On Reversible Markov Chains and Maximization of Directed Information

Siva K. Gorantla and Todd P. Coleman
ECE Department
Coordinated Science Laboratory
University of Illinois
{sgorant2,coleman}@illinois.edu

Abstract—In this paper, we consider a dynamical system, whose state is an input to a memoryless channel. The state of the dynamical system is affected by its past, an exogenous input, and causal feedback from the channel’s output. We consider maximizing the directed information between the input signal and the channel output, over all exogenous input distributions and/or dynamical system policies. We demonstrate that under certain conditions, reversibility of a Markov chain implies directed information is maximized. With this, we develop achievability theorems for channels with (infinite) memory as well as optimality conditions for sequential estimation of Markov processes through dynamical systems with causal feedback. We provide examples, which includes the exponential server timing channel and the trapdoor channel.

I. INTRODUCTION

Time reversibility plays an important role in disciplines concerning dynamical systems, e.g. in physics (conservation laws); statistical mechanics (in terms of equilibrium states); stochastic processes (e.g. queuing networks [1] and convergence rates of Markov chains [2, ch 20]); and biology (e.g. trans paths in ion channels [3]). However, its use in providing information-theoretic fundamental limits appears to be somewhat limited 1. In queuing systems, the celebrated Burke’s theorem [1], [6] uses time reversibility to show that, in a certain stochastic dynamical system - an M/M/1 queue in steady-state - the state of the system (queue) at time t is independent of all outputs (departures) before time t. This observation has been used in proving achievability theorems for queuing timing channels [7], [8],[9]. In feedback information theory, statistical independence of one random variable at time t from others up to and including time t plays an important role - for example, tightness conditions in the converse to the channel coding with feedback 2. This begs the question: what might be the role of time reversibility in characterizing information-theoretic fundamental limits of stochastic systems with dynamics? Or in other words, what are the conditions when time reversibility is sufficient in characterizing fundamental limits through via implications of statistical independence? Recent developments at the intersection of information theory and control have demonstrated how directed information [12], [13], [14], [15] characterizes fundamental limitations of stochastic systems with dynamics. Here, we consider a class of stochastic systems for which time reversibility implies saturation of fundamental limits that are characterized by directed information.

A. Problem Setup

Consider a dynamical system with exogenous inputs \( Z = (Z_i \in \mathbb{Z} : i \geq 0) \) where the state \( X_i \in \mathcal{X} \) is an input to a discrete memoryless channel (DMC) with output \( Y_i \in \mathcal{Y} \):

\[
P_{Y_i|X_i,Z_i}(y_i|x_i,z_i) = P_{Y_i|X_i}(y_i|x_i), \quad \forall i \geq 0. \tag{1}
\]

The policy \( \mu = (\mu_i : \mathcal{X} \times \mathcal{Z}^i \rightarrow \mathcal{Y}, \ i \geq 1) \) of the dynamical system governs how the state is affected by past exogenous inputs, and channel outputs:

\[
x_1 \equiv z_0; \quad x_i = \mu_i(z_{i-1}, y_{i-1}), \quad i \geq 1 \tag{2}
\]

The statistics of all processes are governed by \( P_Z, P_{Y_i|X_i} \), and \( \mu \) - so we denote the system as \( (P_Z, \mu, P_{Y_i|X_i}) \). See Figure 1. Examples of such systems in this setup could include:

![Figure 1](image.png)

- A queuing system, where \( X_i \) is the number of customers at time \( i \) in the queue, \( Z \) is the arrival process, \( Y \) is the departure process, and \( \mu \) encodes the update laws of a queuing system’s dynamics. Note that in the context of communication without encoder feedback, feedback appears because this is a channel with memory; it is the objective to \( P_Z \) (Section IV). Or, given a fixed Poisson process \( P_Z \) for the arrival process, the policy \( \mu \) can be optimized, to minimize the difference distortion in timings between the input and output processes and minimize the average queue lengths (Section V).

---

1 although Mitter et al [4], [5] have related Markov chain reversibility to entropy flow and equilibrium states in thermodynamic systems

2 which is part of the “posterior matching principle” for optimal communication with feedback over a DMC [10, Sec III],[11]
• A biological system [16] where a ‘transmitter’ signals by adding molecules into a system. The receiver samples signals from the channel and withdraws a molecule, which in the process modifies the state (e.g. chemical concentration) of the system.

• A brain-machine interface [17], where Z is a Markov process representing a human’s intended trajectory of some smooth path. X_i represents the subsequent ‘intended’ thought the user specifies to the interface. The interface observes a noisy version, Y_i and independent of Z. (Section V).

The directed information
\[ I(Z^n \to Y^n) \]
where (5a) follows from (2); (5b) follows from (1); (5c) follows because conditioning reduces entropy; and (5d) follows from (3). The equalities in (5c) and (5d) hold if and only if:

\[ I(Y_i; Y^{i-1}) = 0 \iff I(X_i; Y^{i-1}) = 0 \] (6)

\[ P_{X_i} \sim P_{X}(\eta, P_{Y_i\mid X}; b) \iff \arg \max_{P_X} I(P_X', P_{Y_i\mid X}) \] (7)

Conditions in (6) are equivalent because of memoryless property of the DMC (1). Thus the directed information of the system \((P_{Z}, \mu, P_{Y_i\mid X})\) with state constraint is maximized if the conditions (6) and (7) are satisfied. (6) implies statistical independence of \(X_i\) with \(Y^{i-1}\) and is analogous to the statement made by Burke’s theorem in Sec I - which was proven using reversible Markov chains. This indicates a possible connection between reversible Markov chains and saturating the fundamental limits of certain stochastic systems.

Section III provides the main theorem - that weds reversibility of Markov chains with maximization of directed information - by providing a sufficient condition for (5d) to be tight in terms of an underlying reversible Markov chain. We then begin to explore what the tightness of (5d) means within the context of the system \((P_{Z}, \mu, P_{Y_i\mid X})\) described in Figure 1, (i) providing the cost-constrained capacity for a class of infinite-memory communication channels with input \(Z\) and output \(Y\), (Section IV); and (ii) providing a sufficient condition for optimality of linear-time-invariant estimator pairs for sequential estimation of Markov processes through dynamical systems with causal feedback (Section V). In both cases, time-reversibility is used to prove optimality. Section VI provides examples, which include the exponential server timing channel (ESTC) [7], [8] and the trapdoor channel [20],[21].

II. DEFINITIONS AND NOTATIONS

- We write \(Z \cup W\) if \(Z\) and \(W\) are independent.

- Define \(\mathcal{S}\) as the set of stationary Markov policies, i.e.

\[ x_i = \mu_i (x_{i-1}, y_{i-1}) = (x_{i-1}, z_{i-1}, y_{i-1}), \quad i \geq 1 \] (8)

- A random process \(W = \{W_i : i \geq 1\}\) is a Markov chain with matrix \(P\) if

\[ P_{W_{k+i+1}|W_k}(w_{k+i+1}|w_k) = P_{W_{k+i+1}|W_k}(w_{k+1}|w_k) \]. If the Markov chain is positive Harris recurrent (PHR), then a unique invariant distribution \(\pi\) exists, which satisfies \(\pi^T P = \pi^T\).

- If \(Z_0^\infty\) are i.i.d., then from (1) and (8), \(X\) is a Markov chain. Suppose \(\mu \in \mathcal{S}\). Define \(\Psi(\mu, P_{Y_i\mid X})\) to be the set of \(P_Z\) s.t. \(\mathcal{Z}_0^\infty\) are i.i.d. \(P_Z\) and independent of \(Z_0\), (b) the subsequent Markov chain \(X\) is PHR, and (c) \(Z_0 = X_1 \sim \pi \iff \pi \in \mathcal{S}(P_Z, \mu, P_{Y_i\mid X})\).

- Denote \(\mathcal{R}(W)\) as the set of reversible Markov chains distributions on \(W\), namely the set of \(P_W\) for which (a) \(W\) is a PHR Markov chain, and (b) the forward and reverse time processes are statistically indistinguishable:

\[ (W_j : 1 \leq j \leq n) \overset{d}{=} (W_{n-j+1} : 1 \leq j \leq n) \] (9a)

\[ \iff \pi_j P_{ij} = \pi_j P_{ji} \quad \forall i, j \in W \] (9b)
III. MAIN THEOREM

In this section, we will show that a policy with a linear, time-invariant Markov structure and i.i.d. exogenous source $P_Z$ results in (5c)-(5d) being tight - provided an underlying Markov chain is reversible.

**Definition 3.1:** Denote $\mathcal{L}$ to be the set of systems $(P_{Z, \mu}, P_{Y|X})$ s.t.:

(a) For some field $\mathbb{F}$, $Z = X = Y = \mathbb{F}$ and $\mu$ satisfies:

\[
x_k = x_{k-1} + a_2 z_{k-1} + a_3 y_{k-1}, \quad (\text{10a})
\]

\[
\bar{x}_k = \bar{x}_{k-1} + a_2 z_{k-1}, \quad (\text{10b})
\]

(b) $P_{Z} \in \Psi(\mu, P_{Y|X})$.

(c) $P_{Z, \bar{X}} \in \mathcal{R}(X \times X)$.

(d) $\pi(P_{Z, \mu}, P_{Y|X}) \sim P_{X}(\eta, P_{Y|X}, b)$

Note that $\bar{x}$ in condition (a) is the update to the state after the output of the channel is taken into consideration and before the source $z$ is updated to the state. Condition (b) means that the source distribution consists of i.i.d. inputs. Condition (c) is the key time reversibility condition on $(X, \bar{X})$, which in many scenarios is equivalent to a reversibility condition on $X$ (e.g. when $\bar{X}$ is a birth-death chain [22]). Lastly, condition (d) simply states that each $X_i \sim P_X(\eta, P_{Y|X}, b)$. With this definition, we have the following theorem:

**Theorem 3.2:** If $(P_{Z, \mu}, P_{Y|X}) \in \mathcal{L}$, then $I_n(P_{Z, \mu}, P_{Y|X}) = C(\eta, P_{Y|X}, b)$.

**Proof:** Following [6], [22], note from (10c) that:

\[
y_k = a_3^{-1}(\bar{x}_k - x_k), \quad (\text{11})
\]

\[
z_{k-1} = a_2^{-1}(x_k - \bar{x}_{k-1}), \quad (\text{12})
\]

As such, we have

\[
I(\bar{X}_i; Y^{i-1}) = I(\bar{X}_i; (\bar{x}_k - x_k); 1 \leq k \leq i) \quad (\text{13})
\]

\[
= I(\bar{X}_i; (\bar{x}_{k-1} - x_k); i + 1 \leq k \leq 2i) \quad (\text{14})
\]

\[
= I(\bar{X}_i; Z_i^{2i-1}) \quad (\text{15})
\]

\[
= 0, \quad (\text{16})
\]

where (13) follows from (11); (14) from Defn 3.1(d) and (9a); and (15) from (12); and (16) from Defn 3.1 (c). Thus $\bar{X}_i \perp Y^{i-1}$. From Defn 3.1 (c) and (10c), this implies that $X_i \perp Y^{i-1}$. Also since $Z_i \sim P_X(\eta, P_{Y|X}, b)$ from Defn 3.1(d), equations (6) and (7) are satisfied and thus $I_n(P_{Z, \mu}, P_{Y|X}) = C(\eta, P_{Y|X}, b)$.

IV. FIXING $\mu$ AND OPTIMIZING $P_Z$

In this section, consider fixing $\mu = \mu^*$ of structure (10) and optimizing over $P_Z$ to maximize $I_n(P_{Z, \mu}, P_{Y|X})$, as in Figure 2. We demonstrate here that not only is the upper bound $I_n(P_{Z, \mu}, P_{Y|X}) \leq C(\eta, P_{Y|X}, b)$ tight when reversibility holds, but reliable communication from $Z$ to $Y$ - a channel with infinite memory - up to capacity $C(\eta, P_{Y|X}, b)$ is possible. The proof again uses the fact that reversibility implies statistical independence to invoke a SLLN for Markov chains.

**Lemma 4.1:** Consider communication over the infinite memory channel with $\mu$ fixed, given by (10). If there exists a
output of the noisy channel, to help guide the estimator \( d = (d_i : Y^i \rightarrow Z, \ i \geq 1) \). So we have two costs that we have to attempt to make low: (1) the cost of bad reproductions, represented with a time-invariant distortion measure given by \( \rho : Z \times Z \times Z \rightarrow \mathbb{R}_+ \), and (2) the state cost function that relates this to control and stability, measured by \( \eta(x) \). So in short, our expected cost is given by

\[
\rho_n(z^n, \tilde{z}^n) = \frac{1}{n} \sum_{i=1}^{n} \rho(z_i, \hat{z}_i, \tilde{z}_{i-1}) + \alpha \eta(x_i). \tag{22}
\]

where \( \alpha > 0 \) balances the importance of control of \( X \) and estimation of \( Z \). We must optimize the (possibly non-linear, time-varying) policy, decoder pair:

\[
(d_n^*, \mu_n^*) = \arg \min_{d \in \mathcal{G}, \mu_n} \mathbb{E}_{Z^n}[\rho_n(Z^n, \hat{Z}_n^1)] \tag{23}
\]

The sequential rate distortion function is [12, Ch. 5]:

\[
\hat{R}_n(D) \triangleq \inf_{P_n \in \mathcal{P}_n(D)} \frac{1}{n} I(Z^n \rightarrow \hat{Z}^n) \tag{25}
\]

\[
\bar{R}_n(D) \triangleq \inf_{P_n} \mathbb{E}[\rho_n(Z^n, \hat{Z}_n^1)] \leq D, \tag{26}
\]

\[
P_{\hat{Z}_n^1|Z^n} = P_{\hat{Z}_n^{i-1} | Z^{i-1}} \quad \forall i \in \{1, \ldots, n\} \tag{27}
\]

(27) encodes the physical constraints of causal encoding/decoding imposed on each \( d_i \). It is already known (c.f. [12, Lemma 5.4.2]) the minimizer of \( \hat{R}_n(D) \) will have an added benefit: it will satisfy \( P_{\hat{Z}_n^1|Z^n} = P_{\hat{Z}_n^{i-1} | Z^{i-1}} \) for any Markov process \( Z \). This hints at a possible time-invariant structure for the policy, estimator pair.

For any \( \mu \) and \( d \) s.t. \( P_X \) satisfies the stability constraint \( \mathbb{E}[\eta(X)] \leq b \) and \( P_{\hat{Z}_n^1|Z^n} \in \mathcal{P}_n(D) \), then:

\[
\hat{R}_n(D) \leq \frac{1}{n} I(Z^n \rightarrow \hat{Z}^n) \leq \frac{1}{n} I(Z^n \rightarrow Y^n) \leq C(\eta, P_{Y|X}, b) \tag{28a}
\]

(28a) follows from the data processing inequality, and (28b) follows from (5d) and the definition of \( C(\eta, P_{Y|X}, b) \). Note that this holds with equality if and only if: (a) \( d_i \) is information-lossless; (b) \( X_i \perp Y^{i-1} \) for all \( i \), and (c) \( X_i \sim P_X(\eta, P_{Y|X}, b) \).

A policy-estimator pair \((\mu, d)\) is LTI if for some \( \tilde{c}_3 \neq 0 \):

\[
x_i = x_{i-1} + (z_i - \hat{z}_{i-1}) - (\hat{z}_{i-1} - \hat{z}_{i-2}) \tag{30a}
\]

\[
x_{i-1} + c_2 \hat{z}_i - c_3 y_{i-1} \tag{30b}
\]

and so (29a) is equivalent to (10) operating on the i.i.d. process \( \hat{Z} \) - so this machinery is equivalent to the stationary Markov policy setup in Section III.

**Definition 5.1:** We say an LTI policy-estimator pair \((\mu^*, d^*)\) are optimal if they minimize (22) for some \( \alpha > 0 \).

In other words, this definition means no other policy-estimator pair \((\mu, d)\) could result in smaller distortion without increasing expected state cost.

**Theorem 5.2:** Consider a Markov source \( Z \) defined as (21), DMC \( P_{Y|X} \), fixed state cost function \( \eta \) and sequence of distortion functions \( \rho \). Then under the LTI policy-estimator pair \((\mu^*, d^*)\) is optimal if for some \( \alpha_1, \alpha_2 > 0 \) and \( \rho' : X \rightarrow \mathbb{R} \):

(a) the \( Y^n \) are i.i.d.

(b) \( \eta(x) = \alpha_1 D(P_{Y|X=x} \| P_Y) + \eta_0 \)

(c) \( \rho(z, \hat{z}, z') = -\alpha_2 \log P_{Y|X} \left( \frac{\hat{z} - z'}{c_3} \right) (z - z') + \rho'(z) \)

**Proof:** Note that from (29), \( \hat{Z}^n \) and \( Y^n \) form a bijection. By virtue of (29) and the DMC (1), it follows that

\[
P_{\hat{Z}_n^i|Z_n, Z_{n-1}} (\hat{z}_i | \hat{Z}_i^{i-1}, z^n) = P_{\hat{Z}_n^i|Z_n, Z_{n-1}} (\hat{z}_i | \hat{Z}_i^{i-1}, z^{i-1})(31)
\]

\[
= P_{Y|x} \left( \frac{\hat{z}_i - z'}{c_3} \right) (z - z') \tag{32}
\]

where (31) follows from (27), which holds for any Markov source with causal encoder [12]; (32) follows from (29)(1). Note as a consequence that

\[
\frac{1}{n} \sum_{i=1}^{n} \rho(z_i, \hat{z}_i, \tilde{z}_{i-1}) = -\alpha_2 \log P_{Z^n, \hat{Z}_n} (z^n, \hat{Z}_n) + \rho'(z^n) \tag{33}
\]

\[
\frac{1}{n} \sum_{i=1}^{n} \eta(x_i) = \alpha_0 D(P_{Y_n|X=x} \| P_Y) + \eta_0 \tag{34}
\]

(33) where the latter holds because \( Y^n \) are i.i.d. Thus, since \( Y^n \) are i.i.d., for this setting, \( \hat{R}_n(D) = C(\eta, P_{Y|X}, b) \). The rest of the proof follows directly from [18].

This leads to a corollary linking reversibility and optimality:

**Corollary 5.3:** If an LTI policy-estimator pair leads to \( Z \in \psi(\mu, P_{Y|X}) \), then it is optimal with respect to \( \eta \) and \( \rho \) specified by Thm 5.2 conditions (b)-(c).

So in these algebraic settings, the distortion measures that are sufficient for optimality are time-invariant.

**VI. EXAMPLES AND DISCUSSION**

**A. Gaussian Channels**

Consider the problem in Section V with \( \hat{Z}_i \) are i.i.d. Gaussian, \( \eta(x) = x^2 \), an additive Gaussian channel, and squared error distortion. We recover the optimality of using an LTI state-estimator pair, with \( \hat{c}_3 \) playing the role of the sequential MMSE estimator [25],[12, Chap 6] by invoking Theorem 5.2 directly.
B. Queuing Timing Channels

Let $\mathcal{F} = \mathcal{Q}$, $\eta(x) = +\infty 1_{(x \not\in \mathbb{Q})}$. Consider the policy $\mu$ given by (10) with $c_2 = 1, c_3 = 1$ and the DMC with outputs $y \in \{0, 1\}$ specified by the $Z$ channel:

$$P_{Y|X}(1|x) = \begin{cases} 0 & x = 0 \\ \mu \Delta & x > 0 \end{cases}$$  \hspace{1cm} (35)$$

where $\Delta \ll 1$. This represents the ESTC of rate $\mu$: $Y_i = 1$ when a departure occurs, and $Z_i$ represents the number of arrivals in the $i$th length-$\Delta$ time slot, and service times are exponentially distributed (see [7, Fig 4]). With $Z^\infty$ i.i.d. with $P(Z_i = 1) = p$, the three-state Markov chain is birth-death with steady-state distribution $\pi = [p^2, 2p(1-p), (1-p)^2]$ and $I(\pi, P_{Y|X}) = h_0(\pi_0 + \frac{\mu}{\Delta}) - \pi_1$. Thm 3.2 and reversibility imply that $R < I(\pi, P_{Y|X})|_{p=1} = 0.5$ is achievable - coinciding with [20].

\section{VII. Acknowledgements}

The authors thank T. Basar, T. Bretl, B. Hajek, N. Kiyavash, M. Raginsky, and R. Srikant for useful discussion.

REFERENCES


\section{Introduction}

The Chemical (Trapdoor) Channel

\section{C. The Chemical (Trapdoor) Channel}

Let $\mathcal{F} = \mathcal{Q}$, $\eta(x) = +\infty 1_{(x \not\in \mathbb{Q})}$. Consider the policy $\mu$ given by (10) with $c_2 = 1, c_3 = 1$ and the DMC with outputs $y \in \{0, 1\}$ specified by the ‘inverted’ $Z$ channel:

$$P_{Y|X}(1|x) = \begin{cases} 0 & x = 0 \\ \frac{1}{2} & x = 1 \\ 1 & x = 2 \end{cases}$$  \hspace{1cm} (36)$$

Here, the ‘chemical channel’ [16], [20], [21] has a state $X \in \{0, 1, 2\}$ represents the number of red balls in a system that can contain two balls, each of which is either red or blue. At each time, one of the two trapdoors is opened randomly, with $Y_i = 1$ if a red ball departs and 0 otherwise. Each entering ball $Z_i$ is either red (1) or blue (0). With $Z^\infty$ i.i.d. with $P(Z_i = 1) = p$, the three-state Markov chain is birth-death with steady-state distribution $\pi = [p^2, 2p(1-p), (1-p)^2]$ and $I(\pi, P_{Y|X}) = h_0(\pi_0 + \frac{\mu}{\Delta}) - \pi_1$. Thm 3.2 and reversibility imply that $R < I(\pi, P_{Y|X})|_{p=1} = 0.5$ is achievable - coinciding with [20].

Also, if we let the state cost function $\eta(x) = 1$ for all non-negative integers and $\infty$ otherwise, then any $b \in \{0, 1\}$, places a constraint on the queue utilization - given by $\frac{1}{\mu}$. This equates to a constraint on the rate of the input process $Z$ and recovers the rate-constrained capacity $C(\lambda, \mu) = \lambda \log \frac{1}{\mu}$.