Data Rate Theorem for Stabilization over
Time-Varying Feedback Channels

Workshop on Frontiers in Distributed
Communication, Sensing and Control

Massimo Franceschetti, UCSD

(joint work with P. Minero, S. Dey and G. Nair)
Motivation
Abstraction

DYNAMICAL SYSTEM → OBSERVER/ESTIMATOR → CHANNEL → CONTROLLER/ACTUATOR

Quality channel

Time
Packet loss point of view
Packet loss point of view

- Rate process:

\[ R_k = \begin{cases} \infty & \text{w.p. } 1 - p \\ 0 & \text{w.p. } p \end{cases} \]

- There is a critical dropout probability \( p \) for estimation (Sinopoli et al 2004) and control (Schenato et al 2007, Gupta et al 2007):

\[ p < \frac{1}{\max_i |\lambda_i|^2}. \]
Rate limited feedback point of view
Rate limited feedback point of view

• Rate process:

\[ R_k = R \]

\[ R \]

Rate Process

Time

• Data rate theorem (Tatikonda-Mitter 2002):

\[ R > \sum_{u} m_u \log_2 |\lambda_u| \]
Can we merge the two approaches?
Can we merge the two approaches?

- **Stochastic time-varying rate limited feedback**: rate process \( \{ R_k \} \) i.i.d. according to a random variable \( R \).

- **Special cases**: \( R \) is deterministic, \( R \) is 0 or \( \infty \)
Problem formulation

- A set of unstable eigenmodes \( |\lambda_1| \geq 1, \cdots, |\lambda_n| \geq 1 \).

- Problem: conditions on \( R \) to achieve second moment stability:

\[
\sup_k \mathbb{E} \left[ \| x_k \|^2 \right] < \infty.
\]
Problem formulation

- Distributions of \( x_0 \) and system disturbances (\( v \) and \( w \)):
  - Unbounded support
  - A bounded moment \( > 2 \) (e.g. Gaussian distribution)
Problem formulation

- Distributions of $x_0$ and system disturbances ($v$ and $w$):
  - Unbounded support
  - A bounded moment $> 2$ (e.g. Gaussian distribution)

- Rate process:
  - At time $k$, coder and decoder have knowledge of $\{R_i\}_{i=0}^k$.
  - The case of $R_k = 0$ can be thought as a packet loss with acknowledgment at the transmitter.
Problem formulation

- Distributions of $x_0$ and system disturbances ($v$ and $w$):
  - Unbounded support
  - A bounded moment $> 2$ (e.g. Gaussian distribution)

- Rate process:
  - At time $k$, coder and decoder have knowledge of $\{R_i\}_{i=0}^k$.
  - The case of $R_k = 0$ can be thought as a packet loss with acknowledgment at the transmitter.

- Coder has access to more information than the decoder (Classical information structure).
### Related works

<table>
<thead>
<tr>
<th></th>
<th>Decoding errors</th>
<th>Disturbances</th>
<th>Time-varying rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tatikonda-Mitter ((Tran AC 02))</td>
<td>No</td>
<td>Bounded</td>
<td>No</td>
</tr>
<tr>
<td>Nair-Evans ((SIAM J.C. 04))</td>
<td>No</td>
<td>Unbounded</td>
<td>No</td>
</tr>
<tr>
<td>Yuksel-Basar ((CDC 05))</td>
<td>Yes</td>
<td>Unbounded</td>
<td>No</td>
</tr>
<tr>
<td>Martins-Dahleh-Elia ((Tran AC 06))</td>
<td>No</td>
<td>Bounded</td>
<td>Yes</td>
</tr>
<tr>
<td>Sahai-Mitter ((Tran IT 06))</td>
<td>Yes</td>
<td>Bounded</td>
<td>Yes</td>
</tr>
<tr>
<td>This work</td>
<td>No</td>
<td>Unbounded</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Summary of results

- **Scalar case:**
  - Necessary and sufficient condition for second moment stabilizability.
Summary of results

- **Scalar case:**
  - Necessary and sufficient condition for second moment stabilizability.

- **Vector case:** we ‘sandwich’ the stabilizability region between two polytopes.
  - **Necessity:**
    - The outer polytope has a special geometric structure.
  - **Sufficiency:**
    - The inner and outer polytopes coincide in some special cases.
Scalar case

- Unstable system, so $|\lambda| \geq 1$.

**Theorem 1.** Necessary and sufficient condition for stabilization in the mean square sense is that

$$\mathbb{E} \left[ \frac{|\lambda|^2}{2^2 R} \right] < 1.$$
Scalar case: \( \mathbb{E} \left[ \frac{|\lambda|^2}{2^2R} \right] < 1 \)

- At each time step, uncertainty volume

\[
\uparrow |\lambda|^2, \quad \downarrow \frac{1}{2^2R}
\]

If the ratio, on average, is \( \geq 1 \), then the information sent cannot compensate the system dynamics.
Scalar case: $\mathbb{E} \left[ \frac{|\lambda|^2}{2^R} \right] < 1$

- At each time step, uncertainty volume

$$\uparrow |\lambda|^2, \quad \downarrow \frac{1}{2^R}$$

If the ratio, on average, is $\geq 1$, then the information sent cannot compensate the system dynamics.

- **Deterministic rate**: we recover the data rate theorem

$$R > \log_2 |\lambda|.$$
Scalar case: $\mathbb{E} \left[ \frac{\left| \lambda \right|^2}{2^2R} \right] < 1$

- At each time step, uncertainty volume

$$\uparrow \left| \lambda \right|^2, \quad \downarrow \frac{1}{2^2R}$$

If the ratio, on average, is $\geq 1$, then the information sent cannot compensate the system dynamics.

- **Deterministic rate**: we recover the data rate theorem

$$R > \log_2 \left| \lambda \right|.$$  

- **Packet loss**: we recover the critical dropout probability

$$\mathbb{E} \left[ \frac{\left| \lambda \right|^2}{2^2R} \right] = p \frac{\left| \lambda \right|^2}{2^0} + (1 - p) \frac{\left| \lambda \right|^2}{2^2r} < 1 \Rightarrow p < \frac{1}{\left| \lambda \right|^2}, \quad \text{as } r \to \infty$$
Scalar case: proofs

- **Necessity:** Using the entropy power inequality we establish the following recursion

\[
\mathbb{E}[x_k^2] \geq \mathbb{E} \left[ \frac{|\lambda|^2}{2^{2R}} \right] \mathbb{E}[x_{k-1}^2] + \text{const.},
\]

Thus,

\[
\sup_k \mathbb{E}[x_k^2] < \infty \Rightarrow \mathbb{E} \left[ \frac{|\lambda|^2}{2^{2R}} \right] < 1.
\]

- **Sufficiency:** Design an adaptive quantizer, avoid saturation, high resolution through successive refinements.
Successive refinements

- Divide time into cycles of *fixed* duration $\tau$ (of our choice)
- Observe the system at the beginning of each cycle and send an initial estimate
- During the remaining of a cycle, ‘refine’ the initial estimate.

Number of bits per cycle is a random variable, dependent on the rate process.
Successive refinements

- Quantizer can be generated *recursively* from a 2-bit quantizer, as follows:

- $\rho$ is a parameter chosen dependent on the distribution of the disturbances.
Successive refinements: example

- Suppose we need to quantize a positive real value \( x \).

- At time \( k = 1 \), suppose \( R_1 = 1 \).

- With one bit of information, the decoder knows that \( x > 0 \).
Successive refinements: example

- At time $k = 2$, suppose $R_2 = 2$: coder and decoder partition the real axis according to a 3-bit quantizer.

- However, coder and decoded only ‘label’ the partitions in the positive real line (2 bits suffice).

- After receiving ‘01’, the decoder knows that $x \in \left[ \frac{1}{2}, 1 \right)$. The initial estimate $x \in [0, \infty)$ has been refined.

- The scheme works as if we had known ahead of time that $R_1 + R_2 = 3$. 
Sufficiency: analysis

We find a recursion for $E[x_{k\tau}^2]$

$$E[x_{k\tau}^2] \leq \text{const} \left( E \left[ \frac{|\lambda|^2}{2^2R} \right] \right)^\tau E[x_{(k-1)\tau}^2] + \text{const.}$$

Thus, for $\tau$ large enough

$$E \left[ \frac{|\lambda|^2}{2^2R} \right] < 1 \Rightarrow \text{const} \left( E \left[ \frac{|\lambda|^2}{2^2R} \right] \right)^\tau < 1.$$
Vector case

- $n$ unstable sub-mode associated to eigenvalues $|\lambda_1| \geq 1$, $|\lambda_2| \geq 1$, $\cdots$, $|\lambda_n| \geq 1$.

- Sub-mode $i$ has multiplicity $m_i$.

- $(A, B)$ is reachable and $(C, A)$ observable.
Theorem 2. Necessary condition for stabilizability in the mean square sense is that \((\log_2 |\lambda_1|, \ldots, \log_2 |\lambda_n|) \in \mathbb{R}^n_+\) are inside the region determined by the following set of inequalities

\[
\sum_{i=1}^{n} s_i \log_2 |\lambda_i| < -\sum_{i=1}^{n} s_i \log_2 \mathbb{E} \left[ 2^{-\sum_{i=1}^{n} s_i R} \right],
\]

where \(s = [s_1, \cdots, s_n]^T \neq 0\), \(s_i \in \{0, \ldots, m_i\}\), \(i = 1, \cdots, n\).
Vector case: necessity

Example 1. Suppose

\[ A = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} , \]

so \( m_1 = 2 \) and \( m_2 = 1 \).

\[ \log |\lambda_2| < -\frac{1}{2} \log_2 \mathbb{E} \left[ 2^{-2R} \right] \]

\[ 2 \log |\lambda_1| + \log |\lambda_2| < -\frac{3}{2} \log_2 \mathbb{E} \left[ 2^{-\frac{2}{3}R} \right] \]

\[ 2 \log |\lambda_1| < -\log_2 \mathbb{E} \left[ 2^{-R} \right] \]
Special case: deterministic rate

• We recover the necessity of the data rate theorem

\[ 2 \log |\lambda_1| + \log |\lambda_2| < R. \]

• Polyhedron reduces to a hyperplane
Special case: packet erasure

- We recover the necessity on the critical dropout probability

\[ p < \frac{1}{\max_i |\lambda_i|^2}. \]

- Polyhedron reduces to a hypercube
Vector case: sufficiency

- From the scalar case result we can achieve the two red points.
Vector case: sufficiency

- We can then time share between the two red points.

- Can we do better?

\[ \log_2 |\lambda_1| \]

\[ \log_2 |\lambda_2| \]

Stabilizable region

Not stabilizable
Sufficiency: rate allocation

**Theorem 3.** Sufficient condition for stabilizability in the mean square sense is that \((\log_2 |\lambda_1|, \ldots, \log_2 |\lambda_n|) \in \mathbb{R}_+^n\) are inside the convex hull of the regions determined by the following \(n\) inequalities

\[
\mathbb{E} \left[ \frac{|\lambda_i|^2}{2^2 \alpha_i(R)} \right] < 1, \quad i = 1, \ldots, n,
\]

where the rate allocation vector \(\alpha(R) := [\alpha_1(R), \ldots, \alpha_n(R)]^T\) satisfies

\[
\begin{cases}
\alpha_i(r) \in [0, 1] \\
\frac{r}{m_i} \alpha_i(r) \in \mathbb{N} \\
\sum_{i=1}^n \alpha_i(r) \leq 1
\end{cases}
\]

for all possible values \(r\) that \(R\) can take (excluding 0).
Sufficiency: rate allocation

Example 2. Suppose the rate process is given by

\[ R = \begin{cases} 6 \text{ w.p. } 1 - p \\ 0 \text{ w.p. } p, \end{cases} \]

and, as in the previous example,

\[ A = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}. \]

There are four (dominant) possible allocations \( \alpha(R) = [\alpha_1(R), \alpha_2(R)]^T \):

\[ [\alpha_1(6), \alpha_2(6)]^T \in \left\{ \begin{bmatrix} 6 \\ 6' \end{bmatrix}, \begin{bmatrix} 4 \\ 6, 6' \end{bmatrix}, \begin{bmatrix} 2 \\ 6, 6' \end{bmatrix}, \begin{bmatrix} 0, 6' \end{bmatrix} \right\}, \]
Sufficiency: example

\[ \log_2 |\lambda_2| \]

\[ \log_2 |\lambda_1| \]

- One point on the dominant face of the pentagon is optimal
Sufficiency: remarks

- One point on the dominant face of the pentagon is optimal

- If $R = \begin{cases} r \text{ w.p. } 1 - p \\ 0 \text{ w.p. } p \end{cases}$, scheme asymptotically optimal as $r \to \infty$

- Scheme is optimal if pentagon reduces to hypercube or hyperplane
Remarks

- Above theorems provide polytopic inner and outer bounds to stabilizability region.

- Improved coding scheme for the case of an erasure channel.
Concluding Remarks

• Summary:
  ○ Unified approach to information-theoretic and packet loss models.
  ○ Second moment stabilization with stochastic time-varying rates.
    - Scalar case: complete characterization.
    - Vector case: geometric structure, scheme is optimal in some limiting cases.

• Future directions:
  ○ Close the gap in the vector case
  ○ Model decoding errors