Keeping Your Story Straight: Truthtelling and Liespotting

Johannes Hörner[*], Xiaosheng Mu[†], Nicolas Vieille[‡]

April 11, 2026

Abstract

An agent privately observes a Markov chain online and reports it to a designer. To what patterns in the reported data should the designer pay attention? We show that, in general, keeping track of the empirical frequency of transition counts in the agent's reports is insufficient, despite the true state being Markovian. Nonetheless, we derive conditions under which any deviation that can be distinguished from truth-telling by checking the frequency of strings of an arbitrary (finite) size can be detected by "checking pairs." Further, we find that some undetectable deviations cannot be profitable, independent of the agent's preferences. Hence, we provide weaker sufficient conditions that ensure that the agent finds honesty to be the best strategy. We explore the implications of these results for the literature on (i) linking incentives, (ii) dynamic implementation, and (iii) repeated games and agency models.

Keywords: Detectability, testing, Markov chains, implementation, repeated games.

JEL codes: C72, C73

---

[*] CNRS (TSE), 1, Esplanade de l'Université, 31000 Toulouse, France, `joh.horner@gmail.com`.

[†] Princeton University, 20 Washington Rd, Princeton, NJ 08540, USA `xmu@princeton.edu`.

[‡] HEC, 1, rue de la Libération, 78351 Jouy en Josas, France, `vieille@hec.fr`.

1 Introduction

A liar should have a good memory.
—Quintilian

If you tell the truth, you don't have to remember anything.
—Mark Twain

The least initial deviation from the truth is multiplied later thousandfold.
—Aristotle

Time and repetition provide opportunities to screen. To do so well, one must interpret the evidence that accumulates over time. Persistence in private information makes this especially valuable, as it constrains the scope for lying. However, it also raises a challenge: what patterns in the reports should one look for? An agent observes a Markov chain online and reports the states as they arrive to a designer. Preferences are summarized by the agent's reward function over state and report; technology is summarized by the Markov matrix. The designer seeks to elicit truth-telling in every round. She merely decides which report sequences are acceptable. In particular, she can require that the agent's reports be consistent with specific zero-one laws.

For instance, the designer might constrain the agent to choosing from those infinite sequences that are consistent with the long-run frequency of each state. Indeed, testing the state frequency is common in the literature (see below). It is both a simple and a surprisingly powerful test. Yet, because the underlying process is Markovian, it is natural to check for the empirical frequency of *pairs* of consecutive reports, requiring that this frequency matches the theoretical transition count of states. One might hope that such a test would outperform the former and be without loss of generality. Our first example shows that it performs better, but it does not perform best. Prevarication adapts to liespotting: when pairs are checked, the agent gains from employing more sophisticated strategies, which satisfy the test of pairs but pay no heed to richer statistics.

We allow the designer to test for the frequency of k-tuples of reports ("testing k-tuples"), for any positive integer k, and address the following three questions.

First, what are the undetectable distributions (over states and reports) that the agent can engineer, for a given test? Second, when is truth-telling attainable, given the agent's preferences? These first two questions admit simple answers: the set of such distributions is a convex polytope, with a distinguished vertex associated with truth-telling. Checking whether truth-telling is attainable with a given test amounts to checking whether the agent's expected utility is maximized by this distinguished vertex—a simple matter of linear programming.

Third, for each of these two questions, is it without loss to consider states, pairs, or more generally, k-tuples, in the sense that considering longer strings will not help in reducing the set of undetectable distributions or in inducing truth-telling, independent of the agent's preferences?

We show that testing (singleton) states is without loss of generality in some cases that are both special yet commonly assumed in economics: for instance, when there are two possible states only, or when the Markov chain is renewal. However, we prove that, provided that the Markov chain has three states or more, testing for singleton states only is with loss for almost all transition matrices. As suggested above, testing pairs is also not without loss. For some Markov chains, testing triples is better: doing so further reduces the set of undetectable distributions and expands the set of preferences for which speaking the truth is optimal.

There is nothing special about pairs. The same holds for k-tuples ($k \geq 2$). For some Markov chains, testing $k+1$-tuples shrinks the set of undetectable distributions that the agent can engineer and expands the set of preferences for which honesty is best. When restricting attention to three states, we obtain two further unexpected findings. First, while testing pairs (or k-tuples) is generally not without loss, there is an open positive-measure set of Markov matrices for which it is, unlike for singleton states. Second, there is an open, positive-measure set of Markov matrices for which testing longer strings (rather than pairs) affects the set of undetectable distributions *yet* does not affect the range of preferences for which honesty is best. For such Markov chains, refining the test reduces the scope for lying, but the lies that get pruned out are irrelevant, independent of the agent's preferences. To be clear, even when pairs, or k-tuples, suffice to test the joint distributions over states and reports (a sufficient statistic for payoffs with time-separable preferences, if the agent is patient), more complicated tests might still reduce the set of distributions over the processes themselves.

Our focus is on how the designer uses the information available to her, not on the instrument at her disposal. Yet, this instrument matters for the choice of relevant information, whether this is authority, as in our baseline model, or money, as in Section 4.1. There, we establish a duality between money and authority: restricting the agent to sequences that fit the theoretical frequency of k-tuples is equivalent to using transfers that only depend on the last k reports.[1] For

---

[1] Transfers with bounded memory satisfy desirable properties in related contexts. See, for instance, Bhaskar, Mailath and Morris (2013).

instance, forcing the agent to match the theoretical frequency of each reported state amounts to assuming that the designer is only able to condition transfers on the latest report. Hence, money buys time: there is no need to wait literally forever to pass judgment.

This paper has implications for implementation. In Section 4.2, we show the relationship between our results and the existing ones with respect to static implementation. More precisely, our results extend those of Rochet (1987) to a dynamic environment, couched in terms of cyclical monotonicity. By determining the set of incentive-feasible allocation functions, implementation is the first step in the design of the optimal mechanism given a particular objective function, such as efficiency or revenue maximization. Baron and Besanko (1984), Battaglini (2005), and other related works, are examples of such specific design problems, where the state (the valuation of the agent) follows a Markov chain. Examples of such hidden-knowledge agency problems arise frequently in regulation (*e.g.*, rate setting by public firms that have superior information regarding the current state of demand).

Our results also have immediate consequences for dynamic agency problems and Bayesian games. To the extent that there is no "universal" test –that is, no integer k that would be without loss, independent of the environment– there is also no hope of extending the celebrated "Bellman-type" characterizations of the equilibrium payoff set to environments with persistent information and interdependent values.[2] As we explain in Section 4.3, keeping track of more complicated statistics of past evidence implies that, in general, the dynamic game cannot be summarized by a one-shot (or, more generally, k-shot) game.

Section 2 introduces an example that illustrates the main ideas. Section 3 generalizes the example and contains the main results. Section 4 develops what they mean for the role of money in dynamic relationships, for implementation, dynamic games and agency.

Related Literature: Many papers examine the power of "linking incentive constraints" under imperfect observability. In the context of dynamic agency, early examples include Radner (1981), Townsend (1982) and Rubinstein and Yaari (1983). In the context of dynamic games, Abreu, Milgrom and Pearce (1991) shows how tying together multiple draws help sustain cooperation (see also Samuelson and Stacchetti, 2017). In the static context, Fang and Norman (2006), Jackson and Sonnenschein (2007) and Matsushima, Miyazaki and Yagi (2010) develop similar ideas. All these papers focus on the i.i.d. case and use a version of the frequency test for singleton

---

[2] Such characterizations appear for instance in Shapley (1953) and Abreu, Pearce and Stacchetti (1990) for a given discount factor, and in Fudenberg and Levine (1994) for the patient limit.

states (variously referred to as "review strategies" or "quota mechanisms").[3] In our Markovian setting, additionally imposing future reports to be consistent with the current report makes the agent less willing to lie in the fear of future punishments. Nonetheless, this conclusion implicitly relies on the assumption that the agent does not know future states in advance (even though he has an informational advantage). As we show in Lemma 4, checking singleton state frequencies would be sufficient if the agent were to know future states.[4]

To the best of our knowledge, Escobar and Toikka (2013) is the first to consider a test that goes beyond singleton states. Theirs is a variation on the test of pairs. As they argue in Section 2, testing pairs improves on testing singletons in their Markovian environment.[5] More broadly, our paper is closely related to the literature on repeated games with incomplete information that follows Aumann and Maschler (1995), in particular, Renault, Solan and Vieille (RSV, 2013) and Hörner, Takahashi and Vieille (2015). See also Athey and Bagwell (2008) and Barron (2017). There are two major differences between our work and these contributions. First, theirs are games without commitment, and hence, the solution concepts differ. Second, this literature focuses on the characterization of the set of equilibrium payoffs.[6] In our environment, the range of possible payoffs for the agent is trivial: his payoff is minimized by a report-independent allocation function, and his maximum payoff is certainly incentive-compatible. Our focus is instead on whether truthful reporting can be implemented, and our results in turn have implications for dynamic Bayesian games with interdependent values (*i.e.*, when the designer's payoff is also of interest).[7] Some of our techniques and ideas are clearly related to a large literature. In particular, the representation of undetectable distributions in terms of copulas when the state

---

[3] Here, we blend together games of moral hazard and adverse selection. Repeated games with adverse selection are often viewed as a special class of repeated games with imperfect monitoring (with product structure), see, *e.g.*, Mailath and Samuelson (2006). Therefore, it is no surprise that quota mechanisms have been successfully applied in both contexts. But of course, this is only possible when the type is i.i.d., which is perhaps the least interesting case considered here.

[4] This does not mean that more complicated tests would not be useful in terms of rates of convergence, as shown by Cohn (2010) in the setup of Jackson and Sonnenschein (2007).

[5] The reason why such a test is useful in their environment is related, but slightly different: what matters is to cross-check a player's report with her opponent's previous report, to ensure that a player cannot engineer and benefit from a profitable correlation between her report and the opponent's likely concurrent state. Therefore, an important difference between their set-up and ours is that the past signal against which a player's report is tested is out of her control, since it is her opponent's report.

[6] Indeed, the main result of these papers is that under certain assumptions, the equilibrium payoff set can be described by a relatively simple recursive equation (involving, at most, pairs of states) or obtained by strategies that rely on testing at most pairs.

[7] In a dynamic mechanism design environment, Athey and Segal (2013) establish the existence of efficient, incentive-compatible and under certain assumptions budget-balanced mechanisms. Here also, the focus on efficiency and private values implies that the mechanism can be represented in a recursive fashion.

follows a pseudo-renewal chain is directly borrowed from RSV.

Our paper yields a dynamic counterpart of Rochet (1987)'s characterization that cyclical monotonicity is a necessary and sufficient condition for implementation.[8] Rahman (2010) provides a general (extensive-form) version of the result that implementation is equivalent to undetectable deviations being unprofitable. In our context, the entire difficulty is to determine which deviations affect the distribution over outcomes but are undetectable. Despite its name, the literature on repeated implementation is less related. Lee and Sabourian (2011, 2015) and Mezzetti and Renou (2017) provide sufficient conditions on an allocation function to guarantee *full* implementation, whereas our focus is to exactly characterize implementation in the single-agent Markovian case.[9] We note that with a single agent, our notion of truthful implementation is equivalent (up to indifferences) to full implementation. But it is an interesting question for future work how our characterization here may generalize to multiple agents, and whether it may inform full implementation even when agents do not have complete information.

We do not explore here what additional structure a specific objective function for the designer might impose. In the case of revenue maximization, Battaglini and Lamba (2017) shows the difficulty of characterizing optimal mechanisms in Markovian environments with three or more states, unless the environment has a special, non-generic structure (*e.g.*, AR(1) processes). Their analysis contrasts with the simple mechanism obtained by Battaglini (2005), which focuses on two states only. The analogy between this dichotomy in their results and ours has its limitations, however: Battaglini's mechanism can be described by a simple state variable, but has infinite memory nonetheless.

Our paper is also related to some literature in statistics. The problem of identifying hidden Markov chains has a long tradition in econometrics; see e.g. Blackwell and Koopmans (1957), Rothschild (1982) and Connault (2016). Our problem is quite different, however, as there is no exogenous signal about the hidden Markov chain but instead a report, which is strategically chosen as a function of the entire history, and hence, does not satisfy the usual assumptions imposed on the signal distribution. We cannot simply restrict the agent to using Markovian strategies: the more complicated the test applied to the agent is, the more sophisticated his best-reply. Finally, our study of undetectable deviations is related to examples of non-Markovian processes that nevertheless satisfy the Chapman-Kolmogorov equation of a given Markov chain; see Feller (1959) and Rosenblatt (1960). We go beyond these examples by demonstrating which

---

[8] See also Rockafellar (1966) and Afriat (1967).

[9] Renou and Tomala (2015) allows for Markovian states, but does not attempt a characterization.

deviations matter for the joint distribution of states and reports, or for incentives.

2 An Example

Here, we develop an example to introduce the main ideas. We gloss over some of the formal or technical details, which are revisited in later sections.

In each round $n = 1, 2, \ldots$, an agent (he) observes the realization of a Markov chain, with no foresight. The Markov chain takes values in $S = \{s_1, s_2, s_3\}$ according to the transition matrix

$$\begin{array}{c|ccc} & s_1' & s_2' & s_3' \\ \hline s_1 & 1/2 & 1/2 & 0 \\ s_2 & 0 & 3/4 & 1/4 \\ s_3 & 1/2 & 0 & 1/2 \end{array}$$

That is, state s_2' follows s_1 with probability (w.p.) 1/2, etc. Note that the ergodic distribution λ assigns equal weight to s_1 and s_3 and more weight to s_2: state 2 has invariant probability 1/2, while states 1 and 3 have probability 1/4 each.[10]

In round n, the agent is invited to make a report a_n regarding the prevailing state. For now, we assume that this report is an element from a copy of S, denoted A, to distinguish reports from states. Report a_i, $i = 1, 2, 3$, is interpreted as referring to state s_i.

The goal of the designer or principal (she) is to elicit truth-telling from the agent. Unfortunately, she has access to rather indigent information, as she observes the agent's reports only. Her instruments are equally primitive: her only option is to impose constraints on the sequence of reports that the agent is allowed to produce. Neither the agent nor the designer is in a rush, and hence, the fulfillment of these constraints can be evaluated "at the end of time."

For instance, she can require the agent to report s_2 one-half of the time and the other two states one-quarter of the time. An honest agent passes this test almost surely (a.s.). Indeed, a large literature has focused on (some version of) this test, showing that it implements any *ex-ante efficient* social choice function (see Townsend, 1982, Jackson and Sonnenschein, 2007). With this test, the designer is as effective as if she had transfers at her disposal, in a one-shot interaction.

---

[10] While this is irrelevant, we may think of the initial state as being drawn according to this ergodic distribution.

However, our designer may not wish to maximize the agent's payoff only. For instance, she might take into account the welfare of other (uninformed and unmodeled) agents whose payoffs also depend on the state. This might not suit the agent. Truth-telling might be beyond reach, whatever constraint is placed on the report sequence. The question, then, is the following: assuming that it is possible to keep the agent honest, what is a simple way to do so?

We bypass the underlying decision problem that maps reports into outcomes by positing some utility function for the agent over state and report, $r : S \times A \to \mathbf{R}$. His realized payoff over N rounds is then

$$\frac{1}{N}\sum_{n=1}^{N} r(s_n, a_n),$$

and his goal is to maximize the (lower) limit as $N \to \infty$ of his expected payoff over admissible sequences. Equivalently, the agent chooses a distribution $\mu \in \Delta(S \times A)$ over states and reports, interpreted as the limit frequency of these pairs arising under some reporting policy. Viewed in this way, the agent maximizes

$$\mathbf{E}_\mu[r(s,a)],$$

over those distributions μ that he can engineer through a reporting policy that abides by the constraints imposed by the designer. To be concrete, let us focus on a subset of such distributions. For $x \in [0, 1/4]$, consider the distribution μ_x over states and reports given by

$$\begin{array}{c c} & \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 1/4 - x & x & 0 \\ 0 & 1/2 - x & x \\ x & 0 & 1/4 - x \end{pmatrix} \end{array}$$

When emulating such a distribution, the agent "exaggerates" the state a fraction $3x$ of the time, pretending in those rounds that the prevailing state is $s_{i+1} \pmod 3$ when it is s_i. A high x would be attractive to the agent if his preferences were, say, $r(s_i, a_{i+d}) = d$, $i = 1, 2, 3$, $d = -1, 0, 1$. Note that telling the truth corresponds to the distribution with $x = 0$, and we set $\mu^{tt} := \mu_0$.

2.1 Analysis

What distributions μ_x can the agent bring about while passing the singleton state frequency test? He can engineer *any* desired distribution in this class. Here is a strategy that does the trick: in state s_i, he mixes between reporting a_i and a_{i+1} with suitable probabilities, independent of his private history; *e.g.*, in s_1, he reports a_1 w.p. $1-4x$, and a_2 w.p. $4x$. This policy passes the singleton test because the marginal distribution of μ_x on A (the sum over each column) is equal to the invariant distribution λ. For each $\mu \in \mathcal{M}_0 = \{\mu \in \Delta(S \times A) : \text{marg}_S\mu = \text{marg}_A\mu = \lambda\}$, one can construct such a policy.

Yet it is easy to spot the agent's lie when emulating $\mu_{\frac{1}{4}}$, that is, the distribution

$$\begin{array}{cc} & \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 0 & 1/4 & 0 \\ 0 & 1/4 & 1/4 \\ 1/4 & 0 & 0 \end{pmatrix} \end{array}$$

Indeed, to achieve $\mu_{\frac{1}{4}}$, the agent has no choice but to report a_1 when the state is s_3. However, state s_3 sometimes occurs after s_2, when the agent is willing to report a_2 one-half of the time without yet knowing what the next state will be. Hence, the agent must sometimes report a_2 followed by a_1, a sequence that cannot occur under the true state transitions. Checking pairs of consecutive reports suffices to detect such a lie. Intuitively, because states are not i.i.d., it is useful to check the frequency of pairs of states, rather than singletons. As we will see, this intuition is essentially correct: checking singleton states does as well as pairs (or triples, etc.) *only* for i.i.d. chains, or rather, for a slightly broader class of Markov chains (Theorem 1).

Because the underlying state follows a Markov chain, it is natural to keep track of the frequency of pairs (as in Escobar and Toikka, 2013). However, is it enough? Consider the distribution $\mu_{\frac{1}{6}}$ given by

$$\begin{array}{cc} & \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 1/12 & 1/6 & 0 \\ 0 & 1/3 & 1/6 \\ 1/6 & 0 & 1/12 \end{pmatrix} \end{array}$$

Given a report a_1 or a_3, the "lower" state (s_3 or s_2) occurs twice as often as the reported state.

Suppose that the previous report was a_3 and the current state is s_2. What should the agent report? Reporting a_1 is not an option, as the entry (s_2, a_1) of $\mu_{\frac{1}{6}}$ is assigned probability 0. He also cannot report a_2, as he reported a_3 previously: state s_3 cannot be followed by s_2. Hence, he has no choice but to report a_3 again. Note that, according to the joint distribution he emulates, when reporting a_3 the state is s_2 two-thirds of the time (*i.e.*, $p(s_2|a_3) = 2/3$), and state s_2 repeats itself w.p. $^3/_4$. Thus, state s_2 occurs after report a_3 at least $\frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2}$ of the time, and in these situations the agent must report a_3 again. But $^1/_2$ is the frequency with which state s_3 repeats itself; hence, after a_3, the agent cannot report a_3 if any other state is realized, lest this be reflected by an excessive frequency of the pair of reports (a_3, a_3). To summarize, when the last report is a_3, the agent reports a_3 if, and only if, the current state is s_2; otherwise he reports a_1 (since a_2 cannot come after a_3). This holds for any reporting policy that generates the correct frequency over pairs (a_3, a_3) and is consistent with the desired joint distribution $\mu_{\frac{1}{6}}$.

However, suppose that our designer keeps track of *triples*. What if she observes two consecutive reports of a_3? Given the argument above, she can deduce that the current state must be s_2, and thus, the next state must also be s_2 w.p. $^3/_4$, whereupon the agent must report a_3 again. Hence, she can predict that the next report will again be a_3 w.p. $^3/_4$, which contradicts the true frequency with which the state s_3 repeats itself: the policy of the agent is then distinguished from truth-telling. This shows the agent cannot engineer $\mu_{\frac{1}{6}}$ while passing the triple test.[11]

The intuition is the following. To have the correct frequency of pairs, the agent must not only take into account his current state (here, s_2) when selecting the appropriate report but also consider yesterday's report (a_3). A liar must indeed take advantage of his memory, unlike an honest agent. As a result, today's state statistically depends on yesterday's report after conditioning on the current report (even though they are independent without conditioning). Since today's state affects tomorrow's state and hence tomorrow's report, tomorrow's report and yesterday's report are not independent conditional on today's report –a violation of the Markov property that can be detected by checking triples.

We have shown that the agent cannot achieve $\mu_{\frac{1}{6}}$ when the designer keeps track of triples. In order to conclude that checking triples does *strictly* better than pairs, we now demonstrate that the agent can get away with the distribution $\mu_{\frac{1}{6}}$ while satisfying the correct frequency of report pairs. As argued above, the agent has a (unique) policy after a report of a_3 to generate the correct frequency over pairs (a_3, a_3). What about pairs (a_1, a_1) and (a_2, a_2)? Suppose that,

---

[11] The argument thus far hinges on the fact that following a_3, the agent reports a_3 if, and only if, the current state is s_2. However, as long as s_2 is more persistent than s_3 and $p(s_2|s_3)$ is small enough, the probability of reporting a_3 following a_3 and s_2 would be too high and a similar contradiction would follow. Hence, the example is robust to small perturbations of the transition probabilities.

conditional on a prior report a_1, his reports are given by[12]

$$\begin{array}{cc} & \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 1/3 & 2/3 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

while, conditional on the previous report a_2, his report follows the rule

$$\begin{array}{cc} & \begin{array}{cc} a_2 & a_3 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 1 & 0 \\ 7/8 & 1/8 \\ 0 & 1 \end{pmatrix} \end{array}$$

It is immediate to verify that this policy achieves $\mu_{\frac{1}{6}}$.

How likely is a_1 to be followed by a_1? Recall that, according to the joint distribution $\mu_{\frac{1}{6}}$, when a_1 is reported the state is either s_3 or s_1, the former being twice as likely as the latter. In the former case, the next state is equally likely to be s_3, which triggers a_1 given the above policy, or s_1, which leads to a_1 w.p. $1/3$. In the latter case, only if the next state is also s_1 (which occurs w.p. $1/2$) does report a_1 occur (w.p. $1/3$). Overall, given report a_1 the next report is a_1 again w.p.

$$\frac{2}{3}\left(\frac{1}{2}\cdot 1+\frac{1}{2}\cdot\frac{1}{3}\right)+\frac{1}{3}\cdot\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{2},$$

which is the correct frequency.

What about the probability that a_2 is followed by another report a_2? A similar decomposition reveals that it is equal to

$$\underbrace{\frac{1}{3}}_{s^{-1}=s_1}\left(\underbrace{\frac{1}{2}}_{s=s_1|s^{-1}=s_1}\cdot\underbrace{1}_{a=a_2|s=s_1}+\underbrace{\frac{1}{2}}_{s=s_2|s^{-1}=s_1}\cdot\underbrace{\frac{7}{8}}_{a=a_2|s=s_2}\right)+\underbrace{\frac{2}{3}}_{s^{-1}=s_2}\cdot\underbrace{\frac{3}{4}}_{s=s_2|s^{-1}=s_2}\cdot\underbrace{\frac{7}{8}}_{a=a_2|s=s_2}=\frac{3}{4},$$

---

[12] In this matrix, rows are states, columns are reports, and entries are probabilities of the report given the state (and the previous report).

where s^{-1} is the past state, and underbraces indicate the event the probability of which is being given (omitting the conditioning on $a^{-1} = a_2$).

The above computations suggest that the candidate reporting policy yields the correct frequency over pairs of identical reports. Note that we have also required the frequency over pairs (a_i, a_{i-1}) to be zero. Hence, in this example the frequency over all pairs of reports is indistinguishable from truth-telling. This proves that by checking triples in addition to pairs, the designer is able to eliminate certain joint distributions the agent can emulate.

Yet, it is not difficult to see that checking triples is not the ultimate test; keeping track of quadruples further reduces the scope of lies, and so forth.

2.2 Discussion

Several remarks are in order. First, the surprising result (if any) is not that some deviating *reporting strategies* become detectable once triples, rather than pairs, are scrutinized. Rather, it is that checking triples limits the set of *joint distributions* over states and reports that the agent can generate in an undetectable fashion. To appreciate the difference, note that with an i.i.d. chain, the singleton test cannot be improved upon. This implies that no further test affects the set of undetectable joint distributions; yet, simple restrictions on pairs would prune, say, perfectly persistent reports. Restrictions on triples would eliminate more reporting strategies, for instance, the one that is persistent in even and odd rounds separately.

Second, we have insisted that the agent tell the truth in all rounds. This is in contrast with many papers, which only require truth-telling in "most" rounds, with high enough probability – certainly a more reasonable demand if the horizon were finite, for instance. Yet our conclusion is robust, in the following sense. Given any $\mu \in \Delta(S \times A)$ close enough to $\mu_{\frac{1}{6}}$ in the example, there is no reporting policy that approximately achieves μ as the joint distribution and that has approximately correct frequency of all report triples. Hence, checking triples improves on checking pairs, even if only approximate truth-telling is required.

Finally, we have studied the example from the perspective of a statistician, who is interested in minimizing the scope of undetectable lies (joint distributions). However, the conclusions would be different from the perspective of an economist concerned with enforcing truth-telling. Note that the agent's payoff is written as

$$\mathbf{E}_\mu[r(s,a)] = \mu_x \cdot r = \mu^{tt} \cdot r + xJ \cdot r,$$

where

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

The linearity in x means that truth-telling is best within the class of distributions $\{\mu_x : x \leq 1/4\}$ if, and only if, $J \cdot r \leq 0$, a property that is *independent of* x. For this reason, checking pairs rather than singletons does *not* enlarge the set of preferences for which truth-telling is optimal, within this class. This suggests that the set of undetectable deviations is not necessarily the appropriate benchmark for the economist. As we argue below, what matters for truthful reporting is the cone spanned by this set, and the difference matters.[13]

3 Main Results

This section generalizes the example, addressing the following questions:

1. What are the distributions over states and reports that the agent can engineer, when the designer checks singletons (pairs, triples, etc.)?
2. When is truth-telling incentive compatible, given such a test and the agent's preferences?
3. In each case, what is the simplest test that is without loss?

For the sake of clarity, we retain the simplest setup and abstract for now from important but auxiliary issues (transfers, discounting, etc.), which are relegated to Section 4 and Appendix B.2.

3.1 Setup

We begin with a time-homogeneous, irreducible and aperiodic Markov chain $(s_n)_{n\geq 1}$, taking values in the finite set S. Transition probabilities are denoted $p(s'|s)$, or $p_{ss'}$, and $P = (p(s'|s))_{s,s'}$ is the transition matrix. We let $\lambda \in \Delta(S)$ denote the invariant probability vector. In an abuse of notation, we also write $\lambda \in \Delta(S^{k+1})$, $k \in \mathbf{N}$, for the invariant probability vector of strings of

---

[13] Our distinction between what is of interest to the statistician versus the economist makes for dramatic effect, but there are good reasons that an economist might also be interested in the set of undetectable distributions. In particular, if truth-telling is out of reach, it matters for the second-best solution. See also Section 4.3.

length $k+1$, namely,

$$\lambda(s^{-k}, \ldots, s) = \lambda(s^{-k})p(s^{-(k-1)}|s^{-k}) \cdots p(s|s^{-1}).$$

The initial state s_1 is drawn from a distribution q.

As before, an agent privately observes the Markov chain online and makes unverifiable reports to the designer. Reports are elements from A, a copy of S. Because preference is over the current state and report (see below), the agent does not benefit from conditioning reports on past realizations of the Markov chain. Hence, a reporting policy is a map $\sigma = (\sigma_n)_{n \geq 1}$, with $\sigma_n : A^{n-1} \times S \to \Delta(A)$, mapping the agent's past reports and the current state into a possibly random report. We denote by $\mathbf{P}_q$ the law of the entire sequence (s_n) of states and by $\mathbf{P}_{q,\sigma}$ the law of the sequence (s_n, a_n) under the policy σ.

The agent has flow utility function $r : S \times A \to \mathbf{R}$. His realized payoff over N rounds is

$$\frac{1}{N} \sum_{n=1}^{N} r(s_n, a_n),$$

and he seeks to maximize the lower limit of its expectation under $\mathbf{P}_{q,\sigma}$ over policies σ.[14]

The above average payoff can also be written as $\mathbf{E}_{\mu^N_{q,\sigma}}[r(s,a)]$, where $\mu^N_{q,\sigma}$ denotes the expected empirical distribution of states and reports over the first N rounds.

The designer wishes to incentivize truth-telling. That is, she wishes the agent to use σ^{tt}, the policy that always reports the true state, independent of past reports (that is, on and off-path, so truth-telling is sequentially rational). Under truth-telling, the agent's expected payoff is simply

$$\mathbf{E}_{\mu^{tt}}[r(s,a)],$$

where $\mu^{tt} \in \Delta(S \times A)$ is the distribution under truth-telling (*i.e.*, $\mu^{tt}(s_i, a_i) = \lambda(s_i)$ for all $s_i \in S$).

To achieve her goal, the designer can require the agent's reports to satisfy certain statistical constraints. Fix an integer k. For $n \geq k+1$, let $f^k_n(a^{-k}, \ldots, a)$ denote the empirical frequency over rounds $i \leq n$, in which the $k+1$ most recent reports are $(a^{-k}, \ldots, a)$. Under truth-telling, (f^k_n) converges to λ as $n \to +\infty$, with probability 1. Hence, a natural way of using the limit statistics of reports is to check whether (f^k_n) converges to λ. This test is captured in the following

---

[14] As an alternative, slightly weaker criterion, the agent maximizes the expectation of the lower limit. Our propositions do not rely on this choice.

definition.

Definition 1 *Σ_k is the set of σ such that $f_n^k \to \lambda$ as $n \to +\infty$, $\mathbf{P}_{q,\sigma}$-a.s.*

There are certainly other tests that are not subsumed by a frequency test on strings (for instance, one could run some automaton). However, for any particular choice of tests, one could presumably obtain results that parallel ours.

The next definition formalizes whether a frequency test enforces truth-telling.

Definition 2 *Truth-telling is k-limit optimal if*

$$\liminf_{N\to\infty} \mathbf{E}_{\mu^N_{q,\sigma}}[r(s,a)] \leq \mathbf{E}_{\mu^{tt}}[r(s,a)],$$

for every $\sigma \in \Sigma_k$. Truth-telling is limit optimal if it is k-limit optimal for some k.

Hence, truth-telling is k-limit optimal if it suffices to require the agent's reports to satisfy the correct frequency over strings of length $k+1$.[15]

The limit of means criterion has the benefit of delivering particularly clean results. However, our results have counterparts for low discounting.[16] The sets of distributions defined in this paper ($\mathcal{M}_k$ and $\mathcal{C}_k$ as defined below) still play a role for enforcing truth-telling under discounting, as we show in Appendix B.2.

3.2 Undetectable Deviations

In the example in Section 2, we saw how the agent can engineer any joint distribution $\mu \in \Delta(S \times A)$ that passes the test based on singleton states, provided that μ has the correct marginal distribution on reports. Recall the following:

Definition 3 *Let $\mathcal{M}_0$ be the set of distributions $\mu \in \Delta(S \times A)$ such that $marg_S\mu = marg_A\mu = \lambda$.*

We now generalize this definition to tests based on longer strings of reports. Fix $k \geq 0$, and a *stationary* reporting policy $\sigma : A^k \times S \to \Delta(A)$ mapping the last k reports and the current state into a report (alongside some arbitrary specification for the first k rounds). For obvious

---

[15] We use the $\liminf$ criterion to simplify the proof of Proposition 2. If we considered vanishing discounting instead, the choice of $\liminf$ vs. $\limsup$ would be irrelevant.

[16] See Frankel (2016) for an analysis of quota mechanisms, in the spirit of Jackson and Sonnenschein (2007), with discounting.

reasons, we say such a policy has *memory k*. Any such policy induces a Markov chain over the set $A^k \times S \times A$, which may admit multiple invariant measures.

Definition 4 *Fix $k \geq 0$. Let $\mathcal{M}_k$ be the set of distributions $\mu \in \Delta(S \times A)$ such that there exists a stationary reporting policy $\sigma : A^k \times S \to \Delta(A)$ and an invariant distribution $\nu \in \Delta(A^k \times S \times A)$ for σ with the following two properties:*

- *(i) the marginal of ν over $S \times A$ is equal to μ.*
- *(ii) the marginal of ν over A^{k+1} is equal to λ, the invariant distribution of $(s_{n-k}, \ldots, s_n)$.*

Intuitively, this is the set of joint distributions that can be generated by a policy indistinguishable from truth-telling, when strings of length $k+1$ are scrutinized. The set $\mathcal{M}_k$ is the relevant object for the statistician, as it summarizes the joint distributions that cannot be detected given a test based on such strings. Plainly, the sequence $(\mathcal{M}_k)_{k \in \mathbf{N}}$ is nested. In Section 3.4, we ask whether it is eventually constant (for the example in Section 2, it is not).

Our first result below explains the interest in $\mathcal{M}_k$, which turns out to be the limit set of expected empirical joint distributions under a reporting policy that passes the memory-k test.

Proposition 1 *For every $\mu \in \mathcal{M}_k$ there exists $\sigma \in \Sigma_k$ such that $\lim_{N\to\infty} \mu^N_{q,\sigma} = \mu$ for each q. Conversely, for every $\sigma \in \Sigma_k$ such that $\mu := \lim\limits_{N\to\infty} \mu^N_{q,\sigma}$ exists for some q, one has $\mu \in \mathcal{M}_k$.*

The (second-half) result has bite because reporting policies in Σ_k can be arbitrary, while the definition of $\mathcal{M}_k$ only considers stationary policies.

How does one compute $\mathcal{M}_k$? Its definition invokes the existence of a stationary policy σ with certain properties. The next lemma provides an alternative characterization in terms of the invariant distribution ν.

Lemma 1 *For every k, $\mathcal{M}_k$ is a convex polytope.*[17] *It is the set of distributions $\mu \in \Delta(S \times A)$ such that the following linear system in $\nu(a^{-k}, \ldots, a^{-1}, s, a)$ has a solution:*

$$\sum_{a^{-k},\ldots,a^{-1}} \nu(a^{-k}, \ldots, a^{-1}, s, a) = \mu(s, a). \tag{1}$$

---

[17] See Gale (1960), for instance.

$$\sum_{s} \nu(a^{-k}, \ldots, a^{-1}, s, a) = \lambda(a^{-k}) p(a^{-(k-1)}|a^{-k}) \cdots p(a|a^{-1}). \tag{2}$$

$$\sum_{a} \nu(a^{-k}, \ldots, a^{-1}, s, a) = \sum_{s^{-1}} \left(p(s|s^{-1}) \cdot \sum_{a^{-(k+1)}} \nu(a^{-(k+1)}, \ldots, a^{-2}, s^{-1}, a^{-1}) \right). \tag{3}$$

$$\nu(a^{-k}, \ldots, a^{-1}, s, a) \geq 0. \tag{4}$$

Nonetheless, the polytope $\mathcal{M}_k$ is not easy to describe. In the special case $|S| = 3$ (which we focus on below), there exists a direct description of the set $\mathcal{M}_1$ by means of "dual" constraints on μ.[18] In general, we can show $\mathcal{M}_k$ is "well-behaved" in the following sense.

Lemma 2 *For all k, the polytope $\mathcal{M}_k$ has dimension $(|S|-1)^2$ and varies continuously with the transition matrix.*

Lemma 2 implies that the inequality $\mathcal{M}_1 \neq \mathcal{M}_2$ obtained in the example in Section 2 is robust to small changes in the transition matrix. In particular, the inequality still holds if all transition probabilities are strictly positive.

3.3 Unprofitable Deviations

Given Proposition 1, the following should come as no surprise.

Proposition 2 *Truth-telling is k-limit optimal if, and only if,*

$$\mathbf{E}_{\mu}[r(s,a)] \leq \mathbf{E}_{\mu^{tt}}[r(s,a)] \text{ for all } \mu \in \mathcal{M}_k. \tag{5}$$

However, the sets $\mathcal{M}_k$ do not provide a tight characterization of truth-telling for the following reason. It could be that $\mathcal{M}_k \neq \mathcal{M}_{k+1}$, while all rays spanned with vertices adjacent to μ^{tt} (a vertex that both sets have in common) are identical. That is, these polytopes might define the same cone pointed at μ^{tt}. Formally, we define these cones as follows.

Definition 5 *Given $k \geq 0$, let*

$$\mathcal{C}_k := \{\mu \in \Delta(S \times A) : \mu = \mu^{tt} + \alpha(\mu' - \mu^{tt}), \text{for some } \alpha \geq 0, \mu' \in \mathcal{M}_k\}.$$

---

[18] That is, there exists a finite set of linear inequalities on $\mu(s,a)$ (without existential quantifiers) that characterizes $\mathcal{M}_1$.

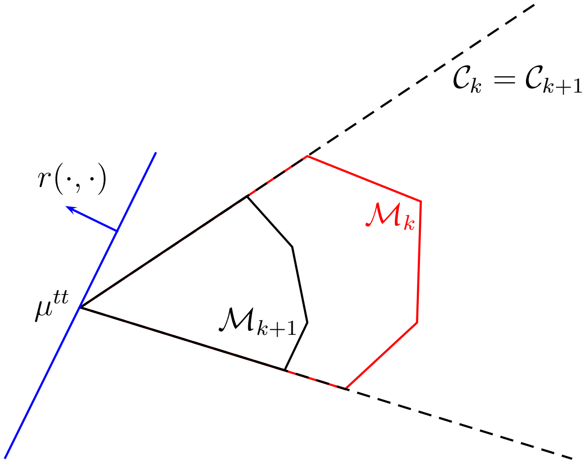

Figure 1: An example in which $\mathcal{M}_k \neq \mathcal{M}_{k+1}$, yet $\mathcal{C}_k = \mathcal{C}_{k+1}$.

Plainly, $\mathcal{C}_k \supset \mathcal{C}_{k+1}$. See Figure 1 for an illustration. In that case, despite the proper set inclusion, one cannot find a utility function $r(\cdot,\cdot)$ for which truth-telling maximizes expected utility within $\mathcal{M}_{k+1}$ but not within $\mathcal{M}_k$.

Such a situation appears improbable, but it is not impossible. In our leading example, $\mu_{\frac{1}{4}} \notin \mathcal{M}_1$, but it lies on the ray spanned by μ^{tt} and $\mu_{\frac{1}{6}} \in \mathcal{M}_1$ since we can write $\mu_{\frac{1}{4}} = \mu^{tt} + \frac{3}{2} \cdot (\mu_{\frac{1}{6}} - \mu^{tt})$. Thus, (at least) that ray remains the same as we move from $\mathcal{M}_0$ to $\mathcal{M}_1$, despite the fact that $\mathcal{M}_1$ contains a smaller segment of the ray.[19] Intuitively, the joint distribution $\mu_{\frac{1}{6}}$ can be viewed as a mixed strategy, mixing between truth-telling and $\mu_{\frac{1}{4}}$. So for any utility function, whether truth-telling is better than $\mu_{\frac{1}{6}}$ is equivalent to whether it is better than $\mu_{\frac{1}{4}}$. Incentives are unaffected in such a case.

From the perspective of incentives, the cone $\mathcal{C}_k$ is the proper object of study. By a standard separation argument,

$$\mathcal{C}_k \neq \mathcal{C}_{k+1} \Leftrightarrow \exists\, r \in \mathbf{R}^{|S|\times|A|} : \mathbf{E}_{\mu^{tt}}[r(s,a)] = \max_{\mu \in \mathcal{M}_{k+1}} \mathbf{E}_{\mu}[r(s,a)] < \max_{\mu \in \mathcal{M}_k} \mathbf{E}_{\mu}[r(s,a)].$$

That is, if the cones differ, there are preferences such that truth-telling is $k+1$-limit optimal but

---

[19] In fact, this is a more general feature of that example: as k increases, the set $\{\mu_x : \mu_x \in \mathcal{M}_k\}$ shrinks but never reduces to μ^{tt}.

not k-limit optimal.

3.4 How Much Memory?

Depending on the agent's preferences, it can be impossible to induce the agent to tell the truth. Nevertheless, we may ask when the restriction to Σ_k is without loss *independent of preferences*, in the sense that truth-telling is limit optimal if, and only if, it is k-limit optimal. Given the discussion before, we need to study whether $\mathcal{C}_k = \mathcal{C}_{k'}$ for all $k' > k$. The same can be asked about the set $\mathcal{M}_k$, the set of undetectable distributions. These are properties of the Markov chain alone.

3.4.1 Testing Singleton States

We begin with the simplest test, namely, singleton states. We define a particular class of Markov chains.

Definition 6 *The chain* (s_n) *is* pseudo-renewal *if* $p(s|s') = p(s|s'')$ *for any three distinct states* s, s', s''.

That is, a chain is pseudo-renewal if the probability of a change to a given state is independent of the initial state, provided that they are distinct. Constant chains are pseudo-renewal chains, and so are i.i.d. processes. Pseudo-renewal chains are linear (although not necessarily convex) combinations thereof. When the chain belongs to one of these two special classes, it is intuitively clear that the dynamic problem becomes static and checking singleton states is sufficient. The following theorem shows that pseudo-renewal chains characterize the reduction.

Theorem 1 $\mathcal{C}_0 = \mathcal{C}_k$ *holds for all* k *if, and only if,* (s_n) *is pseudo-renewal.*

Indeed, as shown in Renault, Solan and Vieille (2013), if (s_n) is pseudo-renewal, then for each $\mu \in \mathcal{M}_0$, there exists a reporting policy inducing μ, such that the distribution of the entire sequence (a_n) is equal to that of the sequence (s_n); hence $\mathcal{M}_0 = \mathcal{M}_k$ and $\mathcal{C}_0 = \mathcal{C}_k$ for all k.[20] Relative to their construction, our Theorem 1 provides the converse result: if (s_n) is not pseudo-renewal, then $\mathcal{C}_0 \supsetneq \mathcal{C}_1$. This implies the existence of a preference such that truth-telling is 1-limit, but not 0-limit, optimal.

---

[20] To be clear, their result is even stronger than $\mathcal{M}_0 = \mathcal{M}_k$ for all k because it covers arbitrary tests, regardless of whether they are based on strings.

How common are pseudo-renewal chains? In what follows, genericity statements are relative to the class of Markov chains of a given size $|S|$, identified (via the transition matrix) with a subset of $\mathbf{R}^{|S|\times(|S|-1)}$. The following is immediate. Therefore, its proof is omitted.

Lemma 3 *Any Markov chain with* $|S| = 2$ *is pseudo-renewal. The set of pseudo-renewal chains is non-generic for* $|S| \geq 3$.

It is tempting to interpret Lemma 3 as saying that testing singletons is never robustly optimal. However, we stress that the conclusion does not only rely on the Markov chain, but also on the information structure. Below we consider a *prophetic* agent, who knows the entire realized sequence (s_n) before being asked to make reports. In that case, it turns out that independent of the Markov chain, the designer cannot do better than checking state frequencies. One might surmise that knowing states in advance would increase the number of incentive constraints and call for more complex statistical tests, but our next result establishes that a prophetic agent can foil all but the simplest tests.

Lemma 4 *Suppose that the agent is prophetic and the Markov chain has full support.*[21] *Then, truth-telling can be achieved if, and only if, it can be achieved when checking singleton states.*

That is, foresight suffices to nullify any potential benefit from exploiting the correlation of the states.[22] Jackson and Sonnenschein (2007)'s test is not only natural but, in this context (with an infinite number of linked decisions), without loss of generality for implementation, even when states have a Markovian correlation structure.

3.4.2 Testing Pairs and More

Given Lemma 3, we turn our attention to $|S| \geq 3$. Our main result, Theorem 2 below, is negative. The question of whether $\mathcal{C}_k$ is constant for $k \geq 1$ depends on fine details of the transition matrix, and finite memory is not enough very generally, whether or not one is interested in detectability or profitability. Note that even though we state this negative result for the case of $|S| = 3$, it easily generalizes.

---

[21] As the proof establishes, it suffices that the agent knows the next $|S| + 1$ states ahead of time.

[22] The role of correlation in our main results might bring to mind those of Crémer and McLean (1988), which suggest that correlated signals/states call for payments that depend on the entire vector of reports. However, the role of time is different, as elucidated by Lemma 4. Another difference is that Crémer and McLean (1988) rely on multiple agents telling the truth about one another, while we have a single agent.

Theorem 2 *Let* $|S| = 3$*. For every* $k \geq 1$*,* $\mathcal{C}_0 \subsetneq \mathcal{C}_1 \subsetneq \cdots \subsetneq \mathcal{C}_k$ *for an open set of transition matrices.*

The proof, in appendix, establishes a stronger version of this result (see Appendix C) suggesting that the strict inclusion is the norm rather than the exception.

Theorem 2 implies that in some cases, checking triples (resp., quadruples, etc.) allows truth-telling when pairs (resp., triples, etc.) would not. However, we do not know if, for any given Markov chain, some string length suffices.

Given that the set $\mathcal{M}_k$ is described by finitely many inequalities involving polynomials in the transition probabilities $p(s' \mid s)$, it is tempting to conjecture that, generically, the extreme points of the polytopes $\mathcal{M}_k$ and $\mathcal{M}_{k+1}$ would differ. Surprisingly, this is not so. In a separate paper (Hörner, Mu and Vieille, 2025), we construct an open set of transition matrices, with $\mathcal{M}_1 = \mathcal{M}_k$ for all $k > 1$ (and therefore, $\mathcal{C}_k = \mathcal{C}_1$ for each $k \geq 1$). In fact, for transition matrices that are arbitrarily small perturbations of the (constant) matrix for which (s_n) are i.i.d and uniform, one may *robustly* have $\mathcal{M}_k = \mathcal{M}_1$ for all k, or $\mathcal{M}_k \neq \mathcal{M}_1$ for some k.

Possibly even more surprisingly, we prove the existence of an open set of transition matrices for which $\mathcal{M}_1 \neq \mathcal{M}_k$ yet $\mathcal{C}_1 = \mathcal{C}_k$ for each $k > 1$. This extends the observation from our leading example that detectability and profitability are distinct problems.

There are also cases in which $\mathcal{M}_1 = \mathcal{M}_2 \supsetneq \mathcal{M}_3$. Hence, just because testing some longer string does not help does not imply that no longer string helps. One may also wonder about the focus on strings of *consecutive* reports. However, there are Markov chains for which testing the frequencies of (a^{-1}, a), (a^{-2}, a), and (a^{-2}, a^{-1}) (that is, pairs of consecutive and non-consecutive reports) leads to a strictly larger set of undetectable distributions than when testing triples (a^{-2}, a^{-1}, a).[23]

It is worth stressing that Theorem 2 does not depend on exact truth-telling, as opposed to some weaker notion. To see this, consider a transition function for which $\mathcal{C}_2$ is a strict subset of $\mathcal{C}_1$. Then, there exists a reward function r such that

$$\max_{\mu \in \mathcal{M}_2} \mathbf{E}_\mu[r\,(s, a)] = \mathbf{E}_{\mu^{tt}}[r\,(s, a)] < \max_{\mu \in \mathcal{M}_1} \mathbf{E}_\mu[r\,(s, a)]. \tag{6}$$

---

[23] This is not to say that checking pairs of consecutive reports is necessarily the best test based on pairs. For instance, depending on the transition matrix, it might be better or worse to keep track of pairs with one round in between, as numerical examples show. That is, making the dependence on the Markov chain explicit, the set $\mathcal{M}_1(P)$ is sometimes but not always included in the set $\mathcal{M}_1(P^2)$. Hence, against a designer with bounded memory, more persistence does not necessarily harm the informed agent. This contrasts with the result of Pęski and Toikka (2017) in the related context of dynamic Bayesian games.

By the equality in (6), truth-telling can be obtained when triples are checked. However, the inequality implies that when only pairs are checked, some perfectly undetectable deviation improves upon any reporting policy that is approximately truthful. Thus truth-telling fails to be even approximately incentive-compatible when the designer only checks pairs.[24]

4 Implications and Applications

4.1 Money vs. Time

In many economic problems, the designer has no direct control over the agent's reports, but she can influence them via transfers. Here, let us assume that the designer can choose a sequence (t_n) of transfers, measurable and bounded functions of the sequence of reports. For reasons that will become clear, we slightly change the environment and assume that the horizon is doubly infinite. States in rounds $n \leq 0$ are "publicly observed." However, s_n is privately observed by the agent in round $n \geq 1$, as before. The environment is otherwise unchanged, with the agent maximizing the expectation of

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} (r(s_n, a_n) + t_n). \tag{7}$$

We stress that the designer can no longer impose a direct constraint on the sequence of reports.

Given $k \in \mathbf{N}$, a transfer function is a *transfer with memory k*, if all t_n are functions of the last report a_n, in addition to the k most recent ones, $a_{n-k}, \ldots, a_{n-1}$. It is *stationary* if this function is independent of calendar time. The requirement that the transfer function has memory k is *a priori* restrictive: for instance, if the transfer function has memory 0, only the current report matters, and all statistical evidence is discarded. We have the following equivalence:

Proposition 3 *Truth-telling is k-limit optimal if, and only if, truth-telling is optimal for some stationary transfer with memory k.*

The reader might wonder about the importance of stationarity. Under the discounted payoff criterion, if ϕ is implementable via a direct mechanism with some memory-k transfer, then this

---

[24] In this discussion, we are holding fixed the memory k and considering a very fine level of approximation. An alternative question would be to fix the level of approximation, and ask whether there exists k such that any utility function truthfully implementable with some finite memory can be approximately implemented with memory k. We conjecture the answer to be yes.

can be achieved with a stationary memory-k transfer.[25] Whether it is also without loss in the undiscounted case is an open question.

4.2 Implementation

Having the agent tell the truth is only useful to the extent that it has payoff-relevant consequences. In this section we study such consequences within the framework of implementation. As primitives we are given an outcome set Y and state-dependent preferences captured by some utility function $u : S \times Y \to \mathbf{R}$. The question in the standard static setup is this: for which maps $\phi : S \to Y$ can one find transfers $t : A \to \mathbf{R}$ such that truth-telling is optimal, given the payoff $u(s, \phi(a)) + t(a)$? This connects to our problem once transfers are introduced as in Section 4.1, and once we define $r(s, a) = u(s, \phi(a))$ given the map ϕ.

The case in which transfers do not rely on past reports (mechanisms with memory 0) has been extensively studied in prior work. For such mechanisms, the agent's problem becomes separable over time. Hence, it reduces to the static problem, for which implementability has been characterized by Rochet (1987). Given $\phi : S \to Y$, u is *cyclically monotone* if, for every finite sequence $s_0, s_1, \ldots, s_m = s_0$ of states, it holds that

$$\sum_{i=0}^{m-1} (u(s_i, \phi(s_{i+1})) - u(s_i, \phi(s_i))) \leq 0.$$

Rochet proves that ϕ is implementable if, and only if, u is cyclically monotone.

Taking the dual of Rochet's characterization immediately yields that ϕ is implementable via a stationary direct mechanism with memory 0 if, and only if,

$$\mathbf{E}_{\mu}[u(s, \phi(a))] \leq \mathbf{E}_{\mu^{tt}}[u(s, \phi(a))] \text{ for all } \mu \in \mathcal{M}_0,$$

which is precisely the special case $k = 0$ of Proposition 2.[26]

More generally, we have the following *dual characterization of implementability* as an immediate corollary of Proposition 2 and Proposition 3.

Corollary 1 *Let $\phi : S \to Y$. The map ϕ is implementable via transfers of memory k if, and*

---

[25] A proof is available upon request.

[26] In the static setup, this dual formulation appears well known but difficult to be attributed to a specific author.

only if,

$$\mathbf{E}_{\mu}[u(s,\phi(a))] \leq \mathbf{E}_{\mu^{tt}}[u(s,\phi(a))] \text{ for all } \mu \in \mathcal{M}_k.$$

There is also a primal version of Corollary 1, which generalizes cyclical monotonicity to the dynamic setting. See Appendix B.3 for detailed discussion of why the primal becomes less tractable with dynamics.

4.3 Recursive Representations

Many results in dynamic economics rely on a recursive representation of the payoff set. Promised utility is used as a state variable; then, a simple Bellman-type equation characterizes the equilibrium payoff correspondence, as a function of this variable (and possibly other state variables). This representation holds quite generally in the context of agency models (see Spear and Srivastava, 1987, and Thomas and Worrall, 1990), as well as in repeated games with restriction to public strategies (see Shapley, 1953, and Abreu, Pearce and Stacchetti, 1990).

An important implication of our results is that such a representation does not readily extend to the case of persistent private information. More precisely, with only one agent (as assumed thus far), the payoff set is still easy to compute: the agent can be expected to report truthfully when the designer's objective is to maximize his payoff; and to minimize his payoff, the designer can ignore his report. However, given Theorem 2, we do not have a simple characterization of implementable choice functions. As a result, there is no *universal* recursive representation of the equilibrium payoff set when there is more than one agent and values are interdependent. Note that such a representation exists in the case of independent private values; see Hörner, Takahashi and Vieille (2015).[27]

The next example, building on the example in Section 2, makes this point. Consider a three-player infinitely repeated game. Player 1 is informed of the state, while players 2 and 3 are not. The state follows the Markov chain of our leading example. Player 1 has three actions $a^{(1)} \in \{1,2,3\}$, while player 2 has a binary action $a^{(2)} \in \{0,M\}$, taken immediately after player 1's (here, $M > 0$ is a large parameter to be determined). Player 3's action is also binary,

---

[27] Interdependence arises naturally in repeated games or repeated agency problems in which actions are hidden (*e.g.*, in repeated insurance problems as in Rubinstein and Yaari, 1983). Our negative result may be bad news for the tractability of dynamic models without transfers, but it is good news for the power of long-run relationships to support cooperation. This is because it implies that continuation play is not simply an imperfect substitute for transfers: in a Markovian environment, it can improve on transfers, to the extent that equilibrium payoffs can be achieved that would not be attainable in a static setup with contractual transfers.

$a^{(3)} \in \{0, M\}$, taken immediately after player 2's. Let A denote the set of action profiles. Payoffs are

$$r^1(s_j, a) = a^{(2)} + \begin{cases} 0 & \text{if } a^{(1)} = j, \\ 1 & \text{if } a^{(1)} = j+1, \\ -L & \text{if } a^{(1)} = j-1, \end{cases}$$

where indices are modulo 3, and

$$r^2(s_j, a) = -a^{(2)} - a^{(3)} + \begin{cases} 1 & \text{if } a^{(1)} = j, \\ -1 & \text{otherwise.} \end{cases}$$

Finally, $r^3(s_j, a) \equiv 0$. Actions are observed, but payoffs are not. We note that player 2's action is *de facto* a transfer, which is irrelevant for efficiency (*i.e.*, maximizing $r^1 + r^2 + r^3$). Also, we can use player 3 (a dummy) to incentivize player 2. The challenge is to use the action of player 2 to induce player 1 to tell the truth, which is necessary for efficiency.

It can be checked, using notations from Section 2, that $x_k := \max\{x : \mu_x \in \mathcal{M}_k\}$ is strictly positive and converges to zero as $k \to \infty$ (see Appendix **??**). From the specification of payoffs, it also holds that for each k, there is some large parameter L such that the distribution $\mu \in \mathcal{M}_k$ that maximizes $\mu \cdot r^1$ is precisely μ_{x_k}.[28] Consequently, for large parameters L and M, the above game has an equilibrium that results in a joint payoff arbitrarily close to 1.[29]

On the other hand, fix $k \geq 1$. Given a function $t : A^k \to \mathbf{R}^3$, define the k-repeated game $\Gamma^k(t)$ as follows. Player 1 privately observes online $s_1, \ldots, s_k$, drawn according to the invariant distribution λ in the initial round and following the Markov chain afterwards. All players perfectly monitor actions. Player i's expected payoff given strategy profile σ is

$$v^i(\sigma) = \mathbf{E}_{\lambda,\sigma}\left[\frac{1}{k}\sum_{n=1}^{k} r^i(s_n, a_n) + t^i(a_1, \ldots, a_k)\right],$$

where the transfers t^i are subject to the balanced-budget constraint $\sum_i t^i(a_1, \ldots, a_k) \leq 0$ for all

---

[28] For sufficiently large L, player 1 does not want to put any weight on $a^{(1)} = j - 1$ when the state is s_j. And given that, he maximizes the weights on $a^{(1)} = j + 1$.

[29] Indeed, we first choose k such that x_k is close to zero, so the joint payoff at μ_{x_k} is close to one. Next pick L such that $\mu \cdot r^1$ is maximized at μ_{x_k} for $\mu \in \mathcal{M}_k$. By Proposition 3, there are stationary transfers with memory k such that the optimal reporting policy of player 1 implements μ_{x_k}. Finally pick M sufficiently large (and use the dummy player 3) to sustain these transfers as equilibrium strategies of player 2.

$(a_1, \ldots, a_k)$.

Let $E^{k,L,M}(t)$ denote the set of Nash equilibria of this finite-stage game with transfers. Then, define the maximal surplus as

$$S^{k,L,M} = \sup_{\substack{t: A^k \to \mathbf{R}^3 \\ \sigma \in E^{k,L,M}(t)}} \sum_i v^i(\sigma).$$

It follows from our main results (and a continuity argument found in the Appendix) that

$$\sup_{L,M} S^{k,L,M} < 1.$$

Intuitively, player 1 wants to deviate to a strategy inducing $\mu_{x_{k-1}}$, which is undetectable in the k-repeated game by Propositions 2 and 3. This deviation pushes the joint payoff away from efficiency.

To summarize, for each k, there are parameters L and M such that computing the set of equilibrium payoffs would require considering a reduced-form with more than k stages.[30] However, we suspect that L cannot be chosen uniformly in k; hence, for a given game, there may be a recursive representation in terms of a k-shot game with a suitably large k.

5 Concluding Comments

It would be of interest to assume that the agent does not directly observe the state. What happens if the agent's information itself is only a signal about the underlying state? A major difficulty, then, is that the sequence of observations is not itself a Markov chain.

Related environments, such as continuous-time Markov chains, remain entirely unexplored: what is the meaning of a round in continuous time? Considering the uniformized discrete-time Markov chain does not seem to be as helpful here as in related problems because the exact time elapsed since the last reported switch is a piece of information that the designer might want to use. What if the agent observes a diffusion process, such as a Brownian motion?

---

[30] Our claim is not that strategies with bounded memory are with loss in repeated games (a well-known fact), but that the equilibrium payoff set in general Bayesian games cannot be solved for by analyzing a related one-shot game with transfers, as is done in Abreu, Pearce and Stacchetti (1990). We also do not claim that no one-shot game exists with the same equilibrium payoff set. We have fixed the payoff function and the Markov chain to be the same as in the infinite-horizon game; otherwise, the question becomes meaningless.

From an economic perspective, it is desirable to go beyond the rather extreme objective to achieve truth-telling, and achieve it for all preferences. If truth-telling is not achievable, we may still ask how close one can get to it, and how memory affects this distance. At an abstract level, using the revelation principle, the agent would truthfully report a belief (rather than a state), coinciding with the Bayesian posterior that the designer forms upon hearing this report. Given a test, one then computes the distance between the resulting empirical distribution over posterior beliefs and the true distribution of the state (this calls for preferences r over $S \times \Delta(S)$). Conversely, insisting on truth-telling, one might be interested in identifying the class of preferences for which a given test suffices, or in finding conditions under which implementability is simple to verify.[31] In the dynamic setting, truth-telling depends both on the preferences and the Markov chain. Our dual approach is promising in this regard: since the bi-dual is just the primal, the dual inequality constraints for $\mathcal{M}_k$ (discussed in the appendix) characterize the extremal reward functions that are k-implementable. Such an analysis is an important step toward bridging the gap between results on implementation, such as ours, and the literature on dynamic mechanism design.

References

[1] Abreu, D., D. Pearce and P. Milgrom (1991). “Information and Timing in Repeated Partnerships,” *Econometrica*, **59**, 1713–1733.

[2] Abreu, D., D. Pearce and E. Stacchetti (1990). “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica*, **58**, 1041–1063.

[3] Afriat, S.N. (1966). “The Construction of Utility Functions from Expenditure Data,” *International Economic Review*, **8**, 67–77.

[4] Altman, E. (1999). *Constrained Markov Decision Processes*, Chapman& Hall/CRC.

[5] Athey, S. and K. Bagwell (2008). “Collusion with Persistent Shocks,” *Econometrica*, **76**, 493–540.

[6] Athey, S. and I. Segal (2013). “An Efficient Dynamic Mechanism,” *Econometrica*, **81**, 2463–2485.

---

[31] This would be in the spirit of the Spence-Mirrlees condition, under which cyclical monotonicity reduces to standard monotonicity. See Rochet (1987).

[7] Aumann, R.J. and M. Maschler (1995). *Repeated Games with Incomplete Information*, MIT Press.

[8] Baron, D. and D. Besanko (1984). "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, **1**, 447–470.

[9] Barron, D. (2017). "Attaining Efficiency with Imperfect Public Monitoring and Markov Adverse Selection," *Theoretical Economics*, **12**, 957–978.

[10] Battaglini, M. (2005). "Long-Term Contracting With Markovian Consumers," *American Economic Review*, **95**, 637–658.

[11] Battaglini, M. and R. Lamba (2019). "Optimal Dynamic Contracting: the First-Order Approach and Beyond," *Theoretical Economics*, **14**, 1435–1482.

[12] Bhaskar, V., G.J. Mailath and S. Morris (2013). "A Foundation for Markov Equilibria in Sequential Games with Finite Social Memory," *Review of Economic Studies*, **80**, 925–948.

[13] Billingsley, P. (1961). "Statistical Methods in Markov Chains," *Annals of Mathematical Statistics*, **32**, 12–40.

[14] Blackwell, D. and L. Koopmans (1957). "On the Identifiability Problem for Functions of Finite Markov Chains," *Annals of Mathematical Statistics*, **28**, 1011–1015.

[15] Cohn, Z. (2010). "A Note on Linked Bargaining," *Journal of Mathematical Economics*, **46**, 238–247.

[16] Connault, B. (2016). "Hidden Rust Models," working paper, University of Pennsylvania.

[17] Crémer, J. and R.P. McLean (1988). "Full Extraction of the Surplus in Bayesian and Dominant policy Auctions," *Econometrica*, **56**, 1247–1257.

[18] Escobar, P. and J. Toikka (2013). "Efficiency in Games with Markovian Private Information," *Econometrica*, **81**, 1887–1934.

[19] Fang, H. and P. Norman (2006). "To Bundle or Not To Bundle," *RAND Journal of Economics*, **37**, 946–963.

[20] Feller, W. (1960). “Non-Markovian Processes with the Semi-group Property,” *Annals of Mathematical Statistics*, **30**, 1252–1253.

[21] Frankel, A. (2016). “Discounted Quotas,” *Journal of Economic Theory*, **166**, 396–444.

[22] Fudenberg, D. and D. Levine (1994). “Efficiency and Observability in Games with Long-Run and Short-Run Players,” *Journal of Economic Theory*, **62**, 103–135.

[23] Fudenberg, D., D. Levine and E. Maskin (1994). “The Folk Theorem with Imperfect Public Information,” *Econometrica*, **62**, 997–1040.

[24] Gale, D. (1960). *The Theory of Linear Economic Models*, University Of Chicago Press.

[25] Hörner, J., X. Mu and N. Vieille (2025). “Testing Truth-Telling,” working paper, Columbia, HEC Paris and TSE.

[26] Hörner, J., S. Takahashi and N. Vieille (2015). “Truthful Equilibria in Dynamic Bayesian Games,” *Econometrica*, **83**, 1795–1848.

[27] Jackson, M.O. and H.F. Sonnenschein (2007). “Overcoming Incentive Constraints by Linking Decisions,” *Econometrica*, **75**, 241–258.

[28] Judd, K.L., S. Yeltekin and J. Conklin (2003). “Computing Supergame Equilibria,” *Econometrica*, **71**, 1239–1254.

[29] Laraki, R. (2004). “On the Regularity of the Convexification Operator on a Compact Set,” *Journal of Convex Analysis*, **11**, 209–234

[30] Lee, J. and H. Sabourian (2011). “Efficient Repeated Implementation,” *Econometrica*, **79**, 1967–1994.

[31] Lee, J. and H. Sabourian (2015). “Complexity and Repeated Implementation,” *Journal of Economic Theory*, **158**, 259–292.

[32] Mailath, G.J., and L. Samuelson (2006). *Repeated Games and Reputations*, Oxford University Press.

[33] Matsushima, H., K. Miyazaki and N. Yagi (2010). “Role of Linking Mechanisms in Multitask Agency with Hidden Information,” *Journal of Economic Theory*, **145**, 2241–2259.

[34] Mezzetti, C. and L. Renou (2017). "Repeated Nash Implementation," *Theoretical Economics*, **12**, 249–285.

[35] Pęski, M. and J. Toikka (2017). "Value of Persistent Information," *Econometrica*, **85**, 1921–1948.

[36] Radner, R. (1981). "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship," *Econometrica*, **49**, 1127–1148.

[37] Rahman, D. (2010). "Dynamic Implementation," working paper, University of Minnesota.

[38] Renault, J., E. Solan and N. Vieille (2013). "Dynamic Sender-Receiver Games," *Journal of Economic Theory*, **148**, 502–534.

[39] Renou, L. and T. Tomala (2015). "Approximate Implementation In Markovian Environments," *Journal of Economic Theory*, **159**, 401–442.

[40] Rochet, J.-C. (1987). "A Necessary and Sufficient Condition for Rationalizability in a Quasilinear Context," *Journal of Mathematical Economics*, **16**, 191–200.

[41] Rockafellar, T. (1966). "Characterization of the Subdifferentials of Convex Functions," *Pacific Journal of Mathematics*, **17**, 497–510.

[42] Rosenblatt, M. (1960). "An Aggregation Problem for Markov Chains," in *Information and Decision Processes*, R.E. Machol, ed., McGraw-Hill, New York, 87–92.

[43] Rothschild, M. (1982). "A Note on Partially Observed Markov Systems," *Economics Letters*, **10**, 207–209.

[44] Rubinstein, A. and M. Yaari (1983). "Repeated Insurance Contracts and Moral Hazard," *Journal of Economic Theory*, **30**, 74–97.

[45] Samuelson, L. and E. Stacchetti (2017). "Even Up: Maintaining Relationships," *Journal of Economic Theory*, **169**, 170–217.

[46] Shapley, L.S. (1953). "Stochastic Games," *Proceedings of the National Academy of Sciences of the U.S.A.*, **39**, 1095–1100.

[47] Smith, G. and J. Dawson (1979). “The Determination of Necessary and Sufficient Conditions for the Existence of a Solution to the $3 \times 3 \times 3$ Multi-index Problem,” *Aplikace Matematiky*, **24**, 201–208.

[48] Spear, S. and S. Srivastava (1987). “On Repeated Moral Hazard with Discounting,” *Review of Economics Studies*, **54**, 599–617.

[49] Thomas, J. and T. Worrall (1990). “Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem,” *Journal of Economic Theory*, **51**, 367–390.

[50] Townsend, R. (1982). “Optimal Multi-Period Contracts and the Gain from Enduring Relationships under Private Information,” *Journal of Political Economy*, **90**, 1166–1186.

The structure of the Appendix mostly replicates that of the paper.

A Undetectability and Unprofitability

A.1 Proof of Proposition 1

Let $\mu \in \mathcal{M}_k$ be given. Pick a stationary policy $\sigma_S : A^k \times S \to \Delta(A)$, and an invariant distribution $\nu \in \Delta(A^k \times S \times A)$ for σ_S, with marginals λ and μ (see Definition 4). Because the Markov chain induced by σ_S over the set $A^k \times S \times A$ of (reports, state) pairs need not be unichain, empirical frequencies $(f^k_n)_n$ need not converge to ν w.p.1. To circumvent this, we construct a policy σ that plays σ_S over blocks of increasing length, separated by shuffling blocks. Formally, each phase $m \geq 1$ consists of a shuffling phase of random duration, and of a σ_S-phase. A starting configuration $(a^k, s)_{(m)} \in A^k \times S$ is first drawn according to ν. In the shuffling phase, σ picks reports uniformly at random. This shuffling phase ends as soon as the k most recent reports and current state coincide with $(a^k, s)_{(m)}$. The σ_S-phase then lasts for m periods, during which σ selects reports as σ_S. By construction, the empirical distributions of (state,report)-pairs in the successive σ_S-phases are independent, with expectation ν. Since the duration of the σ_S-phases increases to $+\infty$, and since the durations of the shuffling phases are independent with a finite expectation, the asymptotic density of the set of stages which belong to some shuffling phase is zero, w.p. 1. This implies that $\sigma \in \Sigma_k$ and that $(\mu^N_{q,\sigma})$ converges to μ, as desired.

We next prove a slightly stronger version of the second statement, that applies to a discounted version of the model. Given a reporting policy σ and $\delta < 1$, we denote by $\nu^\delta_\sigma \in \Delta(A^k \times S \times A)$ the expected discounted frequencies of the various reports-state sequences and by $\mu^\delta_\sigma \in \Delta(S \times A)$ and $\lambda^\delta_\sigma \in \Delta(A^{k+1})$ its marginals.[32] If σ is a stationary policy, and $\nu_0 \in \Delta(A^k \times S)$ is some 'extended' initial distribution over previous reports and current state,[33] we also write $\nu^\delta_{\nu_0,\sigma}$, $\mu^\delta_{\nu_0,\sigma}$ and $\lambda^\delta_{\nu_0,\sigma}$.

Lemma 5 *Let $\sigma \in \Sigma_k$. Any limit point of (μ^δ_σ) as $\delta \to 1$ belongs to $\mathcal{M}_k$.*

If $\lim_{N\to\infty} \mu^N_{q,\sigma}$ exists, then it is equal to $\lim_{\delta\to 1} \mu^\delta_\sigma$ and the second statement in Proposition 1 follows.

---

[32] Omitting the initial distribution q.

[33] A stationary policy depends on previous reports, which do not exist in the first stages. Hence the need for such extended initial distributions.

Proof. We introduce the set $\mathcal{M}$ of distributions $\mu \in \Delta(S \times A)$ that are limit points of μ_σ^δ for some $\sigma \in \Sigma_k$, as $\delta \to 1$. We will prove that

$$\sup_{\mu \in \mathcal{M}} \alpha \cdot \mu \leq \sup_{\mu \in \mathcal{M}_k} \alpha \cdot \mu,$$

for each $\alpha \in \mathbf{R}^{S \times A}$. Since $\mathcal{M}_k$ is a compact polytope, this will imply that $\mathcal{M} \subset \mathcal{M}_k$, as desired.

Fix $\alpha \in \mathbf{R}^{S \times A}$ and $\bar{\mu} \in \mathcal{M}$. Let $\bar{\sigma} \in \Sigma_k$ and a sequence (δ_n) converging to one, such that $\bar{\mu} = \lim_{n \to \infty} \mu_{\bar{\sigma}}^{\delta_n}$. For given $\delta < 1$ and $\eta > 0$, consider the following constrained optimization problem: find a reporting policy σ and $\nu_0 \in \Delta(A^k \times S)$ with marginal q, that maximizes $\alpha \cdot \mu_{\nu_0, \sigma}^\delta$, among the pairs (ν_0, σ) such that $\|\lambda_{\nu_0, \sigma}^\delta - \lambda\| \leq \eta$. By Chapter 3 in Altman (1999), there is an optimal pair $(\nu_0(\delta, \eta), \sigma(\delta, \eta))$ for which $\sigma(\delta, \eta)$ is stationary.

Since $\bar{\sigma} \in \Sigma_k$, one has $\lim_n f_n^k = \lambda$ and therefore $\lim_{\delta \to 1} f_\delta^k = \lambda$, $\mathbf{P}_{\bar{\sigma}}$-a.s. Taking expectations, this implies that, for fixed $\eta > 0$, the pair $(\bar{\nu}_0, \bar{\sigma})$ is feasible in the constrained optimization problem provided δ is close enough to one, for any distribution $\bar{\nu}_0$ with marginal q. In particular, there is $\delta(\eta) < 1$ such that

$$\delta > \delta(\eta) \Rightarrow \alpha \cdot \mu_{\bar{\sigma}}^\delta \leq \alpha \cdot \mu_{\nu_0(\delta, \eta), \sigma(\delta, \eta)}^\delta. \tag{8}$$

Consider now any sequences (δ_m) and (η_m) such that $\lim_m \delta_m = 1$, $\lim_m \eta_m = 0$ and $\delta_m > \delta(\eta_m)$ for each m. Assume w.l.o.g. that the sequences $\sigma_m := \sigma(\delta_m, \eta_m)$ and $\nu_m := \nu_{\nu_0(\delta_m, \eta_m), \sigma(\delta_m, \eta_m)}^{\delta_m} \in \Delta(A^k \times S \times A)$ converge, with limits σ and ν. By Lemma 6 below, ν is invariant for σ.

Since the pair $(\nu_0(\delta_m, \eta_m), \sigma(\delta_m, \eta_m))$ is feasible in the constrained optimization problem defined by (δ_m, η_m), the marginal $\lambda_m \in \Delta(A^{k+1})$ of ν_m is such that $\|\lambda_m - \lambda\| \leq \eta_m$. Thus, the marginal of ν on A^{k+1} coincides with λ, hence the limit μ of the marginal $\mu_m \in \Delta(S \times A)$ of ν_m belongs to $\mathcal{M}_k$. Applying (8) with (δ_m, η_m) and letting $m \to \infty$, one obtains $\alpha \cdot \bar{\mu} \leq \alpha \cdot \mu$, as desired. ■

The above proof uses the observation below, whose proof is omitted.

Lemma 6 *Let $\delta_m < 1$ and $\sigma_m : A^k \times S \to \Delta(A)$ be given for $m \geq 1$, s.t.* $\lim_m \delta_m = 1$ *and* $\lim_m \sigma_m = \sigma$. *For $m \geq 1$, let an initial distribution $\nu_{0,m} \in \Delta(A^k \times S)$ be given, and denote by $\nu_m \in \Delta(A^k \times S \times A)$ the expected δ_m-discounted frequency of (state, reports)-pairs, given the initial distribution $\nu_{0,m}$ and the reporting policy σ_m. Then any limit point of $(\nu_m)_{m \geq 1}$ is invariant for the Markov chain over $A^k \times S \times A$ induced by σ.*

A.2 Proof of Lemma 1

Generic elements of $A^k \times S \times A$ are denoted $(a^{-k}, \dots, a^{-1}, s, a)$, with the interpretation that s and a are the current state and report, while a^{-i} is the report that was submitted i rounds earlier. Given $\omega = (a^{-k}, \dots, a^{-1}) \in A^k, s \in S$ and $a \in A$, the transition function of the chain induced by σ over $A^k \times S \times A$ is

$$\pi(\omega', s' \mid \omega, s) := \sigma(a \mid \omega, s) p(s' \mid s),$$

where $\omega' = (a^{-(k-1)}, \dots, a^{-1}, a)$ is obtained by shifting the entries of ω, and appending a.

Therefore, a distribution $\nu \in \Delta(A^k \times S \times A)$ is invariant for (the Markov chain induced by) σ if for each $\omega = (a^{-k}, \dots, a^{-1}) \in A^k$ and $s \in S$,

$$\nu(\omega, s) = \sum_{a^{-(k+1)} \in A, s^{-1} \in S} \nu(\omega^{-1}, s^{-1}) \sigma(a^{-1} \mid \omega^{-1}, s^{-1}) p(s \mid s^{-1}),$$

where ω^{-1} stands for $(a^{-(k+1)}, \dots, a^{-2})$, and $\nu(\omega, s) := \sum_{a \in A} \nu(\omega, s, a)$ is the marginal of ν over $A^k \times S$. Equivalently, a distribution $\nu \in \Delta(A^k \times S \times A)$ is invariant for *some* reporting policy σ if, and only if,

$$\sum_a \nu(\omega, s, a) = \sum_{a^{-(k+1)} \in A, s^{-1} \in S} \nu(\omega^{-1}, s^{-1}, a^{-1}) p(s \mid s^{-1}), \tag{9}$$

for all ω and s. Indeed, ν is then invariant for the policy σ defined by

$$\sigma(a \mid a^{-k}, \dots, a^{-1}, s) = \frac{\nu(a^{-k}, \dots, a^{-1}, s, a)}{\sum_{a' \in A} \nu(a^{-k}, \dots, a^{-1}, s, a')},$$

if the denominator is non-zero (and arbitrarily otherwise).

We denote by $\mathcal{N}_k$ the compact convex set of such distributions ν. From Definition 4, $\mathcal{M}_k$ is the set of joint distributions $\mu \in \Delta(S \times A)$, for which there exists $\nu \in \mathcal{N}_k$, such that (i) the marginal of ν over $S \times A$ is μ and (ii) the marginal of ν over $A^k \times A$ is equal to the invariant distribution of $(s_{n-k}, \dots, s_n)$:

$$\sum_{s \in S} \nu(a^{-k}, \dots, a^{-1}, s, a) = \lambda(a^{-k}) p(a^{-(k-1)} \mid a^{-k}) \cdots p(a \mid a^{-1}),$$

for all $a^{-k}, \dots, a$. Since (9) is the same as (3) in the statement of Lemma 1, this completes the proof.

A.3 Proof of Lemma 2

For this proof, we will write $\mathcal{M}_k(P)$ to highlight the dependence on the transition matrix P. First consider the statement about dimension. Obviously $\mathcal{M}_k(P) \subset \mathcal{M}_0(P)$, which has dimension $(|S|-1)^2$ due to the marginal constraints. To show this maximal dimension is achieved, we recall the linear system (1)–(4) defining $\mathcal{M}_k$. By Lemma 9 in the Online Appendix, for any $\mu \in \mathcal{M}_0(P)$ there is a (not necessarily positive) solution ν, which is zero whenever λ is.

Let $\mu^{ind}(s,a) = \lambda(s) \cdot \lambda(a)$ be the joint distribution achieved by reporting an identical Markov chain independent of the states. This policy induces a $\nu^{ind}(a^{-k}, \dots, a^{-1}, s, a)$ that is positive and strictly positive whenever $\lambda(a^{-k}, \dots, a^{-1}, a) > 0$. Thus if (ν, μ) solves (1)–(4), then for sufficiently small positive ε, $((1-\varepsilon)\nu^{ind} + \varepsilon\nu, (1-\varepsilon)\mu^{ind} + \varepsilon\mu)$ also satisfies the linear system. Since $(1-\varepsilon)\nu^{ind} + \varepsilon\nu \geq 0$, this implies $(1-\varepsilon)\mu^{ind} + \varepsilon\mu \in \mathcal{M}_k(P)$. Hence $\mathcal{M}_k(P)$ has the same dimension as $\mathcal{M}_0(P)$.

Next we prove continuity. U.h.c. is immediate because the constraints (1)–(4) vary continuously with p. To show l.h.c., we characterize $\mathcal{M}_k(P)$ by a dual linear program. Specifically, we can write (1)–(4) abstractly as

$$C(P) \cdot \nu = D(p, \mu) \;\;\&\;\; \nu \geq 0.$$

Here $C(P)$ is a $c_1 \times c_2$ matrix whose entries are linear combinations of entries in p; $D(p,\mu)$ is a $c_1 \times 1$ column vector whose entries are either $\mu(s,a)$ or $\lambda(a^{-k})p(a^{-(k-1)}|a^{-k}) \cdots p(a|a^{-1})$ or 0. To get rid of redundant constraints, let us only include those $\mu(s,a)$ where s, a belong to the first $|S|-1$ states. In other words, we only consider $(|S|-1)^2$ equations of type (1). Let $\tilde{\mu} \in \mathbf{R}^{(|S|-1)^2}$ be the projection of μ onto the first $|S|-1$ states and reports. Similarly define $\tilde{\mathcal{M}}_k(P)$. Henceforth we write $D(p, \tilde{\mu})$ in place of $D(p, \mu)$.

By Farkas' lemma, there exists a non-negative solution ν to $C(P) \cdot \nu = D(p, \tilde{\mu})$ if, and only if, for any $y \in \mathbf{R}^{c_1}$,

$$y' \cdot C(P) \geq 0 \implies y' \cdot D(p, \tilde{\mu}) \geq 0. \tag{10}$$

Due to homogeneity, we can restrict attention to those y whose coordinates lie in $[-1, 1]$. Then the condition (10) simplifies to

$$y' \cdot D(p, \tilde{\mu}) \geq 0, \ \forall y' \in W(P), \tag{11}$$

where $W(P)$ is the bounded polytope $\{y \in \mathbf{R}^{c_1} : |y_j| \leq 1, y' \cdot C(P) \geq 0\}$, which is u.h.c. with respect to p. The above condition characterizes $\tilde{\mathcal{M}}_k(P)$ via a family of linear constraints on $\tilde{\mu}(s,a)$. In fact, only a finite collection of constraints matters, because $W(P)$ has finitely many vertices.

To prove l.h.c., we now fix p as well as a sequence $p_n \to p$. Fix $\tilde{\mu} \in int(\tilde{\mathcal{M}}_k(P))$, whose existence is guaranteed by dimension-counting. Then the finitely many relevant constraints in (11) cannot be binding at $\tilde{\mu}$. This implies the existence of $\varepsilon > 0$ such that

$$y' \cdot D(p, \tilde{\mu}) > \varepsilon, \ \forall y' \in W(P).$$

Take any $\tilde{\mu}_n \to \tilde{\mu}$ in $\mathbf{R}^{(|S|-1)^2}$. By the continuity of $D(p, \tilde{\mu})$ in both arguments, as well as the upper-hemicontinuity of $W(P)$, we deduce that for all large n,

$$y'_n \cdot D(p_n, \tilde{\mu}_n) > 0, \ \forall y'_n \in W(p_n).$$

By condition (11), this implies $\tilde{\mu}_n \in int(\tilde{\mathcal{M}}_k(p_n))$ for all large n. Thus the interior of $\tilde{\mathcal{M}}_k(P)$ satisfies l.h.c. The result follows because l.h.c. is preserved under closure.

A.4 Proof of Proposition 2

Denote $f(\sigma) := \liminf_N \mathbf{E}_{q,\sigma}\left(\frac{1}{N}\sum_{n=1}^{N} r(s_n, a_n)\right) = \liminf_N \mathbf{E}_{\mu_N^{q,\sigma}}[r(s,a)]$ the objective of the agent. We need to prove that

$$(f(\sigma) \leq f(\sigma_{tt}) \text{ for all } \sigma \in \Sigma_k) \iff \mathbf{E}_{\mu}[r(s,a)] \leq \mathbf{E}_{\mu_{tt}}[r(s,a)] \text{ for all } \mu \in \mathcal{M}_k.$$

Observe that $f(\sigma_{tt}) = \mathbf{E}_{\mu_{tt}}[r(s,a)]$ since (s_n) is irreducible.

For one direction, let $\mu \in \mathcal{M}_k$ be given. By Proposition 1, there is $\sigma \in \Sigma_k$ s.t. $\mu = \lim_N \mu_N^{q,\sigma}$, hence $f(\sigma) = \mathbf{E}_{\mu}[r(s,a)]$. The desired inequality $\mathbf{E}_{\mu}[r(s,a)] \leq \mathbf{E}_{\mu_{tt}}[r(s,a)]$ thus follows from $f(\sigma) \leq f(\sigma_{tt})$.

For the other direction, let $\sigma \in \Sigma_k$. By Lemma 5, all limit points μ of $(\mu_\delta^{q,\sigma})$ belong to $\mathcal{M}_k$ hence $\mathbf{E}_\mu[r(s,a)] \le \mathbf{E}_{\mu_{tt}}[r(s,a)]$. This implies $\liminf_{\delta\to 1} \mathbf{E}_{\mu_\delta^{q,\sigma}}[r(s,a)] \le \mathbf{E}_{\mu_{tt}}[r(s,a)]$ and, *a fortiori*,

$$\liminf_{N\to 1} \mathbf{E}_{\mu_N^{q,\sigma}}[r(s,a)] \le \mathbf{E}_{\mu_{tt}}[r(s,a)].$$

Thus, $f(\sigma) \le f(\sigma_{tt})$, as desired.

A.5 Proof of Theorem 1

One direction is proven in Renault, Solan, and Vieille (2013). They show that if (s_n) is pseudo-renewal, then for any $\mu \in \mathcal{M}_0$ there is a reporting policy such that (i) the law of the entire sequence of reports (a_n) is equal to the law of the sequence (s_n) of states, and (ii) the distribution of the state-report pair (s_n, a_n) is equal to μ for each n. It follows that $\mu \in \mathcal{M}_k$, so that $\mathcal{M}_0 = \mathcal{M}_k$ for each k.

For the other direction, assume that the Markov chain is not pseudo-renewal. For concreteness, we set $S = \{s_1, \ldots, s_{|S|}\}$ and write $\lambda_i = \lambda(s_i)$ as a shorthand. Up to a relabelling of the states, we may assume that $p(s_3 \mid s_1) \neq p(s_3 \mid s_2)$. Consider then the permutation matrix Π that permutes the two states s_1 and s_2. For $\varepsilon \le \min(\lambda_1, \lambda_2)$, the joint distribution

$$\mu_\varepsilon := diag(\lambda) + \varepsilon(\Pi - I) = \begin{pmatrix} \lambda_1 - \varepsilon & \varepsilon & & & \\ \varepsilon & \lambda_2 - \varepsilon & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_{|S|} \end{pmatrix}$$

belongs to $\mathcal{M}_0$.

We claim that for any $\varepsilon > 0$, μ_ε fails to belong to $\mathcal{M}_1$. Else, there would exist a reporting policy $\sigma : A \times S \to \Delta(A)$ and an invariant distribution $\nu \in \Delta(A \times S \times A)$, such that properties (i) and (ii) in Definition 4 hold.

By (ii), the frequency of rounds in which the agent reports successively (a_1, a_3) must be $\lambda(a_1, a_3) = \lambda_1 \cdot p(s_3 \mid s_1)$. By (i), and since $a = a_3$ if, and only if, $s = s_3$, this frequency is also equal to

$$\begin{aligned}\nu(a_1, a_3) &= \mathbf{P}_\sigma(s_1, a_1, a_3) + \mathbf{P}_\sigma(s_2, a_1, a_3) \\ &= \mathbf{P}_\sigma(s_1, a_1) \cdot p(s_3 \mid s_1) + \mathbf{P}_\sigma(s_2, a_1) \cdot p(s_3 \mid s_2) \\ &= (\lambda_1 - \varepsilon) \cdot p(s_3 \mid s_1) + \varepsilon \cdot p(s_3 \mid s_2).\end{aligned}$$

This is different from $\lambda_1 \cdot p(s_3 \mid s_1)$, leading to a contradiction.

Thus, the half-line starting at μ^{tt} and pointing in the direction $\Pi - I$ only intersects $\mathcal{M}_1$ at μ^{tt}, but does intersect $\mathcal{M}_0$ along a non-trivial segment. In other words, $\mathcal{C}_0 \supsetneq \mathcal{C}_1$, as we need to show.

A.6 Proof of Lemma 4

Fix $\vec{s} = (s^1, \dots, s^p)$ to be any list of $p \leq |S|$ distinct states. We identify the list $\vec{s}$ with a permutation π over S, which maps s^j to s^{j+1} for each j and leaves fixed all remaining states. Denote by Π the permutation matrix associated with π. Under the assumption that the agent is $|S|+1$-prophetic, we will construct below a reporting policy σ that has the following properties:

P1 the sequence (a_n) of reports has the same probability distribution as the sequence (s_n) of states,

P2 the joint distribution of the state-report pair (s_n, a_n) in a given round n is equal to μ^{tt} if n is a multiple of $p+1$, and is equal to $\mu^{tt} + \varepsilon(\Pi - I)$ otherwise (for some positive ε).

For truth-telling to be optimal, the expectation of $r(s,a)$ under Π cannot exceed the expectation under I. That is,

$$\sum_{j=1}^{p} r(s^j, s^{j+1}) \leq \sum_{j=1}^{p} r(s^j, s^j).$$

Since this is true for each list $\vec{s}$, we would be able to conclude that ϕ is 0-IC.

To construct such a reporting policy σ, we denote by $\vec{s}^j$ the list $(s^j, \dots, s^p, s^1, \dots, s^{j-1})$ obtained from $\vec{s}$, when starting from its j-th element. Pick $\varepsilon > 0$ small enough so that $\mathbf{P}(s, \vec{s}^j, s') \geq \varepsilon$ for all j and $s, s' \in S$. The policy σ operates in blocks of size $p+1$: At a period n that is a multiple of $p+1$, if the sequence of forthcoming states $s_{n+1}, \dots, s_{n+p}$ is equal to $\vec{s}^j$ for some j, the agent reports $(s_n, \vec{s}^{j+1})$ with a probability $\dfrac{\varepsilon}{\mathbf{P}(\vec{s}^j \mid s_n, s_{n+p+1})}$, and reports truxthfully $(s_n, \vec{s^j})$ with the residual probability. If the sequence of forthcoming states fails to coincide with any permutation $\vec{s^j}$ of $\vec{s}$, the agent reports truthfully on this block.

Conditional on the first states s_n and s_{n+p+1} of two consecutive blocks, the probability of reporting $\vec{s^{j+1}}$ rather than $\vec{s^j}$ is equal to ε. Hence, conditional on s_n and s_{n+p+1}, the sequence of reports within the block has the same distribution as the distribution of states, and this distribution is independent of the reports submitted in the earlier blocks. This implies **P1**.

By construction, σ reports truthfully in the first round of each block. For any other round n, one has either $s_n = a_n$ or $(s_n, a_n) = (s^j, s^{j+1})$ for some j. The latter occurs with probability ε, for each j. This establishes **P2** and proves Lemma 4.

B Implications and Applications

B.1 Proof of Proposition 3

We start with the reverse implication. Assume that truth-telling is optimal for some stationary transfer $t : A^{k+1} \to \mathbf{R}$ with memory k. Choose an arbitrary $\mu \in \mathcal{M}_k$ and let σ be the associated stationary reporting policy, with invariant measure $\nu \in \Delta(A^k \times S \times A)$. Because truth-telling is optimal,

$$\mathbf{E}_\nu \left[r(s,a) + t(a^{-k}, \ldots, a^{-1}, a) \right] \leq \mathbf{E}_{\nu^{tt}} \left[r(s,a) + t(a^{-k}, \ldots, a^{-1}, a) \right].$$

Since $\mu \in \mathcal{M}_k$, expected transfers are the same on both sides. So the above inequality rewrites, as desired,

$$\mathbf{E}_\mu[r(s,a)] \leq \mathbf{E}_{\mu^{tt}}[r(s,a)].$$

The proof of the direct implication relies on a minmax theorem. Assume that truth-telling is k-limit optimal, so that $\mathbf{E}_\mu[r(s,a)] \leq \mathbf{E}_{\mu^{tt}}[r(s,a)]$ for each $\mu \in \mathcal{M}_k$. Consider the zero-sum game in which the designer chooses a transfer $t : A^{k+1} \to \mathbf{R}$ such that $\mathbf{E}_\lambda[t(a^{-k}, \ldots, a^{-1}, a)] = 0$, the agent chooses an invariant distribution $\nu \in \mathcal{N}_k$ that arises under some stationary policy σ, and the payoff to the agent is

$$g(\nu, t) := \mathbf{E}_\nu \left[r(s,a) + t(a^{-k}, \ldots, a^{-1}, a) \right].$$

Both pure strategy sets are convex, and the agent's strategy set is compact. Since the payoff function is affine in each strategy, the game has a value in pure strategies. In addition, the agent has an optimal pure policy, by Sion Theorem.

Next, we claim that the value $V = \max_\nu \inf_t g(\nu, t)$ of the game is equal to $\mathbf{E}_{\mu^{tt}}[r(s,a)]$. Plainly, the agent can guarantee this amount by reporting truthfully. On the other hand, fix $\nu \in \mathcal{N}_k$. If the marginal of ν over A^{k+1} coincides with λ, then by definition the marginal of ν over $S \times A$ (denoted as μ) belongs to $\mathcal{M}_k$. In this case, one has

$$g(\nu, t) = \mathbf{E}_\mu [r(s,a)] \leq \mathbf{E}_{\mu^{tt}}[r(s,a)].$$

Assume now that the marginal of ν over A^{k+1} is not equal to λ. By a separation argument, there exists $t : A^{k+1} \to \mathbf{R}$ such that $\mathbf{E}_\lambda[t(a^{-k}, \ldots, a^{-1}, a)] = 0$ and $\mathbf{E}_\nu \left[t(a^{-k}, \ldots, a^{-1}, a)\right] < 0$. In that case, $\lim_{c \to +\infty} g(\nu, ct) = -\infty$. This concludes the proof that $V = \mathbf{E}_{\mu^{tt}}[r(s,a)]$.

We next show that the designer has an optimal policy, which will imply that truth-telling is optimal for some transfer function. Given $\mu \in \Delta(S \times A)$, we denote by $d(\mu, \mathcal{M}_k)$ its (Euclidean) distance to the convex set $\mathcal{M}_k$. And for any $\nu \in \mathcal{N}_k$, we denote by $\mu(\nu)$ and $\lambda(\nu)$ its marginals over $S \times A$ and A^{k+1} respectively. The next technical claim follows from $\mathcal{M}_k$ and $\mathcal{N}_k$ being polytopes, and the fact that $\mu(\nu) \in \mathcal{M}_k$ if, and only if, $\lambda(\nu) = \lambda$.

Claim 1 *There exists $c > 0$ such that $d(\mu(\nu), \mathcal{M}_k) \leq c \cdot \|\lambda(\nu) - \lambda\|$ (Euclidean norm) for every $\nu \in \mathcal{N}_k$.*

This claim implies that for each ν, there is a pure policy t of the designer with $\|t\| \leq c \cdot \|r\|$ (where $\|r\|$ is the Euclidean norm of the payoff function $r(s,a)$), such that $g(\nu, t) \leq V$. Indeed, if $\lambda(\nu) = \lambda$ then $t = 0$ does the job. If not, we set

$$t := -c \cdot \|r\| \cdot \frac{\lambda(\nu) - \lambda}{\|\lambda(\nu) - \lambda\|}.$$

Denoting by $\mu \in \mathcal{M}_k$ the point in $\mathcal{M}_k$ that is closest to $\mu(\nu)$, then it holds that

$$\begin{aligned} r \cdot (\mu(\nu) - \mu) &\leq \|r\| \cdot \|\mu(\nu) - \mu\| \\ &\leq c \cdot \|r\| \cdot \|\lambda(\nu) - \lambda\| \\ &= t \cdot (\lambda - \lambda(\nu)) \\ &= \mathbf{E}_\lambda(t) - \mathbf{E}_\nu(t) = -\mathbf{E}_\nu(t), \end{aligned}$$

where the first inequality uses the Cauchy-Schwarz inequality, and the second one follows from Claim 1. Reorganizing terms, we then obtain

$$\begin{aligned} \mathbf{E}_{\nu}\left[r(s,a)+t(a^{-k},\ldots,a)\right] &= \mathbf{E}_{\mu(\nu)}\left[r(s,a)\right]+\mathbf{E}_{\nu}\left[t\right] \\ &\leq \mathbf{E}_{\mu^{tt}}\left[r(s,a)\right]=V, \end{aligned}$$

as desired.

Thus, the zero-sum game in which the designer is restricted to the compact set of pure strategies t such that $\|t\| \leq c \cdot \|r\|$ still has value V in pure strategies. So the designer has an optimal strategy t, completing the proof of Proposition 3.

B.2 Discounting

In this appendix we show that our results for the undiscounted model (as presented in the main text) have analogues in the discounted patient limit. Recall that $\mu^{\delta}_{q,\sigma} \in \Delta(S \times A)$ denotes the expected, δ-discounted distribution of state and report pairs given a reporting policy σ. In Lemma 5, we have shown that any limit point of $\mu^{\delta}_{q,\sigma}$ as $\delta \to 1$ belongs to $\mathcal{M}_k$ (see the proof of Proposition 1). The following is a counterpart to Proposition 3:

Proposition 4 *Assume that*

$$\mathbf{E}_{\mu}[r(s,a)] < \mathbf{E}_{\mu^{tt}}[r(s,a)] \text{ for all } \mu \in \mathcal{M}_k \setminus \{\mu^{tt}\}.$$

Then for every $\varepsilon > 0$, when δ is sufficiently close to one there exists a transfer function $t_\delta : A^{k+1} \to \mathbf{R}$, such that

$$\|\mu^{\delta}_{q,\sigma} - \mu^{tt}\| < \varepsilon,$$

for all optimal reporting policies σ in the δ-discounted Markov decision problem induced by t_δ.

In particular, the discounted frequency of rounds in which the agent reports truthfully exceeds $1-\varepsilon$.

We here prove Proposition 4 by following the proof of Proposition 3. Note that in the zero-sum game considered in that proof, any optimal policy ν_* of the agent is such that $\lambda(\nu_*) = \lambda$. Thus, under the stronger assumption that $\mathbf{E}_{\mu}[r(s,a)] < \mathbf{E}_{\mu^{tt}}[r(s,a)]$ holds strictly for each $\mu \in \mathcal{M}_k \backslash \{\lambda\}$, truth-telling is the *unique* optimal policy of the agent in this zero-sum game.

Recall that $\nu^\delta_{q,\sigma} \in \Delta(A^k \times S \times A)$ is the expected discounted frequency given (q, σ). Let us set $\mathcal{N}^\delta_k$ be the convex hull of all $\nu^\delta_{q,\sigma}$. Then we have the following technical lemma:

Lemma 7 *As $\delta \to 1$, $\mathcal{N}^\delta_k$ converges to $\mathcal{N}_k$ (see the proof of Lemma 1) in the Hausdorff distance.*

Proof. Let $\nu \in \mathcal{N}_k$ be given, and σ be a stationary reporting policy which admits ν as an invariant measure. For each δ, if the agent follows σ and the initial distribution is ν, then expected discounted frequencies coincide with ν. Thus $\mathcal{N}_k$ is a subset of $\lim_{\delta \to 1} \mathcal{N}^\delta_k$. The reverse inclusion follows from Lemma 6. ■

Given $\delta < 1$, consider the zero-sum game in which the action set of the agent is $\mathcal{N}^\delta_k$, the action set of the principal is those transfers t with $\|t\| \leq c \cdot \|r\|$ and $\mathbf{E}_\lambda[t(s^{-k}, \ldots, s)] = 0$, and the payoff is defined as in the proof of Proposition 3. This game has a value V^δ, and we denote by ν^δ_* an arbitrary optimal policy of the agent. It follows from the Hausdorff convergence of $\mathcal{N}^\delta_k$ to $\mathcal{N}_k$ that V^δ converges to V, and that any limit point of ν^δ_* as $\delta \to 1$ is an optimal policy of the agent in the zero-sum game of Proposition 3. As stressed above, truth-telling is the unique optimal policy in that game. Therefore, $\nu^\delta_* \to \nu^{tt}$ and its marginal converges to μ^{tt} as desired.

B.3 Primal Version of Corollary 1

The primal version of Corollary 1 generalizes of cyclical monotonicity to the dynamic setting. We sketch this generalization below, quoting results from the literature on Markov decision processes (MDPs). Let transfers $t : A^{k+1} \to \mathbf{R}$ be such that truth-telling is optimal in the induced undiscounted MDP over the state space $A^k \times S$, and denote by V the value of the MDP. Then, there exist so-called *relative values* $h : A^k \times S \to \mathbf{R}$, with the following properties:
(i) for each $(\omega, s) \in A^k \times S$, one has

$$V + h(\omega, s) = \max_{a \in A} \left(u(s, \phi(a)) + t(\omega, a) + \sum_{s' \in S} p(s' \mid s) h((\omega^{+1}, a), s') \right) \qquad (12)$$

(denoting by (ω^{+1}, a) the sequence obtained by dropping the first entry of ω and appending a);
(ii) the maximum is achieved for $a = s$.

Consequently, if ϕ is implementable with transfers of memory k, then truth-telling is optimal in the *static* problem defined by the RHS of (12). By Rochet (1987), the function

$$g_\omega(s,a) := u(s,\phi(a)) + \sum_{s' \in S} p(s' \mid s) h((\omega^{+1},a),s')$$

is cyclically monotone for each $\omega \in A^k$.

Conversely, the existence of V and h such that (i) and (ii) hold implies the optimality of truth-telling when transfers are set to t, so that ϕ is implementable with transfers of memory k. If the above function g_ω is cyclically monotone for all ω, then there exists t such that truthful reporting achieves the maximum on the RHS of (12). If, in addition, t can be chosen in such a way that the maximum is equal to $V + h(\omega, s)$ for some V, then truth-telling would be optimal in the MDP.

This last condition is what complicates the dynamic setting relative to the static problem: the existence of a function h such that all functions g_ω are cyclically monotone is *not* sufficient for implementability. In the dynamic case, unlike in the static one, the primal version is "self-referential," as (12) relates the relative values h to the transfers t, the existence of which is precisely in question.[34]

B.4 Finite-horizon Implementation

As an alternative to our focus on infinite-horizon Markovian implementation problems, one might hope that long, but finite horizon versions of the implementation problem would be more tractable. They are, but unfortunately, with a finite horizon, implementation is only possible if it is possible in a one-shot interaction. This extends to more sophisticated schemes such as *review phases*, which apply to the infinite-horizon case. That is, following Radner (1981), it is customary in the literature to segment the infinite horizon into non-overlapping "phases" of $k+1$ rounds and define transfers on each phase separately, such that the agent's problem becomes separable across phases. Assume that reports are submitted in each round, but transfers are made every $k+1$ rounds, as a function of the last $k+1$ reports only. Say that ϕ is k-implementable in phases if there is $t : A^{k+1} \to \mathbf{R}$ such that truth-telling is optimal given transfers t.

Note that each such phase is independent of any other. Thus, in round $k+1$, irrespective of the reports $a_1, \ldots, a_k$ submitted thus far, truthful reporting is optimal in the *static* problem

---

[34] One might hope that long, but finite-horizon versions of the implementation problem would be more tractable. They are, but unfortunately, implementation with a finite horizon is only possible if it is possible in a one-shot interaction. This follows from the incentives in the last round.

when transfers are given by $t(a_1, \ldots, a_k, \cdot)$. This means ϕ is implementable with transfers of memory 0.

Thus, considering finite-horizon versions of the implementation problem (or using review phases more generally) is a poor way of understanding what happens with an infinite horizon: those versions collapse to the static problem. Equivalently, this says that requiring *exact* truth-telling in finite-horizon problems is extraordinarily demanding.

C Proof of Theorem 2

We prove the following, much stronger result.

Theorem 2' *Fix any k. There exists $\varepsilon > 0$ such that if the Markov chain is not pseudo-renewal, has full-support and $p_{ii} > 1 - \varepsilon, \forall i$, then $C_1 \supsetneq C_2 \supsetneq \cdots \supsetneq C_k \supsetneq C_{k+1}$.*

Below we first assume $k = 1$ to illustrate the methods, and then proceed to general k.

C.1 Iterative Linear System for $\mathcal{M}_k$

From Lemma 1, the set $\mathcal{M}_k$ is given by those joint distributions μ that satisfy (1)-(4), together with some ν. Given $\mu \in \mathcal{M}_0$, solving this system is one direct way of determining whether $\mu \in \mathcal{M}_k$. Alternatively, we can also check *iteratively* whether some undetectable distribution $\nu(a^{-(j-1)}, \ldots, a^{-1}, s, a)$ under memory $j-1$ might be *extended to* some $\nu(a^{-j}, \ldots, a^{-1}, s, a)$ that remains undetectable under memory j. This is formally expressed in the next lemma:

Lemma 8 *$\mu \in \mathcal{M}_k$ if, and only if, there exists $\nu : \cup_{j=0}^{k}(A^j \times S \times A) \to \mathbf{R}_+$, such that $\nu(s, a) = \mu(s, a)$, and for $1 \leq j \leq k$ the following holds:*

$$\sum_{a^{-j}} \nu(a^{-j}, \ldots, a^{-1}, s, a) = \nu(a^{-(j-1)}, \ldots, a^{-1}, s, a). \tag{13}$$

$$\sum_{s} \nu(a^{-j}, \ldots, a^{-1}, s, a) = \lambda(a^{-j}) p(a^{-(j-1)}|a^{-j}) \cdots p(a|a^{-1}). \tag{14}$$

$$\sum_{a} \nu(a^{-j}, \ldots, a^{-1}, s, a) = \sum_{s^{-1}} p(s|s^{-1}) \cdot \nu(a^{-j}, \ldots, a^{-2}, s^{-1}, a^{-1}). \tag{15}$$

Proof. Given $\nu(a^{-k}, \ldots, a^{-1}, s, a)$ that satisfies (1) to (4), simply define

$$\nu(a^{-j}, \ldots, a^{-1}, s, a) = \sum_{a^{-k}, \ldots, a^{-(j+1)}} \nu(a^{-k}, \ldots, a^{-1}, s, a).$$

Conditions (13) to (15) are checked in a straightforward manner. ■

This result will be very useful in our proofs below, because the linear system (13) to (15) allows us to solve for $\nu(a^{-j}, \ldots, a^{-1}, s, a)$ on the LHS in terms of $\nu(a^{-(j-1)}, \ldots, a^{-1}, s, a)$ on the RHS, a procedure that can be iterated. Importantly, note that we can hold $a^{-(j-1)}, \ldots, a^{-1}$ as fixed and consider (13) to (15) as a *3-dimensional transportation problem* in the variables a^{-j}, s and a, see Smith and Dawson (1979). This way, we reduce the original large linear system (with $|S|^{k+2}$ variables) to a collection of smaller linear systems (each with $|S|^3$ variables).

As an easy corollary, we have

Lemma 9 *For any $\mu \in \mathcal{M}_0$ there exists $\nu(a^{-k}, \ldots, a^{-1}, s, a)$, not necessarily positive, that solves (1) to (3). One may further have $\nu(a^{-k}, \ldots, a^{-1}, s, a) = 0$ whenever $\lambda(a^{-k}, \ldots, a^{-1}, a) = 0$.*

Proof. Let us first ignore the requirement regarding $\nu = 0$. By the previous lemma, it suffices to show that for fixed $\nu(a^{-(j-1)}, \ldots, a^{-1}, s, a)$, we can solve for $\nu(a^{-j}, \ldots, a^{-1}, s, a)$ (not necessarily positive) from equations (13) to (15). Holding fixed $a^{-(j-1)}, \ldots, a^{-1}$, the linear system is a 3-dimensional transportation problem. A solution exists if, and only if, certain "add-up constraints" involving the RHS of (13) to (15) are satisfied. Specifically, we need

$$\sum_{s} \nu(a^{-(j-1)}, \ldots, a^{-1}, s, a) = \sum_{a^{-j}} \lambda(a^{-j}) p(a^{-(j-1)}|a^{-j}) \cdots p(a|a^{-1}).$$

$$\sum_{a} \nu(a^{-(j-1)}, \ldots, a^{-1}, s, a) = \sum_{a^{-j}} \sum_{s^{-1}} p(s|s^{-1}) \cdot \nu(a^{-j}, \ldots, a^{-2}, s^{-1}, a^{-1}).$$

$$\sum_{a} \lambda(a^{-j}) p(a^{-(j-1)}|a^{-j}) \cdots p(a|a^{-1}) = \sum_{s} \sum_{s^{-1}} p(s|s^{-1}) \cdot \nu(a^{-j}, \ldots, a^{-2}, s^{-1}, a^{-1}).$$

These follow from the induction hypothesis that $\nu(a^{-(j-1)}, \ldots, a^{-1}, s, a)$ is a solution to the linear system defining $\mathcal{M}_{j-1}$.

To incorporate the requirement regarding $\nu = 0$, we show that if $\nu(a^{-(j-1)}, \ldots, a^{-1}, s, a) = 0$ whenever $\lambda(a^{-(j-1)}, \ldots, a^{-1}, s, a) = 0$, then for fixed $a^{-(j-1)}, \ldots, a^{-1}$ there exists a solution $\nu(a^{-j}, \ldots, a^{-1}, s, a)$ to the transportation problem, and this solution is zero whenever λ is. Indeed, if $\lambda(a^{-(j-1)}, \ldots, a^{-1}) = 0$, then we set $\nu(a^{-j}, \ldots, a^{-1}, s, a)$ to be zero for all a^{-j}, s, a. If

$\lambda(a^{-(j-1)}, \ldots, a^{-1}) \neq 0$, then we set $\nu(a^{-j}, \ldots, a^{-1}, s, a) = 0$ whenever $p(a^{-(j-1)} \mid a^{-j}) = 0$ or $p(a \mid a^{-1}) = 0$. In other words, for some subsets A_1, A_2 of S, we set $\nu = 0$ when $a^{-j} \in A_1$ or $a \in A_2$. This way we can ignore the value of ν at these points and treat equations (13) to (15) as a $|S - A_1| \times |S| \times |S - A_2|$ transportation problem. The add-up constraints are still satisfied, so such a ν exists. ■

C.2 The Case of $k = 1$

Since the process is not renewal, we may without loss assume:

$$p_{31} > p_{21}, \tag{16}$$

where p_{ij} denotes the ij-th entry of the matrix P. This assumption is maintained throughout this appendix.

We are going to consider those joint distributions μ that have the correct marginals and $\mu(1,2) = \mu(3,1) = 0$; here and later, $\mu(1,2)$ is a shorthand for $\mu(s = 1, a = 2)$ and similarly for ν. These μ form a polygon on a 2-dimensional plane, thus the restriction of $\mathcal{C}_k$ onto this plane is determined by two rays, which correspond to minimum and maximum values of the ratio $\frac{\mu(1,3)}{\mu(3,2)}$. To be more specific, all μ on this plane can be parameterized by

$$\mu(s,a) = \begin{pmatrix} \lambda_1 - y & 0 & y \\ y & \lambda_2 - x & x - y \\ 0 & x & \lambda_3 - x \end{pmatrix}$$

with $x \geq y \geq 0$. Any ray on this plane is determined by the ratio $\frac{\mu(1,3)}{\mu(3,2)} = \frac{y}{x}$.

Suppose $\nu(a^{-1}, s, a)$ is a solution to the linear system for $\mu \in \mathcal{M}_1$, then we clearly have

$$\mu(1,3) \geq \nu(2,1,3). \tag{17}$$

From $\mu(1,2)=\mu(3,1)=0$ we obtain $\nu(2,1,2)=\nu(2,3,1)=0$. Thus

$$\begin{aligned}\nu(2,1,3) &= [\nu(2,1,1)+\nu(2,1,2)+\nu(2,1,3)]-[\nu(2,1,1)+\nu(2,2,1)+\nu(2,3,1)]+\nu(2,2,1)\\ &= [\mu(2,2)p_{21}+\mu(3,2)p_{31}]-[(\mu(2,2)+\mu(3,2))p_{21}]+\nu(2,2,1)\\ &= \mu(3,2)(p_{31}-p_{21})+\nu(2,2,1).\end{aligned} \tag{18}$$

(17) and (18) together imply that a *necessary* condition for $\mu \in \mathcal{C}_1$ is

$$\mu(1,3) \geq \mu(3,2)\cdot(p_{31}-p_{21}), \forall \mu \in \mathcal{C}_1. \tag{19}$$

The following claim shows this condition is "tight" for characterizing $\mathcal{C}_1$, but a stronger necessary condition is needed for $\mu \in \mathcal{C}_2$. As a result, $\mathcal{C}_1 \neq \mathcal{C}_2$.

Claim 2 *Suppose $p_{11}+p_{22}+p_{33}>2$. Then there exists $\mu \in \mathcal{C}_1$ such that $\mu(1,3)=\mu(3,2)(p_{31}-p_{21})$. But for every $\mu \in \mathcal{C}_2$ it holds that $\mu(1,3)>\mu(3,2)(p_{31}-p_{21})$.*

Proof. Let us first handle the second part. Suppose $\nu(a^{-2},a^{-1},s,a)$ solves the linear system for $\mathcal{M}_2$ together with some μ. Define $\nu(a^{-1},s,a)$ to be the marginal of ν in those coordinates, then we have

$$\mu(1,3) \geq \nu(2,1,3)+\nu(2,3,1,3). \tag{20}$$

From (18) we have $\nu(2,1,3) \geq \mu(3,2)(p_{31}-p_{21})$. On the other hand, from $\nu(2,3,1,2)=\nu(2,3,3,1)=0$ we obtain

$$\begin{aligned}\nu(2,3,1,3) &= \sum_a \nu(2,3,1,a)-\nu(2,3,1,1)\\ &= \sum_a \nu(2,3,1,a)-\sum_s \nu(2,3,s,1)+\nu(2,3,2,1)\\ &= \nu(2,1,3)(p_{11}-p_{31})-\nu(2,2,3)(p_{31}-p_{21})+\nu(2,3,2,1).\end{aligned} \tag{21}$$

Next we compute $\nu(2,2,3)$ as follows:

$$\begin{aligned}\nu(2,2,3) &= \sum_a \nu(2,2,a) - \nu(2,2,2) - \nu(2,2,1) \\ &= \sum_a \nu(2,2,a) - \left(\sum_s \nu(2,s,2) - \nu(2,3,2)\right) - \nu(2,2,1) \\ &= \mu(3,2)(p_{32} - p_{22}) + \nu(2,3,2) - \nu(2,2,1) \\ &= \mu(3,2)(1 - p_{22} + p_{32}) - \nu(1,3,2) - \nu(3,3,2) - \nu(2,2,1) \\ &\leq \mu(3,2)(1 - p_{22} + p_{32}) - \nu(3,3,2).\end{aligned} \tag{22}$$

Plugging (18) and (22) into (21), we obtain

$$\nu(2,3,1,3) \geq \mu(3,2)(p_{31}-p_{21})(p_{11}-p_{31}-1+p_{22}-p_{32}) = \mu(3,2)(p_{31}-p_{21}) \cdot (p_{11}+p_{22}+p_{33}-2).$$

By (20), we conclude that a necessary condition for $\mu \in \mathcal{C}_2$ is

$$\mu(1,3) \geq \mu(3,2)(p_{31} - p_{21})(p_{11} + p_{22} + p_{33} - 1), \forall \mu \in \mathcal{C}_2. \tag{23}$$

By assumption $p_{11} + p_{22} + p_{33} - 1 > 1$, thus the above implies $\mu(1,3) > \mu(3,2)(p_{31} - p_{21})$ as desired.[35]

It remains to prove that $\mu(1,3) = \mu(3,2)(p_{31} - p_{21})$ is attainable for some $\mu \in \mathcal{C}_1$. In order for (19) to hold equal, we need $\nu(1,1,3) = \nu(3,1,3) = \nu(2,2,1) = 0$ from (17) and (18). Next we distinguish between two possibilities:

1. $p_{13} \geq p_{23}$. In this case we will solve for ν under the additional assumption that $\nu(1,3,2) = \nu(3,3,2) = 0$. Then ν has the following *configuration of zeros*:

$$\nu = \begin{pmatrix} + & 0 & 0 \\ + & + & + \\ 0 & 0 & + \end{pmatrix} \begin{pmatrix} + & 0 & + \\ 0 & + & + \\ 0 & + & + \end{pmatrix} \begin{pmatrix} + & 0 & 0 \\ + & + & + \\ 0 & 0 & + \end{pmatrix}$$

For example, the “0” in the third row and second column of the first matrix means

---

[35] The constant “2” appearing in the assumption $p_{11}+p_{22}+p_{33} > 2$ cannot be reduced. We can show (available from the authors) that if $p_{11} = p_{22} = p_{32} = 1-2w$ and other transition probabilities are equal to w, then $\mathcal{C}_1 = \mathcal{C}_k$ for every k. For such a non-renewal process, $p_{11} + p_{22} + p_{33} = 2 - 3w$ which can be made arbitrarily close to 2.

$\nu(1,3,2)=0$. We can compute ν as follows:

$$\nu(1,1,1)=\sum_a \nu(1,1,a)=\mu(1,1)p_{11}+\mu(2,1)p_{21}.$$

$$\nu(1,2,1)=\sum_s \nu(1,s,1)-\nu(1,1,1)=\mu(2,1)(p_{11}-p_{21}).$$

$$\nu(1,2,2)=\sum_s \nu(1,s,2)=\lambda_1 p_{12}.$$

$$\nu(1,3,3)=\sum_a \nu(1,3,a)=\mu(1,1)p_{13}+\mu(2,1)p_{23}.$$

$$\nu(1,2,3)=\sum_s \nu(1,s,3)-\nu(1,3,3)=\mu(2,1)(p_{13}-p_{23}).$$

$$\nu(2,1,1)=\sum_s \nu(2,s,1)=\lambda_2 p_{21}.$$

$$\nu(2,1,3)=\sum_{a^{-1}} \nu(a^{-1},1,3)=\mu(1,3).$$

$$\nu(2,3,2)=\sum_{a^{-1}} \nu(a^{-1},3,2)=\mu(3,2).$$

$$\nu(2,2,2)=\sum_s \nu(2,s,2)-\nu(2,3,2)=\mu(2,2)p_{22}-\mu(3,2)(1-p_{22}).$$

$$\nu(2,2,3)=\sum_a \nu(2,2,a)-\nu(2,2,2)=\mu(2,2)p_{22}+\mu(3,2)p_{32}-\nu(2,2,2)=\mu(3,2)(1-p_{22}+p_{32}).$$

$$\nu(2,3,3)=\sum_a \nu(2,3,a)-\nu(2,3,2)=\mu(2,2)p_{23}-\mu(3,2)(1-p_{33}).$$

$$\nu(3,1,1)=\sum_a \nu(3,1,a)=\mu(1,3)p_{11}+\mu(2,3)p_{21}+\mu(3,3)p_{31}.$$

$$\nu(3,2,1)=\sum_{a^{-1}} \nu(a^{-1},2,1)-\nu(1,2,1)=\mu(2,1)(1-p_{11}+p_{21}).$$

$$\nu(3,2,2)=\sum_s \nu(3,s,2)=\lambda_3 p_{32}.$$

$$\nu(3,3,3)=\sum_a \nu(3,3,a)=\mu(1,3)p_{13}+\mu(2,3)p_{23}+\mu(3,3)p_{33}.$$

$$\nu(3,2,3) = \sum_{s} \nu(3,s,3) - \nu(3,3,3) = \mu(1,3)(p_{33} - p_{13}) + \mu(2,3)(p_{33} - p_{23}).$$

Here $\nu(1,2,1) \geq 0$ because $p_{11} > 2 - p_{22} - p_{33} > 1 - p_{22} > p_{21}$. Also, $\nu(3,2,3) \geq 0$ because similarly $p_{33} \geq p_{13}, p_{23}$. While $\nu(2,2,2)$ and $\nu(2,3,3)$ could be negative, they do not cause any issue because they are on the diagonal (so that we can mix ν with ν^{tt} to make these entries positive). Thus ν is positive and $\mu \in \mathcal{C}_1$.

2. $p_{23} > p_{13}$. Here we instead assume $\nu(1,2,3) = \nu(3,3,2) = 0$ (the difference is that $\nu(1,3,2)$ is no longer zero), leading to the following configuration of zeros:

$$\nu = \begin{pmatrix} + & 0 & 0 \\ + & + & 0 \\ 0 & + & + \end{pmatrix} \begin{pmatrix} + & 0 & + \\ 0 & + & + \\ 0 & + & + \end{pmatrix} \begin{pmatrix} + & 0 & 0 \\ + & + & + \\ 0 & 0 & + \end{pmatrix}$$

Again, we can solve ν in terms of μ. The solution differs from the previous case in that $\nu(1,2,3), \nu(1,3,2), \nu(2,2,2), \nu(3,3,3)$ are increased by $\Delta = \mu(2,1)(p_{23} - p_{13})$, whereas $\nu(1,2,2), \nu(1,3,3), \nu(2,2,3), \nu(2,3,2)$ are decreased by Δ. We need to show that the resulting ν is still positive. It suffices to check the four entries that are decreased, which further reduces to checking the off-diagonal entires $\nu(2,2,3)$ and $\nu(2,3,2)$:

$$\nu(2,2,3) = \mu(3,2)(1-p_{22}+p_{32}) - \mu(2,1)(p_{23}-p_{13}) = \mu(3,2)(1-p_{22}+p_{32}) - \mu(1,3)(p_{23}-p_{13}).$$

$$\nu(2,3,2) = \mu(3,2) - \mu(1,3)(p_{23} - p_{13}).$$

Both are positive because $\mu(3,2) = \mu(1,3) + \mu(2,3) \geq \mu(1,3)$. Thus again $\mu \in \mathcal{C}_1$.

This completes the proof of Claim 2 and thus of Theorem 2' when $k = 1$. ■

C.3 General k: Lower Bound on $\frac{\mu(1,3)}{\mu(3,2)}$

Here we develop the analogue of (23) for $\mu \in \mathcal{C}_k$. Suppose $\nu(a^{-k}, \ldots, a^{-1}, s, a)$ together with μ solves the linear system defining $\mathcal{M}_k$. For $1 \leq m < k$, let $\nu(a^{-m}, \ldots, a^{-1}, s, a)$ denote the marginal of ν in those coordinates. Note that it solves the linear system for $\mathcal{M}_m$.

By the positivity of ν, we must have

$$\mu(1,3) \geq \sum_{m=1}^{k} \nu(a^{-m} = 2, a^{-m-1} = \cdots = a^{-1} = 3, s = 1, a = 3). \tag{24}$$

As before, we will give a lower bound to the RHS sum in terms of μ. Starting from $\nu(2,1,3)$ and $\nu(2,2,3)$, we have the following recursive relations regarding $\nu(2,3,\ldots,3,1,3)$ and $\nu(2,3,\ldots,3,2,3)$:

$$\begin{aligned}
&\nu(a^{-m} = 2,3,\ldots,3, s = 1, a = 3) - \nu(a^{-m} = 2,3,\ldots,3, s = 2, a = 1) \\
=&\sum_{a} \nu(a^{-m} = 2,3,\ldots,3, s = 1, a) - \sum_{s} (a^{-m} = 2,3,\ldots,3, s, a = 1) \\
=&\nu(a^{-(m-1)} = 2,3,\ldots,3, s = 1, a = 3)(p_{11} - p_{31}) - \nu(a^{-(m-1)} = 2,3,\ldots,3, s = 2, a = 3)(p_{31} - p_{21}).
\end{aligned}$$

$$\begin{aligned}
&\nu(a^{-m} = 2,3,\ldots,3, s = 2, a = 3) + \nu(a^{-m} = 2,3,\ldots,3, s = 2, a = 1) - \nu(a^{-m} = 2,3,\ldots,3, s = 3, a = 2) \\
=&\sum_{a} \nu(a^{-m} = 2,3,\ldots,3, s = 2, a) - \sum_{s} \nu(a^{-m} = 2,3,\ldots,3, s, a = 2) \\
=&\nu(a^{-(m-1)} = 2,3,\ldots,3, s = 1, a = 3)(p_{12} - p_{32}) + \nu(a^{-(m-1)} = 2,3,\ldots,3, s = 2, a = 3)(p_{22} - p_{32}).
\end{aligned}$$

Simple induction leads to the following lemma:

Lemma 10 *Define the first-order recursive sequences* $\{\gamma_m\}$ *and* $\{\eta_m\}$*:*

$$\begin{aligned}
\gamma_1 &= 1; \\
\eta_1 &= 0; \\
\gamma_{m+1} &= (p_{11} - p_{31})\gamma_m - (p_{31} - p_{21})(p_{12} - p_{32})\eta_m; \\
\eta_{m+1} &= \gamma_m + (p_{22} - p_{32})\eta_m.
\end{aligned} \tag{25}$$

Suppose $\nu(a^{-k},\ldots,a^{-1},s,a)$ *and* μ *solves the linear system for* $\mathcal{M}_k$*, with* $\mu(1,2) = \mu(3,1) = 0$*.*

Then for $1 \leq m \leq k$ *it holds that*

$$\begin{aligned}\nu(a^{-m} = 2,3,\ldots,3, s = 1, a = 3) = \quad & \nu(2,1,3)\gamma_m - \nu(2,2,3)(p_{31} - p_{21})\eta_m \\ & + \sum_{j=2}^{m} \nu(a^{-j} = 2,3,\ldots,3, s = 2, a = 1)(\gamma_{m+1-j} + (p_{31} - p_{21})\eta_{m+1-j}) \\ & - \sum_{j=2}^{m} \nu(a^{-j} = 2,3,\ldots,3, s = 3, a = 2)(p_{31} - p_{21})\eta_{m+1-j}.\end{aligned} \tag{26}$$

Similarly

$$\begin{aligned}& \nu(a^{-m} = 2,3,\ldots,3, s = 2, a = 3) \\ = \quad & \nu(2,1,3)(p_{12} - p_{32})\eta_m + \nu(2,2,3)(\gamma_m + (p_{22} - p_{32} - p_{11} + p_{31})\eta_m) \\ & - \sum_{j=2}^{m} \nu(a^{-j} = 2,3,\ldots,3, s = 2, a = 1)(\gamma_{m+1-j} + (p_{22} - p_{12} - p_{11} + p_{31})\eta_{m+1-j}) \\ & + \sum_{j=2}^{m} \nu(a^{-j} = 2,3,\ldots,3, s = 3, a = 2)(\gamma_{m+1-j} + (p_{22} - p_{32} - p_{11} + p_{31})\eta_{m+1-j}).\end{aligned} \tag{27}$$

Plugging (26) and (27) into the inequality (24), we obtain:

$$\begin{aligned}\mu(1,3) \geq \quad & \nu(2,1,3) \cdot \sum_{j=1}^{k} \gamma_j - \nu(2,2,3) \cdot (p_{31} - p_{21}) \sum_{j=1}^{k} \eta_j \\ & + \sum_{j=2}^{k} \nu(a^{-j} = 2,3,\ldots,3, s = 2, a = 1) \cdot \sum_{m=j}^{k} (\gamma_{m+1-j} + (p_{31} - p_{21})\eta_{m+1-j}) \\ & - \sum_{j=2}^{k} \nu(a^{-j} = 2,3,\ldots,3, s = 3, a = 2) \cdot (p_{31} - p_{21}) \sum_{m=j}^{k} \eta_{m+1-j}.\end{aligned} \tag{28}$$

This enables us to deduce the following lower bounds on $\frac{\mu(1,3)}{\mu(3,2)}$.

Lemma 11 *Suppose* $\gamma_j, \eta_j \geq 0, \forall 1 \leq j \leq k$. *If* $p_{13} \geq p_{23}$, *then a necessary condition for* $\mu \in \mathcal{C}_k$ *and* $\mu(1,2) = \mu(3,1) = 0$ *is:*

$$\mu(1,3) \geq \mu(3,2) \cdot (p_{31} - p_{21}) \left[\sum_{j=1}^{k} \gamma_j - (1 - p_{22} + p_{32}) \sum_{j=1}^{k} \eta_j \right], \forall \mu \in \mathcal{C}_k. \tag{29}$$

If instead $p_{23} > p_{13}$*, then the corresponding necessary condition is:*

$$\mu(1,3)\cdot\left(1-(p_{31}-p_{21})(p_{23}-p_{13})\sum_{j=1}^{k}\eta_j\right) \geq \mu(3,2)\cdot(p_{31}-p_{21})\left[\sum_{j=1}^{k}\gamma_j-(1-p_{22}+p_{32})\sum_{j=1}^{k}\eta_j\right], \forall\mu\in\mathcal{C}_k. \tag{30}$$

Proof. From (18) and (22) above, we have $\nu(2,1,3) \geq \mu(3,2)\cdot(p_{31}-p_{21})$ and $\nu(2,2,3) \leq \mu(3,2)(1-p_{22}+p_{32})-\nu(3,3,2) \leq \mu(3,2)\cdot(1-p_{22}+p_{32})-\sum_{j=2}^{k}\nu(a^{-j}=2,3,\ldots,3,s=3,a=2)$. Plugging these into (28), we deduce (29) with extra positive terms on the RHS.

If $p_{23} > p_{13}$, then as mentioned $\nu(1,3,2)$ is no longer zero. In fact,

$$\begin{aligned}\nu(1,3,2) &= \sum_{a}\nu(1,3,a)-\nu(1,3,3)\\ &= \mu(1,1)p_{13}+\mu(2,1)\cdot p_{23}-\sum_{s}\nu(1,s,3)+\nu(1,1,3)+\nu(1,2,3)\\ &= \mu(2,1)(p_{23}-p_{13})+\nu(1,1,3)+\nu(1,2,3)\\ &= \mu(1,3)(p_{23}-p_{13})+\nu(1,1,3)+\nu(1,2,3).\end{aligned}$$

Plugging this into the penultimate line of (22), we obtain the stronger bound

$$\nu(2,2,3) \leq \mu(3,2)(1-p_{22}+p_{32})-\mu(1,3)(p_{23}-p_{13})-\nu(3,3,2). \tag{31}$$

This stronger inequality together with (18) and (28) enables us to conclude (30). ■

C.4 General k: Attaining the Lower Bound

The following result shows that the lower bound (29) or (30) is attainable and increasingly more restrictive as the memory increases. It will imply Theorem 2'.

Lemma 12 *Fix* $k \geq 2$ *and suppose* p_{ii} *is sufficiently close to* 1 *for each* i*. Then* $\gamma_j > (1-p_{22}+p_{32})\eta_j > 0, \forall 1 \leq j \leq k$*. Moreover there exists* $\mu\in\mathcal{C}_k$ *such that* $\mu(1,2)=\mu(3,1)=0$*, and (29) or (30) holds with equality depending on whether* p_{13} *or* p_{23} *is larger.*

As $p_{ii}\to 1$, it is clear from (25) that $\gamma_j \to 1$ and $\eta_j \to j-1$ for $1\leq j\leq k$. Thus we do have $\gamma_j > (1-p_{22}+p_{32})\eta_j > 0$.

Below we prove in detail that when $p_{13}\geq p_{23}$, the relevant necessary condition (29) is achieved with equality at some $\mu\in\mathcal{C}_k$. The other case $p_{23} > p_{13}$ can be handled by almost the same arguments, which we omit.

We begin by investigating necessary conditions on $\nu(a^{-k}, \ldots, a^{-1}, s, a)$ in order for (29) to hold with equality. From the proof of Lemma 11, we see that the following ensures the equivalence between (28) and (29):

$$\begin{aligned} &\nu(1,3,2) = \nu(3,3,2) = 0. \\ &\nu(a^{-j} = 2, 3, \ldots, 3, s = 2, a = 1) = 0, \ \forall 1 \le j \le k. \end{aligned} \tag{32}$$

Since (28) comes from (24), we also impose the following to achieve (24) with equality:

$$\nu(a^{-j} = 1, 3, \ldots, 3, s = 1, a = 3) = 0, \ \forall 1 \le j \le k \tag{33}$$

$$\nu(a^{-k} = \cdots = a^{-1} = 3, s = 1, a = 3) = 0. \tag{34}$$

We will use these conditions to iteratively solve for $\nu(a^{-m}, \ldots, a^{-1}, s, a)$, for $1 \le m \le k$. The base case $\nu(a^{-1}, s, a)$ is straightforward:

Claim 3 *Suppose p_{ii} is sufficiently close to one, then there exists $\nu(a^{-1}, s, a) \ge 0$ with $\nu(1,1,3) = \nu(1,3,2) = \nu(2,2,1) = \nu(3,3,2) = 0$ that solves the linear system for $\mu \in \mathcal{M}_1$.*

Proof. Such a ν can be uniquely determined, and the solution is mostly identical to that given in the proof of Claim 2. The difference here results from $\mu(1,3)$ being strictly larger than $\mu(3,2)(p_{31} - p_{21})$, and so only a few entries change as follows:

$$\nu(2,1,3) = \mu(3,2)(p_{31} - p_{21}).$$

$$\nu(3,1,3) = \mu(1,3) - \nu(2,1,3) = \mu(1,3) - \mu(3,2)(p_{31} - p_{21}).$$

$$\begin{aligned} \nu(3,2,3) &= \mu(1,3)(p_{33} - p_{13}) + \mu(2,3)(p_{33} - p_{23}) - \mu(1,3) + \mu(3,2)(p_{31} - p_{21}) \\ &= \mu(3,2)(p_{22} - p_{32}) - \mu(1,3)(1 + p_{13} - p_{23}). \end{aligned}$$

Since μ satisfies (29) with equality, we have $\frac{\mu(1,3)}{\mu(3,2)} \to 0$ as $p_{ii} \to 1$. Thus $\nu(3,2,3) > 0$, completing the proof. ■

The induction step is summarized by the following claim:

Claim 4 *Suppose that, for some $2 \le m \le k$, we have found $\nu(a^{-(m-1)}, \ldots, a^{-1}, s, a) \ge 0$ that solves the linear system for $\mathcal{M}_{m-1}$, obeys (32) to (34) and satisfies the following proportionality* *condition:*

$$\frac{\nu(a^{-(m-1)},\dots,a^{-2},a^{-1}=2,s=1,a=3)}{\nu(a^{-(m-1)},\dots,a^{-2},a^{-1}=2,s=2,a=3)}=\frac{\nu(2,1,3)}{\nu(2,2,3)},\forall a^{-(m-1)},\dots,a^{-2}. \tag{35}$$

Then this ν can be extended to $\nu(a^{-m},\dots,a^{-1},s,a)\geq 0$ that solves the linear system for $\mathcal{M}_m$, obeys (32) to (34) and satisfies the corresponding proportionality condition for all $a^{-m},\dots,a^{-2}$.

Proof. By our discussion of the iterated linear system in Appendix C.1, we shall determine $\nu(a^{-m},\dots,a^{-1},s,a)$ from the following system of equations:

$$\sum_{a^{-m}}\nu(a^{-m},\dots,a^{-1},s,a)=\nu(a^{-(m-1)},\dots,a^{-1},s,a).$$

$$\sum_{s}\nu(a^{-m},\dots,a^{-1},s,a)=\lambda(a^{-m})p(a^{-(m-1)}|a^{-m})\cdots p(a|a^{-1}).$$

$$\sum_{a}\nu(a^{-m},\dots,a^{-1},s,a)=\sum_{s^{-1}}\nu(a^{-m},\dots,a^{-2},s^{-1},a^{-1})p(s|s^{-1}).$$

We take $a^{-(m-1)},\dots,a^{-1}$ as fixed and view the above as a $3\times3\times3$ transportation problem. First assume $a^{-1}=1$. Because $\nu(1,1,2)=\nu(1,1,3)=\nu(1,3,1)=\nu(1,3,2)=0$, we must have $\nu(a^{-m},\dots,1,1,2)=\cdots=0$. Then we can uniquely solve for $\nu(a^{-m},\dots,a^{-1}=1,s,a)$ from the required row and column sums. The only off-diagonal entry remaining is

$$\nu(a^{-m},\dots,a^{-1}=1,s=2,a=1)=\nu(a^{-m},\dots,a^{-1}=2,s^{-1}=2,a^{-1}=1)(p_{11}-p_{21}).$$

This is positive, so we are done. As before, diagonal entries being negative is not an issue because we are concerned with the cone $\mathcal{C}_1$.

Next assume $a^{-1}=2$. In this case we have $\nu(2,1,2)=\nu(2,2,1)=\nu(2,3,1)=0$, from which we obtain

$$\nu(a^{-m},\dots,a^{-1}=2,s=1,a=3)=\nu(a^{-m},\dots,a^{-2},s^{-1}=3,a^{-1}=2)(p_{31}-p_{21}).$$

From the proportionality condition we deduce

$$\nu(a^{-m},\dots,a^{-1}=2,s=2,a=3)=\nu(a^{-m},\dots,a^{-2},s^{-1}=3,a^{-1}=2)(p_{31}-p_{21})\frac{\nu(2,2,3)}{\nu(2,1,3)}.$$

The remaining off-diagonal entry is thus given by

$$\begin{aligned}&\nu(a^{-m},\dots,a^{-1}=2,s=3,a=2)-\nu(a^{-m},\dots,a^{-1}=2,s=2,a=3)\\=&\sum_s\nu(a^{-m},\dots,a^{-1}=2,s,a=2)-\sum_a\nu(a^{-m},\dots,a^{-1}=2,s=2,a)\\=&\nu(a^{-m},\dots,a^{-2},s^{-1}=3,a^{-1}=2)(p_{22}-p_{32}).\end{aligned}$$

Therefore $\nu(a^{-m},\dots,a^{-1}=2,s=2,a=3)$ and $\nu(a^{-m},\dots,a^{-1}=2,s=3,a=2)$ are both positive as desired.

Lastly we turn to $a^{-1}=3$. Here we know $\nu(3,1,2)=\nu(3,3,1)=\nu(3,3,2)=0$. If $a^{-m},\dots,a^{-2}$ are not all equal to state "3", we let t denote the smallest index such that $a^{-t}\neq 3$. There are three possibilities to consider:

1. $a^{-t}=1$. From (33) we have $\nu(a^{-m},\dots,a^{-1}=3,s=1,a=3)=0$. We can then solve for the other entries of ν:

$$\begin{aligned}&\nu(a^{-m},\dots,a^{-1}=3,s=2,a=1)\\=&\sum_s\nu(a^{-m},\dots,a^{-1}=3,s,a=1)-\sum_a\nu(a^{-m},\dots,a^{-1}=3,s=1,a)\\=&\nu(a^{-m},\dots,a^{-2},s^{-1}=2,a^{-1}=3)(p_{31}-p_{21}).\end{aligned}\tag{36}$$

   In the last step, we used $\nu(a^{-m},\dots,a^{-2},s^{-1}=1,a^{-1}=3)=0$, which is due to $a^{-t}=1,a^{-(t-1)}=\cdots=a^{-2}=3$ and the induction hypothesis (33). Similarly

$$\nu(a^{-m},\dots,a^{-1}=3,s=2,a=3)=\nu(a^{-m},\dots,a^{-2},s^{-1}=2,a^{-1}=3)(p_{33}-p_{23}).\tag{37}$$

   These are both positive, as desired.

2. $a^{-t}=2$. From (32) we know that $\nu(a^{-m},\dots,a^{-1}=3,s=2,a=1)=0$. The remaining off-diagonal entries $\nu(\cdots,3,1,3)$ and $\nu(\cdots,3,2,3)$ can then be solved as follows:

$$\begin{aligned}&\nu(a^{-m},\dots,a^{-1}=3,s=1,a=3)\\=&\sum_a\nu(a^{-m},\dots,a^{-1}=3,s=1,a)-\sum_s\nu(a^{-m},\dots,a^{-1}=3,s,a=1)\\=&\nu(a^{-m},\dots,a^{-2},s^{-1}=1,a^{-1}=3)(p_{11}-p_{31})-\nu(a^{-m},\dots,a^{-2},s^{-1}=2,a^{-1}=3)(p_{31}-p_{21}).\end{aligned}\tag{38}$$

To show this is positive, note that the proportionality condition (35) generalizes to the following:

$$\frac{\nu(a^{-m},\dots,a^{-t}=2,3,\dots,3,s^{-1}=1,a^{-1}=3)}{\nu(a^{-m},\dots,a^{-t}=2,3,\dots,3,s^{-1}=2,a^{-1}=3)}=\frac{\nu(a^{-t}=2,3,\dots,3,s^{-1}=1,a^{-1}=3)}{\nu(a^{-t}=2,3,\dots,3,s^{-1}=2,a^{-1}=3)}. \tag{39}$$

The case $t=2$ is exactly the proportionality condition assumed in the statement of the claim. The general result follows by exploiting recursive formulae of the form (38).

Given (39), the RHS of (38) is positive if $\nu(a^{-t}=2,3,\dots,3,s^{-1}=1,a^{-1}=3)(p_{11}-p_{31})\geq \nu(a^{-t}=2,3,\dots,3,s^{-1}=2,a^{-1}=3)(p_{31}-p_{21})$. Plugging in (26) and (27) as well as (32) and (33),[36] we simply need to show $\nu(2,1,3)\gamma_t\geq\nu(2,2,3)(p_{31}-p_{21})\eta_t$. We have computed that $\nu(2,1,3)=\mu(3,2)(p_{31}-p_{21})$ and $\nu(2,2,3)=\mu(3,2)(1-p_{22}+p_{32})$. Thus, the desired inequality reduces to $\gamma_t\geq(1-p_{22}+p_{32})\eta_t$, which we have shown earlier.

Similarly we have

$$\begin{aligned}&\nu(a^{-m},\dots,a^{-1}=3,s=2,a=3)\\=&\sum_a\nu(a^{-m},\dots,a^{-1}=3,s=2,a)-\sum_s\nu(a^{-m},\dots,a^{-1}=3,s,a=2)\\=&\nu(a^{-m},\dots,a^{-2},s^{-1}=1,a^{-1}=3)(p_{12}-p_{32})+\nu(a^{-m},\dots,a^{-2},s^{-1}=2,a^{-1}=3)(p_{22}-p_{32}).\end{aligned}$$

Using (39) and plugging in (26) and (27) again, we see that the RHS above is positive if $\nu(2,1,3)(p_{32}-p_{12})\eta_t\leq\nu(2,2,3)(\gamma_t+(p_{22}-p_{32}-p_{11}+p_{31})\eta_t))$. This further simplifies to $(p_{31}-p_{21})(p_{32}-p_{12})\eta_t\leq(1-p_{22}+p_{32})(\gamma_t+(p_{22}-p_{32}-p_{11}+p_{31})\eta_t))$, which holds because $0<p_{31}-p_{21}\leq 1-p_{11}+p_{31}$, $p_{32}-p_{12}<1-p_{22}+p_{32}$ and $\gamma_t\geq(1-p_{22}+p_{32})\eta_t$. Hence we have proved in this case that ν is positive off the diagonal.

3. $a^{-m}=\cdots=a^{-1}=3$, so that t does not exist. Here the off-diagonal entries are $\nu(\cdots,3,1,2)=\nu(\cdots,3,3,1)=\nu(\cdots,3,3,2)=0$, as well as $\nu(\cdots,3,1,3)$, $\nu(\cdots,3,2,1)$

---

[36] These are valid since $\nu(a^{-t},\dots,a^{-2},s^{-1},a^{-1})$ solves the linear system for $\mathcal{M}_{t-1}$ by the induction hypothesis.

and $\nu(\cdots,3,2,3)$. Using (33) and (26), we compute that

$$\begin{aligned}
&\nu(a^{-m}=\cdots=a^{-1}=3,s=1,a=3)\\
=&\mu(1,3)-\sum_{j=1}^{m}\nu(a^{-j}=2,a^{-(j-1)}=\cdots=a^{-1}=3,s=1,a=3)\\
=&\mu(1,3)-\mu(3,2)(p_{31}-p_{21})\sum_{j=1}^{m}(\gamma_j-(1-p_{22}+p_{32})\eta_j).
\end{aligned}$$

This is positive by the assumption that (29) holds equal and $\gamma_j > (1-p_{22}+p_{32})\eta_j$. In particular, at $m=k$ we have $\nu(a^{-k}=\cdots=a^{-1}=3,s=1,a=3)=0$, as required by (34). From (32) we similarly obtain

$$\begin{aligned}
&\nu(a^{-m}=\cdots=a^{-1}=3,s=2,a=1)\\
=&\nu(3,2,1)-\sum_{j=2}^{m}\nu(a^{-j}=1,a^{-(j-1)}=\cdots=a^{-1}=3,s=2,a=1)\\
=&\mu(1,3)(1-p_{11}+p_{21})-\sum_{j=1}^{m-1}\nu(a^{-j}=1,a^{-(j-1)}=\cdots=a^{-1}=3,s=2,a=3)(p_{31}-p_{21})\\
=&\mu(1,3)(1-p_{11}+p_{21})-\mu(1,3)\cdot(p_{13}-p_{23})(p_{31}-p_{21})\sum_{j=1}^{m-1}(p_{33}-p_{23})^{j-1}.
\end{aligned}$$

In the last two steps above we used recursive formulae of the form (36) and (37) and $\nu(1,2,3)=\mu(1,3)(p_{13}-p_{23})$. Since $1-p_{11}+p_{21} > p_{13}-p_{23}$ and $(p_{31}-p_{21})\sum_{j=1}^{m-1}(p_{33}-p_{23})^{j-1} < \frac{p_{31}-p_{21}}{1-p_{33}+p_{23}} < 1$, this ν is also positive.

Finally,

$$\begin{aligned}
&\nu(a^{-m}=\cdots=a^{-1}=3,s=2,a=3)\\
=&\mu(2,3)-\sum_{j=1}^{m}\nu(a^{-j}=1,3,\ldots,3,s=2,a=3)-\sum_{j=1}^{m}\nu(a^{-j}=2,3,\ldots,3,s=2,a=3)\\
=&\quad\mu(3,2)-\mu(1,3)-\mu(1,3)(p_{13}-p_{23})\frac{1-(p_{33}-p_{23})^m}{1-p_{33}+p_{23}}-\mu(3,2)(1-p_{22}+p_{32})\sum_{j=1}^{m}\gamma_j\\
&-\mu(3,2)\left((p_{31}-p_{21})(p_{12}-p_{32})+(1-p_{22}+p_{32})(p_{22}-p_{32}-p_{11}+p_{31})\right)\sum_{j=1}^{m}\eta_j.
\end{aligned}$$

In the last step we used the calculation just now for $\sum_{j=1}^{m} \nu(a^{-j} = 1, 3, \ldots, 3, s = 2, a = 3)$ as well as (27). Now recall that as $p_{ii} \to 1$, $\gamma_j \to 1$ while $\eta_j \to j - 1$. Thus $(1 - p_{22} + p_{32}) \sum_{j=1}^{m} \gamma_j$ and $((p_{31} - p_{21})(p_{12} - p_{32}) + (1 - p_{22} + p_{32})(p_{22} - p_{32} - p_{11} + p_{31})) \sum_{j=1}^{m} \eta_j$ both vanish. Furthermore, since we assume (29) holds equal,

$$\frac{\mu(1,3)}{\mu(3,2)(1 - p_{33} + p_{23})} < \frac{\mu(1,3)}{\mu(3,2)(p_{31} - p_{21})} < \sum_{j=1}^{k} \gamma_j$$

remains bounded. These observations together imply that $\nu(a^{-m} = \cdots = a^{-1} = 3, s = 2, a = 3)$ is positive as $p_{ii} \to 1$.

With that we have verified $\nu(a^{-m}, \ldots, a^{-1}, s, a)$ is everywhere positive. This completes the proof of the claim. Lemma 12 follows, and so does Theorem 2'. ■

D Omitted Proofs in Section 4.3

With the notations of Section 4.3, we here prove that $\sup_{L,M} S^{k,L,M} < 1$, for each k.

The intuition for the proof is straightforward, as discussed in the main text. The only challenge is the following. To get a payoff arbitrarily close to 1 in the finitely repeated game, the equilibrium strategy of player 1 must be close to truth-telling (in a sense made precise below), but it need not be exactly truth-telling. The strategy that implements μ^{tt} (the candidate for a profitable deviation by player 1) is indistinguishable from exact truth-telling, but it need not be indistinguishable from a strategy close to it. Hence, one must show that for every strategy close to truth-telling, there is a strategy close to the one implementing μ^{tt} that is indistinguishable from it. That this is possible is a consequence of Lemma 14 below, which builds on Lemma 13.

Lemma 13 *Let T and B be finite sets, and $\mu \in \Delta(T \times B)$. For each $\tilde{\mu}|_B \in \Delta(B)$, there exists a distribution $\hat{\mu} \in \Delta(T \times B)$ such that*

$$\hat{\mu}|_T = \mu|_T,\ \hat{\mu}|_B = \tilde{\mu}|_B \text{ and } \|\mu - \hat{\mu}\|_1 = \|\mu|_B - \tilde{\mu}|_B\|_1.$$

In this statement, $\mu|_T$ stands for the marginal of μ on T, and $\|\cdot\|_1$ represents the L_1 norm. Lemma 13 is a finite-dimensional version of Lemma 8.2 in Laraki (2004). For completeness, we provide a proof below.

Proof. The distribution $\hat{\mu}$ will be obtained as the solution to a linear program. Define $B_+ = \{b, \tilde{\mu}|_B(b) \geq \mu|_B(b)\}$ and $B_- = \{b, \tilde{\mu}|_B(b) \leq \mu|_B(b)\}$. Elements of B_+ (resp., B_-) should see their weight increase (resp., decrease). Consider the linear program

$$\sup \left(\sum_{t \in T, b \in B_+} \delta_{tb} - \sum_{t \in T, b \in B_-} \delta_{tb} \right),$$

subject to the constraints:

C1 $\delta_{tb} \geq 0$ for all $b \in B_+$ and $\delta_{tb} \leq 0$ for all $b \in B_-$, all $t \in T$.

C2 $0 \leq \mu(t,b) + \delta_{tb} \leq 1$, all $t \in T$, $b \in B$.

C3 $\sum_{b \in B} \delta_{tb} = 0$, all $t \in T$.

C4 $\mu|_B(b) + \sum_{t \in T} \delta_{tb} \leq \tilde{\mu}|_B(b)$, for all $b \in B_+$, and $\mu|_B(b) + \sum_{t \in T} \delta_{tb} \geq \tilde{\mu}|_B(b)$ for all $b \in B_-$.

This problem is feasible ($\delta = 0$) and bounded (by **C2**), hence has an optimal solution $\hat{\delta}$. We will show that $\hat{\mu}(t,b) := \mu(t,b) + \hat{\delta}_{tb}$ is the desired distribution.

Conditions **C2** and **C3** ensure that $\hat{\mu} \in \Delta(T \times B)$, with $\hat{\mu}|_T = \mu|_T$. Condition **C1** implies that $\|\hat{\mu} - \mu\|_1 = \sum_{t, b \in B_+} \hat{\delta}_{tb} - \sum_{t, b \in B_-} \hat{\delta}_{tb}$. Condition **C4** implies that $\hat{\mu}|_B(b) \leq \tilde{\mu}|_B(b)$ if, and only if, $b \in B_+$. Hence, the value of the program cannot exceed $\|\mu|_B - \tilde{\mu}|_B\|_1$ and equality holds iff $\hat{\mu}|_B = \tilde{\mu}|_B$. In that case, it follows that $\|\mu - \hat{\mu}\|_1 = \|\mu|_B - \tilde{\mu}|_B\|_1$.

Thus, it suffices to check that $\hat{\mu}|_B = \tilde{\mu}|_B$. Assume to the contrary that $\hat{\mu}|_B(b) < \tilde{\mu}|_B(b)$ and $\hat{\mu}|_B(b') > \tilde{\mu}|_B(b')$ for some $b \in B_+$ and $b' \in B_-$. Choose t such that $\hat{\mu}(t,b') > 0$, and thus $\hat{\mu}(t,b) < 1$. Increasing slightly $\hat{\mu}(t,b)$ at the expense of $\hat{\mu}(t,b')$ improves upon $\hat{\mu}$, leading to a contradiction. ■

Lemma 14 *Let σ_1 be a strategy of player 1 in $\Gamma^k(t)$, and let $\tilde{\lambda}_k \in \Delta(A^k)$ be arbitrary. Then there exists a strategy $\hat{\sigma}_1$ such that*

$$\mathbf{P}_{\lambda,\hat{\sigma}_1}|_{A^k} = \tilde{\lambda}_k \text{ and } \|\mathbf{P}_{\lambda,\hat{\sigma}_1} - \mathbf{P}_{\lambda,\sigma_1}\|_1 \leq C_k \|\mathbf{P}_{\lambda,\sigma_1} - \tilde{\lambda}_k\|_1,$$

for some constant C_k.

In this statement, $\mathbf{P}_{\lambda,\sigma_1}$ is the distribution of the sequence of player 1's actions induced by σ_1. For notational simplicity, we abstract from the actions of player 2 and 3, whose strategies are fixed.

Proof. We use induction on k. For $k = 1$, the claim coincides with the conclusion of Lemma 13. Let now a strategy σ_1 in $\Gamma^{k+1}(t)$ and $\tilde{\lambda}_{k+1} \in \Delta(A^{k+1})$ be given, and denote by $\sigma_{1,k}$ and $\tilde{\lambda}_k$ their restrictions to the first k rounds. Using the induction hypothesis, let $\hat{\sigma}_{1,k}$ be a strategy over the first k rounds such that $\mathbf{P}_{\lambda,\sigma_{1,k}}|_{A^k} = \tilde{\lambda}_k$ and $\|\mathbf{P}_{\lambda,\sigma_{1,k}} - \mathbf{P}_{\lambda,\hat{\sigma}_{1,k}}\|_1 \leq C_k \|\mathbf{P}_{\lambda,\sigma_{1,k}}|_{A^k} - \tilde{\lambda}_k\|_1$.

Let $\bar{\sigma}_1$ be the strategy in $\Gamma^{k+1}(t)$ which coincides with $\hat{\sigma}_{1,k}$ in the first k rounds, and with σ_1 in the last one.[37] Fix $\vec{a} \in A^k$. Applying Lemma 13 with $T = S^{k+1}$, $B = A$, $\mu := \mathbf{P}_{\lambda,\bar{\sigma}_1}(\cdot \mid \vec{a})$ and $\tilde{\mu}|_A = \tilde{\lambda}_{k+1}(\cdot \mid \vec{a}) \in \Delta(A)$, we get $\hat{\mu}_{\vec{a}} \in \Delta(T \times B)$ with the distributional properties stated there.

We then define $\hat{\sigma}_1$ in round $k+1$ by setting $\hat{\sigma}_1(\cdot \mid \vec{a}, \vec{s}) := \mu_{\vec{a}}(\cdot \mid \vec{s}) \in \Delta(A)$ for each $\vec{s} \in S^{k+1} = T$. By construction and since $\mathbf{P}_{\lambda,\hat{\sigma}_1}$ coincides on A^k with $\tilde{\lambda}_k$, the marginal of $\mathbf{P}_{\lambda,\hat{\sigma}_1}$ on A^{k+1} is equal to $\tilde{\lambda}_{k+1}$.

In addition, because σ_1 and $\bar{\sigma}_1$ coincide in round $k+1$ and by the induction hypothesis, one has

$$\|\mathbf{P}_{\lambda,\sigma_1} - \mathbf{P}_{\lambda,\bar{\sigma}_1}\|_1 = \|\mathbf{P}_{\lambda,\sigma_1}|_{(S\times A)^k} - \mathbf{P}_{\lambda,\bar{\sigma}_1}|_{(S\times A)^k}\|_1 \leq C_k \|\mathbf{P}_{\lambda,\sigma_1}|_{A^k} - \tilde{\lambda}_k\|_1. \tag{40}$$

On the other hand,

$$\begin{aligned}
\|\mathbf{P}_{\lambda,\hat{\sigma}_1} - \mathbf{P}_{\lambda,\bar{\sigma}_1}\|_1 &= \sum_{\vec{a}} \mathbf{P}_{\lambda,\bar{\sigma}_1}(\vec{a}) \times \sum_{a\in A, \vec{s}\in S^{k+1}} |\mathbf{P}_{\lambda,\hat{\sigma}_1}(a, \vec{s} \mid \vec{a}) - \mathbf{P}_{\lambda,\bar{\sigma}_1}(a, \vec{s} \mid \vec{a})| \\
&= \sum_{\vec{a}} \mathbf{P}_{\lambda,\bar{\sigma}_1}(\vec{a}) \times \|\hat{\mu}_{\vec{a}} - \mu\|_1 \\
&= \sum_{\vec{a}} \mathbf{P}_{\lambda,\bar{\sigma}_1}(\vec{a}) \times \|\mu|_A - \tilde{\lambda}_{k+1}(\cdot \mid \vec{a})\|_1 \\
&= \|\mathbf{P}_{\lambda,\bar{\sigma}_1}|_{A^{k+1}} - \tilde{\lambda}_{k+1}\|_1 \\
&\leq \|\mathbf{P}_{\lambda,\sigma_1}|_{A^{k+1}} - \mathbf{P}_{\lambda,\bar{\sigma}_1}|_{A^{k+1}}\|_1 + \|\mathbf{P}_{\lambda,\sigma_1}|_{A^{k+1}} - \tilde{\lambda}_{k+1}\|_1 \\
&= \|\mathbf{P}_{\lambda,\sigma_1}|_{A^k} - \mathbf{P}_{\lambda,\bar{\sigma}_1}|_{A^k}\|_1 + \|\mathbf{P}_{\lambda,\sigma_1}|_{A^{k+1}} - \tilde{\lambda}_{k+1}\|_1 \\
&\leq (C_k + 1)\|\mathbf{P}_{\lambda,\sigma_1}|_{A^{k+1}} - \tilde{\lambda}_{k+1}\|_1,
\end{aligned}$$

where the first equality is an identity, the second one follows from the definition of $\hat{\sigma}_1$ in round $k+1$, the third one follows since $\hat{\mu}_{\vec{a}}$ satisfies the conclusion of Lemma 13, the first inequality follows from the triangle inequality, and the rest follows from the same lines as (40).

---

[37] In particular, the distribution of the first k actions of player 1 is $\tilde{\lambda}_k$.

Adding the last inequality and (40) yields the result, with $C_{k+1} = 2C_k + 1$. ■

We proceed to the proof that $\sup_{L,M} S^{k,L,M} < 1$. Assume to the contrary that for each $\varepsilon > 0$, there is $L_\varepsilon, M_\varepsilon$, t_ε, and an equilibrium $\sigma_\varepsilon \in E^{k,L_\varepsilon,M}(t_\varepsilon)$ such that $\sum_i v^i(\sigma_\varepsilon) \geq 1 - \varepsilon$. Since total payoffs equal to 1 can only be obtained when player 1 reports truthfully, this implies that $\mathbf{P}_{\lambda,\sigma_\varepsilon}$ converges to the distribution on $(S \times A)^k$ induced by truth-telling, as $\varepsilon \to 0$. This in turn implies that the equilibrium payoff vector $(v^i(\sigma_\varepsilon))_i$ (net of transfers) converges to $\mu^{tt} \cdot r$.

Now, fix a stationary reporting policy σ_* and an invariant distribution ν_* implementing μ_{x_k}. The reporting policy σ_* can be viewed as a strategy σ_1 of player 1 in $\Gamma^k(t_\varepsilon)$, when letting player 1 draw in round 1 a fictitious past according to $\nu_*(\cdot \mid s_1)$. In particular, the expected payoff vector (net of transfers) induced by $(\sigma_1, \sigma_{-1,\varepsilon})$ is equal to $\mu_{x_k} \cdot r$.

Next, we apply Lemma 14 with σ_1 and $\tilde{\lambda} := \mathbf{P}_{\lambda,\sigma_\varepsilon}|_{A^k}$ to get a new strategy $\hat{\sigma}_{1,\varepsilon}$. Since $\mathbf{P}_{\lambda,\hat{\sigma}_{1,\varepsilon}}|_{A^k} = \mathbf{P}_{\lambda,\sigma_{1,\varepsilon}}|_{A^k}$, the expected transfers to player 1 do not change when deviating from $\sigma_{1,\varepsilon}$ to $\hat{\sigma}_{1,\varepsilon}$. Since $\|\mathbf{P}_{\lambda,\sigma_1} - \mathbf{P}_{\lambda,\hat{\sigma}_{1,\varepsilon}}\|_1 \leq C_k \|\mathbf{P}_{\lambda,\sigma_{1,\varepsilon}}|_{A^k} - \lambda\|_1$ converges to zero as $\varepsilon \to 0$, expected payoffs net of transfers under $(\hat{\sigma}_{1,\varepsilon}, \sigma_{-1,\varepsilon})$ converge to $\mu_{x_k} \cdot r$.

Since $\mu_{x_k} \cdot r^1 > \mu^{tt} \cdot r^1$, it follows that $\hat{\sigma}_{1,\varepsilon}$ is a profitable deviation upon $\sigma_{1,\varepsilon}$ for ε small enough. This contradiction proves $\sup_{L,M} S^{k,L,M} < 1$.