to constant deviation strategies is relevant when a deviating node has a limited deviation capability in the sense that it can reset its transmission probability only at the beginning.

We build a protocol based on a review strategy. When a node uses a review strategy, it starts from a review phase for which it transmits with probability p_{c} and collects signals. When the review phase ends, the node performs a statistical test whose null hypothesis is that every node transmitted with probability p_{c} during the review phase, using the collected signals. Then, the node moves to a reciprocation phase for which it transmits with probability p_{c} (cooperation phase) if the test is passed, and with probability 1 (punishment phase) if the test fails. When the reciprocation phase ends, a new review phase begins. A review strategy, denoted by σ^{r}, can be characterized by three elements, (R, L, M), where R is a statistical test and L and M are natural numbers that represent the lengths of a review phase and a reciprocation phase, respectively. Thus, we sometimes write σ^{r} as $\sigma^{\mathrm{r}}(R, L, M)$. With a protocol based on review strategy $\sigma^{\mathrm{r}}(R, L, M)$, each node performs the statistical test R after slot $l(L+M)+L$ based on the signals $(z_i^{l(L+M)+1}, \ldots, z_i^{l(L+M)+L})$ collected in the recent review phase, for $l = 0, 1, \ldots$. A schematic presentation of a review strategy with private signals is provided in Fig. 1.

B. Analysis of Protocols With Private Signals

1) Existence of Deviation-Proof Protocols: For the sake of analysis, we consider a fixed constant deviation strategy $\sigma^{\mathrm{d}} \in \Sigma_{\mathrm{c}}$ and the corresponding deviation probability $p_{\mathrm{d}} \in (p_{\mathrm{c}}, 1]$. Given a symmetric protocol that prescribes a review strategy, we can compute two probabilities of errors.

- False punishment probability $P_{\mathrm{f}}(R, L)$: probability that there is at least one node whose test fails after a review phase when nodes follow a protocol σ^{r}.

- Miss detection probability $P_{\rm m}(R, L; p_{\rm d})$: probability that there is no node among those following $\sigma^{\rm r}$ whose test fails after a review phase when there is exactly one node deviating to $\sigma^{\rm d}$.

Since the payoff of every node is zero when there are two or more punishing nodes, we need to have a small false punishment probability to achieve a small efficiency loss. On the other hand, in order to punish a deviating node effectively, we need to have a small miss detection probability. Indeed, as will be shown in Proposition 2, achieving small $P_{\rm f}$ and $P_{\rm m}$ is sufficient to design a near-optimal DP protocol.

The payoff of a node when every node follows a review strategy $\sigma^{\rm r}$ is given by

$$U(\sigma^{\rm r};\sigma^{\rm r}) = \frac{(1-p_{\rm c})^{N-1}}{L+M}\Bigg(p_{\rm c}L + p_{\rm c}\left(1-P_{\rm f}\right)M + \left((1-P_{\rm f})^{\frac{N-1}{N}}\left(1-(1-P_{\rm f})^{1/N}\right)\right)M\Bigg).$$

The payoff of a node choosing deviation strategy $\sigma^{\rm d}$ while other nodes follow $\sigma^{\rm r}$ is given by

$$U(\sigma^{\rm d};\sigma^{\rm r}) = \frac{p_{\rm d}(1-p_{\rm c})^{N-1}}{L+M}\left(L + P_{\rm m}M\right).$$

By Definition 1, $\sigma^{\rm r}$ is DP against $\sigma^{\rm d}$ if and only if

$$U(\sigma^{\rm r};\sigma^{\rm r}) \geq U(\sigma^{\rm d};\sigma^{\rm r}). \tag{2}$$

The following theorem provides a necessary and sufficient condition for a review strategy to be DP against $\sigma^{\rm d}$.

Theorem 1: Given $p_{\rm d} \in (p_{\rm c}, 1]$, protocol $\sigma^{\rm r}(R, L, M)$ is DP against $\sigma^{\rm d}$ if and only if $g(R, L; p_{\rm d}) > 0$ and $M \geq M_{\min}(R, L; p_{\rm d})$, where

$$g(R,L;p_{\rm d}) \triangleq (1-P_{\rm f}(R,L))^{\frac{N-1}{N}} - (1-p_{\rm c})\left(1-P_{\rm f}(R,L)\right) - p_{\rm d}P_{\rm m}(R,L;p_{\rm d}) \tag{3}$$

and

$$M_{\min}(R,L;p_{\rm d}) \triangleq \frac{(p_{\rm d}-p_{\rm c})L}{g(R,L;p_{\rm d})}.$$

Proof: The net payoff gain from deviating to the deviation strategy $\sigma^{\rm d}$ is given by

$$U(\sigma^{\rm d};\sigma^{\rm r}) - U(\sigma^{\rm r};\sigma^{\rm r}) = \frac{(1-p_{\rm c})^{N-1}}{L+M}\Big((p_{\rm d}-p_{\rm c})L - g(R,L;p_{\rm d})M\Big). \tag{4}$$

The first term in (4) is the gain during a review phase, while the second term is the loss during a reciprocation phase. By (2), $\sigma^{\rm r}$ is DP against $\sigma^{\rm d}$ if and only if $(p_{\rm d}-p_{\rm c})L \leq g(R, L; p_{\rm d})M$. It is easy to check that $g(R, L; p_{\rm d}) > 0$ and $M \geq M_{\min}(R, L; p_{\rm d})$ imply $(p_{\rm d}-p_{\rm c})L \leq g(R, L; p_{\rm d})M$. Suppose that $(p_{\rm d}-p_{\rm c})L \leq g(R, L; p_{\rm d})M$. Since $(p_{\rm d}-p_{\rm c})L > 0$, we must have $g(R, L; p_{\rm d}) > 0$, which in turn implies $M \geq M_{\min}(R, L; p_{\rm d})$. ■

Theorem 1 shows that for a given statistical test R, we can construct a DP protocol based on the test if and only if there exists a natural number L such that $g(R, L; p_{\rm d}) > 0$. Once we find such L, we can use it as the length of a review phase and then choose a natural number M satisfying $M \geq M_{\min}(R, L; p_{\rm d})$ to determine the length of a reciprocation phase. An immediate consequence of Theorem 1 is that if protocol $\sigma^{\rm r}(R, L, M)$ is DP against $\sigma^{\rm d}$, then protocol $\sigma^{\rm r}(R, L, M')$ with $M' \geq M$ is also DP against $\sigma^{\rm d}$. Thus, $M_{\min}(R, L; p_{\rm d})$ can be interpreted as the minimum length of a reciprocation phase to make $\sigma^{\rm r}(R, L, M)$ DP against $\sigma^{\rm d}$. The following result provides a sufficient condition on R under which we can find L such that $g(R, L; p_{\rm d}) > 0$ and thus a DP protocol based on R can be constructed.

Corollary 1: Given $p_{\rm d} \in (p_{\rm c}, 1]$, suppose that R satisfies $\lim_{L\to\infty} P_{\rm f}(R, L) = 0$ and $\lim_{L\to\infty} P_{\rm m}(R, L; p_{\rm d}) = 0$. Then, there exists L such that $g(R, L; p_{\rm d}) > 0$.

Proof: By (3), $\lim_{L\to\infty} P_{\rm f}(R, L) = 0$ and $\lim_{L\to\infty} P_{\rm m}(R, L; p_{\rm d}) = 0$ imply that $\lim_{L\to\infty} g(R, L; p_{\rm d}) = p_{\rm c} > 0$. Thus, $g(R, L; p_{\rm d}) > 0$ for sufficiently large L. ■

Combining Theorem 1 and Corollary 1, we can see that if test R is "asymptotically perfect" in the sense that the two probabilities of errors converge to zero as the test is performed using more signals, then we can always design a review strategy based on R that is DP against $\sigma^{\rm d}$.

2) Near-Optimal Deviation-Proof Protocols: Suppose that every node follows a review strategy $\sigma^{\rm r}$. Since signals provide only imperfect information about the transmission probabilities of other nodes, it is possible that a punishment is triggered, which results in an efficiency loss as confirmed in the following proposition. We use $\Sigma_{\rm r}$ to denote the set of all review strategies with private signals.

Proposition 1: $C(\sigma^{\rm r}) \geq 0$ for all $\sigma^{\rm r} \in \Sigma_{\rm r}$ (with equality if and only if $P_{\rm f} = 0$).

Proof: Fix a protocol $\sigma^{\rm r}(R, L, M) \in \Sigma_{\rm r}$. By (1), we can express the efficiency loss of $\sigma^{\rm r}$ as

$$C(\sigma^{\rm r}) = \frac{NM}{L+M}(1-p_{\rm c})^{N-1} \times \left[p_{\rm c}P_{\rm f} - (1-P_{\rm f})^{\frac{N-1}{N}} + (1-P_{\rm f})\right]. \tag{5}$$

Since $(1-P_{\rm f})^{\frac{N-1}{N}}$ is concave, we have $(1-P_{\rm f})^{\frac{N-1}{N}} \leq 1 - \frac{N-1}{N}P_{\rm f}$ for $P_{\rm f} \in [0, 1]$, with equality if and only if $P_{\rm f} = 0$. Using $p_{\rm c} = 1/N$, we obtain the result. ■

Proposition 1 says that there is always a positive efficiency loss resulting from a review strategy unless there is a perfect statistical test in the sense that punishment is never triggered when every node follows $\sigma^{\rm r}$ (i.e., $P_{\rm f} = 0$). Punishment results in an efficiency loss because the system payoff is the same as $Nu^{\rm PO}$ when there is only one punishing node, while it is zero when there are two or more. Hence, a longer punishment induces a larger efficiency loss. As can be seen from (5), for given R and L, $C(\sigma^{\rm r})$ is nondecreasing (and increasing if $P_{\rm f} > 0$) in M. Therefore, if we find (R, L) such that $g(R, L; p_{\rm d}) > 0$, choosing $M = \lceil M_{\min}(R, L; p_{\rm d})\rceil$ minimizes the efficiency loss while having $\sigma^{\rm r}(R, L, M)$ DP against $\sigma^{\rm d}$, where $\lceil\cdot\rceil$ denotes the ceiling function. This observation allows us to reduce the design choice from (R, L, M) to (R, L).

We can see from (5) that there are two ways to achieve a small efficiency loss. One is to have large ratio L/M, and the other is to have $P_{\rm f}$ close to zero. If we do not require deviation-proofness, we can achieve an arbitrarily

small efficiency loss by increasing L even when $P_{\rm f}$ does not converge to zero. However, if we impose deviation-proofness, we cannot make L/M arbitrarily large because of the relationship $M = \lceil M_{\min}(R, L; p_{\rm d})\rceil$. We have $L/M \approx g(R, L; p_{\rm d})/(p_{\rm d} - p_{\rm c})$, and since $g(R, L; p_{\rm d})$ is bounded above, L/M is bounded above as well. Hence, with deviation-proofness, increasing L itself does not guarantee a small efficiency loss. To guarantee a small efficiency loss as L increases, we need $P_{\rm f}$ converging to zero. The following proposition provides a sufficient condition on the statistical test for constructing a near-optimal DP protocol.

Proposition 2: Given $p_{\rm d} \in (p_{\rm c}, 1]$, suppose that R satisfies $\lim_{L\to\infty} P_{\rm f}(R, L) = 0$ and $\lim_{L\to\infty} P_{\rm m}(R, L; p_{\rm d}) = 0$. Then, for any $\delta > 0$, there exist L and M such that $\sigma^{\rm r}(R, L, M)$ is DP against $\sigma^{\rm d}$ and δ-PO.

Proof: Since $\lim_{L\to\infty} g(R, L; p_{\rm d}) = p_{\rm c} > 0$, there exists L_1 such that $g(R, L; p_{\rm d}) > 0$ for all $L \geq L_1$. By Theorem 1, $\sigma^{\rm r}(R, L, \lceil M_{\min}(R, L; p_{\rm d})\rceil)$ is DP against $\sigma^{\rm d}$ for all $L \geq L_1$. Since $C(\sigma^{\rm r})$ is nondecreasing in M, we have

$$\begin{aligned} 0 \leq\ & C(\sigma^{\rm r}) \\ \leq\ & \frac{N\left(M_{\min}(R, L; p_{\rm d}) + 1\right)}{L + \left(M_{\min}(R, L; p_{\rm d}) + 1\right)}(1 - p_{\rm c})^{N-1} \\ & \times \left[p_{\rm c}P_{\rm f} - (1 - P_{\rm f})^{\frac{N-1}{N}} + (1 - P_{\rm f})\right]. \end{aligned} \tag{6}$$

Note that $\lim_{L\to\infty} M_{\min}(R, L; p_{\rm d})/L = (p_{\rm d} - p_{\rm c})/p_{\rm c}$, and thus the right-hand side of (6) converges to zero as L goes to infinity, which implies $\lim_{L\to\infty} C(\sigma^{\rm r}) = 0$. Therefore, there exists L_2 such that $C(\sigma^{\rm r}) < \delta$ for all $L \geq L_2$. Choose $L \geq \max\{L_1, L_2\}$ and $M = \lceil M_{\min}(R, L; p_{\rm d})\rceil$ to obtain a protocol with the desired properties. ■

Proposition 2 shows that the efficiency loss of a DP protocol can be made arbitrarily small when there is an asymptotically perfect statistical test. It also points out a tradeoff between optimality and implementation cost. In order to make the efficiency loss within a small desired level, L should be chosen sufficiently large, which requires large M by the relationship $M = \lceil M_{\min}(R, L; p_{\rm d})\rceil$. At the same time, as L and M become larger, each node needs to maintain longer memory to execute a review strategy, which can be considered as higher implementation cost.

We note that the constructed DP protocols are DP against a more general class of deviation strategies with which a permanent deviation to $p_{\rm d}$ occurs in an arbitrary slot (determined deterministically or randomly). A deviating node cannot gain starting from a review phase after a deviation occurs, and without discounting its temporary gain is always smaller than the perpetual loss.

IV. Protocols Based on the ACK Ratio Test

A. Description of Protocols Based on the ACK Ratio Test

In this section, we illustrate the results in Section III by considering the ACK signal structure, introduced in Example 1 in Section II-B. We propose a particular statistical test called the ACK ratio test. The test statistic of the ACK ratio test is the ratio of the number of ACK signals obtained in a review phase to the length of a review phase, i.e., $\sum_{k=1}^{L} \chi\{z_i^{\tau+k} = S\}/L$, where χ is an indicator function and $\tau + 1$ represents a slot when a review phase begins. The test is passed if the statistic exceeds a threshold value, $q_{\rm c} - B$, where $q_{\rm c} \triangleq p_{\rm c}(1 - p_{\rm c})^{N-1}$ and $B \in (0, q_{\rm c})$, and fails otherwise. Note that $q_{\rm c}$ is the expected value of the ACK ratio when every node transmits with probability $p_{\rm c}$. If there is a deviating node, the ACK ratio tends to be smaller because its expected value is reduced to $q_{\rm d} \triangleq p_{\rm c}(1 - p_{\rm c})^{N-2}(1 - p_{\rm d})$. The ACK ratio test is designed to distinguish between these two events statistically while having B as a "margin of error." Since the ACK ratio test can be identified with B, we use B instead of R to represent the ACK ratio test.

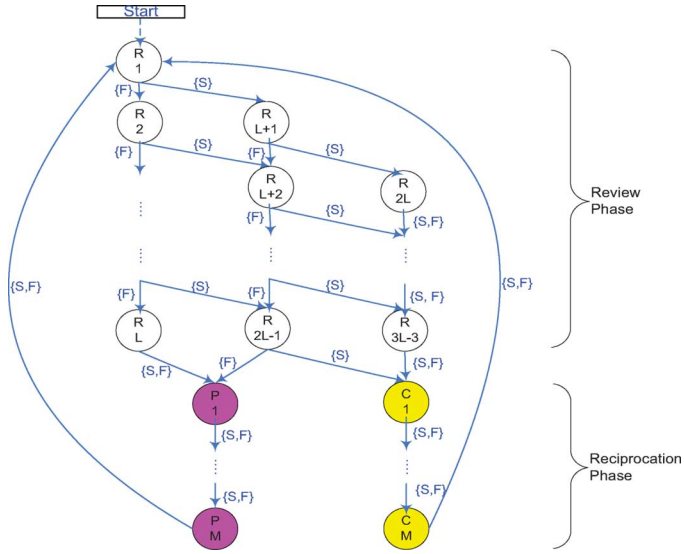

Fig. 2. Automaton representation of a review strategy based on the ACK ratio test with parameters satisfying $1 \leq L(q_{\rm c} - B) < 2$.

A review strategy based on the ACK ratio test, $\sigma^{\rm r}(B, L, M)$, can be represented formally as follows:

$$\sigma^{\rm r}(h_i^t) = \begin{cases} p_{\rm c}, & t \in [l(L+M)+1, l(L+M)+L], \\ 1, & t \in [l(L+M)+L+1, (l+1)(L+M)], \\ & \sum_{k=l(L+M)+1}^{l(L+M)+L} \frac{\chi\{z_i^k = S\}}{L} \leq q_{\rm c} - B \\ p_{\rm c}, & t \in [l(L+M)+L+1, (l+1)(L+M)], \\ & \sum_{k=l(L+M)+1}^{l(L+M)+L} \frac{\chi\{z_i^k = S\}}{L} > q_{\rm c} - B \end{cases}$$

for $l = 0, 1, \ldots$. Fig. 2 shows an automaton representation of the review strategy $\sigma^{\rm r}$ for $1 \leq L(q_{\rm c} - B) < 2$ so that a node triggers punishment if it obtains less than two successes in a review phase. Each state transition is labeled by the set of signals that induce the transition. In a reciprocation phase, a node goes through either states P1 to PM (punishment phase) or states C1 to CM (cooperation phase) depending on the number of ACK signals obtained in the review phase. Note that the number of states in the automaton representation of protocol $\sigma^{\rm r}(B, L, M)$ is given by $N_{\rm s}(\sigma^{\rm r}) \triangleq kL - k(k-1)/2 + 2M$, where $k \geq 2$ is the natural number satisfying $k - 2 \leq L(q_{\rm c} - B) < k - 1$.

B. Analytical Results

Let $F(y; n, p)$ be the cumulative distribution function of a binomial random variable with parameters n and p, i.e.,

$$F(y; n, p) = \sum_{m=0}^{\lfloor y \rfloor} \binom{n}{m} p^m (1 - p)^{n-m}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Suppose that every node transmits with probability $p_{\rm c}$ in a review phase. Then, the number of ACK signals that a node receives in the review phase follows a binomial distribution with parameters L and $q_{\rm c}$. Thus, the probability that a punishment is triggered by node i is given by

$$\Pr\left\{\sum_{k=1}^{L}\chi\{z_i^{\tau+k}=S\}/L \le q_{\rm c}-B\right\} = F(L(q_{\rm c}-B);L,q_{\rm c})$$

and the false punishment probability is given by

$$P_{\rm f}(B,L) = 1-[1-F(L(q_{\rm c}-B);L,q_{\rm c})]^N.$$

Suppose that there is exactly one deviating node using $\sigma^{\rm d}$, i.e., transmitting with probability $p_{\rm d}$. Then, the second parameter of the binomial distribution changes to $q_{\rm d}$, and thus the miss detection probability is given by

$$P_{\rm m}(B,L;p_{\rm d}) = [1-F(L(q_{\rm c}-B);L,q_{\rm d})]^{N-1}.$$

The monotonicity of $P_{\rm f}$ and $P_{\rm m}$ with respect to the test parameter B is readily obtained.

Proposition 3: Given $p_{\rm d} \in (p_{\rm c},1]$ and L, $P_{\rm f}(B,L)$ and $P_{\rm m}(B,L;p_{\rm d})$ are nonincreasing and nondecreasing in $B \in (0,q_{\rm c})$, respectively.

Proof: The proof is straightforward by noting that $F(L(q_{\rm c}-B);L,q_{\rm c})$ and $F(L(q_{\rm c}-B);L,q_{\rm d})$ are nonincreasing in $B \in (0,q_{\rm c})$. ■

As the margin of error is larger, it is more likely that the test is passed, yielding a smaller false punishment probability and a larger miss detection probability. The following lemma examines the asymptotic properties of $P_{\rm f}$ and $P_{\rm m}$ as L becomes large.

Lemma 1: Given $p_{\rm d} \in (p_{\rm c},1]$, $\lim_{L\to\infty} P_{\rm f}(B,L)=0$ for all $B \in (0,q_{\rm c})$, $\lim_{L\to\infty} P_{\rm m}(B,L;p_{\rm d})=0$ for all $B \in (0,q_{\rm c}-q_{\rm d})$, and $\lim_{L\to\infty} P_{\rm m}(B,L;p_{\rm d})=1$ for all $B \in (q_{\rm c}-q_{\rm d},q_{\rm c})$.

Proof: Since $\chi\{z_i^{\tau+k}=S\}$, for $k=1,\ldots,L$, can be considered as L i.i.d. random variables, we can apply the strong law of large numbers to the ACK ratio [23]. When every node transmits with probability $p_{\rm c}$, the ACK ratio converges almost surely to $q_{\rm c}$ as L goes to infinity, which implies that the false punishment probability goes to zero for all $B>0$. When there is exactly one node transmitting with probability $p_{\rm d}$, the ACK ratio of a node transmitting with probability $p_{\rm c}$ converges almost surely to $q_{\rm d}$ as L goes to infinity. Hence, if $q_{\rm d} < q_{\rm c}-B$ (resp. $q_{\rm d} > q_{\rm c}-B$), the miss detection probability goes to zero (resp. one). ■

Lemma 1 provides a sufficient condition on the ACK ratio test to apply Proposition 2.

Proposition 4: Suppose that $B \in (0,q_{\rm c}-q_{\rm d})$. For any $\delta>0$, there exist L and M such that $\sigma^{\rm r}(B,L,M)$ is DP against $\sigma^{\rm d}$ and δ-PO.

Proof: The proposition follows from Lemma 1 and Proposition 2. ■

Proposition 4 states that for given $p_{\rm d} \in (p_{\rm c},1]$, we can construct a protocol $\sigma^{\rm r}$ that is DP against $\sigma^{\rm d}$ and achieves an arbitrarily small efficiency loss by setting B such that $0<B<q_{\rm c}-q_{\rm d}=p_{\rm c}(1-p_{\rm c})^{N-2}(p_{\rm d}-p_{\rm c})$. Note that as $p_{\rm d}$ is larger, it is easier to detect a deviation, and thus we have a wider range of B that renders deviation-proofness and near-optimality.

So far, we have considered a constant deviation strategy $\sigma^{\rm d}$ prescribing a fixed deviation probability $p_{\rm d}$ and designed a protocol that is DP against $\sigma^{\rm d}$. However, it is natural to regard $p_{\rm d}$ as a choice of a deviating node, and thus in principle it can be any probability. Now we allow the possibility that a deviating node can use any constant deviation strategy, and we obtain the following result.

Theorem 2: For any $\epsilon>0$ and $\delta>0$, there exist B, L, and M such that $\sigma^{\rm r}(B,L,M)$ is robust ϵ-DP and δ-PO.

Proof: The proof is relegated to Appendix A. ■

We can interpret ϵ and δ as performance requirements. Requiring smaller ϵ makes protocols more robust while requiring smaller δ results in a higher system payoff. In addition to the tradeoff between optimality and implementation cost already mentioned following Proposition 2, we can identify a similar tradeoff between robustness and implementation cost in that smaller ϵ in general requires larger L and M to construct a robust ϵ-DP protocol.

C. Numerical Results

To obtain numerical results, we consider a network with five nodes, i.e., $N=5$ and $p_{\rm c}=1/N=0.2$. Fig. 3 plots $P_{\rm f}(B,L)$ and $P_{\rm m}(B,L;p_{\rm d})$ while varying L for $B=0.04$, 0.06. Fig. 3(a) shows that $P_{\rm f}(B,L)$ exhibits a decreasing tendency as L increases, with discontinuities occurring at the points where the floor function of $L(q_{\rm c}-B)$ has a jump. We can also see that $P_{\rm f}$ is smaller for larger B, consistent with Proposition 3. The upper threshold for the parameter B to yield $\lim_{L\to\infty} P_{\rm m}(R,L;p_{\rm d})=0$ in Lemma 1 is $q_{\rm c}-q_{\rm d}=0.0512$ for $p_{\rm d}=0.7$. We can see from Fig. 3(b) that $P_{\rm m}(B,L;p_{\rm d})$ has a decreasing (resp. increasing) tendency as L increases when B is smaller (resp. larger) than this threshold. Fig. 3(b) also shows that, for fixed B, $P_{\rm m}$ is smaller for larger $p_{\rm d}$. That is, as the deviation becomes greedier, it is more likely to be detected.

Fig. 4 plots the relationship between the length of a review phase L and the minimum length of a reciprocation phase $\lceil M_{\min}(B,L;p_{\rm d})\rceil$ to have a DP protocol for different values of B and $p_{\rm d}$. In Fig. 4(a), we fix $p_{\rm d}=0.7$ and consider $B=0.04,0.06$. Note that when $B=0.06$, some values of L result in large minimum values of M, which are not displayed in Fig. 4(a). Also, the values of L with which no DP protocol can be constructed for given B and $p_{\rm d}$ (i.e., $g(B,L;p_{\rm d}) \le 0$) are indicated with $\lceil M_{\min}(B,L;p_{\rm d})\rceil = 0$ in Fig. 4(a). For example, we cannot construct a DP protocol using L such that $42 \le L \le 45$ or $84 \le L \le 91$ when $B=0.06$ and $p_{\rm d}=0.7$. When $B=0.04$, we can construct a DP protocol using any $L \ge 10$. In Fig. 4(b), we fix $B=0.04$ and consider $p_{\rm d}=0.7$, 0.85. For the considered values of $p_{\rm d}$, we observe that the minimum length of a reciprocation phase is increasing in $p_{\rm d}$ for the most values of L. Also, in general, a longer review phase requires a longer reciprocation phase for fixed $p_{\rm d}$ although a reverse relationship may be obtained, especially when L is small. Note that L and $\lceil M_{\min}(B,L;p_{\rm d})\rceil$ have a linear relationship in the limit since $\lim_{L\to\infty} M_{\min}(B,L;p_{\rm d})/L=(p_{\rm d}-p_{\rm c})/p_{\rm c}$.

Fig. 5 plots efficiency loss $C(\sigma^{\rm r})$ against L while setting $M=\lceil M_{\min}(B,L;p_{\rm d})\rceil$ for different values of B and $p_{\rm d}$. The

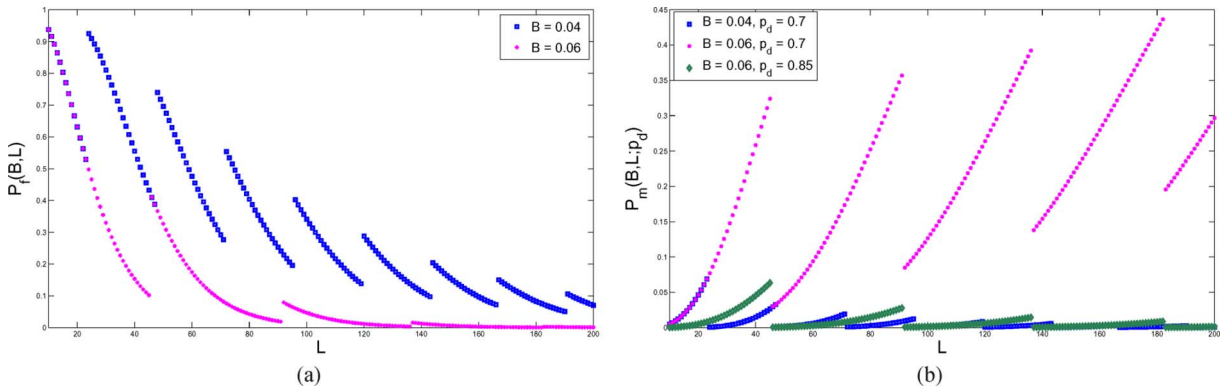

Fig. 3. Error probabilities (a) $P_{\mathrm{f}}(B, L)$ and (b) $P_{\mathrm{m}}(B, L; p_{\mathrm{d}})$ versus the length of a review phase L.

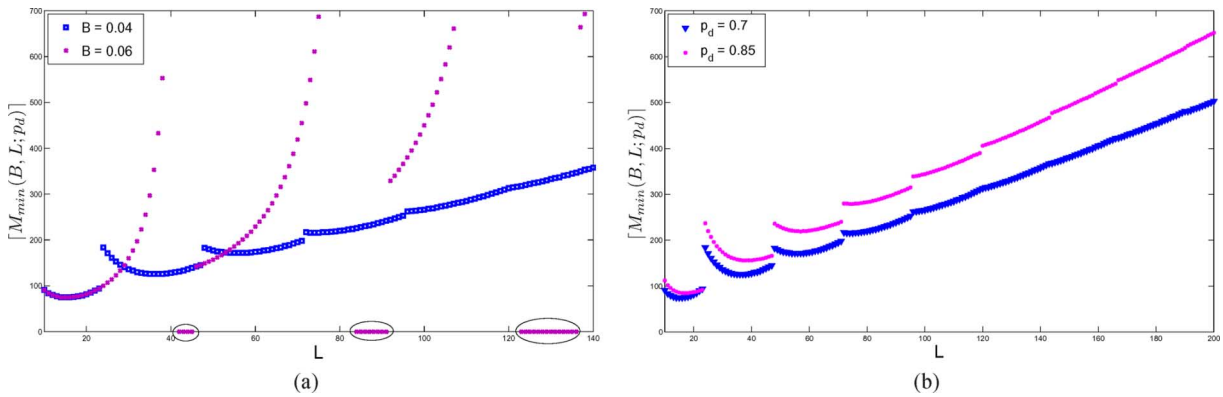

Fig. 4. Minimum length of a reciprocation phase $\lceil M_{\min}(B, L; p_{\mathrm{d}})\rceil$ versus the length of a review phase L: (a) $p_{\mathrm{d}} = 0.7$ and (b) $B = 0.04$.

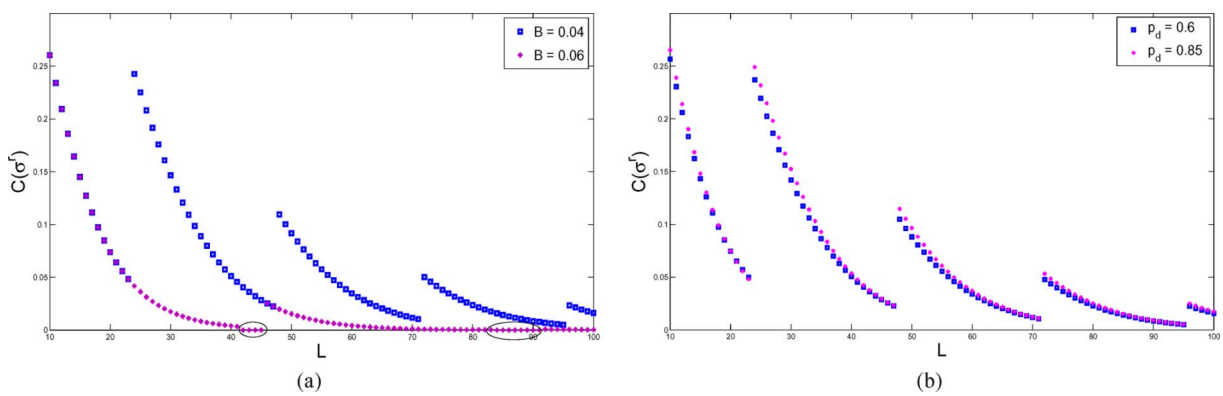

Fig. 5. Efficiency loss $C(\sigma^{\mathrm{r}})$ versus the length of a review phase L: (a) $p_{\mathrm{d}} = 0.7$ and (b) $B = 0.04$.

points where efficiency loss is shown as 0 in Fig. 5(a) are where no DP protocol exists for the given parameters. We can observe that as L increases, efficiency loss tends to decrease to 0, which is consistent with Proposition 4. Fig. 5(a) shows that for fixed $p_{\mathrm{d}} = 0.7$, efficiency loss is smaller when $B = 0.06$ than when $B = 0.04$. This is because the false punishment probability of the former case is smaller than that of the latter case as shown in Fig. 3(a). Fig. 5(b) shows that efficiency loss is almost the same for the two considered deviation probabilities when $B = 0.04$.

Lastly, we provide numerical results on Theorem 2. We consider $\epsilon = 0.015, 0.025$ and $\delta = 0.06, 0.08$, and compute minimum (L, M) to have a robust ϵ-DP and δ-PO protocol while using $B = 0.004$ and 0.007 when $\epsilon = 0.015$ and 0.025, respectively. For $(\epsilon, \delta) = (0.015, 0.06)$, $(0.015, 0.08)$, $(0.025, 0.06)$, $(0.025, 0.08)$, we obtain $(L, M) = (1912, 1245)$, $(1424, 1169)$, $(1013, 1073)$, $(707, 1053)$, respectively. We can see that as we require smaller ϵ and δ, we need to have longer review and punishment phases.

D. Deviation-Proof Protocols With Complexity Considerations

We mention briefly how to incorporate complexity considerations in the protocol design problem. One approach to measure the complexity of a repeated game strategy is to use the number of the states of the smallest automaton that can implement the strategy [24]. With this approach, we can formulate the

TABLE II
PARAMETERS AND THE EFFICIENCY LOSS OF OPTIMAL PROTOCOLS

p_d	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
(L, M)	(22,101)	(23,101)	(23,94)	(23,91)	(23,90)	(23,92)	(23,96)	(23,102)	(22,106)
$C(\sigma^r)$	0.0570	0.0490	0.0483	0.0480	0.0479	0.0481	0.0485	0.0490	0.0575

following protocol design problem, assuming that the deviation strategy is fixed as σ^{d}:

$$\begin{aligned} \text{minimize} \quad & C(\sigma^{\mathrm{r}}(B, L, M)) \\ \text{subject to} \quad & \sigma^{\mathrm{r}} \text{ is DP against } \sigma^{\mathrm{d}} \\ & N_{\mathrm{s}}(\sigma^{\mathrm{r}}) \leq \overline{N}_{\mathrm{s}}. \end{aligned} \tag{7}$$

$N_{\mathrm{s}}(\sigma^{\mathrm{r}})$ is the number of states in the automaton representation of the review strategy σ^{r}, as defined in Section IV-A, and $\overline{N}_{\mathrm{s}}$ is the imposed upper bound on the number of states. The second constraint can be interpreted as a complexity constraint that bounds the number of states in the automaton representation of σ^{r}. Without a complexity constraint, efficiency loss can be made arbitrarily small while satisfying the first constraint by choosing sufficiently large L, as shown in Proposition 4. Thus, the second constraint prevents optimal L from growing without bound.

We propose the following method to find an optimal protocol that solves the protocol design problem (7).

- *Step 1*. Determine a finite set $\mathcal{B} \subset (0, q_{\mathrm{c}})$ as the set of possible values of B.
- *Step 2*. Fix $B \in \mathcal{B}$. Identify the set of feasible (L, M) in the sense that (L, M) satisfies the second constraint of (7) given B.
- *Step 3*. Fix feasible L, and check whether $g(B, L; p_{\mathrm{d}})$ in (3) is positive. If so, choose M as the smallest feasible value of M larger than or equal to $M_{\min}(B, L; p_{\mathrm{d}})$, which we denote by $M(B, L)$, if such a value exists. Then, $\sigma^{\mathrm{r}}(B, L, M(B, L))$ is a protocol that satisfies both constraints of (7).
- *Step 4*. By varying B and L, obtain protocols that satisfy both constraints. Among these protocols, choose a protocol that yields the smallest efficiency loss.

As an illustrative example, we consider $N = 5$ and set $\overline{N}_{\mathrm{s}} = 2^8 = 256$ so that protocols can be implemented using 8-bit memory. For simplicity, we fix B at 0.04, i.e., $\mathcal{B} = \{0.04\}$. Table II presents the parameters (L, M) and the efficiency loss of optimal protocols for different deviation probabilities. We can see that the optimal protocols have different parameters for different values of p_{d}. Due to jumps in efficiency loss with respect to L as shown in Fig. 5, the optimal protocols do not necessarily have the longest possible review phase.

V. DEVIATION-PROOF PROTOCOLS WHEN SIGNALS ARE PUBLIC

A. Motivation

When signals are private, nodes do not know the results of the test performed by other nodes. Hence, without a cooperation phase, nodes cannot distinguish a deviating node from a punishing node and thus cannot coordinate to begin a new review phase in case some node moves to a punishment phase. However, the existence of a cooperation phase creates a weakness that can be exploited by a deviating node. A deviating node can cooperate in a review phase to avoid punishment and then defect in a reciprocation phase to obtain a payoff gain. To exclude such a deviation, in Sections III and IV we have focused on constant deviation strategies when designing DP protocols. When signals are public, we can dispense with a cooperation phase. With public signals, the result of the test is the same across nodes, and thus nodes can coordinate whether to move to a punishment phase or to begin a new review phase. Thus, with public signals, it becomes possible to construct a protocol that is DP against a larger class of deviation strategies. This added robustness of protocols with public signals can be regarded as the *value of public signals* when the signal structure is a design choice.[7]

B. Description of Protocols With Public Signals

When signals are public, nodes receive a common signal, and thus we use z^t, without subscript i, to denote the signal in slot t. A review strategy with public signals is the same as the one described in Section III-A except that there is no cooperation phase. That is, a new review phase begins immediately if the statistical test is passed. If the test fails, a punishment phase occurs as before. Since we focus on symmetric protocols, all nodes use the same statistical test and perform the test based on the same signals. Hence, all nodes obtain the same result of the test, and thus they are always in the same phase. We use $\tilde{\sigma}^{\mathrm{r}}(R, L, M)$ to denote the review strategy with public signals that uses test R and has L and M as the lengths of a review phase and a punishment phase, respectively.

C. Analysis of Protocols With Public Signals

We first consider a fixed deviation strategy $\tilde{\sigma}^{\mathrm{d}}$ that has the same structure as the prescribed review strategy $\tilde{\sigma}^{\mathrm{r}}$. That is, a deviating node transmits with probability p_{d} in a review phase and with p_{r} in a punishment phase. Since no node obtains a positive payoff in a punishment phase, the choice of p_{r} does not affect the analysis, and thus only p_{d} matters for analysis. For the same reason as in Section III, we focus on the case where $p_{\mathrm{d}} > p_{\mathrm{c}}$.

As in the case of private signals, we can compute two probabilities of errors: the false punishment probability $\tilde{P}_{\mathrm{f}}(R, L)$ and the miss detection probability $\tilde{P}_{\mathrm{m}}(R, L; p_{\mathrm{d}})$. Since a punishment phase occurs with probability $\tilde{P}_{\mathrm{f}}$ and results in zero payoff for every node when all nodes follow a review strategy, we have

$$U(\tilde{\sigma}^{\mathrm{r}}; \tilde{\sigma}^{\mathrm{r}}) = \frac{L q_{\mathrm{c}}}{L + \tilde{P}_{\mathrm{f}} M}.$$

[7]Even when signals are private, we can dispense with a cooperation phase if a node can broadcast the failure of its test, as in [18]. Then, all nodes can move to a punishment phase if the test of some node fails, and begin a new review phase otherwise. However, broadcasting the result of the test requires communications among nodes, which we do not allow in the considered distributed protocols.

Note that $(L+\tilde{P}_{\rm f}M)$ is the average length of an epoch, defined as a review phase and the following punishment phase if one exists, and $Lq_{\rm c}$ is the accumulated expected payoff for a node in an epoch. The payoff of a node choosing deviation strategy $\tilde{\sigma}^{\rm d}$ while other nodes follow $\tilde{\sigma}^{\rm r}$ is given by

$$U(\tilde{\sigma}^{\rm d};\tilde{\sigma}^{\rm r})=\frac{Lq_{\rm d}}{L+(1-\tilde{P}_{\rm m})M}.$$

The efficiency loss of $\tilde{\sigma}^{\rm r}$ can be computed as

$$C(\tilde{\sigma}^{\rm r})=\frac{N\tilde{P}_{\rm f}Mq_{\rm c}}{L+\tilde{P}_{\rm f}M} \tag{8}$$

which is always nonnegative (positive if $\tilde{P}_{\rm f}>0$). Note that the nonnegativity of the efficiency loss does not require $p_{\rm c}=1/N$, unlike in the case of private signals (see the proof of Proposition 1). The following theorem is an analogue of Theorem 1 for the case of public signals.

Theorem 3: Given $p_{\rm d}\in(p_{\rm c},1]$, protocol $\tilde{\sigma}^{\rm r}(R,L,M)$ is DP against $\tilde{\sigma}^{\rm d}$ if and only if $\tilde{g}(R,L;p_{\rm d})>0$ and $M\geq\tilde{M}_{\min}(R,L;p_{\rm d})$, where

$$\tilde{g}(R,L;p_{\rm d})\triangleq p_{\rm c}\big(1-\tilde{P}_{\rm m}(R,L;p_{\rm d})\big)-p_{\rm d}\tilde{P}_{\rm f}(R,L)$$

and

$$\tilde{M}_{\min}(R,L;p_{\rm d})\triangleq\frac{(p_{\rm d}-p_{\rm c})L}{\tilde{g}(R,L;p_{\rm d})}.$$

Proof: $U(\tilde{\sigma}^{\rm r};\tilde{\sigma}^{\rm r})\geq U(\tilde{\sigma}^{\rm d};\tilde{\sigma}^{\rm r})$ if and only if $(p_{\rm d}-p_{\rm c})L\leq\tilde{g}(R,L;p_{\rm d})M$. Note that $(1-p_{\rm c})^{N-1}(p_{\rm d}-p_{\rm c})L$ is the gain from deviation in a review phase, while $(1-p_{\rm c})^{N-1}\tilde{g}(R,L;p_{\rm d})M$ is the expected loss from deviation in a punishment phase. The result can be obtained by using a similar argument as in the proof of Theorem 1. ■

Theorem 3 shows that for a given statistical test R, we can construct a DP protocol based on the test if and only if there exists a natural number L such that $\tilde{g}(R,L;p_{\rm d})>0$. As in the case of private signals, we can reduce the design choices for a review strategy from (R,L,M) to (R,L) by setting $M=\lceil\tilde{M}_{\min}(R,L;p_{\rm d})\rceil$. The next result is an analog of Proposition 2, showing that if an asymptotically perfect statistical test is available, we can construct a near-optimal DP protocol.

Proposition 5: Given $p_{\rm d}\in(p_{\rm c},1]$, suppose that R satisfies $\lim_{L\to\infty}\tilde{P}_{\rm f}(R,L)=0$ and $\lim_{L\to\infty}\tilde{P}_{\rm m}(R,L;p_{\rm d})=0$. Then, for any $\delta>0$, there exist L and M such that $\tilde{\sigma}^{\rm r}(R,L,M)$ is DP against $\tilde{\sigma}^{\rm d}$ and δ-PO.

Proof: The proof is similar to that of Proposition 2, and thus is omitted for brevity. ■

VI. Protocols Based on the Idle Slot Ratio Test

A. Description of Protocols Based on the Idle Slot Ratio Test

To illustrate the results in Section V, we consider the ternary signal structure, introduced in Example 2 in Section II-B. We consider a review strategy with which nodes use the fraction of idle slots in a review phase, or the idle slot ratio, as the test statistics. If every node transmits with probability $p_{\rm c}$, the expected value of the idle slot ratio is $\tilde{q}_{\rm c}\triangleq(1-p_{\rm c})^N$. On the other hand, if there is exactly one deviating node that transmits with probability $p_{\rm d}$ during a review phase, the expected value is reduced to $\tilde{q}_{\rm d}\triangleq(1-p_{\rm d})(1-p_{\rm c})^{N-1}$. The idle slot ratio test is passed if the idle slot ratio, $\sum_{k=1}^{L}\chi\{z^{\tau+k}=0\}/L$, exceeds a threshold value, $\tilde{q}_{\rm c}-B$, and fails otherwise. Fig. 6 shows an automaton representation of a review strategy $\tilde{\sigma}^{\rm r}$ whose parameters satisfying $1\leq L(\tilde{q}_{\rm c}-B)<2$. State transition occurs depending on the received signals, as depicted in Fig. 6. When a review phase ends, nodes either start a new review phase or move to a punishment phase depending on whether the number of idle slots in the review phase exceeds $L(\tilde{q}_{\rm c}-B)$ or not.

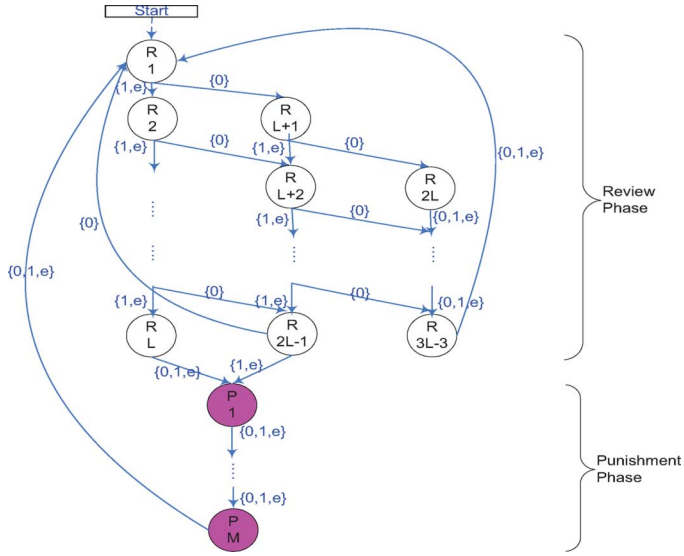

Fig. 6. Automaton representation of a review strategy based on the idle slot ratio test with parameters satisfying $1\leq L(\tilde{q}_{\rm c}-B)<2$.

B. Analytical Results

Suppose that every node follows a review strategy based on the idle slot ratio test, $\tilde{\sigma}^{\rm r}(B,L,M)$. Then, every node transmits with probability $p_{\rm c}$ in a review phase, and the number of idle slots occurring in a review phase follows a binomial distribution with parameters L and $\tilde{q}_{\rm c}$. Thus, the false punishment probability is given by

$$\tilde{P}_{\rm f}(B,L)=F(L(\tilde{q}_{\rm c}-B);L,\tilde{q}_{\rm c}).$$

Since a deviating node using transmission probability $p_{\rm d}$ changes the second parameter of the binomial distribution from $\tilde{q}_{\rm c}$ to $\tilde{q}_{\rm d}$, the miss detection probability is given by

$$\tilde{P}_{\rm m}(B,L;p_{\rm d})=1-F(L(\tilde{q}_{\rm c}-B);L,\tilde{q}_{\rm d}).$$

The monotonicity of $\tilde{P}_{\rm f}$ and $\tilde{P}_{\rm m}$ with respect to the margin of error B is stated as follows.

Proposition 6: Given $p_{\rm d}\in(p_{\rm c},1]$ and L, $\tilde{P}_{\rm f}(B,L)$ and $\tilde{P}_{\rm m}(B,L;p_{\rm d})$ are nonincreasing and nondecreasing in $B\in(0,\tilde{q}_{\rm c})$, respectively.

Proof: The proof is straightforward by noting that $F(L(\tilde{q}_{\rm c}-B);L,\tilde{q}_{\rm c})$ and $F(L(\tilde{q}_{\rm c}-B);L,\tilde{q}_{\rm d})$ is nonincreasing in $B\in(0,\tilde{q}_{\rm c})$. ■

The next lemma examines the asymptotic properties of $\tilde{P}_{\rm f}$ and $\tilde{P}_{\rm m}$ as L becomes large.

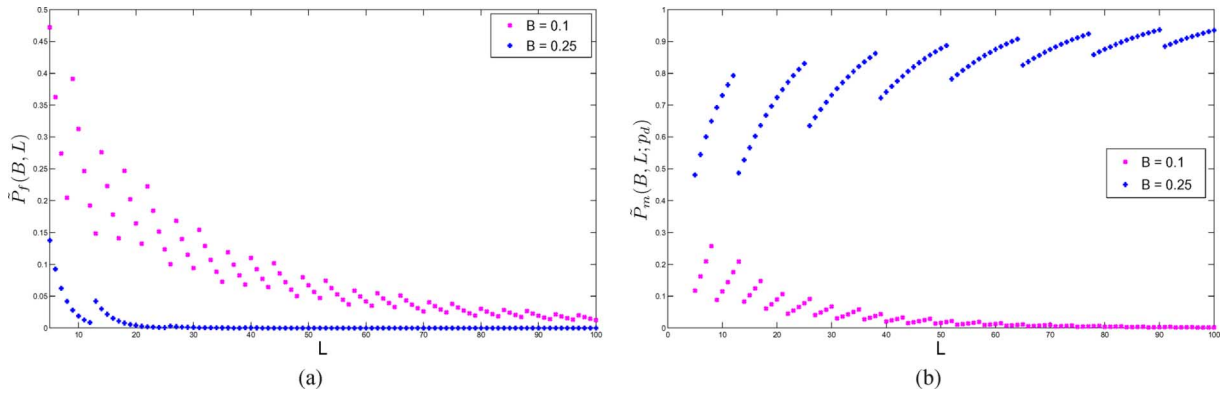

Fig. 7. Error probabilities (a) $\tilde{P}_{\rm f}(B,L)$ and (b) $\tilde{P}_{\rm m}(B,L;p_{\rm d})$ versus the length of a review phase L when $p_{\rm d}=0.7$.

Lemma 2: Given $p_{\rm d}\in(p_{\rm c},1]$, $\lim_{L\to\infty}\tilde{P}_{\rm f}(B,L)=0$ for all $B\in(0,\tilde{q}_{\rm c})$, $\lim_{L\to\infty}\tilde{P}_{\rm m}(B,L;p_{\rm d})=0$ for all $B\in(0,\tilde{q}_{\rm c}-\tilde{q}_{\rm d})$, and $\lim_{L\to\infty}\tilde{P}_{\rm m}(B,L;p_{\rm d})=1$ for all $B\in(\tilde{q}_{\rm c}-\tilde{q}_{\rm d},\tilde{q}_{\rm c})$.

Proof: The proof is similar to that of Lemma 1, and thus is omitted for brevity. ■

Lemma 2 gives a sufficient condition on the idle slot ratio test to apply Proposition 5.

Proposition 7: Suppose that $B\in(0,\tilde{q}_{\rm c}-\tilde{q}_{\rm d})$. For any $\delta>0$, there exist L and M such that $\tilde{\sigma}^{\rm r}(B,L,M)$ is DP against $\tilde{\sigma}^{\rm d}$ and δ-PO.

Proof: The proposition follows from Lemma 2 and Proposition 5. ■

Proposition 7 states that for given $p_{\rm d}\in(p_{\rm c},1]$, we can always construct a protocol based on the idle slot ratio test that is DP against $\tilde{\sigma}^{\rm d}$ and achieves an arbitrarily small efficiency loss by choosing B such that $0<B<\tilde{q}_{\rm c}-\tilde{q}_{\rm d}=(p_{\rm d}-p_{\rm c})(1-p_{\rm c})^{N-1}$. As in the case of the ACK ratio test, we have a wider range of B that renders deviation-proofness as $p_{\rm d}$ is larger.

We have considered deviation strategies that prescribe a constant transmission probability in a review phase. We now consider the case where a deviating node can use any strategy in Σ, which includes strategies that adjust transmission probabilities depending on the signals obtained in the current review phase. The following theorem shows that we can construct a protocol based on the idle slot ratio test that is approximately NE and near-optimal.

Theorem 4: For any $\epsilon>0$ and $\delta>0$, there exist B, L, and M such that $\tilde{\sigma}^{\rm r}(B,L,M)$ is ϵ-NE and δ-PO.

Proof: The proof is relegated to Appendix B. ■

The interpretation of ϵ and δ as performance requirements as well as the tradeoff between performance and implementation cost, as discussed following Theorem 2, is still valid in the case of public signals.

Remark (Protocols With Sliding Windows): Suppose that more than $L(\tilde{q}_{\rm c}-B)$ idle slots have occurred before the end of a review phase. Then, a deviating node, knowing that a punishment will not occur regardless of the outcome in the remaining slots of the review phase, can increase its transmission probability for the remainder of the review phase to obtain a payoff gain. We can make a protocol based on a review strategy robust to such a manipulation by having sliding windows for review phases. In a review strategy with sliding windows, a review phase begins in each slot unless there is a new or ongoing punishment. Once the idle slot ratio test based on the recent L signals fails, a review stops and punishment occurs for M slots. Once a punishment phase ends, a review phase begins in each slot until another punishment occurs. A detailed analysis of protocols with sliding windows is left for future research.

C. Numerical Results

To obtain numerical results on DP protocols with public signals, we again consider a network with $N=5$ and $p_{\rm c}=1/N=0.2$ while varying $p_{\rm d}$ and the protocol parameters.

Fig. 7 plots $\tilde{P}_{\rm f}$ and $\tilde{P}_{\rm m}$ against L for $B=0.1,0.25$ and $p_{\rm d}=0.7$. As in the case of private signals, $\tilde{P}_{\rm f}$ tends to decrease with L and approaches zero for large L. Also, $\tilde{P}_{\rm f}$ is smaller for a larger margin of error B. The upper threshold for B to yield $\lim_{L\to\infty}\tilde{P}_{\rm m}(B,L;p_{\rm d})=0$ in Lemma 2 is $\tilde{q}_{\rm c}-\tilde{q}_{\rm d}=0.2048$. We can see that when B is larger than this threshold, $\tilde{P}_{\rm m}$ tends to increase with L and approaches 1 for large L. On the contrary, when B is smaller than the threshold, $\tilde{P}_{\rm m}$ approaches zero for large L, making the test asymptotically perfect.

Fig. 8 plots the minimum length of a reciprocation phase $\lceil\tilde{M}_{\min}(B,L;p_{\rm d})\rceil$ to have a DP protocol as a function of the length of a review phase L. We can see that for fixed $p_{\rm d}=0.7$, a longer reciprocation phase is needed for larger B, except when L is small, and that DP protocols cannot be constructed with some small values of L when $B=0.1$ (displayed as $\lceil\tilde{M}_{\min}(B,L;p_{\rm d})\rceil=0$). Also, when $B=0.1$, a longer reciprocation phase is needed for $p_{\rm d}=1$ than for $p_{\rm d}=0.7$. The efficiency loss of DP protocols with the minimum length of a reciprocation phase is shown in Fig. 9. We can see that larger B results in a smaller efficiency loss, because $\tilde{P}_{\rm f}$ is smaller for larger B as shown in Proposition 6. Also, efficiency loss approaches zero as L becomes large, which is consistent with Proposition 7.

Lastly, we provide numerical results on Theorem 4. We set the parameters in the proof of Theorem 4 as $\beta=0.4$, $\rho=0.8$, and $\mu=20$, which determine B and M as functions of L. For the four considered pairs of $(\epsilon,\delta)=$ (0.01,0.01), (0.01, 0.05), (0.02, 0.01), (0.02, 0.05), we obtain $(L,M)=(41,820)$, (19, 380), (34, 680), (16, 320), respectively. We can see that we need to have longer review and punishment phases as ϵ and δ are smaller.

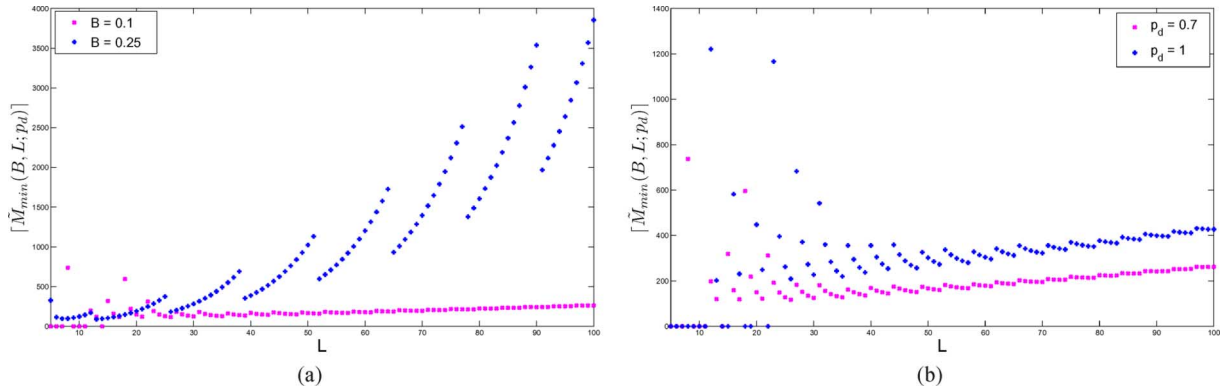

Fig. 8. Minimum length of a punishment phase $\lceil \tilde{M}_{\min}(B, L; p_{\mathrm{d}}) \rceil$ versus the length of a review phase L: (a) $p_{\mathrm{d}} = 0.7$ and (b) $B = 0.1$.

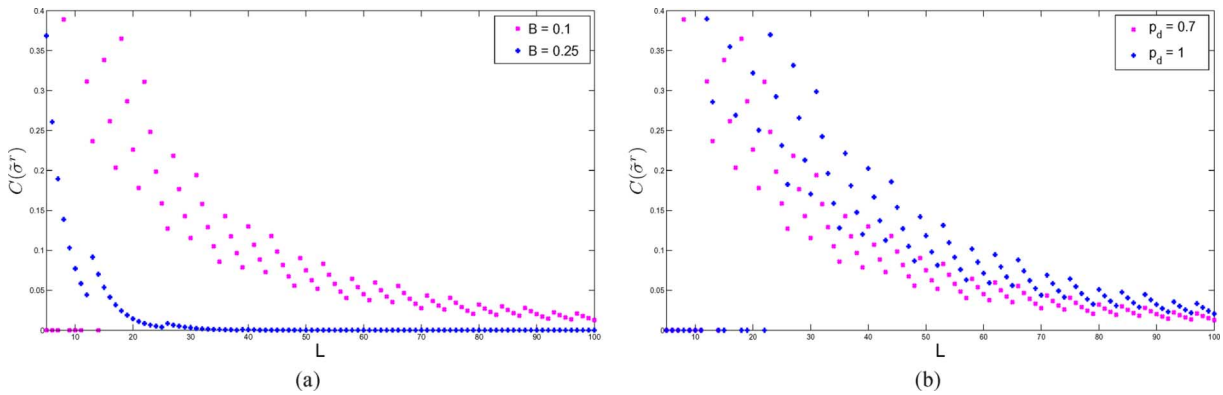

Fig. 9. Efficiency loss $C(\tilde{\sigma}^{\mathrm{r}})$ versus the length of a review phase L: (a) $p_{\mathrm{d}} = 0.7$ and (b) $B = 0.1$.

VII. Extension to a CSMA/CA Network With Selfish Nodes

In this section, we discuss how the proposed protocols based on a review strategy can be modified for a CSMA/CA network. As in [9], we consider a CSMA/CA network in which selfish nodes use a fixed contention window size. The authors of [9] show a discrepancy between NE and Pareto optimum. The contention window size of each node at the unique symmetric PO outcome is denoted by W^*, which results in a transmission probability $p_{\mathrm{c}} = 2/(W^* + 1)$ in the steady state. The optimal payoff u^{PO}, i.e., the throughput at Pareto optimum, can be computed using [9, Eq. (1)], which is based on the model of [25].

A review strategy in a CSMA/CA network can be described as follows, assuming private signals (i.e., sensing information is private). At the beginning, nodes are synchronized to start a review phase. In a review phase, which lasts for L time period, each node sets its window size at W^*. After a review phase, each node i computes its actual throughput, denoted by τ_i, and compares it to u^{PO}, the expected throughput when no node has deviated from W^*. A deviating node chooses its window size W^{d} smaller than W^* in order to increase its transmission probability from p_{c} and thus to obtain a higher throughput. Since a deviation decreases the throughput of the well-behaved nodes, we can design a test such that the test performed by node i is passed if and only if $\tau_i > u^{\mathrm{PO}} - B$ for some constant $B \in (0, u^{\mathrm{PO}})$. If the test of node i is passed, node i moves to a cooperation phase during which it continues to set its window size at W^*. Otherwise, it moves to a punishment phase during which it sets its window size at the minimum value 1. A reciprocation phase lasts for M time period, and a new review phase begins after a reciprocation phase.

As in a slotted multiaccess network, τ_i converges almost surely to u^{PO} as L goes to infinity, and thus the proposed test can be made asymptotically perfect by choosing an appropriate value of B. Hence, when window sizes take discrete values, we can construct a protocol that is DP against any constant deviation strategy and achieves a small efficiency loss, following a similar approach to Theorem 2. We omit the details due to space limitations.

VIII. Conclusion

It is well known that the decentralized operation of multiple-access communication systems with selfish nodes often results in an inefficient use of the shared medium. To overcome this problem, we have proposed a new class of slotted MAC protocols that are robust to selfish manipulation while achieving near-optimality. The proposed protocols are based on the idea of a review strategy in the theory of repeated games. We have provided conditions under which we can design deviation-proof protocols with a small efficiency loss and illustrated general results with particular statistical tests. Our framework and design methodology are not limited to multiple-access communications. They can be applied to other networking and communication scenarios in which agents obtain imperfect signals about the decisions of other agents and a deviation influences the distribution of signals. An important future direction is to design

deviation-proof protocols in multihop networks, where we need to consider incentives for nodes to relay others' packets.

APPENDIX A
PROOF OF THEOREM 2

Choose arbitrary $\epsilon > 0$ and $\delta > 0$. Define $p_\epsilon \triangleq p_c + \epsilon/(1-p_c)^{N-1}$. Note that p_ϵ is the minimum deviation probability with which a deviating node gains at least ϵ in a slot when other nodes transmit with probability p_c. Choose $B \in (0, \epsilon/(N-1))$. Note that $q_c - q_d \geq \epsilon/(N-1)$ for all $p_d \in [p_\epsilon, 1]$. Define

$$\hat{g}(B,L) \triangleq (1-P_f(B,L))^{\frac{N-1}{N}} - (1-p_c)(1-P_f(B,L)) - P_m(B,L;p_\epsilon).$$

Since $P_m(B,L;p_d)$ is nonincreasing in p_d, we have $g(B,L;p_d) \geq \hat{g}(B,L)$ for all $p_d \in [p_\epsilon, 1]$, where $g(B,L;p_d)$ is defined in (3). Also, by Lemma 1, we have $\lim_{L\to\infty} P_f(B,L) = 0$ and $\lim_{L\to\infty} P_m(B,L;p_\epsilon) = 0$. Therefore, $\lim_{L\to\infty} \hat{g}(B,L) = p_c$, and thus there exists L_1 such that $g(B,L;p_d) > 0$ for all $p_d \in [p_\epsilon, 1]$, for all $L \geq L_1$. Define $\hat{M}(L) \triangleq \lceil (1-p_c)L/\hat{g}(B,L) \rceil$. Since $\hat{M}(L) \geq M_{\min}(B,L;p_d)$ for all $p_d \in [p_\epsilon, 1]$, protocol $\sigma^r(B,L,\hat{M}(L))$ is DP against all constant strategies using $p_d \in [p_\epsilon, 1]$, for all $L \geq L_1$.

Since $C(\sigma^r)$ is nondecreasing in M, we have

$$\begin{aligned} 0 &\leq C(\sigma^r(B,L,\hat{M}(L))) \\ &\leq \frac{N\left((1-p_c)L/\hat{g}(B,L)+1\right)}{L+\left((1-p_c)L/\hat{g}(B,L)+1\right)} \\ &\quad \times (1-p_c)^{N-1}\left[p_c P_f - (1-P_f)^{\frac{N-1}{N}} + (1-P_f)\right]. \end{aligned}$$

Therefore, $\lim_{L\to\infty} C(\sigma^r) = 0$, and there exists L_2 such that $C(\sigma^r) < \delta$ for all $L \geq L_2$. Choose $L \geq \max\{L_1, L_2\}$ and $M = \hat{M}(L)$. Then, $\sigma^r(B,L,M)$ is DP against all constant strategies using $p_d \in [p_\epsilon, 1]$ and satisfies $C(\sigma^r) < \delta$. Finally, note that the payoff gain from deviating to a constant strategy using $p_d \in [0, p_\epsilon)$ is bounded above by ϵ. Hence, $\sigma^r(B,L,M)$ is robust ϵ-DP and δ-PO. This completes the proof.

APPENDIX B
PROOF OF THEOREM 4

Consider the problem of a deviating node maximizing its payoff given that all the other nodes use a review strategy $\tilde{\sigma}^r(B,L,M)$, i.e., $\max_{\sigma\in\Sigma} U(\sigma;\tilde{\sigma}^r)$. We can define a state space with total $L(L+1)/2 + M$ states, where a state is a pair consisting of the slot position and the number of idle slots since the beginning of the current review phase in the case of a review phase while it is the slot position in the case of a punishment phase. Given the structure of a review strategy, the relevant elements in histories are summarized in the states, and thus the problem can be considered as a dynamic programming problem with no discounting. By [26, Theorem 5], there exists an optimal policy which is stationary, and we use σ^* to denote the stationary optimal policy.

Suppose that the deviating node uses σ^* while all the other nodes follow $\tilde{\sigma}^r$. Let p_t be the expected value of the transmission probability of the deviating node in slot t of a review phase (conditional on null history). Let $I_t = \chi\{z^t = 0\}$. Since $E[I_t] = (1-p_t)(1-p_c)^{N-1}$, we have

$$\begin{aligned} U(\sigma^*;\tilde{\sigma}^r) &= \frac{(1-p_c)^{N-1}\sum_{t=\tau+1}^{\tau+L} p_t}{L + P_f^* M} \\ &= \frac{L(1-p_c)^{N-1} - E\left[\sum_{t=\tau+1}^{\tau+L} I_t\right]}{L + P_f^* M} \end{aligned} \quad (9)$$

where $\tau+1$ is the first slot of a review phase and P_f^* is the punishment probability, i.e., $P_f^* = \Pr\left\{\sum_{t=\tau+1}^{\tau+L} I_t \leq L(\tilde{q}_c - B)\right\}$. Since $\sum_{t=\tau+1}^{\tau+L} I_t \geq 0$, using Markov's inequality, we have

$$E\left[\sum_{t=\tau+1}^{\tau+L} I_t\right] \geq (1-P_f^*)L(\tilde{q}_c - B). \quad (10)$$

Combining (9) and (10), we obtain

$$U(\sigma;\tilde{\sigma}^r) \leq \frac{Lq_c + P_f^*(1-p_c)^N L + (1-P_f^*)BL}{L + P_f^* M} \quad (11)$$

for all $\sigma \in \Sigma$.

Choose arbitrary $\epsilon > 0$ and $\delta > 0$. As in [17], we relate the choice of M and B to L by $B = \beta L^{\rho-1}$ and $M = \mu L$ for some $\beta > 0$, $\rho \in (1/2, 1)$, and $\mu > N-1$. By Chebychev's inequality

$$\tilde{P}_f(B,L) \leq \frac{\tilde{q}_c(1-\tilde{q}_c)}{B^2 L} = \frac{\tilde{q}_c(1-\tilde{q}_c)}{\beta^2 L^{2\rho-1}}. \quad (12)$$

Also, note that

$$C(\tilde{\sigma}^r) = \frac{N\tilde{P}_f M q_c}{L + \tilde{P}_f M} = \frac{N\tilde{P}_f \mu q_c}{1 + \tilde{P}_f \mu}. \quad (13)$$

Since $\tilde{P}_f(B,L)$ in (12) converges to zero as L goes to infinity, we can achieve an arbitrarily small efficiency loss in (13) by choosing sufficiently large L. In other words, for any $\kappa > 0$, there exists L_κ such that $C(\tilde{\sigma}^r) \leq \kappa$ for all $L \geq L_\kappa$. With $\mu > N-1$, we can show that the upper bound on the deviation payoff in (11)

$$\frac{q_c + P_f^*(1-p_c)^N + (1-P_f^*)\beta L^{\rho-1}}{1 + P_f^* \mu}$$

is decreasing in P_f^*. Thus, the deviation payoff is bounded above by $q_c + \beta L^{\rho-1}$.

Choose L such that

$$L \geq \max\left\{L_\delta, L_{N\epsilon/2}, \left(\frac{2\beta}{\epsilon}\right)^{\frac{1}{1-\rho}}\right\}.$$

Since $L \geq L_{N\epsilon/2}$, we have

$$q_c - \epsilon/2 \leq U(\tilde{\sigma}^r;\tilde{\sigma}^r) \leq U(\sigma^*;\tilde{\sigma}^r). \quad (14)$$

Since $L \geq (2\beta/\epsilon)^{1/(1-\rho)}$, we have

$$U(\sigma^*; \tilde{\sigma}^{\mathrm{r}}) \leq q_{\mathrm{c}} + \beta L^{\rho-1} \leq q_{\mathrm{c}} + \frac{\epsilon}{2}. \tag{15}$$

Then, by (14) and (15), we obtain

$$U(\sigma^*; \tilde{\sigma}^{\mathrm{r}}) - U(\tilde{\sigma}^{\mathrm{r}}; \tilde{\sigma}^{\mathrm{r}}) \leq \epsilon$$

which proves that $\tilde{\sigma}^{\mathrm{r}}(B, L, M)$ is an ϵ-NE. Lastly, since $L \geq L_\delta$, we have $C(\tilde{\sigma}^{\mathrm{r}}) \leq \delta$, and thus $\tilde{\sigma}^{\mathrm{r}}(B, L, M)$ is δ-PO.

REFERENCES

[1] *Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE 802.11, Aug. 1999.

[2] Y. E. Sagduyu, R. Berry, and A. Ephremides, "MAC games for distributed wireless network security with incomplete information of selfish and malicious user types," in *Proc. IEEE GameNets*, May 2009, pp. 130–139.

[3] L. Chen, S. H. Low, and J. C. Doyle, "Random access game and medium access control design," *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1303–1316, Aug.. 2010.

[4] J. Park and M. van der Schaar, "Medium access control protocols with memory," *IEEE/ACM Trans. Netw.*, vol. 18, no. 6, pp. 1921–1934, Dec. 2010.

[5] R. T. Ma, V. Misra, and D. Rubenstein, "An analysis of generalized slotted-Aloha protocols," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 936–949, Jun. 2009.

[6] E. Altman, R. El Azouzi, and T. Jimenez, "Slotted Aloha as a game with partial information," *Comput. Netw.*, vol. 45, no. 6, pp. 701–713, Aug. 2004.

[7] L. Buttyán and J.-P. Hubaux, *Security and Cooperation in Wireless Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[8] G. Tan and J. Guttag, "The 802.11 MAC protocol leads to inefficient equilibria," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005, vol. 1, pp. 1–11.

[9] M. Cagalj, S. Ganeriwal, I. Aad, and J.-P. Hubaux, "On selfish behavior in CSMA/CA networks," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005, vol. 4, pp. 2513–2524.

[10] P. Kyasanur and N. H. Vaidya, "Selfish MAC layer misbehavior in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 4, no. 5, pp. 502–516, Sep. 2005.

[11] A. B. MacKenzie and S. B. Wicker, "Stability of multipacket slotted Aloha with selfish users and perfect information," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003, vol. 3, pp. 1583–1590.

[12] Y. Jin and G. Kesidis, "Equilibria of a non-cooperative game for heterogeneous users of an Aloha network," *IEEE Commun. Lett.*, vol. 6, no. 7, pp. 282–284, Jul. 2002.

[13] J. Park and M. van der Schaar, "Stackelberg contention games in multiuser networks," *EURASIP J. Adv. Signal Process.*, vol. 2009, p. 305978, 2009.

[14] Y. Jin and G. Kesidis, "A pricing strategy for an Aloha network of heterogeneous users with inelastic bandwidth requirements," in *Proc. CISS*, Princeton, NJ, Mar. 2002, pp. 1030–1033.

[15] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Repeated open spectrum sharing game with cheat-proof strategies," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1922–1933, Apr. 2009.

[16] R. J. La and V. Anantharam, "Optimal routing control: Repeated game approach," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 437–450, Mar. 2002.

[17] R. Radner, "Repeated principal-agent games with discounting," *Econometrica*, vol. 53, no. 5, pp. 1173–1198, Sep. 1985.

[18] R. Radner, "Repeated partnership games with imperfect monitoring and no discounting," *Rev. Econ. Stud.*, vol. 53, no. 1, pp. 43–57, Jan. 1986.

[19] L. G. Roberts, "Aloha packet system with and without slots and capture," *Comput. Commun. Rev.*, vol. 5, no. 2, pp. 28–42, Apr. 1975.

[20] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1992.

[21] B. Hajek and T. van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Autom. Control*, vol. AC-27, no. 3, pp. 559–569, Jun. 1982.

[22] M. Kandori, "Introduction to repeated games with private monitoring," *J. Econ. Theory*, vol. 102, no. 1, pp. 1–15, Jan. 2002.

[23] P. Billingsley, *Probability and Measure*. New York: Wiley, 1995.

[24] E. Kalai and W. Stanford, "Finite rationality and interpersonal complexity in repeated games," *Econometrica*, vol. 56, no. 2, pp. 397–410, Mar. 1988.

[25] G. Bianchi, "Performance analysis of the IEEE 802. 11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[26] D. Blackwell, "Discrete dynamic programming," *Ann. Math. Statist.*, vol. 33, no. 2, pp. 719–726, Jun. 1962.

Khoa Tran Phan received the B.Sc. degree in telecommunications (First Class Honors) from the University of New South Wales, Sydney, Australia, in 2005, the M.Sc. degree in electrical engineering from the University of Alberta, Edmonton, AB, Canada, in 2008, and the M.Sc. degree in electrical engineering from the California Institute of Technology (Caltech), Pasadena, in 2009. He was a Ph.D. student with the Electrical Engineering Department, University of California, Los Angeles, from 2009 to 2011.

He is now with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada. His current research interests are efficient protocol design, scheduling in wireless networks, network economics, and applications of game theory and mechanism design in communications networks.

Jaeok Park received the B.A. degree in economics from Yonsei University, Seoul, Korea, in 2003, and the M.A. and Ph.D. degrees in economics from the University of California, Los Angeles, in 2005 and 2009, respectively.

He is currently an Assistant Professor with the School of Economics, Yonsei University. From 2009 to 2011, he was a Postdoctoral Scholar with the Electrical Engineering Department, University of California, Los Angeles. From 2006 to 2008, he served in the Republic of Korea Army. His research interests include game theory, mechanism design, network economics, and wireless communication.

Mihaela van der Schaar (M'98–SM'04–F'10) received the M.S. and Ph.D. degrees in electrical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 1996 and 2001, respectively.

She is the Chancellor's Professor of Electrical Engineering with the University of California, Los Angeles. Her research interests include multimedia networking, communication, processing, and systems, multimedia stream mining, dynamic multiuser networks and system designs, online learning, network economics and game theory

Prof. van der Schaar is a Distinguished Lecturer of the IEEE Communications Society for 2011–2012, the Editor-in-Chief of the IEEE TRANSACTIONS ON MULTIMEDIA, and a member of the Editorial Board of the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING. She received an NSF CAREER Award in 2004, the Best Paper Award from the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 2005, the Okawa Foundation Award in 2006, the IBM Faculty Award in 2005, 2007, and 2008, the Most Cited Paper Award from EURASIP: *Image Communications Journal* in 2006, the Gamenets Conference Best Paper Award in 2011, and the 2011 IEEE Circuits and Systems Society Darlington Award.