TECHNISCHE
UNIVERSITÄT
DARMSTADT

The Asymptotic Behaviour of Moments of the Causal Estimator for the Binary Input in a Poisson Channel

A report submitted as a part of Virtual Internship 2020 at
Bio- Inspired Communication Systems Lab
TECHNISCHE UNIVERSITAT DARMSTADT

Aabesh Bhattacharya
Abhinandan Dalal
Navonil Deb

Instructor: Prof. Heinz Koeppl and Mark Sinzger
Programme : Virtual Internship Summer 2020
Institution: TU Darmstadt
Date: November, 2020

Contents

1 Introduction

The Poisson Channel is a model for detection of events such as mRNA synthesis and analyze the consequent minimal gene expression. With a stochastic intensity $X(t)$ as input of the process, the output is denoted by $Y(t)$. For instance, in mRNA synthesis the stochastic switching between the active state (say x_1) and the inactive state (x_2) is modeled by $X(t)$, a binary, time homogeneous and stationary stochastic process (BMP), and $Y(t)$ denotes the mRNA synthesis rate which is a Poisson random variable depending on $X(t)$. By $X_{[0,T]}$ we denote the trajectory $\{X(t)\}_{0\leq t\leq T}$ of the time varying input signal with transmission duration T. The sojourn times σ_1 in x_1 and σ_2 in x_2 are exponentially distributed with parameters c_1 and c_2 such that $\mathbb{E}[\sigma_1] = c_1$ and $\mathbb{E}[\sigma_2] = c_2$. The output of the channel is a Poisson process with intensity parameter $c_3X(t)$, where c_3 is a constant also known as *channel gain*. We consider the mutual information (MI) rate which is defined as, $\bar{\mathbb{I}}(X,Y) := \lim_{T\to\infty} T^{-1}\mathbb{I}(X_{[0,T]}, Y_{[0,T]})$, where $\mathbb{I}(X_{[0,T]}, Y_{[0,T]})$ is the path-wise mutual information up to time T.

The computation of mutual information (MI) between time-varying, bio-molecular signals is complex, and analytical solutions so far relied on Gaussian or steady-state approximations. Exact computation of the MI along the Poisson Channel has been done in very few literature including [1]. This path-wise MI is essential to determine the information rate capacity of the Poisson Channel. The information rate capacity can be obtained as the constrained supremum (over c_1 and c_2) of a function of MI. In our work, we present some properties if the solution(s) of the optimization problem and characterize its higher order moments.

This report is organised as follows. Section 2 focuses some background review done in this field, followed by a glance to some useful mathematical machinery used in this problem in Section 3. Section 4 discusses the problem that we wish to tackle in this report. Section 5 introduces the novel idea of using linear program as a computational technique to approximate the moments, and studies the properties of the bounds created and the solutions we obtain. Section 6 attempts to generalize the framework to a general binary input $X(t)$ taking values x_1 and x_2. Section 7 validates some of the results we have proved using simulations and computations. We finish the report by a discussion in the Epilogue section at the end.

2 Background Review

Let $X(t)$ be a binary, time-homogeneous, stationary Markov process (BMP) switching stochastically between inactivated state x_1 and activated state x_2 and the sojourn (waiting) times σ_1 in x_1 and σ_2 in x_2 are exponentially distributed with inverse mean parameters c_1 and c_2 respectively. Also let $Y(t)$ be an inhomogeneous poison process with intensity $c_3X(t)$. $X_{[0,T]}$ denotes the trajectory $X(t)_{0\leq t\leq T}$ with transmission duration T. The switching times of X are denoted by s_i with $0 < s_1 < \ldots < T$ and the jump times of Y are denoted by t_i with $0 < t_1 < \ldots < T$. The information rate is defined as

$$\bar{\mathbb{I}}(X,Y) := \lim_{T\to\infty} T^{-1}\mathbb{I}(X_{[0,T]}, Y_{[0,T]})$$

where $\mathbb{I}(X_{[0,T]}, Y_{[0,T]})$ is the path-wise Mutual Information. The capacity C_T and the information rate capacity C are defined as follows:

$$C_T = \sup_{c_1, c_2} T^{-1} \mathbb{I}(X_{[0,T]}, Y_{[0,T]}), \quad C = \sup_{c_1, c_2} \bar{\mathbb{I}}(X, Y).$$

The conditional mean $Z(t) = \mathbb{E}[X(t)|Y_{[0,t]}]$ is the optimal causal estimator of the input given the output under a quadratic criterion. The following expression links the path-wise MI with $Z(t)$.

$$\mathbb{I}(X_{[0,T]}, Y_{[0,T]}) = \int_0^T \mathbb{E}[\phi(c_3 X(t))] - \mathbb{E}[\phi(c_3 Z(t))] dt.$$

The conditional mean $Z(t)$ evolves according to the following filtering equation, an ODE with stochastic jumps

$$dZ(t) = A((Z(t) - x_1))dt + g(Z(t-))dY(t) \tag{1}$$

where $A(z) := c_3(c_1 \Delta x / c_3 - \gamma z + z^2)$ with $\gamma := (c_1 + c_2 + \Delta x c_3)/c_3$, $dY(t) = 1$ for jump times $t = t_i$, vanishing otherwise. And $g(Z(t_i)) = \frac{(Z(t_i) - x_1)(x_2 - Z(t_i))}{Z(t_i)}$ is the jump height $dZ(t_i) = Z(t_i+) - Z(t_i-)$ at jump times of Y. We shall denote $Z(\infty)$ as Z_∞ interchangeably with Z.

2.1 Trajectory of Z

For rest of this report, until we reach the generalization framework, we will assume $x_1 = 0, x_2 = 1$. Under this assumption $Y(t)$ increases solely if $X(t) = x_2 = 1$. Hence, $Z(t)$ is reset to $x_2 = 1$ upon jumps of Y. The joint process $\{Z(t), Y(t)\}_t$ is then a piece-wise deterministic Markov process that jumps stochastically from state $\{Z(t_i-), Y(t_i-)\}$ to state $\{Z(t_i+), Y(t_i+)\} = (1, Y(t_i-) + 1)$. Hence, jump times t_i of Y and of Z are identical, and jumps occur with propensity $c_3 Z(t)$. Since the propensity depends only on the first component, the projection onto $Z(t)$ is a piece-wise deterministic Markov process itself. Its probability evolution equation is given by a hybrid generator, composed of the drift (Liouville) and the jump (Poisson) part

$$\frac{\partial}{\partial t} p(t, z) = -\frac{\partial}{\partial z}\{A(z)p(t, z)\} - c_3 z p(t, z), \quad \omega < z < 1, \tag{2}$$

where, stable equilibrium $\omega := \frac{\gamma - \rho}{2}$ with $\rho := \sqrt{\gamma^2 - 4c_1 \Delta x / c_3}$. Since all trajectories $Z(t)$ start in the stationary mean $\mathbb{E}[Z(0)] = \mathbb{E}[X(0)] = \frac{c_1}{c_1 + c_2} = m$, the probability distribution of $Z(0)$ is a point mass. In order to track its progression, we switch to the trajectory-wise perspective 1. Solving the deterministic part of1,

$$\frac{d}{dt} f_a(t) = A(f_a(t)), \quad f_a(0) = a,$$

yields,

$$f_a(t) = \omega + \rho(1 + e^{\rho c_3 t}(\rho/(a - \omega) - 1))^{-1}.$$

Therefore we get,

$$Z(t) = f_m(t) \mathbb{1}_{[0, t_1)}(t) + \sum_i f_1(t - t_i) \mathbb{1}_{[t_i, t_{i+1})}(t).$$

Remark. $t < t_1 \iff Z(t) \le f_1(t) \iff Z(t) = f_m(t).$

2.2 Distribution of Z

Theorem 2.1. *The probability measure $\mu_t : \mathcal{B}(\omega, 1] \to [0,1]$, $\mu_t(B) = P[Z(t) \in B]$, is a hybrid measure*

$$\mu_t(B) = \kappa(t)\delta_{f_m(t)}(B) + \nu_t(B),$$

where,

$$\kappa(t) = \mathbb{P}(t_1 > t) = \exp\Big(- c_3 \int_0^t f_m(s)ds \Big),$$

and ν_t is an absolute continuous measure supported on $(f_1(t), 1]$, with time-independent density π, given by,

$$\pi(z) = \alpha(z-\omega)^{\beta - \frac{3}{2}}(\omega + \rho - z)^{-(\beta + \frac{3}{2})}.$$

Remark. $\nu_t(B) = \mathbb{P}\Big(Z(t) \in B \cap (f_1(t), 1] \Big) = \mathbb{P}\Big(t - \sup\{t_i | t_i \leq t\} \in f_1^{-1}(B), t_1 \leq t \Big)$.

Remark. $p(t,z) = \pi(z)\mathbb{1}_{(f_1(t),1]}(z)$.

Theorem 2.2. *The asymptotic distribution of Z is absolutely continuous, supported on $(\omega, 1]$ with density $\pi(z)$.*

2.3 A Useful Result

The following theorem links the information rate with $\mathbb{E}(\phi(Z))$, where $\mathbb{E}$ is taken over the asymptotic distribution of Z and $\phi(z) = z \ln z$.

Theorem 2.3.

$$\bar{\mathbb{I}}(X,Y) = -c_3 \int_\omega^1 \phi(z)\pi(z)dz.$$

3 Useful mathematical machinery

3.1 Stochastic Integrals

let $M = \{M(t) : t \in [0,\tau]\}$ be a mean zero martingale adapted to the history $\{\mathcal{F}_t\}$. Let $H = \{H(t) : t \in [0,\tau]\}$ be a stochastic process that is predictable. Intuitively this means that for any time t, the value of $H(t)$ is known just before t (possibly apart from unknown parameters). We note that sufficient conditions for H to be predictable are that H is adapted to the history $\{\mathcal{F}_t\}$ and the sample paths of H are left-continuous. We can now introduce the stochastic integral:

$$I(t) = \int_0^t H(s)dM(s).$$

The major interesting fact about a stochastic integral is that $I(t)$ is also a mean zero martingale adapted to the history $\{\mathcal{F}_t\}$. Hence, the martingale property is preserved under stochastic integration.

3.2 Stochastic Integral for Counting Processes

Recall that a counting process $N(t)$ is a right-continuous process with jumps of size 1 at event times and constant in between. Since the counting process is non-decreasing, $N(t) \geq N(s)$ for all $t > s$, and hence,

$$\mathbb{E}(N(t)|\mathcal{F}_s) \geq N(s) \quad \text{for all} \quad t > s.$$

Thus, $N(t)$ is a sub-martingale.

- **Doob-Meyer Decomposition:** Any sub-martingale X can be decomposed uniquely as

  $$X = X^* + M,$$

  where, X^* is a non-decreasing predictable process, often denoted the compensator of X, and M is a mean zero martingale.

Hence by the Doob-Meyer decomposition, there exist a unique non-decreasing predictable process $\Lambda(t)$, called the cumulative intensity process, such that $M(t) = N(t) - \Lambda(t)$ is a mean zero martingale. We will consider the case where the cumulative intensity process is absolutely continuous and the corresponding density will be called the intensity process. Thus in the context of counting processes, the stochastic integral is simple to understand as one can splits the integral in two as follows:

$$I(t) = \int_0^t H(s)dN(s) - \int_0^t H(s)\lambda(s)ds.$$

For given sample paths of the processes, the last integral is simply an ordinary (Riemann) integral. The first integral, however, is to be understood as a sum of the values of H at every jump time of the counting process. Thus

$$\int_0^t H(s)dN(s) = \sum_{T_j \leq t} H(T_j),$$

where $T_1 < T_2 < \ldots$ are the ordered jump times of N.

3.3 Ito Formula for Counting Processes

Let $N(t)$ be a counting processes. Suppose the (multivariate) process $X(t) \in \mathbf{R}^m$ satisfies the following differential equation with stochastic updates

$$dX(t) = a(X(t))dt + b(X(t-))dN(t).$$

Then $Y(t) = F(X(t))$, for some function $F : \mathbf{R}^m \to \mathbf{R}^n$, satisfies the following equation

$$dY(t) = DF(X(t))a(X(t))dt + [F(X(t-) + b(X(t-))) - F(X(t-))]dN(t).$$

The first part is the same as in the ordinary chain rule with the Jacobian $DF \in \mathbf{R}^{n \times m}$. The second part is the jump vector of Y upon jumps of $N(t)$.

4 Proceeding in our situation

4.1 Backdrop

The conditional mean $Z(t) = \mathbb{E}[X(t) \mid Y_{[0,T]}]$ is described in [1] as the optimal causal estimator of the input given the output under a quadratic criterion. In [1], its trajectory-wise evolution is described in equation 3 and the evolution of its probability distribution is described for $x_1 = 0$ in equation 4 and theorem II.1. As $t \to \infty$, the process $Z(t)$ converges in distribution to a random variable Z_∞ in the following. Its distribution for $x_1 = 0$, $x_2 = 1$ is given by equation 7 of [1] The mutual information rate is given as in theorem III.1 by

$$\bar{\mathbb{I}} = \bar{I}(c_1, c_2) = -c_3 \int_\omega^1 \phi(z)\pi(z)\, dz$$

4.2 Focus on moments

We now want to point out that it is favorable to express higher order moments of Z_∞. Since all derivatives of the function $-\phi$ are monotone, decreasing for odd and increasing for even order, we get inequalities

$$-\phi(z) \geq \sum_{k=0}^{2n} -\frac{\phi^{(k)}(\omega)}{k!}(z-\omega)^k \qquad \forall z > \omega \tag{3}$$

and

$$-\phi(z) \leq \sum_{k=0}^{2n-1} -\frac{\phi^{(k)}(\omega)}{k!}(z-\omega)^k \qquad \forall z > \omega \tag{4}$$

If the moments were available, the expectations of the right hand sides could be computed.

By the Ito formula for counting processes, the powers of Z satisfy the stochastic differential equations

$$dZ(t)^n = nZ(t)^{n-1}(c_1 - (c_1+c_2+c_3)Z(t) + c_3 Z(t)^2)\, dt + (1 - Z(t)^n)\, dY(t)$$

By the Doob-Meyer decomposition $dY(t) = dQ(t) + c_3 Z(t)dt$, with $Q(t)$ a zero-mean martingale, we get for the expectation

$$\frac{d}{dt}\mathbb{E}[Z(t)^n] = nc_1\mathbb{E}[Z(t)^{n-1}] - n(c_1+c_2+c_3)\mathbb{E}[Z(t)^n] + (n-1)c_3\mathbb{E}[Z(t)^{n+1}] + c_3\mathbb{E}[Z(t)] \tag{5}$$

so at stationarity the following recursion for $n \geq 2$ holds:

$$\mathbb{E}[Z_\infty^{n+1}] = -\frac{c_1}{(n-1)(c_1+c_2)} + \frac{n}{n-1}\left(\frac{c_1+c_2}{c_3} + 1\right)\mathbb{E}[Z_\infty^n] - \frac{n}{n-1}\frac{c_1}{c_3}\mathbb{E}[Z_\infty^{n-1}] \tag{6}$$

Lemma 4.1. $\lim_{n\to\infty} \mathbb{E}[Z_\infty^n]$ *exists, and equals 0.*

Proof: Since all Z_∞ is bounded by 0 and 1, we have, $\{Z_\infty^n\}_{n\geq 1}$ to be a monotonically non-increasing sequence of random variables. Thus, $\{\mathbb{E}[Z_\infty^n]\}_{n\geq 1}$ also forms a non-increasing sequence. Moreover, this sequence is bounded above by 1 and below by 0. Thus, this sequence $\{E[Z_\infty^n]\}_{n\geq 1}$ has a finite limit. Let this limit be L. Now, using Equation (6), we have,

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[Z_\infty^{n+1}] = -\lim_{n\to\infty} \frac{c_1}{(n-1)(c_1+c_2)} + \lim_{n\to\infty} \frac{n}{n-1}\left(\frac{c_1+c_2}{c_3}+1\right) \lim_{n\to\infty} \mathbb{E}[Z_\infty^n] \\
- \lim_{n\to\infty} \frac{n}{n-1}\frac{c_1}{c_3} \lim_{n\to\infty} \mathbb{E}[Z_\infty^{n-1}]
\end{aligned}
$$

$$
\begin{aligned}
\implies L &= \left(\frac{c_1+c_2}{c_3}+1\right) L - \frac{c_1}{c_3} L \\
&= \left(\frac{c_1}{c_3}+1\right) L \\
\implies L &= 0 \qquad\qquad \because c_1, c_3 > 0
\end{aligned}
$$

■

Conjecture 4.2. $\mathbb{E}[Z_\infty^n] \to 0$ *as* $n \to \infty$ *uniquely determines all moments.*

However, they cannot be solved for easily.

5 The Linear Program Approach

Since the moments cannot be derived directly, we resort to an approach of approximating them. Note that, $\mathbb{P}(X(t) = x_2) = \frac{c_1}{c_1+c_2}$. Therefore, for all $t \geq 0$ we have

$$
\begin{aligned}
\mathbb{E}[Z(t)] = \mathbb{E}\left[\mathbb{E}[X(t) \mid Y_{[0,T]}]\right] = \mathbb{E}[X(t)] = x_1 \cdot \frac{c_2}{c_1+c_2} + x_2 \cdot \frac{c_1}{c_1+c_2} = \frac{c_1x_2+c_2x_1}{c_1+c_2} \\
\implies \mathbb{E}[Z_\infty] = \lim_{t\to\infty} \mathbb{E}[Z(t)] = \frac{c_1x_2+c_2x_1}{c_1+c_2}
\end{aligned}
$$

Thus, for $x_1 = 0,\ x_2 = 1$, $\mathbb{E}[Z_\infty] = \dfrac{c_1}{c_1+c_2}$, and from (6), it is clear that knowing the first two moments is enough to determine all the moments. For simplicity of notation, we write the sequence as $y_n := \mathbb{E}[Z_\infty^n]$. We know the value of y_1, but not y_2. Now, equation (6), if taken upto N equations, gives us a set of equations. We thus formulate the linear program using the constraints:

$$
\begin{aligned}
y_{n+1} &= -\frac{c_1}{(n-1)(c_1+c_2)} + \frac{n}{n-1}\left(\frac{c_1+c_2}{c_3}+1\right) y_n - \frac{n}{n-1}\frac{c_1}{c_3} y_{n-1} \qquad n = 2(1)N \\
y_1 &= \frac{c_1}{c_1+c_2} \\
y_1 &\geq y_2 \geq \cdots \geq y_N \geq y_{N+1} \\
y_{N+1} &\geq 0
\end{aligned} \tag{7}
$$

Clearly these constraints are linear in y_i's. We need to add the monotonicity constraint, as the recursion equation itself does not guarantee that. Now, subject to these constraints in program (7), we solve two linear problems

$$
\max\ y_2 \qquad \text{and} \qquad \min\ y_2
$$

Solving this problem, we get a maximum and minimum value of y_2 depending on the number of recursion equations we decide to use (that is upto moment $N+1$ here), and hence, we can get a handle over what we want to achieve. The following results help us in this goal as well.

Theorem 5.1. *With increasing N, the difference between the maximum and minimum value of y_2 so obtained is non-increasing.*

Proof: Consider the problem formulation of program (7) and note that it's feasible set is a convex compact set in $\mathbb{R}^{N+1}$. Hence the maximum and minimum are always finite and achieved in this set. Now, consider the problem as a linear program in $\mathbb{R}^{\mathbb{N}}$. Then, with increasing N, we are restricting the feasible set of the region more and more, as we introduce an additional constraint in the form of the recursion equation, and another constraint, and this reduces the feasible set. Thus,

$$\max y_2|_{N \text{ moments}} \geq \max y_2|_{N+1 \text{ moments}} \text{ and } \min y_2|_{N \text{ moments}} \leq \min y_2|_{N+1 \text{ moments}}$$

and hence the difference between the maximum and the minimum value of y_2 obtained from the linear program is non-increasing. ■

Corollary 5.1.1. *The sequence of the difference between the maximum and minimum value of y_2 in the linear program converges.*

Proof: Since y_2 is bounded between 0 and 1, the maximum and minimum value of the linear program, for any $N \geq 2$, is between 0 and 1. Hence the sequence of differences is bounded, and from the above theorem, is monotonic. Thus the difference between the maximum and minimum value of y_2 converges to a finite limit. ■

Theorem 5.2. *y_n's are non-decreasing with y_2*

Proof: We shall prove this by induction on n. For the purpose, we strengthen the hypothesis as follows:

Induction Statement $Q(n)$: y_n and $y_n - y_{n-1}$ are both non-decreasing functions of y_2.

We are to show that $Q(n)$ is true for all $n \geq 2$.

Note that, equation (6) allows us to conclude that y_n's are differentiable functions of y_2. Now, for the base case, note that y_1 is constant (ie, equals $\frac{c_1}{c_1+c_2}$), and y_2 is obviously increasing in y_2, and thus so does the difference $y_2 - y_1$. Thus $Q(2)$ is true.

Now let us assume $Q(k)$ is true for all $2 \leq k \leq m$. We shall show $Q(m+1)$ is true as well.

We differentiate equation (6) with respect to y_2, with the notation of y_i's we're using.

$$\begin{aligned}\frac{d}{dy_2}y_{m+1} &= \frac{m}{m-1}\left(\frac{c_1+c_2}{c_3}+1\right)\frac{d}{dy_2}y_m - \frac{m}{m-1}\cdot\frac{c_1}{c_3}\frac{d}{dy_2}y_{m-1} \\ &= \frac{m}{m-1}\frac{c_1}{c_3}\frac{d}{dy_2}(y_m - y_{m-1}) + \frac{m}{m-1}\left(\frac{c_2}{c_3}+1\right)\frac{d}{dy_2}y_m \end{aligned} \tag{8}$$

Both the derivatives on the right hand side are non-negative by the right hand side, and hence, the left hand side is non-negative as well.

Now, again from equation (6), we have,

$$y_{m+1} - y_m = -\frac{c_1}{(m-1)(c_1+c_2)} + \frac{m}{m-1}\frac{c_1}{c_3}(y_m - y_{m-1}) + \frac{m}{m-1}\frac{c_2}{c_3}y_m + \frac{1}{m-1}y_m \tag{9}$$

Differentiating the above with respect to y_2 yields,

$$\frac{d}{dy_2}(y_{m+1} - y_m) = \frac{m}{m-1}\frac{c_1}{c_3}\frac{d}{dy_2}(y_m - y_{m-1}) + \frac{1}{m-1}\left(\frac{mc_2}{c_3} + 1\right)\frac{d}{dy_2}y_m \tag{10}$$

Again the right hand side has both the derivatives non-negative by the induction hypothesis, and hence, the left hand side is non-negative as well. Thus, $y_{m+1} - y_m$ is non-decreasing in y_2.

Thus by the strong form of mathematical induction, $Q(n)$ is true for all $n \geq 2$. This completes our proof. ■

Remark: If one follows the proof carefully, one can see that the argument holds true even to prove that y_n's are strictly increasing in y_2. The exact same argument follows, and henceforth we shall assume when required the stronger form that y_n's are strictly increasing in y_2.

Remark: The above theorem shows that the moment curve (ie, the curve obtained by joining the moments when plotted along n) are non-intersecting, and gradually increases, as we raise our y_2 from the minimum to the maximum value as obtained in the linear program. Since, the true moments of Z_∞ also must satisfy the conditions of program (7), the moment curves obtained by using the maximum and minimum values of y_2 from (7) indeed provide a band in which the true moment curve of Z_∞ lies.

Conjecture 5.3. *There is a unique solution to the recursion equation* (6) *that is monotonic and satisfies* $y_1 = \frac{c_1}{c_1+c_2}$

Note that this conjecture is more strong than Conjecture 4.2, since monotonicity of the recursion implies that its limit would exist and would be 0.

Theorem 5.4. *The sequences of differences of the maximum and minimum values of y_2 obtained from the linear program decreases to 0 if and only if Conjecture 5.3 is true.*

Proof: For the sake of simplicity, in this proof and henceforth, we shall use the notation $y_{2,\max,N}$ and $y_{2,\min,N}$ as the corresponding maximum and minimum values of y_2 obtained from 7.

If part:

Let $y_{2,\max,N} \to y_{2,\max}$ and $y_{2,\min,N} \to y_{2,\min}$, which we know exists by virtue of being bounded and monotonic sequences. Suppose if possible $y_{2,\max} - y_{2,\min} = c > 0$. Now, take $0 < \epsilon < c$ and define $a_2 = y_{2,\min} + \frac{\varepsilon}{2} \in (y_{2\min}, y_{2,\min})$, with $a_1 = \frac{c_1}{c_1+c_2}$ and a_n recursively using (6). By the convexity of the problem, the entire sequence $\{a_n\}$ lies in the feasible set of program 7 for any n, and is hence monotonic. Thus, we have obtained a monotonic solution to the recursion equation. However, we can repeat the argument with another sequence b_n whose $b_2 = y_{2,\min} + \frac{\varepsilon}{3}$ and then continued recursively. This means that we can find two distinct sequences which are monotonic and are solutions to the recursion equation, which contradicts Conjecture 5.3. Hence $y_{2,\max,N} - y_{2,\min,N}$ must converge to 0 for 5.3 to be true.

Only if part:

Given $y_{2,\max,N} - y_{2,\min,N} \to 0$ as $N \to \infty$, and suppose if possible that Conjecture 5.3 fails. Then, there are two distinct monotonic sequences $\{u_n\}$ and $\{v_n\}$ that are solutions of the recursion equation. Now, as the value of u_2 and v_2 uniquely determines the entire sequence, we must have $u_2 \neq v_2$. However, if we truncate $\{u_n\}$ and $\{v_n\}$ to $N+1$ terms, then they both satisfy the conditions of linear program 7, and hence, we have, $y_{2,\max,N} \geq \max\{u_2, v_2\} \geq \min\{u_2, v_2\} \geq y_{2,\min,N}$ and thus, $y_{2,\max,N} - y_{2,\min,N} \geq |u_2 - v_2| \geq 0$. However, this holds for any arbitrary N, and thus taking limit, we have a contradiction to the fact that $y_{2,\max,N} - y_{2,\min,N} \to 0$. This proves the only if part of the proof. ∎

The above results show that our linear program approach would indeed be helpful in finding the moments, and provide interesting scopes for future potential of proving the unproved. Now, we shall explore the solutions we obtain from the linear program.

Theorem 5.5. *If $y_2 = y_{2,\max,N}$, then y_n's obtained from the recursion equation in 7 satisfies $y_{N+1} = y_N$ and $y_2 = y_{2,\min,N}$ satisfies $y_{N+1} = 0$.*

Proof:

Claim: y_2 is increasing in y_{N+1}.

Proof of claim: From the recursion in equation 7, y_{N+1} can be expressed as a function of y_1 and y_2. Since y_1 is fixed, y_{N+1} can be expressed in terms of y_2. Let the range of feasible values of y_2 be denoted by $\mathcal{R}(y_2)$ and the range of feasible values of y_{N+1} be denoted by $\mathcal{R}(y_{N+1})$ and the function that expresses y_{N+1} in terms of y_2 be $f.f : \mathcal{R}(y_2) \mapsto \mathcal{R}(y_{N+1})$, $y_{N+1} = f(y_2)$. From Theorem 4.2, we know that y_n is increasing in y_2 $\forall n$. So, in particular y_{N+1} is increasing in y_2 i.e. f is an increasing function and hence f is $1-1$ and onto. So, f is invertible. Take $y \in \mathcal{R}(y_{N+1})$, then $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$. Since the RHS is positive by virtue of f being increasing, LHS is also positive and hence we can say that f^{-1} is increasing as well.This proves our claim.

According to the conditions of the LP in 7, $0 \leq y_{N+1} \leq y_N$. So in case of maximizing y_2 if at the maxima of y_2, y_{N+1} is not equal to y_N, then we can further increase the value of y_{N+1} and that would result in an increase in the value of y_2 as y_2 is increasing in y_{N+1}. Also, it follows that y_2 achieves its minimum $y_{2,\min,N}$ when $y_{N+1} = 0$ which is the minimum possible value of y_{N+1}. ∎

Finding Closed form expression of $y_{2,\max,N}$:

We can use the result in the above theorem to find out $y_{2,\min,N}$ and $y_{2,\max,N}$ explicitly.We know that $y_1 = \frac{c_1}{c_1+c_2}$. Let the vector $\vec{y}_{\max,N}$ denote the vector of values of $y_2, y_3, \ldots, y_{N+1}$ obtained from the LP by maximizing y_2. then $\vec{y}_{\max,N}$ satisfies the following equations:-

$$y_{n+1} = -\frac{c_1}{(n-1)(c_1+c_2)} + \frac{n}{n-1}\left(\frac{c_1+c_2}{c_3} + 1\right) y_n - \frac{n}{n-1}\frac{c_1}{c_3} y_{n-1} \qquad n = 2(1)N$$

$$y_{N+1} = y_N$$

Thus, we have a set of N equations and N variables to solve for and the solution of y_2 that we obtain from the above set of equations in nothing but $y_{2,\max,N}$ as indicated by the previous theorem. Define the matrix

$A_{\max,N} = ((A_{\max,N,ij}))_{N\times N}$ where

$$A_{\max,N,ij} = \begin{cases} \frac{-2(c_1+c_2+c_2)}{c_3} & \text{if } \ i=1, j=1 \\ 1 & \text{if } \ i=1, j=2 \\ \frac{i+1}{i}\frac{c_1}{c_3} & \text{if } \ i=j, 2\le i\le N-1 \\ -\frac{i+1}{i}(\frac{c_1+c_2}{c_3}+1) & \text{if } \ j=i+1, 2\le i\le N-1 \\ 1 & \text{if } \ j=i+2, 2\le i\le N-1 \\ -1 & \text{if } \ i=N.j=N-1 \\ 1 & \text{if } \ i=N.j=N \\ 0 & \text{otherwise} \end{cases}$$

Define $\vec{B}_{\max,N} = (B_{\max,N,1}, B_{\max,N,2}, \ldots, B_{\max,N,N})^T$ where $B_{\max,N,i} = \begin{cases} -\dfrac{c_1(2c_1+c_3)}{(c_1+c_2)c_3} & \text{if } \ i=1 \\ \dfrac{c_1}{i(c_1+c_2)} & \text{if } \ i=2(1)N-1 \\ 0 & \text{if } \ i=N \end{cases}$

Writing down these equations, in the matrix form, we have,

$$A_{max,N}\cdot \vec{y}_{\max,N} = \vec{B}_{\max,N} \tag{11}$$

Theorem 5.6. *$A_{\max,N}$ is invertible $\forall$ N.*

Proof: Consider the maximization problem in the LP in 7. We know that LP admits at least one solution of the vector $(y_2, y_3, \ldots, y_{N+1})$ because all the variables are bounded above by $y_1 = \frac{c_1}{c_1+c_2}$.Let this solution be denoted by $y'_{\max,N}$. We know that $y'_{\max,N}$ satisfies equation 11 by Theorem 4.5. Thus equation 11 has at least one solution. Now, assume that the matrix $A_{\max,N}$ is not invertible. Then equation 11 has at least 2 different solutions $\vec{y}'_{\max,N}$ and $\vec{y}''_{\max,N}$.But, the value of y_2 corresponding to both of them are the same as both of them are the maximum values of y_2 obtained from the LP. Now, if the two vectors have the same value of y_2, $\vec{y}'_{\max,N} = y''_{\max,N}$ as the other variables satisfy the recursion equation and are fixed once we fix the value of y_2. This contradicts the fact that $\vec{y}'_{\max,N}$ and $\vec{y}''_{\max,N}$ are two different solutions of equation 11. So, $A_{\max,N}$ is invertible. ■

So, we can obtain $\vec{y}_{\max,N} = A^{-1}_{max,N}\vec{B}_{\max,N}$. and hence find a closed form solution of $y_{2,\max,N}$ explicitly.

Finding Closed form expression of $y_{2,\min,N}$

We can use the result in the above theorem to find out $y_{2,\min,N}$ and $y_{2,\max,N}$ explicitly.We know that $y_1 = \frac{c_1}{c_1+c_2}$. Let the vector $\vec{y}_{\min,N}$ denote the vector of values of $y_2, y_3, \ldots, y_{N+1}$ obtained from the LP by minimizing y_2. then $\vec{y}_{\min,N}$ satisfies the following equations:

$$y_{n+1} = -\frac{c_1}{(n-1)(c_1+c_2)} + \frac{n}{n-1}\left(\frac{c_1+c_2}{c_3}+1\right)y_n - \frac{n}{n-1}\frac{c_1}{c_3}y_{n-1} \qquad n=2(1)N$$

$$y_{N+1} = 0$$

Thus, we have a set of N equations and N variables to solve for and the solution of y_2 that we obtain from the above set of equations in nothing but $y_{2,\min,N}$ as indicated by the previous theorem. Define the matrix $A_{\min,N} = ((A_{\min,N,ij}))_{N\times N}$ where

$$A_{\min,N,ij} = \begin{cases} \dfrac{-2\ (c_1+c_2+c_2)}{c_3} & \text{if } \ i=1,\ j=1 \\ 1 & \text{if } \ i=1,\ j=2 \\ \dfrac{i+1}{i}\cdot\dfrac{c_1}{c_3} & \text{if } \ i=j,\ 2\le i\le N-1 \\ -\dfrac{i+1}{i}\left(\dfrac{c_1+c_2}{c_3}+1\right) & \text{if } \ j=i+1,\ 2\le i\le N-1 \\ 1 & \text{if } \ j=i+2,\ 2\le i\le N-1 \\ 1 & \text{if } \ i=N,\ j=N \\ 0 & \text{otherwise} \end{cases}$$

Define $\vec{B}_{\min,N} = (B_{\min,N,1}, B_{\min,N,2}, \ldots, B_{\min,N,N})^T$ where $B_{\min,N,i} = \begin{cases} -\dfrac{c_1(2c_1+c_3)}{(c_1+c_2)c_3} & \text{if } \ i=1 \\ \dfrac{c_1}{i(c_1+c_2)} & \text{if } \ i=2(1)N-1 \\ 0 & \text{if } \ i=N \end{cases}$

Writing down these equations, in the matrix form, we have,

$$A_{min,N}\cdot\vec{y}_{\min,N} = \vec{B}_{\min,N} \tag{12}$$

Theorem 5.7. *$A_{\min,N}$ is invertible $\forall\ N$.*

Proof: Consider the minimization problem in the LP in 7. We know that LP admits at least one solution of the vector $(y_2, y_3, \ldots, y_{N+1})$ because all the variables are bounded below by 0.Let this solution be denoted by $y'_{\min,N}$. We know that $y'_{\min,N}$ satisfies equation 12 by Theorem 4.5. Thus equation 12 has at least one solution. Now, assume that the matrix $A_{\min,N}$ is not invertible. Then equation 12 has at least 2 different solutions $\vec{y}'_{\min,N}$ and $\vec{y}''_{\min,N}$.But, the value of y_2 corresponding to both of them are the same as both of them are the minimum values of y_2 obtained from the LP. Now, if the two vectors have the same value of y_2, $\vec{y}'_{\min,N} = y''_{\min,N}$ as the other variables satisfy the recursion equation and are fixed once we fix the value of y_2. This contradicts the fact that $\vec{y}'_{\min,N}$ and $\vec{y}''_{\min,N}$ are two different solutions of equation 12. So, $A_{\min,N}$ is invertible. ■

So, we can obtain $\vec{y}_{\min,N} = A^{-1}_{min,N}\vec{B}_{\min,N}$. and hence find a closed form solution of $y_{2,\min,N}$ explicitly. Once we find closed form expressions for $y_{2,\min,N}$ and $y_{2,\max,N}$, it provides with a new direction to prove different properties, a notable of which is that $y_{2,\max,N} - y_{2,\min,N} \to 0$.

6 Generalization attempt for arbitrary binary input

In this section, we begin with analyzing the causal estimator $Z(t)$ of $X(t)$ when $X(t)$ can take two values, x_1 with probability $\frac{c_2}{c_1+c_2}$ and x_2 with probability $\frac{c_1}{c_1+c_2}$. Since, $Z(t) = \mathbb{E}[X(t)|Y(t)]$, hence,

$$\mathbb{E}[Z(t)] = \mathbb{E}[\mathbb{E}[X(t)|Y(t)]] = \mathbb{E}[X(t)] = \frac{c_2x_1 + c_1x_2}{c_1+c_2}$$

Now, for the trajectory equation of $Z(t)$ is given by

$$dZ(t) = \left[c_1\Delta x - (c_1+c_2+c_3\Delta x)(Z(t)-x_1) + c_3(Z(t)-x_1)^2\right] dt + g(Z(t-))\; dY(t) \tag{13}$$

where $dY(t) = 1$ for jump times $t = t_i$, and vanishing otherwise. We write $Z(t-)$ as it is measurable upto the filtration till time t. Also $g(Z(t_i)) = \frac{(Z(t_i)-x_1)(x_2-Z(t_i))}{Z(t_i)}$ is the jump height $dZ(t_i) = Z(t_i+) - Z(t_i-)$ at jump times of Y. Following the footsteps as in [1], denoting γ for $\frac{c_1+c_2+c_3\Delta x}{c_3}$, we write the equation as

$$\begin{aligned} dZ(t) &= c_3\left[\frac{c_1\Delta x}{c_3} - \gamma(Z(t)-x_1) + (Z(t)-x_1)^2\right] dt + \frac{(Z(t-)-x_1)(x_2-Z(t-))}{Z(t-)} dY(t) \\ &= c_3\left[\frac{c_1\Delta x}{c_3} - \gamma(Z(t)-x_1) + (Z(t)-x_1)^2\right] dt + \frac{(Z(t-)-x_1)(x_2-Z(t-))}{Z(t-)} [dQ(t) + c_3Z(t-)dt] \end{aligned}$$

Here we have used *Doob Meyer's decomposition* as mentioned in 3.2. We write $dY(t)$ as $dQ(t) + c_3Z(t-) = dQ(t) + c_3Z(t)$ (because $Z(t) = Z(t-)$) where $Q(t)$ is a zero-mean martingale and $Z(t)$ is a predictable process.

Now, (13) writes $Z(t)$ as a stochastic differential equation in terms of the count process $Y(t)$, and we can now use Ito formula on the function $F(x) = (x-x_1)^n$. Therefore the Ito Formula as given in 3.3 upon applying will yield the following form

$$\begin{aligned} d[(Z(t)-x_1)^n] &= n(Z(t)-x_1)^{n-1}\left[c_1\Delta x - c_3\gamma(Z(t)-x_1) + c_3(Z(t)-x_1)^2\right] dt \\ &\quad + \left[F\left(Z(t-) + \frac{(Z(t-)-x_1)(x_2-Z(t-))}{Z(t-)}\right) - F(Z(t-))\right] dY(t) \\ &= n(Z(t)-x_1)^{n-1}\left[c_1\Delta x - c_3\gamma(Z(t)-x_1) + c_3(Z(t)-x_1)^2\right] dt \\ &\quad + \left[\left(\frac{Z(t-)(x_1+x_2) - x_1x_2}{Z(t-)} - x_1\right)^n - (Z(t-)-x_1)^n\right] dY(t) \\ &\approx n(Z(t)-x_1)^{n-1}\left[c_1\Delta x - c_3\gamma(Z(t)-x_1) + c_3(Z(t)-x_1)^2\right] dt \\ &\quad + \left[\left(x_2^n - \frac{nx_1x_2^n}{Z(t-)}\right) - (Z(t-)-x_1)^n\right] [c_3Z(t-)dt + dQ(t)] \\ &\qquad \text{[By using Bernoulli Approximation } (1-x)^n \approx (1-nx) \text{ if } 0 < x \ll 1] \\ &= n(Z(t)-x_1)^{n-1}\left[c_1\Delta x - c_3\gamma(Z(t)-x_1) + c_3(Z(t)-x_1)^2\right] dt \\ &\quad + c_3\left[(x_2^n(Z(t-)-x_1) + x_1x_2^n - nx_1x_2^n) - (Z(t-)-x_1)^{n+1} - x_1(Z(t-)-x_1)^n\right] \\ &\quad + \underbrace{\left[\left(x_2^n - \frac{nx_1x_2^n}{Z(t-)}\right) - (Z(t-)-x_1)^n\right] dQ(t)}_{\text{mean=0 as } Q(t) \text{ is a zero-mean martingale and } Z(t) \text{ is predictable}} \end{aligned} \tag{14}$$

Now we take expectations on both sides of the above expression with $T_n = \mathbb{E}[(Z(t)-x_1)^n], n \in \mathbb{N}$.

$$\begin{aligned}\frac{d}{dt}T_n &= nc_1\Delta x T_{n-1} - nc_3\gamma T_n + nc_3 T_{n+1} + c_3 x_2^n T_1 - (n-1)c_3 x_1 x_2^n - c_3 T_{n+1} - x_1 c_3 T_n \\ &= nc_1\Delta x T_{n-1} - c_3(n\gamma + x_1)T_n + c_3(n-1)T_{n+1} + c_3 x_2^n T_1 - (n-1)c_3 x_1 x_2^n \end{aligned} \tag{15}$$

Thus, at stationarity, taking $\mathbb{E}[(Z_\infty - x_1)^n] = W_n$

$$\begin{aligned} c_3(n-1)W_{n+1} &= -nc_1\Delta x W_{n-1} + c_3(n\gamma + x_1)W_n - \frac{c_3 c_1}{c_1+c_2} x_2^n \Delta x + (n-1)c_3 x_1 x_2^n \\ W_{n+1} &= -\frac{nc_1\Delta x}{c_3(n-1)}W_{n-1} + \frac{n\gamma + x_1}{n-1}W_n - \frac{c_1}{(c_1+c_2)(n-1)} x_2^n \Delta x + x_1 x_2^n \end{aligned} \tag{16}$$

If we put $x_1 = 0,\ x_2 = 1$ i.e. $\Delta x = 1$ in the above equation, then we again obtain equation (6). This serves as a sanity check using the known case, and it is to be noted that the approximation is exact in that case.

Imitating the previous case and denoting $y_n := \mathbb{E}[(Z_\infty - x_1)^n / x_2^n] = W_n / x_2^n$ (scaled moments), we prove the analogous version of Theorem 4.2. First we rewrite the recursion for scaled moments as follows:

$$x_2 y_{n+1} = -\frac{nc_1\Delta x}{c_3(n-1)x_2} y_{n-1} + \frac{n\gamma + x_1}{n-1} y_n - \frac{c_1}{(c_1+c_2)(n-1)}\Delta x + x_1 \tag{17}$$

Theorem 6.1. *For general x_1 and x_2, y_n's are non-decreasing with y_2.*

Proof: Following the calculations of Theorem 4.2, here also we prove by an induction on n. Our strengthened induction hypothesis is:

Induction Statement $Q(n)$: y_n and $x_2 y_n - y_{n-1}\Delta x$ are both non-decreasing functions of y_2.

We will show that $Q(n)$ is true for all $n \geq 2$.

Here also equation (17) allows us to conclude that y_n's are differentiable functions of y_2. For the base case, y_1 is constant, and y_2 is obviously increasing in y_2, and thus so does the difference $x_2 y_2 - y_1 \Delta x$. Thus $Q(2)$ is true.

Now, assume that $Q(k)$ is true for all $2 \leq k \leq m$. We will show $Q(m+1)$ is true as well.

We differentiate equation (17) with respect to y_2, with the notation of y_i's we're using.

$$\begin{aligned} x_2 \cdot \frac{d}{dy_2} y_{m+1} &= -\frac{mc_1\Delta x}{c_3(m-1)x_2}\cdot\frac{d}{dy_2}y_{m-1} + \frac{m\gamma + x_1}{m-1}\cdot\frac{d}{dy_2}y_m \\ &= -\frac{mc_1\Delta x}{c_3(m-1)x_2}\cdot\frac{d}{dy_2}y_{m-1} + \frac{1}{m-1}\left(m\cdot\frac{c_1+c_2+c_3\Delta x}{c_3} + x_1\right)\cdot\frac{d}{dy_2}y_m \\ &= \frac{mc_1}{c_3(m-1)x_2}\cdot\frac{d}{dy_2}(x_2 y_m - y_{m-1}\Delta x) + \frac{1}{m-1}\left(\frac{mc_2}{c_3} + m\Delta x + x_1\right)\cdot\frac{d}{dy_2}y_m \end{aligned} \tag{18}$$

The derivative terms on the right hand side are both positive by virtue of $Q(m)$, therefore $x_2 y_{m+1}$ is non-decreasing in y_2, implying that y_{m+1} is non-decreasing in y_2.

Again, from equation (17),

$$\begin{aligned}\frac{d}{dy_2}(x_2 y_{m+1} - y_m \Delta x) &= \frac{mc_1}{c_3(m-1)x_2} \cdot \frac{d}{dy_2}(x_2 y_m - y_{m-1}\Delta x) \\ &\quad + \frac{1}{m-1}\left(\frac{mc_2}{c_3} + m\Delta x + x_1 - (m-1)\Delta x\right) \cdot \frac{d}{dy_2} y_m \\ &= \frac{mc_1}{c_3(m-1)x_2} \cdot \frac{d}{dy_2}(x_2 y_m - y_{m-1}\Delta x) + \frac{1}{m-1}\left(\frac{mc_2}{c_3} + x_2\right) \cdot \frac{d}{dy_2} y_m \qquad (19)\end{aligned}$$

By $Q(m)$, $(x_2 y_m - y_{m-1}\Delta x)$ and y_m are both non-decreasing in y_2. Therefore, $x_2 y_{m+1} - y_m \Delta x$ is non-decreasing in y_2.

Thus by the strong form of mathematical induction, $Q(n)$ is true for all $n \geq 2$, and the proof in the general case holds true. ■

One can also attempt to prove similar results in the general case as we have done in the previous section, and we expect the proof strategies to follow suit.

7 Simulation Experiments

We have done a few simulations for the case $x_1 = 0$ and $x_2 = 1$ to verify our claims and proofs in section 4. All the simulations have been done till 400 recursions, i.e. by forming the Linear Program considering the first 400 moments.

Figure 1 shows evidence for Theorem 4.1 and suggests that the difference between the maximum and the minimum value of 2nd moment is indeed non-increasing with the number of recursions/constraints that we consider to formulate our linear program. The plot also hints towards the fact that the difference possibly converges to 0 as we increase the number of constraints which is proved in Theorem 4.4. Also, Figure 2 suggests that the maximum value of 2nd moment decreases as we include more constraints, whereas the minimum value of 2nd moment increases with increasing number of recursions.

Figure 3 shows plots of the moment curve with different values of 2nd moment. We see that if we start with the minimum possible value of the 2nd moment as obtained from our optimization problem as a starting point of the recursion, the moment curve lies below the others and if we start from the maximum value of the 2nd moment, the moment curve lies above the others. The other curves are obtained by choosing other possible values of the 2nd moment which lie between the maximum and minimum value.

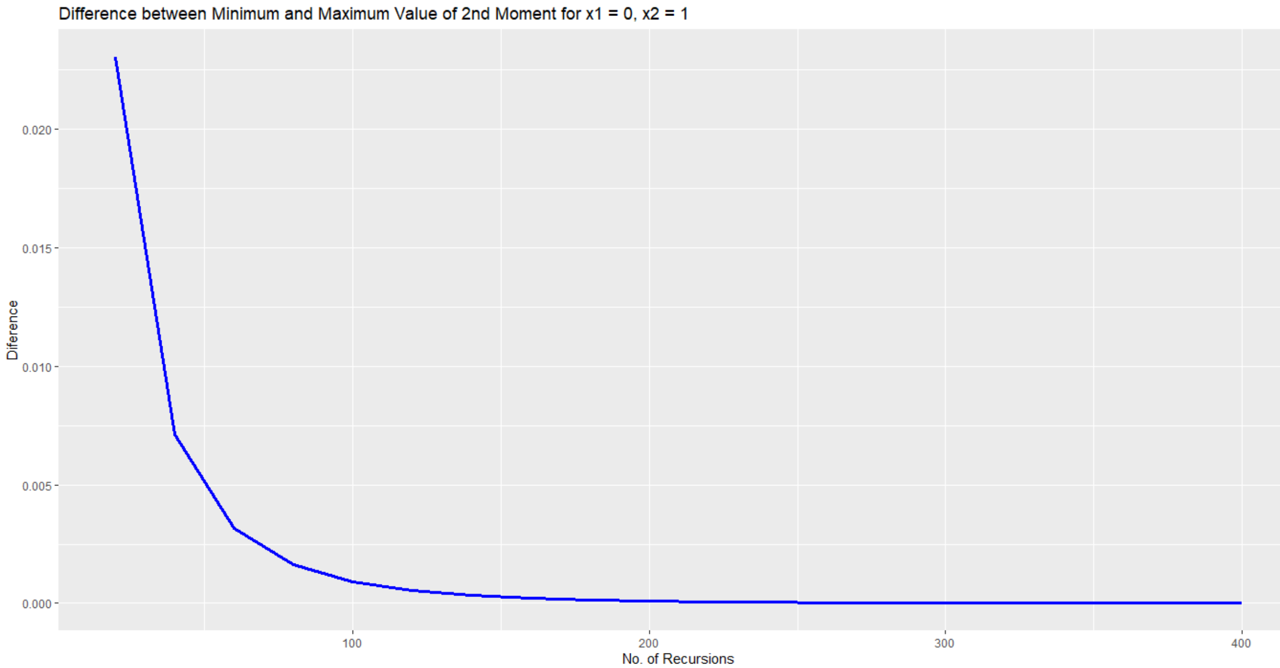

Figure 1: Difference between the maximum and minimum value of 2nd Moment with the No. of Recursions

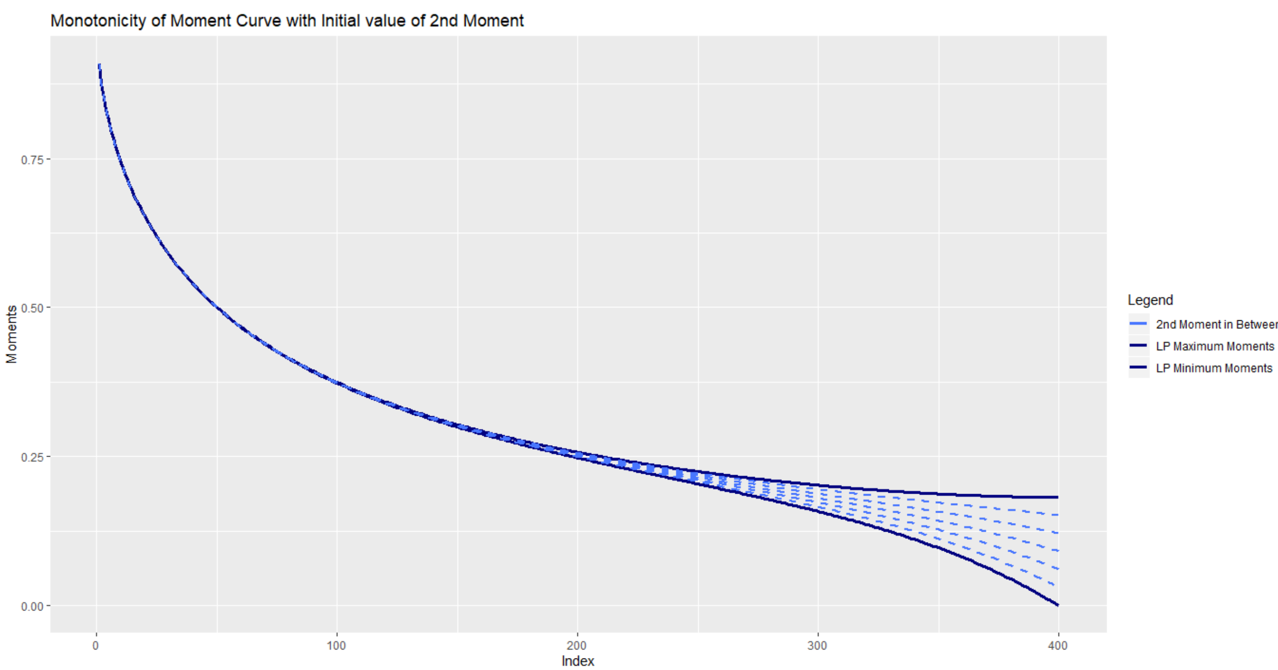

Figure 3: Monotinicity of Moment Curve with values of 2nd Moment

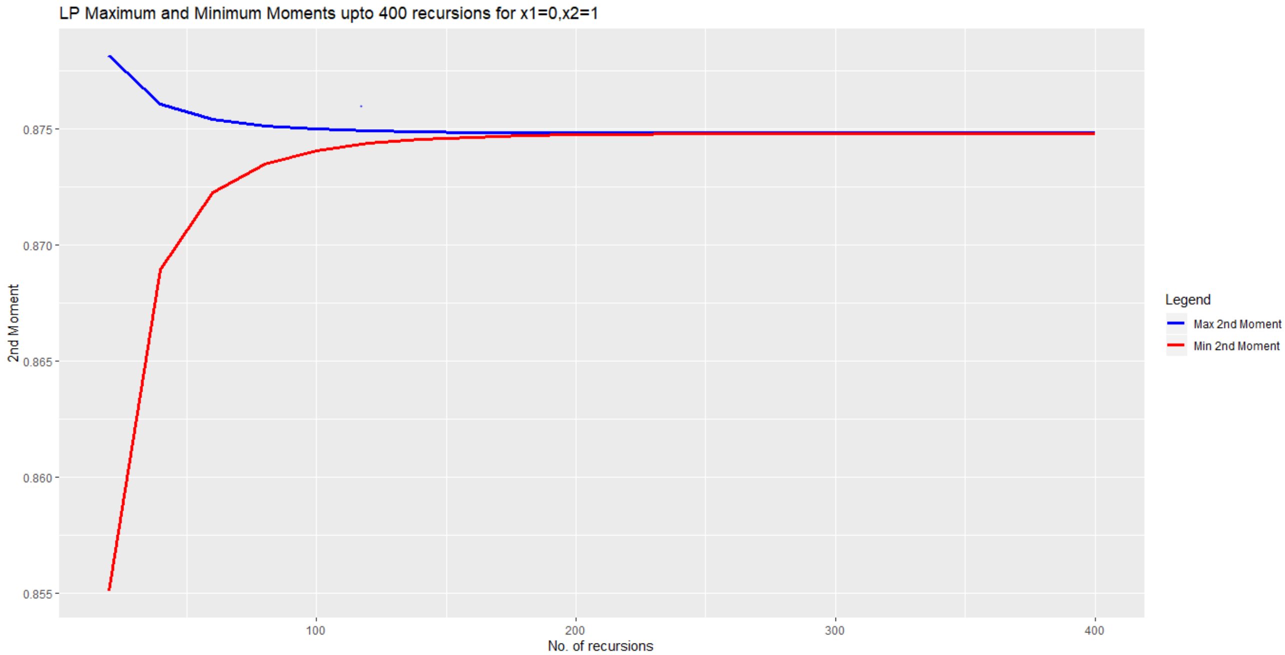

Figure 2: Maximum and minimum value of 2nd Moment with the No. of Recursions

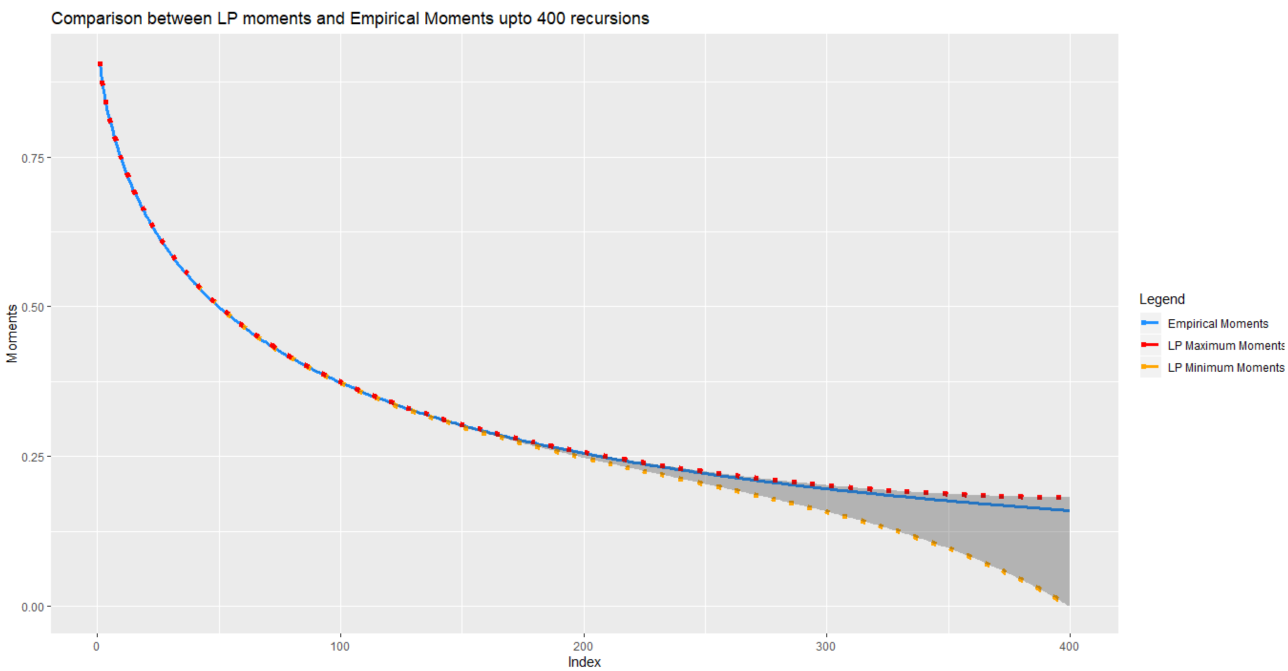

Figure 4: Comparison of Empirical Moment curve and that obtained by our Recursion and LP

Figure 4 illustrates the fact that the actual moment curve lies in between the two curves that are obtained by taking y_2 as the maximum and minimum possible value of the 2nd moment as obtained from our optimization problem and then finding the higher order moments from recursion equation 6. From the above simulations and theorems, we have the following recommendation for finding out the moments of Z_∞ up to a required precision.

1. Choose a user defined ϵ. Since we have proved that $|y_{2,\max,n} - y_{2,\min,n}| \to 0$ as $n \to \infty$, we can find a N such that $|y_{2,\max,N} - y_{2,\min,N}| < \epsilon$. Now choose y_2 to be any value in the interval $[y_{2,\max,N}, y_{2,\min,N}]$.
2. Take $y_1 = \dfrac{c_1}{c_1 + c_2}$.
3. Use these values of y_1 and y_2 to find the other higher order moments by equation 6.

8 Epilogue

In this report we have had a small background review of the existing work, and we worked on an algorithm to approximate the moments, which is computationally efficient. We also attempted to generalize the framework for arbitrary x_1 and x_2 binary input for $X(t)$. We suspect that the difference between the largest and smallest values from the linear programs would indeed converge to 0 with larger number of equations taken into consideration. This is also validated by our simulations. Thus, we can approximate the moments with arbitrary precision.

In fact, since we know that the maximum and minimum values obtained can be obtained directly by inverting a certain matrix for a fixed number of recursion equations, future scopes of study include using these matrices and properties of linear program (arguments using convex polytopes in n dimensional space) to study the convergence between the largest and smallest values of the linear program converging to the same value.

The arbitrary binary input framework uses a Bernoulli approximation, to yield a approximate recursion equation which resembles the original exact recursion for the 0-1 case. This makes us optimistic, and we expect that one can study the nature of approximation provided by this Bernoulli approximation for moments in the arbitrary case. These are indeed future scopes one can explore in this direction.

Acknowledgement

Foremost, we express our sincere gratitude towards the Bioinspired Communications Systems lab, TU Darmstadt for providing us with a scope of pursuing this internship. This opportunity not only marks inclusion of an important research work in our academic experience but it has also allowed us to strengthen interaction among the group members and share research oriented thoughts. We heartily thank Dr. Heinz Koeppl for offering us with intern positions and also allowing us to continue our assigned tasks virtually even amidst the unprecedented pandemic situation. We are also thankful to Ms. Christine Cramer for supervising the logistics and necessary arrangements for the easy-going of our work. To conclude, we cannot forget express our heartfelt thanks to Mr. Mark Sinzger for his constant assistance and careful guidance through the entire span of our work. Without Mr. Mark and his valuable mentorship, the desired results would not have been accomplished.

References

[1] M. Sinzger, M. Gehri and H. Koeppl, “Poisson channel with binary Markov input and average sojourn time constraint, 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, pp. 2873-2878, doi: 10.1109/ISIT44484.2020.9174360, 2020,.

[2] R. S. Liptser and A. N. Shiryaev, “Statistics of random processes II: Applications, vol. 2,” *Springer*, vol. 737, p. 738, 2001.

[3] M. H. Davis, “Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 46, no. 3, pp. 353–376, 1984.

[4] D. L. Snyder and M. I. Miller, *Random point processes in time and space*, ser. Springer texts in Electrical Engineering. Springer-Verlag, 1991.

[5] C. Gardiner, *Stochastic methods: a handbook for the natural and social sciences*, 4th ed. Springer-Verlag, 2009.

[6] O. O. Aalen, Borgan, H. K. Gjessing, *Survival and Event History Analysis: A Process Point of View*, Springer, 2008.