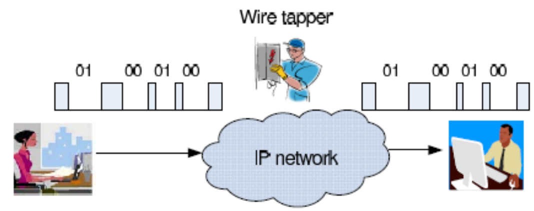
A Memory-less Proof Calculus for Queuing Timing Channels

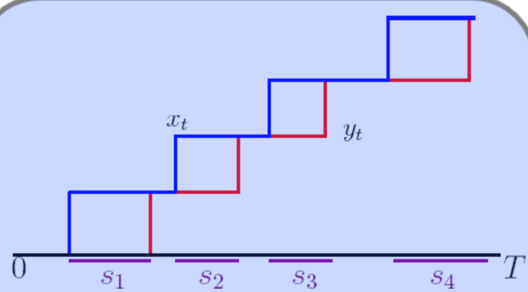
Todd P. Coleman, UIUC

Objective: embed extra (e.g. covert) information in packet *timings* *as they* traverse queuing networks in MANETs



Theory: only the P2P problem solved!

Practice: only recently (Coleman, Kiyavash 08) has practical P2P coding been instantiated



Idea: reason algebraically over the space of counting functions

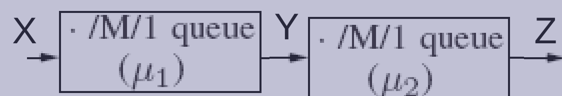
$$P(y^n|x^n,s^n) = \prod_{i=1}^{n} P(y_i|x_i,s_i)$$

$$p(y|x,s) \propto \int_0^T \rho(x_t - y_t)dt$$

$$H(s) = 0$$

Use Point Process Entropy + Algebraic Approach on Counting Functions

A Converse for Tandem Queue Capacity



$$P(z^n|x^n,s_1^n,s_2^n) = \prod_{i=1}^{n} P(y_i|x_i,s_{1,i},s_{2,i})$$

$$P(z|x,s_1,s_2) \propto \sum_{y-x=0}^{z-x} P(z|y,s_2)P(y|x,s_1)$$

$$H(s_1) = H(s_2) = 0$$

How it works: *Two-stage coding scheme.*

1. Encode **counts** (***s***) at 0 rate.
2. Conditioned on ***s***: a **memoryless** channel

Key insight:

H(Z|X,S) **independent** of X for the tandem queue! Simple max-entropy argument (just like AWGN): Poisson inputs!

Assumptions and limitations:

•*Packet losses due to congestion heterogeneity not modeled*

IMPACT

- Introduces new insights into the **algebraic structure** of queuing **timing channels**
- Enables **simple memoryless** (analogous to AWGN) proof of "Bits Through Queues"
- Also enables **first known converse** for the tandem queue

NEXT-PHASE GOALS

- Develop achievability scheme to match the converse
- Extend this approach to more timing-channel MANET architectures

Queuing Timing Channels Afford New Degrees of Freedom in MANETs and are Analogous to AWGN Channels

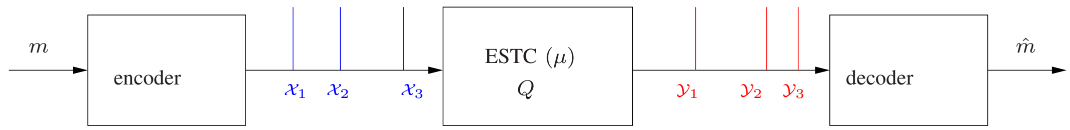


Fig. 1. Conveying information through packet timings in a queueing system.

Single Server Queue:

$$C(\lambda, \mu) = \lambda \log \frac{\mu}{\lambda}, \quad \lambda < \mu \text{ nats/s.} \tag{1}$$

$$C(\mu) = e^{-1}\mu \text{ nats/s,} \tag{2}$$

a) Related Work on Proving $C(\lambda, \mu)$:

- Original Anantharam & Verdu '96: "Bits through Queues": proved closed-form structure of $C(\lambda, \mu)$ by considering the probabilistic dynamics relating n packet arrivals to n packet departures of an ESTC. Required information density arguments
- Bedekar & Azizoglu '98: "The information-theoretic capacity of discrete-time queues": analogous proof technique (requiring information density arguments) as in CT case.
- Prabhakar & Gallager '03: focused on information rates but never explicitly showed achievability of these rates
- Sundarasen & Verdu '06: took a point process approach but still require information densities to illustrate achievability.

A. *Methodology of Our Approach*

We illustrate that one can reason about this channel coding problem completely from a traditional memoryless channel perspective, by exploiting a few key properties:

1) The numerical entropy rate of any finite-rate point process tending to 0 (Lemma 0.2),
2) The memoryless nature of the channel departure likelihood, when conditioned upon a process whose entropy rate tends to 0 (equation (12)).
3) The maximum-entropy nature of the Poisson process (Lemma 0.1).

The reasoning of this paper is completely dual to that of the two-stage lossy compression scheme developed in [7]. Such two-stage compression schemes for non-stationary independent-increments processes date back to Rubin [8] (for Poisson processes) and [9], [10] (for Wiener processes).

B. *Notation on Point Processes*

- Define Γ_T to be the set of all counting functions on $[0,T]$:

$$\Gamma_T \triangleq \{y : [0,T] \to \mathbb{Z}_+, \ \ y \text{ is nondecreasing, and right-continuous } \}. \tag{3}$$

- Point process $\mathcal{Y}$ with occurrence times $\{\mathcal{Y}_1, \mathcal{Y}_2, \ldots\}$. Counting function $(Y_t : t \geq 0)$:

$$Y_t = \sup\{n \in \mathbb{N} : \mathcal{Y}_n \leq t\}.$$

- The *entropy* on $[0,T]$ of a point process $\mathcal{Y}$ with arrival times $\{\mathcal{Y}_1, \mathcal{Y}_2, \ldots\}$ is defined [11] as the sum of its *numerical entropy* and its *positional entropy*:

$$h_T(\mathcal{Y}) := H(Y_T) + h\left(\mathcal{Y}_1, \ldots, \mathcal{Y}_{Y_T} | Y_T\right),$$

  where $H(\cdot)$ is discrete entropy, $h(\cdot)$ is differential entropy, and $\{\mathcal{Y}_1, \ldots, \mathcal{Y}_{Y_T}\}$ are the locations (in time) of the arrivals on $[0,T]$.
- We define the *rate* $r(\mathcal{Y})$ of a point process $\mathcal{Y}$ to be λ if

$$\lim_{T\to\infty} \frac{E[Y_T]}{T} = \lambda.$$

- *Poisson* processes are known to have desirable *extremal entropic* properties [11]:

  Lemma 0.1: The Poisson process of rate λ is maximum-entropy over all rate-λ point processes, and has entropy on $[0,T]$ given in closed form by

$$h_T(\mathcal{Y}) = T\lambda(1 - \log\lambda).$$

-

  Lemma 0.2: For any point process $\mathcal{Y}$ such that $r(\mathcal{Y}) < \infty$,

$$\lim_{T\to\infty} \frac{1}{T} H\left(Y_T\right) = 0.$$

C. *The Likelihood of a Point Process*

Point process $\mathcal{Y}$: $H_t \triangleq \sigma$-algebra generated by $\{Y_\tau : \tau \in [0,t)\}$.

$$\lambda_t \triangleq \lim_{\Delta\to 0} \frac{P\left(Y_{t+\Delta} - Y_t = 1 | H_t\right)}{\Delta}. \tag{4}$$

$$p_T(y) = \exp\left\{\int_0^T \log\lambda_t dy_t - \lambda_t dt\right\} \tag{5}$$

D. Queuing Timing Channels: the ESTC

- service times are i.i.d. and exponentially distributed:

$$\lambda_t = \mu 1_{\{Q_t > 0\}}, \tag{6}$$

$$Q_t \triangleq X_t + Q_0 - Y_t \tag{7}$$

where Q_0 is the initial condition, i.e. the state of the queue at time 0.

For specific realizations of $y \in \Gamma_T$ and $x \in \Gamma_T$:

$$\begin{aligned} p_T\left(y|x, q_0\right) &= \exp\left\{\int_0^T \log\left(\mu 1_{\{q_t>0\}}\right) dy_t - \mu 1_{\{q_t>0\}} dt\right\} \\ &= \begin{cases} \mu^{y_T - y_0} \exp\left\{\int_0^T -\mu 1_{\{x_t + q_0 - y_t > 0\}} dt\right\} & \text{if } x_t + q_0 - y_t \geq 0 \; \forall \, t \in [0,T] \\ 0 & \text{otherwise} \end{cases}. \end{aligned}$$

Note that this can be expressed as

$$p_T\left(y|x, q_0\right) = Z_T(y) \exp\left\{\int_0^T -\mu \rho\left(x_t + q_0 - y_t\right) dt\right\} \tag{8a}$$

$$Z_T(y) \triangleq \mu^{y_T - y_0} \tag{8b}$$

$$\rho(u) \triangleq \begin{cases} 0, & u = 0, \\ 1, & u > 0, \\ \infty, & u < 0 \end{cases} \tag{8c}$$

E. Memoryless Nature of Channel Dynamics Given Queue States

Succinct explanation of ESTC: $\mathcal{X} - Q - \mathcal{Y}$. Specifically, for $x \in \Gamma_{nT}$ and $y \in \Gamma_{nT}$, associate

$$x \Leftrightarrow \tilde{x}^n = (\tilde{x}_1, \ldots, \tilde{x}_n), \;\; \tilde{x}_i \triangleq (x_{t+(i-1)T} : t \in [0,T]) \tag{9}$$

$$y \Leftrightarrow \tilde{y}^n = (\tilde{y}_1, \ldots, \tilde{y}_n), \;\; \tilde{y}_i \triangleq (y_{t+(i-1)T} : t \in [0,T]) \tag{10}$$

$$\tilde{q}^n = (\tilde{q}_1, \ldots, \tilde{q}_n), \;\; \tilde{q}_i \triangleq q_{(i-1)T} - q_0 \tag{11}$$

where $\tilde{x}_i \in \Gamma_T$ and $\tilde{y}_i \in \Gamma_T$.

It follows directly from (8) that given knowledge of $\{\tilde{q}_i\}_{i=1}^n$, this induces a memoryless channel:

$$p_{nT}(\tilde{y}^n|\tilde{x}^n, \tilde{q}^n) = \prod_{i=1}^n p_T(\tilde{y}_i|\tilde{x}_i, \tilde{q}_i) \tag{12}$$

I. A Converse for the Single-Server ESTC

For the single-server ESTC, we first compute the information capacity over all input processes that satisfy the constraint that $r(\mathcal{X}) = \lambda$. Note that for the converse, we can simply use a genie-aided decoder, that has $\tilde{Q}^n$ at the decoder. As a consequence, from (12), it follows that this is a memoryless channel, and thus for any T, from Fano's inequality [14] it must be that

$$\begin{aligned} R &< \frac{1}{nT} I_T(\tilde{X}^n; \tilde{Y}^n | \tilde{Q}^n) \\ &\leq \sup_{\substack{p(\tilde{X}): E[X_T] = \lambda T \\ p(\tilde{Q}): E[\tilde{Q}] < \infty}} \frac{1}{T} I_T(\tilde{X}; \tilde{Y} | \tilde{Q}) \end{aligned}$$

Now allowing $T \to \infty$ so that we are conditioning on less and less genie-adided information $(\tilde{Q}^n)$, we have:

$$\begin{aligned} R &< C(\lambda, \mu), \\ C(\lambda, \mu) &= \sup_{\substack{p(\mathcal{X}): r(\mathcal{X}) = \lambda \\ p(\tilde{Q}): E[Q_0] < \infty}} \liminf_{T \to \infty} \frac{1}{T} I_T(X; Y | Q_0) \\ I_T(X; Y | Q_0) &\triangleq E\left\{ \log \frac{p_T(Y | X, Q_0)}{p_T(Y | Q_0)\, p_T(X | Q_0)} \right\}. \end{aligned}$$

Note that because

$$I_T(X; Y | Q_0) - I_T(X; Y) = I(Q_0; X | Y) - I(Q_0; X) \tag{13}$$

and $\lim_{T \to \infty} \frac{H(Q_0)}{T} = 0$ because $E[Q_0] < \infty$ and Lemma 0.2 below, the probability distribution of the initial state Q_0 will not affect $C(\lambda, \mu)$. So we now assume $Q_0 = 0$ to calculate $C(\lambda, \mu)$:

$$C(\lambda, \mu) = \sup_{p(\mathcal{X}): r(\mathcal{X}) = \lambda} \liminf_{T \to \infty} \frac{1}{T} I_T(X; Y). \tag{14}$$

It will be shown in Section III that $C(\lambda, \mu) = \lambda \log(\frac{\mu}{\lambda})$ and it is achieved with $p(\mathcal{X})$ corresponding to a Poisson process of rate λ.

II. Achievability for the Single-Server ESTC: A Two-Stage Coding Approach

We now show that we can achieve the rate $C(\lambda, \mu)$. We use a simple two-stage coding scheme, using a Poisson process of rate λ as the input.

- For $i \in \{1, \ldots, n\}$: define

$$\tilde{W}_i \triangleq X_{iT} - X_{(i-1)T}. \tag{15}$$

$$C_i \triangleq X_{iT} = \sum_{k=1}^{i} \tilde{W}_i. \tag{16}$$

  First communicate $\tilde{W}^n = \{\tilde{W}_1, \ldots, \tilde{W}_n\} \Leftrightarrow C^n = \{C_1, \ldots, C_n\}$. Note that

$$p(\tilde{y}^n | \tilde{w}^n) = \prod_{i=1}^{n} p(\tilde{y}_i | \tilde{y}_{i-1}, \tilde{w}_i)$$

  and so the capacity of this channel is defined in terms of its information rate.

  Since $r(\mathcal{X}) = \lambda$, for sufficiently large T, each $\tilde{W}_i$ is a non-negative random variable of mean approximately λT. Since we are constraining the process $\mathcal{X}$ such that $r(\mathcal{X}) = \lambda$, Lemma 0.2, it follows directly that

  Corollary 2.1: For each $i \in \{1, \ldots, n\}$, $\lim_{T \to \infty} \frac{H(\tilde{W}_i)}{T} = 0$.

  Thus it follows that we can communicate $\{\tilde{W}_i\}$ with 0 rate using a Poisson-λ input.

- Note that given the information $\tilde{W}^n$ at the decoder, since

$$Q_{iT} = Q_0 + X_{iT} - Y_{iT},$$

  the decoder now has $\tilde{Q}^n$ at its disposal. Constructing X to be a Poisson process, it follows that we can achieve $C(\lambda, \mu)$ as discussed in Section III.

III. Calculation of $C(\lambda, \mu)$ Analogously to the AWGN channel

Define $\mathcal{S}$ as the sequence of induced service times $\{\mathcal{S}_1, \mathcal{S}_2, \ldots\}$ of a an ESTC queuing system with $\mathcal{X}$ as the input and $\mathcal{Y}$ as the output. Specifically,

$$\mathcal{S}_i = \mathcal{Y}_i - \max(\mathcal{Y}_{i-1}, \mathcal{X}_i) \tag{17}$$

For any $\tilde{\mathcal{X}} \in \Gamma_T$ and $\tilde{\mathcal{Y}} \in \Gamma_T$, note that

$$\begin{aligned} I_T(\tilde{\mathcal{X}};\tilde{\mathcal{Y}}) &= h_T(\tilde{\mathcal{Y}}) - h_T(\tilde{\mathcal{Y}}|\tilde{\mathcal{X}}) \\ &= h_T(\tilde{\mathcal{Y}}) - H\left(\tilde{Y}_T|\tilde{\mathcal{X}}\right) - h_T\left(\tilde{\mathcal{Y}}_1, \ldots, \tilde{\mathcal{Y}}_{\tilde{Y}_T}|\tilde{Y}_T, \tilde{\mathcal{X}}\right). \end{aligned} \tag{18}$$

Define $\widehat{I}_T(\tilde{\mathcal{X}};\tilde{\mathcal{Y}})$ and $\tilde{C}(\lambda, \mu)$ as

$$\widehat{I}_T(\mathcal{X};\mathcal{Y}) \triangleq h_T(\tilde{\mathcal{Y}}) - h_T\left(\tilde{\mathcal{Y}}_1, \ldots, \tilde{\mathcal{Y}}_{\tilde{Y}_T}|\tilde{Y}_T, \tilde{\mathcal{X}}\right), \tag{19}$$

$$\tilde{C}(\lambda, \mu) \triangleq \sup_{P_{\tilde{\mathcal{X}}}:r(\tilde{\mathcal{X}})=\lambda} \liminf_{T\to\infty} \frac{1}{T}\widehat{I}_T(\tilde{\mathcal{X}};\tilde{\mathcal{Y}}) \tag{20}$$

Note that by Lemma 0.2, we have that

Corollary 3.1:

$$C(\lambda, \mu) = \tilde{C}(\lambda, \mu).$$

Now note that for any $\mathcal{X}$ such that $r(\mathcal{X}) = \lambda$, we have:

$$\begin{aligned} \widehat{I}_T(\mathcal{X};\mathcal{Y}) &= h_T(\mathcal{Y}) - h_T\left(\mathcal{Y}_1, \ldots, \mathcal{Y}_{Y_T}|Y_T, \mathcal{X}\right) \\ &= h_T(\mathcal{Y}) - h_T\left(\mathcal{S}_1, \ldots, \mathcal{S}_{Y_T}|Y_T, \mathcal{X}\right) \\ &= h_T(\mathcal{Y}) - h_T\left(\mathcal{S}_1, \ldots, \mathcal{S}_{Y_T}|Y_T\right) && (21) \\ &= h_T(\mathcal{Y}) - E[Y_T](1 - \log\mu) && (22) \\ \Rightarrow \liminf_{T\to\infty} \frac{\widehat{I}_T(\mathcal{X};\mathcal{Y})}{T} &= \liminf_{T\to\infty} \frac{h_T(\mathcal{Y})}{T} - \lambda(1 - \log\mu) && (23) \\ &\le \lambda(1 - \log\lambda) - \lambda(1 - \log\mu) && (24) \\ &= \lambda \log\left(\frac{\mu}{\lambda}\right). \end{aligned}$$

(21) follows because the service times are independent of the arrival process in an ESTC; (22) follows because the service times in an ESTC are exponentially distributed of rate μ; (23) follows because for any stable queue, $r(\mathcal{Y}) = r(\mathcal{X}) = \lambda$; and (24) follows from Lemma 0.1; This bound is tight because of Burkes' theorem [15], [16]: a Poisson (λ) arrival process to an ESTC results in a Poisson (λ) departure process.

IV. A CONVERSE FOR THE TANDEM QUEUE

Very difficult timing channel problem... We still enforce the arrival process $\mathcal{X}$ to satisfy $r(\mathcal{X}) = \lambda$. Define

$$Q_{1,t} = Q_{1,0} + X_t - Y_t$$

$$Q_{2,t} = Q_{2,0} + Y_t - Z_t$$

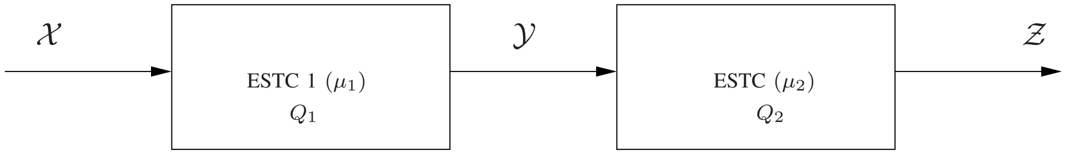


Fig. 2. The tandem queue

We note that analogous to above, the following relationship holds:

$$\mathcal{X} - Q_1 - \mathcal{Y} - Q_2 - \mathcal{Z}.$$

Thus it follows that

$$p_T(z|x, q_{1,0}, q_{2,0}) = \int_{y \in \Gamma_T} p_T(z|y, q_{2,0}) p_T(y|x, q_{1,0}).$$

Note that it must be the case that $x_t \geq y_t$ for all $t \in [0, T]$, and also that $y_t \geq z_t$ for all $t \in [0, T]$. Thus we can re-write this in terms of a convolution:

$$p_T(z|x, q_{1,0}, q_{2,0}) = \int_{y \in \Gamma_T : 0 \leq x - y \leq x - z} p_T(z|y, q_{2,0}) p_T(y|x, q_{1,0}). \quad (25)$$

$$= \int_{x-y=0}^{x-z} p_T(z|y, q_{2,0}) p_T(y|x, q_{1,0}). \quad (26)$$

where (26) follows by denoting, for $x \in \Gamma_T$ and $y \in \Gamma_T$, $x \leq y$ if $x_t \leq y_t$ for all $t \in [0, T]$.

Continuing on, we have:

$$
\begin{aligned}
& \log p_T(z|x, q_{1,0}, q_{2,0}) \\
= \; & \log \left[\int_{x-y=0}^{x-z} p_T(z|y, q_{2,0}) p_T(y|x, q_{1,0}) \right] \\
= \; & \log \left[\int_{x-y=0}^{x-} \mu_2^{z_T} \mu_1^{y_T} \exp \left\{ \int_0^T -\rho_1 \left(x_t + q_{1,0} - y_t \right) - \rho_2 \left(y_t + q_{2,0} - z_t \right) dt \right\} \right] \\
= \; & \log \left[(\mu_1 \mu_2)^{x_T} \int_{x-y=0}^{x-z} \mu_1^{y_T - x_T} \mu_2^{z_T - x_T} \exp \left\{ \int_0^T -\rho_1 \left(x_t + q_{1,0} - y_t \right) - \rho_2 \left(y_t + q_{2,0} - z_t \right) dt \right\} \right] \\
= \; & x_T \log(\mu_1 \mu_2) \\
+ \; & \log \left[\int_{x-y=0}^{x-z} \mu_1^{y_T - x_T} \mu_2^{z_T - x_T} \exp \left\{ \int_0^T -\rho_1 \left(x_t + q_{1,0} - y_t \right) - \rho_2 \left(y_t + q_{2,0} - z_t \right) dt \right\} \right] \\
= \; & x_T \log(\mu_1 \mu_2) + \log \left[\int_{u=0}^{x-z} f_1(u) f_2(x - z - u) \right] \qquad (27) \\
= \; & x_T \log(\mu_1 \mu_2) + \log \left[f_1 * f_2(x - z) \right] \qquad (28)
\end{aligned}
$$

for appropriately defined functions f_1 and f_2 in (27) and (28).

Now it follows as a consequence that

$$
\begin{aligned}
\lim_{T \to \infty} \frac{1}{T} h_T(\mathcal{Z}|\mathcal{X}, Q_{1,0}, Q_{2,0}) \quad &= \quad -\lambda \log(\mu_1 \mu_2) - E \left\{ \log f_1 * f_2(X - Z) \right\} \\
\Rightarrow \lim_{T \to \infty} \frac{1}{T} h_T(\mathcal{Z}|\mathcal{X}, Q_{1,0}, Q_{2,0}) \quad & \quad \text{does } \textbf{not} \text{ depend on } P_X \\
\Rightarrow \text{max-entropy argument:} \quad & \quad \text{Poisson inputs optimal} \\
\Rightarrow \text{precise converse:} \quad & \quad \text{with genie info of zero entropy}
\end{aligned}
$$

V. Discussion

- Conceptually simple, memoryless calculus to view the "Bits Through Queues" problem and solve it
- Calculus enables first known non-trivial converse for tandem queue

REFERENCES

[1] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 4–18, 1996.

[2] S. Verdu, "A general formula for channel capacity," *Information Theory, IEEE Transactions on*, vol. 40, no. 4, pp. 1147–1157, 1994.

[3] A. S. Bedekar and M. Azizoglu, "The information-theoretic capacity of discrete-time queues," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 446–461, 1998.

[4] B. Prabhakar and R. Gallager, "Entropy and the timing capacity of discrete queues," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 357–370, February 2003.

[5] R. Sundaresan and S. Verdú, "Capacity of queues via point-process channels," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2697–2709, June 2006.

[6] B. Prabhakar and N. Bambos, "The entropy and delay of traffic processes in ATM networks," in *Proceedings of the Conference on Information Science and Systems (CISS)*, Baltimore, Maryland, 1995, pp. 448–453.

[7] T. P. Coleman, N. Kiyavash, and V. Subramanian, "The rate-distortion function of a Poisson process with a queuing distortion measure," *IEEE Transactions on Information Theory*, submitted May 2008.

[8] I. Rubin, "Information rates and data-compression schemes for Poisson processes," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 200–210, 1974.

[9] T. Berger, "Information rates of Wiener processes," *Information Theory, IEEE Transactions on*, vol. 16, no. 2, pp. 134–139, 1970.

[10] R. Gray, "Information rates of autoregressive processes," *Information Theory, IEEE Transactions on*, vol. 16, no. 4, pp. 412–421, 1970.

[11] J. McFadden, "The entropy of a point process," *SIAM Journal of Applied Mathematics*, vol. 13, pp. 988–994, 1965.

[12] D. Daley and D. Vere-Jones, *An Introduction to the theory of point processes*. New York: Springer-Verlag, 1988.

[13] P. Bremaud, *Point Processes and Queues: Martingale Dynamics*. New York: Springer-Verlag, 1981.

[14] T. M. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 2006.

[15] R. Gallager, *Discrete Stochastic Processes*. Boston, MA: Kluwer, 1996.

[16] L. Kleinrock, *Queueing Systems. Vol 1: Theory*. New York, NY: Wiley, 1975.