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

Motivation (1)

Communication systems can be impaired by data-independent sources with memory, e.g.,

- phase noise
- multiplicative fading
- a combination of the two previous

In general, these channels are referred to as communication channels with free-running hidden Markov state. We consider the case of no channel state information available at the transmitter.

Estimation of information rates transferred through these channels can be challenging because

- the state space is not finite and it is multidimensional, therefore it cannot be approached by trellis-based techniques based on quantization of the state space, because the number of states of the trellis would be enormous
- the observation can be a nonlinear function of the state, therefore the optimum front-end filter can be a complicated nonlinear function of channel's output

Motivation (2)

Our contribution:

- Upper and lower bounds to the information rate between the hidden state and the measurement based on approximated inference
- Application of these bounds to multiplicative communication channels
- Experimental results for the discrete-time autoregressive moving average (ARMA) phase noise channel

System Model

The dynamical system is based on the state transition equation

$$S_k = f_{k-1}(S_{k-1}, V_{k-1}) \qquad (1)$$

and on the measurement equation

$$Y_k = h_k(S_k, N_k). \qquad (2)$$

V is the process noise, N is the measurement noise, S is the state process, Y is the measurement process, and $\{f_{k-1}(\cdot)\}$ and $\{h_k(\cdot)\}$ are sequences of known functions.

By Markov property of process S, and since process Y is memoryless given S, we have the factorization

$$p(s_0^n, y_1^n) = p(s_0) \prod_{k=1}^{n} p(s_k|s_{k-1}) p(y_k|s_k) \qquad (3)$$

Shannon's mutual information rate between S and Y is

$$I(S;Y) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathrm{E}\left\{ \log_2 \left(\frac{p(Y_k|S_k)}{p(Y_k|Y_1^{k-1})} \right) \right\} = h(Y) - h(Y|S) \qquad (4)$$

Bayesian Inference

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

$$p(s_k|y_1^{k-1}) = \int_S p(s_k|s_{k-1}) p(s_{k-1}|y_1^{k-1})\, \mathrm{d}s_{k-1} \qquad (5)$$

$$p(s_k|y_1^k) = \frac{p(s_k|y_1^{k-1}) p(y_k|s_k)}{p(y_k|y_1^{k-1})} \qquad (6)$$

- If the functions $\{f_{k-1}(\cdot)\}$ and $\{h_k(\cdot)\}$ are affine then the Kalman filter is the solution to the recursion
- In general, the solution to the two-step recursion is unknown, and to make the problem treatable some approximations to the actual probabilities are used
- The normalization factor in (6), $p(y_k|y_1^{k-1})$, is the term needed to compute $h(Y)$

Bounds based on Bayesian Inference

Upper bound based on Bayesian filtering

The upper bound is

$$\overline{I}(S;Y) = \overline{h}(Y) - h(Y|S) \geq I(S;Y) \qquad (7)$$

$$\overline{h}(Y) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log_2 \frac{1}{q(y_k|y_1^{k-1})} \geq h(Y) \qquad (8)$$

$q(y_k|y_1^{k-1})$ is an approximation to $p(y_k|y_1^{k-1})$ of (6), and y_1^n is a sequence drawn according to the actual model (3), i.e., by using (1) and (2).

Lower bound based on Bayesian smoothing

The lower bound is

$$\underline{I}(S;Y) = h(S) - \overline{h}(S|Y) \leq I(S;Y) \qquad (9)$$

$$\overline{h}(S|Y) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log_2 \frac{1}{q(s_k|y_k^{k+l}, s_{k-1})} \geq h(S|Y) \qquad (10)$$

$q(s_k|y_k^{k+l}, s_{k-1})$ is the approximation to $p(s_k|y_k^n, s_{k-1})$ worked out by a lag-l Bayesian smoother initialized from the state s_{k-1} visited by the realization (s_1^n, y_1^n) at time $k-1$, the time lag l being up to the user.

Computing the Bounds by Particle Methods

Due to nonlinearity of functions $\{f_{k-1}(\cdot)\}$ and $\{h_k(\cdot)\}$, the actual distributions (5) and (6) can be multimodal. In Bayesian inference, a particle list $\{(s_k^{(i)}, w_k^{(i)})\}_{i=1}^P$ is a common nonparametric method for representing $q(s_k|y_1^{k-1})$, where P is the number of particles. Specifically, (5) is substituted by

$$s_k^{(i)} \sim p(s_k|s_{k-1}^{(i)}), \qquad i = 1, 2, \ldots, P, \qquad (11)$$

where $\sim$ means *drawn with probability*, and (6) by

$$w_k^{(i)} = \frac{w_{k-1}^{(i)} p(y_k|s_k^{(i)})}{\sum_{j=1}^{P} w_{k-1}^{(j)} p(y_k|s_k^{(j)})}, \qquad i = 1, 2, \ldots, P. \qquad (12)$$

Channels with Free-Running State

Consider a communication channel described by the joint probability

$$p(r_1^n, x_1^n, s_0^n) = p(s_0) \prod_{k=1}^{n} p(s_k|s_{k-1}) p(r_k|x_k, s_k) p(x_k), \qquad (13)$$

where R is the channel output process and X the source process.

Using the chain rule for mutual information we have

$$I(X;R) = I(X;R|S) + I(S;R) - I(S;R|X) \qquad (14)$$

therefore $I(X;R)$ can be sandwiched as

$$\overline{I}(X;R) = I(X;R|S) + \overline{I}(S;R) - \underline{I}(S;R|X) \qquad (15)$$
$$\geq I(X;R) \qquad (16)$$
$$\geq I(X;R|S) + \underline{I}(S;R) - \overline{I}(S;R|X) = \underline{I}(R;X), \qquad (17)$$

or using the differential entropy rates as

$$\overline{I}(X;R) = \overline{h}(R) + \overline{h}(S|X,R) - h(S|X) - h(R|X,S) \qquad (18)$$
$$\geq I(X;R) \qquad (19)$$
$$\geq h(S) + h(R|S) - \overline{h}(S|R) - \overline{h}(R|X) = \underline{I}(R;X). \qquad (20)$$

- $\overline{h}(R)$ and $\overline{h}(R|X)$ are evaluated as in (8)
- $\overline{h}(S|X,R)$ and $\overline{h}(S|R)$ are evaluated as in (10)

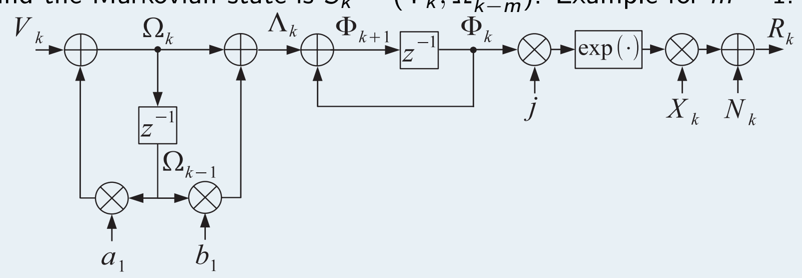
Discrete-Time ARMA Phase Noise Channels

The model is

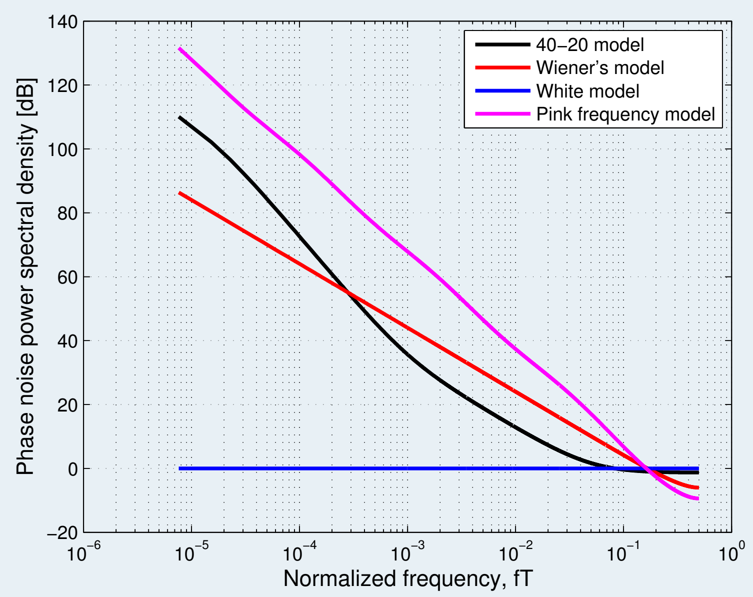
$$R_k = X_k e^{j\Phi_k} + N_k \qquad (21)$$

$$\Phi_{k+1} = \Phi_k + \Omega_k + \sum_{i=1}^{m} b_i \Omega_{k-i}, \qquad \Omega_k = V_k + \sum_{i=1}^{m} a_i V_{k-i}, \qquad (22)$$

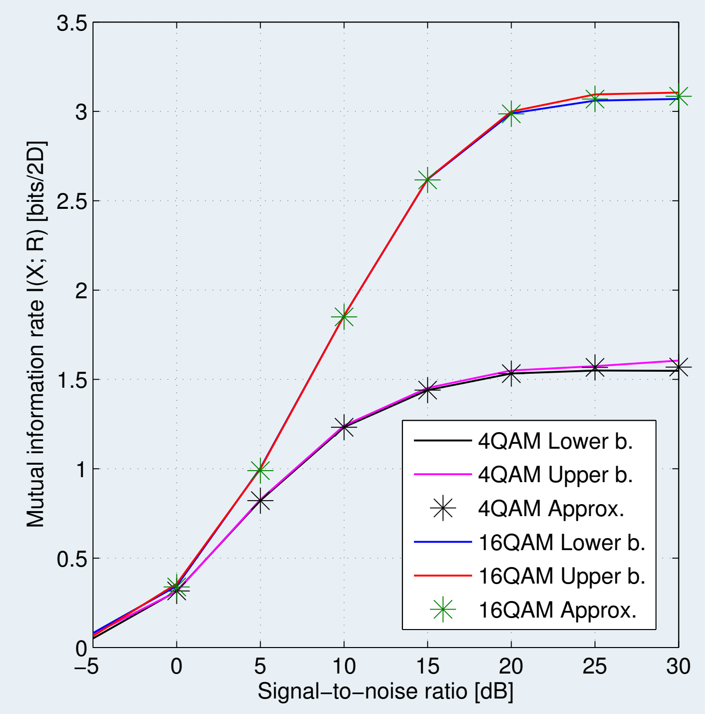
and the Markovian state is $S_k = (\Phi_k, \Omega_{k-m}^{k-1})$. Example for $m = 1$:



Pink frequency model: $a_i = 3 \cdot 4^{-2i}$, $b_i = 3 \cdot 4^{-2i+1}$, $i = 1, \ldots, 4$.



Numerical Results



- Phase noise obtained by accumulation of pink frequency noise with $\mathrm{E}\{V_k^2\} = 0.25$
- Approximation obtained with method of [Dauwels and Loeliger, 2008]

Conclusions

Summary:

- Shannon information between the hidden Markov state process of a dynamical system and the measurement process has been evaluated 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

References

- J. Dauwels and H.-A. Loeliger, "Computation of information rates by particle methods," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 406–409, 2008.
- L. Barletta, M. Magarini, and A. Spalvieri, "The information rate transferred through the discrete-time Wiener's phase noise channel," *J. Lightwave Technol.*, vol. 30, no. 10, pp. 1480–1486, May 2012.
- ——, "Tight upper and lower bounds to the information rate of the phase noise channel," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, July 2013, pp. 2284–2288.
- L. Barletta, M. Magarini, S. Pecorino, and A. Spalvieri, "Upper and lower bounds to the information rate transferred through first-order Markov channels with free-running continuous state," *Submitted to Trans. on Inf. Theory*, 2014.