Bounds on the Information Rate of Markov Channels with Free-Running Continuous State

Luca Barletta†, Maurizio Magarini*, Simone Pecorino*, Arnaldo Spalvieri†
† Technische Universität München, * Politecnico di Milano

System Model

The dynamical system is based on the state transition equation

\[ S_k = f_{S,i}(S_{k-1}, V_{k-1}) \]

and on the measurement equation

\[ Y_k = h(S_k, N_k) \]

where \( V_k \) is the process noise, \( N_k \) is the measurement noise, \( S_k \) is the state process, \( Y_k \) is the measurement process, \( \{ f_{S,i}(\cdot, \cdot) \} \) and \( \{ h(\cdot, \cdot) \} \) are sequences of known functions.

By Markov property of process \( S_k \), and since process \( Y_k \) is memoryless given \( S_k \), we have the factorization

\[ p(S_k, Y_k) = p(S_k) \prod_{i=1}^{k} p(Y_i | S_i, S_{i-1}) \]

Shannon’s mutual information rate between \( S_k \) and \( Y_k \) is

\[ I(S; Y) = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} \left( \log \left( \frac{p(Y_i | S_i, S_{i-1})}{p(Y_i)} \right) - \log \left( \frac{p(S_i | S_{i-1})}{p(S_i)} \right) \right) \]

Bayesian Inference

Based on channel’s observations, one can track the hidden state by a two-step recursion:

\[ p(x_k | y_1^{k-1}) = \int p(x_k | a_k, a_{k-1}) p(a_k | y_1^{k-1}) d a_k \]

\[ p(a_k | y_1^{k-1}) = \frac{p(y_k | a_k, a_{k-1}) p(a_k | y_1^{k-1})}{\sum_{i=1}^{N} p(a_k | y_1^{k-1})} \]

If the functions \( \{ f_{S,i}(\cdot, \cdot) \} \) and \( \{ h(\cdot, \cdot) \} \) are affine then the Kalman filter is the solution to the recursion.

In general, the solution to the two-step recursion is unique, and to make the problem tractable some approximations to the actual probabilities are used.

The normalization factor in (6), \( p(x_k | y_1^{k-1}) \), is the term needed to compute \( h(Y_k) \).

Bounds based on Bayesian Inference

Upper bound based on Bayesian filtering

The upper bound is

\[ \mathcal{T}_S(Y) = \mathcal{T}(Y) - h(Y) S \geq I(S; Y) \]

Lower bound based on Bayesian smoothing

The lower bound is

\[ \mathcal{S}(Y) \leq I(S; Y) \]

Due to nonlinearity of functions \( \{ f_{S,i}(\cdot, \cdot) \} \) and \( \{ h(\cdot, \cdot) \} \), the actual distributions (5) and (6) can be multimodal. In Bayesian inference, a particle list \( \{ w_i^{(k)}, \tilde{a}_i^{(k)} \} \) is a common nonparametric method for representing \( p(x_k | y_1^{k-1}) \), where \( \tilde{a}_i^{(k)} \) is the number of particles. Specifically, (5) is substituted by

\[ \tilde{w}_i^{(k)} = p(y_i | a_k, a_{k-1}) \]

\[ w_i^{(k)} = w_i^{(k)} \tilde{w}_i^{(k)} \sum_{j=1}^{N} \tilde{w}_j^{(k)} \]

where \( \tilde{w}_i^{(k)} \) means drawn with probability, and (6) by

\[ w_i^{(k)} = w_i^{(k)} \tilde{w}_i^{(k)} \sum_{j=1}^{N} \tilde{w}_j^{(k)} \]

Channels with Free-Running State

Consider a communication channel described by the joint probability

\[ p(x_k^N, y_k^N) = p(y_k | \tilde{a}_i^{(k)}), p(y_k | a_k, a_{k-1}) \]

where \( R \) is the channel output process and \( X \) the source process. Using the chain rule for mutual information we have

\[ I(X; R) = I(X; R; S) + I(R; S; X) \]

and

\[ I(X; R; S) = I(X; \tilde{a}_i^{(k)}; R) \]

\[ I(R; S; X) = I(R; S) ; I(R; X) \]

or using the differential entropy rates as

\[ I(X; R) = I(R); I(R; X) - h(S; X) - h(R; X) \]

\[ I(R; S; X) = I(R; S) + I(R; X) - h(R; X) \]

Discrete-Time ARMA Phase Noise Channels

The model is

\[ \Phi_{k+1} = \Phi + \Delta \sum_{i=1}^{N} a_i X_{k-i} \]

and the Markovian state is \( S_k = (\Phi_{k+1}, \Delta, a_{k-1}, \ldots) \). Example for \( m = 1 \):

![Discrete-Time ARMA Phase Noise Channels](attachment:image.png)

Computing the Bounds by Particle Methods

Due to nonlinearity of functions \( \{ f_{S,i}(\cdot, \cdot) \} \) and \( \{ h(\cdot, \cdot) \} \), the actual distributions (5) and (6) can be multimodal. In Bayesian inference, a particle list \( \{ w_i^{(k)}, \tilde{a}_i^{(k)} \} \) is a common nonparametric method for representing \( p(x_k | y_1^{k-1}) \), where \( \tilde{a}_i^{(k)} \) is the number of particles. Specifically, (5) is substituted by

\[ \tilde{w}_i^{(k)} = p(y_i | a_k, a_{k-1}) \]

\[ w_i^{(k)} = w_i^{(k)} \tilde{w}_i^{(k)} \sum_{j=1}^{N} \tilde{w}_j^{(k)} \]

where \( \tilde{w}_i^{(k)} \) means drawn with probability, and (6) by

\[ w_i^{(k)} = w_i^{(k)} \tilde{w}_i^{(k)} \sum_{j=1}^{N} \tilde{w}_j^{(k)} \]

References


Conclusions

Summary:
- Shannon information between the hidden Markov state process of a dynamical system and the measurement process has been computed by the probabilities inferred by Bayesian tracking.
- Upper and lower bounds to the information rate between the hidden state and the measurement can be computed from approximate Bayesian tracking.
- Specific results have been derived for the discrete-time ARMA phase noise channel.

Outlook:
- Bounds for continuous-time channels with free-running continuous state.

Numerical Results

![Numerical Results](attachment:image.png)