RATE-DISTORTION FUNCTION OF POISSON PROCESSES WITH A QUEUING DISTORTION MEASURE

Todd P Coleman and Negar Kiyavash, UIUC
{colemant,kiyavash}@uiuc.edu

Vijay Subramanian, Hamilton Institute
vijay.subramanian@nuim.ie

STATUS QUO

- Consider tracking **times of meaningful events (point processes)** in MANETs
- How is **lossy source coding** of **temporal information** related to traditional notions of compression?
- Verdu developed a notion [1] for Poisson processes x, but quantized **reproductions** x' **not Poisson**

NEW INSIGHTS

- Consider McFadden's ***"Entropy of a point process"*** [2] formulation

The entropy on $[0, T]$ of a point process $\mathcal{P}$ with arrival times $\{P_1, P_2, \ldots\}$ is defined as the sum of its *numerical entropy* and its *positional entropy*:

$$h_T(\mathcal{P}) := H(M) + E_M[h(P_1, \ldots, P_m | M = m)],$$

where M is a random variable denoting the number of arrivals in $[0, T]$, $H(\cdot)$ is discrete entropy, $h(\cdot)$ is differential entropy, and given $M = m$, $\{P_1, \ldots, P_m\}$ are the locations (in time) of the arrivals.

- **Poisson** process has **maximum entropy** $\quad h_T(\mathcal{P}) = T\lambda(1 - \log\lambda)$.
- Gastpar et al [3]: **d(x,x')=-log p(x|x')** for a channel with p(x'|x) is in some sense ideal and canonical.
  - Bernoulli source with Hamming Distortion <-> Binary Symmetric Channel
  - Gaussian source with Squared Error Distortion <-> AWGN Channel

  Both have simple mutual information inequality proofs exploiting maximum entropy distributions, and can be extended to multiterminal problems
- Consider p(x|x') for an Exponential Server Timing Channel (ESTC) queuing system **(Bits through Queues [4])**?
  - Wagner and Anantharam [5] embarked upon understanding the similarities between the BSC, AWGN, and ESTC channels by developing a distance metric for the M/M/1 queuing system that is analogous to Hamming/Euclidean distances for BSC/AWGN channels – in that it is related to the logarithm of the channel's conditional distribution. Using this distance metric they were able to characterize the zero-rate reliability of the ESTC. This led to our conjecture that an analogous rate-distortion problem with an appropriate distortion measure should lead to an elegant set of **mutual information inequalities**, **closed-form lower bound**, and an elegant "**test-channel**" to illustrate achievability.

Distortion Measure

Over $[0, T]$, consider two point processes $\mathcal{X}$ and $\hat{\mathcal{X}}$. Denote the arrival times of $\mathcal{X}$ by $\{X_i\}$ and the arrival times of $\hat{\mathcal{X}}$ by $\{\hat{X}_i\}$. We remind the reader that the associated counting functions of $\mathcal{X}$ and $\hat{\mathcal{X}}$ are $N_X(t)$ and $N_{\hat{X}}(t)$, respectively. For any two point processes $\mathcal{X}$ and $\hat{\mathcal{X}}$ such that $N_X(T) = N_{\hat{X}}(T)$, and $N_{\hat{X}}(t) \geq N_X(t)$, $\forall t \in [0, T]$ define

$$\mathcal{S} = \mathcal{X} \diamond \hat{\mathcal{X}}$$

as the point process with **inter-arrival times** $\{S_i\}$ given by the induced service times of a FCFS queueing system with $\hat{\mathcal{X}}$ as the input and $\mathcal{X}$ as the output as shown in Figure 1. Note the abuse of notation in that for the process $\mathcal{S}$, the $\{S_i\}$s are the inter-arrival times. Specifically, $S_i = X_i - \max(X_{i-1}, \hat{X}_i)$; see Figure 2 for the validity of this relationship. Note that the counting function for $\mathcal{S}$, $N_S(t)$ is uniquely defined on $[0, \sum_{i=1}^{N_X(T)} S_i]$.

With this definition, we can now define the distortion between any two realizations x and $\hat{\mathrm{x}}$ of point processes $\mathcal{X}$ and $\hat{\mathcal{X}}$. Define

$$d_T(\mathrm{x}, \hat{\mathrm{x}}) = \begin{cases} \frac{1}{T}\sum_{i=1} s_i & \text{if } N_X(T) = N_{\hat{X}}(T) \text{ and } N_{\hat{X}}(t) \geq N_X(t),\ \forall t \in [0, T] \\ \infty & \text{otherwise} \end{cases}$$

where $s_i = x_i - \max(x_{i-1}, \hat{x}_i)$ defines a realization s of $\mathcal{S}$.

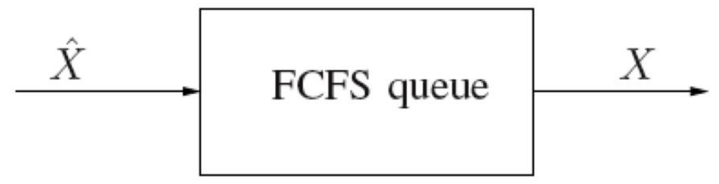

Fig. 1. The Diagram of interpreting $\mathcal{X}$ as the output of a FCFS queueing system with $\hat{\mathcal{X}}$ as the input. The point process $\mathcal{S}$ is defined by treating inter-arrival time S_i as the ith service time.

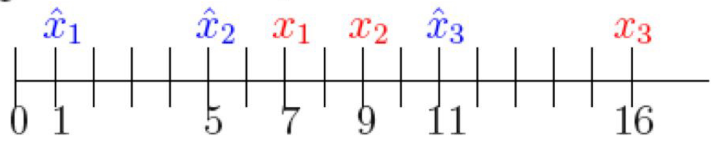

Fig. 2. The arrivals (blue) and departures (red) from a FCFS $\cdot/M/1$ Queue. Note that the service time s_1 for the first packet is given by $x_1 - \hat{x}_1 = 6$. Note that the first packet does not depart from the queue until after the second packet arrives. Thus the service time s_2 for the second packet is given by $s_2 = x_2 - x_1$, because the server starts working on the second packet once the first packet departs. The second packet departs before the third arrival $\hat{x}_3$. Thus the third service time is simply $s_3 = x_3 - \hat{x}_3$. So in general, it follows that $s_i = x_i - \max(\hat{x}_i, x_{i-1})$.

ACHIEVEMENT DESCRIPTION

R(D) of Poisson Processes with Queuing Distortion Measure can be shown in complete analogy to **Bernoulli** source w/ **Hamming distortion** and **Gaussian** source w/ **Euclidean distortion**. Also, **reproduction is a Poisson process**.

$$R(D) = \begin{cases} -\log(\lambda D) \text{ bits/message} & \text{if } 0 < \lambda D \leq 1; \\ 0 & \text{otherwise.} \end{cases}$$

Converse

$$\begin{aligned} I_T(\mathcal{X}; \hat{\mathcal{X}}) &= h_T(\mathcal{X}) - h_T(\mathcal{X}|\hat{\mathcal{X}}) \\ &= \lambda T(1 - \log\lambda) - h_T(\mathcal{X}|\hat{\mathcal{X}}) \qquad (1) \\ &= \lambda T(1 - \log\lambda) - h_T(\mathcal{X} \diamond \hat{\mathcal{X}}|\hat{\mathcal{X}}) \\ &\geq \lambda T(1 - \log\lambda) - h_T(\mathcal{X} \diamond \hat{\mathcal{X}}) \\ &= \lambda T(1 - \log\lambda) - h_T(\mathcal{S}) \\ &\geq \lambda T(1 - \log\lambda) - \mu\left(\frac{\lambda T}{\mu}\right)(1 - \log\mu) \qquad (2) \\ &= \lambda T \log\left(\frac{\mu}{\lambda}\right) = -\lambda T \log(D) \end{aligned}$$

$$\Rightarrow \frac{1}{T} I_T(\mathcal{X}; \hat{\mathcal{X}}) \geq -\lambda \log(D)$$

where (1) follows because $\mathcal{X}$ is a Poisson process with rate λ, and (2) follows because of the argument below.

Note the following properties for $\mathcal{S}$, namely,

1) $\mathcal{S} = \mathcal{X} \diamond \hat{\mathcal{X}}$ is a point process with an average of λT spikes; and
2) $\frac{1}{T} E\left[\sum_i S_i\right] \leq D = \frac{\lambda}{\mu}$, and thus $E\left[\sum_i S_i\right] \leq \frac{\lambda T}{\mu}$.

Thus $h_T(\mathcal{S})$ is upper-bounded by the entropy of the maximum-entropy point process among all point processes of duration $\frac{\lambda T}{\mu}$ and rate at most μ, which is a Poisson process on $[0, \frac{\lambda T}{\mu}]$ of rate μ.

Achievability and Test-Channel

Generate codewords $\hat{\mathcal{X}}$ according to a Poisson process with rate (λ) on $[0, T]$. For any $\hat{\mathcal{X}}$, consider the output of an ESTC with rate $\mu > \lambda$ in steady-state with input process $\hat{\mathcal{X}}$, and denote the output process as $\mathcal{X}$. Note that by Burke's theorem the departure process $\mathcal{X}$ is also a Poisson process with rate λ and thus, by defining $D = \frac{\lambda}{\mu}$, we have that

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T} I_T(\mathcal{X}; \hat{\mathcal{X}}) &= \lim_{T\to\infty} \frac{1}{T}[h_T(\mathcal{X}) - h_T(\mathcal{X}|\hat{\mathcal{X}})] \\ &= \lim_{T\to\infty} \frac{1}{T}[\lambda T(1 - \log\lambda) - h_T(\mathcal{X} \diamond \hat{\mathcal{X}}|\hat{\mathcal{X}})] \\ &= \lambda(1 - \log\lambda) - \lim_{T\to\infty} \frac{1}{T} h_T(\mathcal{S}|\hat{\mathcal{X}}) \\ &= \lambda(1 - \log\lambda) - \lim_{T\to\infty} \frac{1}{T} h_T(\mathcal{S}) \\ &= \lambda(1 - \log\lambda) - \lim_{T\to\infty} \frac{1}{T}\mu\left(\frac{\lambda T}{\mu}\right)(1 - \log\mu) \\ &= \lambda(1 - \log\lambda) - \lambda(1 - \log\mu) = \lambda\log\left(\frac{\mu}{\lambda}\right) = -\lambda\log D. \end{aligned}$$

Note that this immediately suggests the following coding scheme. For a chosen $D \in (0, 1)$ and $T \in (0, +\infty)$, given a realization x of Poisson process $\mathcal{X}$ with rate λ, generate $2^{TR(D)}$ independent realizations of a rate λ Poisson process $\hat{\mathcal{X}}$ denoting the i^{th} realization by $\hat{\mathrm{x}}(i)$ for $i = 1, 2, \ldots, 2^{TR(D)}$. Now choose $\hat{\mathrm{x}}(i^*)$ as the reproduction of x where $i^* \in \arg\min_{i=1,2,\ldots,2^{TR(D)}} d_T(\mathrm{x}, \hat{\mathrm{x}}(i))$.

DISCUSSION

Structure of Rate-Distortion function same as Verdu [1] and Bedekar [6]. However:

- Verdu's result [1] appears to focus more on **extremal properties of the exponential distribution** than on queuing. Reproductions are **not Poisson** or even a renewal process.
- Timing constraints introduced here different than Verdu [1] and Bedekar [6]: reproductions must **lead** original
- Our proof technique, as compared to Bedekar [6], uses point process entropy, the maximum entropy property of the Poisson process. and mutual information inequalities to show similarity to BSC/AWGN problems.

IMPLICATIONS, EXTENSIONS, AND FUTURE WORK

- Extend this Gaussian/Bernoulli analogy to **meaningful multi-user settings** of interest to MANETs
- Explore deeper **isomorphism** of lossy source coding theorems involving **maximum-entropy distributions**
- Develop **practical** lossy source codes of point processes via:

a) **duality** between compression and channel coding

b)Recently developed **practical capacity-approaching codes** [7] for queuing channels

REFERENCES

[1] S. Verdu, "The exponential distribution in information theory", *Problems of Information Transmission*, vol 32, no. 1, pp 86-95, Jan-Mar 1996.

[2] J. McFadden, "The entropy of a point process", *SIAM Journal of Applied Mathematics*, vol 13, pp 988-994, 65.

[3] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication Revisited", *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

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

[5] A. B. Wagner and V. Anantharam, "Zero-rate reliability of the exponential-server timing channel," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 447–465, February 2005.

[6] A. S. Bedekar, "On the information about message arrival times required for in-order decoding," in Proceedings of the *IEEE International Symposium on Information Theory (ISIT)*, Washington, D.C., 2001, p. 227

[7] T. P. Coleman and N. Kiyavash, "Capacity-Approaching Practical Codes for Queueing Channels: An Algebraic, State-Space, Message-Passing Approach", to appear, *UCSD Information Theory & Applications (ITA) Workshop*, Jan 2008