Parameter Estimation Based on Noisy Chaotic Signals in the Weak-Noise Regime

Neri Merhav

The Viterbi Faculty of Electrical & Computer Engineering
Technion—Israel Institute of Technology
Haifa, Israel

ISIT 2024, July 7-12, 2024
Twisted Modulation

Consider the AWGN channel,

\[ y_i = x_i + z_i, \quad i = 1, 2, \ldots, n, \]

where \( z_i \sim \mathcal{N}(0, \sigma^2) \) are i.i.d. and

\[ x = (x_1, \ldots, x_n) = f_n(\theta), \quad \|x\|^2 \leq nQ \]

\( \theta \in [0, 1] \) being a parameter to be estimated at the receiver by \( \hat{\theta} = g_n(y) \).

How well can we estimate \( \theta \) if we have the freedom to choose both the modulator \( f_n(\cdot) \) and the estimator, \( g_n(\cdot) \)?
Twisted Modulation (Cont’d)

- The “waveform communication” problem (Wozencraft & Jacobs, ‘65).
- Source/channel coding: Shannon-Kotel’nikov (‘49; ‘59).
- Hekland (‘07); Floor (‘08+); Hekland, Floor & Ramstad (‘09, ‘23).
- Estimation theory; Cohn, (‘70), Burnashev (‘84, ‘85).
- Linear modulation – Fisher efficient, but limited.
- Nonlinear modulation – flexible, but suffers a threshold effect.
- Most of the literature: total MSE.
- Reasonable to separate Pr\{anomaly\} and weak-noise errors.
- Köken, Günduz & Tuncel (‘17): \( \min MSE \) s.t. \( \Pr\{\text{anomaly}\} \leq \epsilon \).
- Merhav (‘19): exponential \( \Pr\{\text{anomaly}\} \) + matching converse.
- Merhav (‘20): extension to parameter vectors.
- This work: modulators based on chaotic dynamical systems.
Motivations for Studying Chaotic Modulators

- Sensitivity to initial conditions – good weak-noise estimation.
- High degree of flexibility in the design.
- ∃ mature theoretical understanding about chaos.
- Computationally easy to generate the modulated signal.
- ∃ computationally efficient estimation algorithms (halving method).
- Good estimation of initial condition is also good for filtering.
Related Work

Modulators based on chaotic systems have been investigated extensively during the last 3 decades from a variety of aspects:

- Upper/lower bounds on MSE.
- Numerical aspects.
- Algorithmic efficiency.
- System optimization.
- Applications in Turbo coding, hybrid coding, spread spectrum, MIMO, etc.

Chen (‘96); Chen & Wornell (‘98); Cong et al. (‘99); Drake (‘98); Eckmann & Ruelle (‘85); Hen & Merhav (‘04); Kay & Nagesha (‘95); Kennedy & Kolumbán (‘00); Leung et al. (‘06); Pantaleón et al. (‘03); Papadopoulos & Wornell (‘95); Wallinger (‘13); Wang et al. (‘99); Xie et al. (‘09); Yu et al. (‘18), .......
Objectives

The purpose this work is to carry out a systematic study of modulators that are based on certain class of chaotic systems, from the perspective of earlier work on fundamental limits of general modulators:

Given a certain parametric family of modulators, find the one with the best weak-noise error performance s. t. \( \Pr\{\text{anomaly}\} \to 0 \).

We consider a general error performance criterion,

\[
\sup_{\theta \in [0,1]} \mathbb{E} \left\{ \rho(\hat{\theta} - \theta) \right| \text{no anomaly} \} \quad \rho(\cdot) \text{ convex}
\]

and avoid the use of the Cramér-Rao lower bound, which is problematic for systems with discontinuous mappings.
The Chaotic Dynamical System

\[ st = f(st-1) \]

\[ s0 = \theta \]

\[ r = 3 \]

\[ f(s) \]

\[ 1 \]

\[ p(0) \quad p(1) \quad p(2) \quad 1 \]

\[ s \]

\[ DELAY \]
Formulation

Select a positive integer $r$ and a probability vector, $P = \{p(0), p(1), \ldots, p(r - 1)\}$. Define:

$$F(x) = \sum_{x' = 0}^{x-1} p(x'); \quad F(0) = 0, \ F(r) = 1.$$ 

Given $s \in [0, 1]$, let $\phi(s)$ be the value of $x \in \{0, 1, \ldots, r - 1\}$ such that

$$F(x) \leq s < F(x + 1), \quad \phi(1) = 1.$$ 

The non-linear dynamical system is defined by the recursion:

$$x_t = \phi(s_{t-1}), \quad s_0 = \theta \quad \text{itinerary sequence}$$

$$s_t = \frac{s_{t-1} - F(x_t)}{p(x_t)} \quad \text{state sequence}$$

for $t = 1, 2, \ldots$. The channel input is

$$u_t = \sqrt{12Q} \left( s_t - \frac{1}{2} \right).$$
1. **Reconstruction of** $s_0$ **from** $x_1, x_2, \ldots$:

$$s_0 = \sum_{t=1}^{\infty} F(x_t) \prod_{i=1}^{t-1} p(x_i) \triangleq \sum_{t=1}^{\infty} G(x_t) \prod_{i=1}^{t} p(x_i)$$

**Example:** if $P = (1/r, \ldots, 1/r)$,

$$s_0 = \sum_{t=1}^{\infty} x_t r^{-t} = 0.x_1x_2 \ldots.$$ 

2. **$X_t$ as a random process** (with application to process simulation): If $S_0 \sim \text{Unif}[0, 1]$, then $S_t \sim \text{Unif}[0, 1]$ and $\{X_t\}$ is a DMS governed by $P$. (Easy extension to arbitrary processes with memory).

3. **Lyapunov exponent:**

$$\lambda \triangleq \lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left\{ \ln \left| \frac{\partial^{s_n}}{\partial s_0} \right| \right\} = H \quad \text{the entropy of } P.$$
4. **Length of signal locus:**

\[ L_n = \int_0^1 \left\| \frac{\partial}{\partial \theta} u(\theta) \right\| \cdot d\theta = r^n. \]

Therefore \( \ln r \) must be \( < C = \frac{1}{2} \ln(1 + \gamma), \gamma = \frac{Q}{\sigma^2} \) to keep \( \Pr\{\text{anomaly}\} \) small.

5. **Autocorrelation:**

\[
R_S(k) = \mathbb{E}\{S_0 S_k\} = \frac{1}{4} + \frac{1}{12} \cdot \left( \sum_{x=0}^{r-1} p^2(x) \right)^{|k|}.
\]

\[
R_U(k) = \mathbb{E}\{U_0 U_k\} = Q \cdot \left( \sum_{x=0}^{r-1} p^2(x) \right)^{|k|}.
\]
6. Channel input-output mutual information:

\[ C_0 = \lim_{n \to \infty} \frac{I(U^n; Y^n)}{n} \leq \frac{1}{2} \ln \frac{A}{\sigma^2} \triangleq C_1 \]

where

\[ \frac{A}{\sigma^2} = \frac{1}{2} \left[ 1 + \gamma + q^2(1 - \gamma) + \sqrt{(1 + \gamma^2)(1 - q^2)^2 + 2\gamma(1 - q^4)} \right] \]

and

\[ q = \sum_{x=0}^{r-1} p^2(x). \]

7. Ergodic property:

The Lebesgue measure of the set:

\[ \left\{ s_0 : \left| \frac{1}{n} \sum_{t=1}^{n} u_t u_{t+k} - R_U(k) \right| \leq \epsilon \right\} \]

tends to unity as \( n \to \infty. \)
General Lower Bound

Suppose that $\rho(\cdot)$ has the following property: $\forall \ c > 0, \ \rho(e^{-nc}) = e^{-n\zeta(c)}$ with $\zeta(c) > 0$.

For example, if $\rho(\epsilon) = |\epsilon|^a$, then $\zeta(c) = a \cdot c$.

Theorem [Merhav 2019]: For any modulator and estimator,

$$\sup_{0 \leq \theta \leq 1} \mathbb{E} \left\{ \rho(\hat{\theta} - \theta) \left| \text{no anomaly} \right. \right\} \geq \exp \left\{ -n\zeta \left( \frac{1}{2} \ln \gamma \right) \right\}$$

for $\gamma \gg 1$.

Asymptotically achievable by uniform quantization of $\theta$ followed by capacity-achieving channel coding.
**Lower Bound for the Class of Chaotic Modulators**

**Theorem:** For any chaotic modulator from the class defined and any estimator,

\[
\sup_{0 \leq \theta \leq 1} E \left\{ \rho(\hat{\theta} - \theta) \right| \text{no anomaly} \right\} \\
\geq \exp \left( -n\zeta \left[ \min \left\{ C_1, \frac{1}{2} \ln \gamma - \frac{1}{2} \ln \left( \frac{2\pi e}{12} \right) \right\} \right) 
\]

for \( \gamma \gg 1 \).

The term \( \frac{1}{2} \ln \left( \frac{2\pi e}{12} \right) \) = shaping loss since \( \{u_t\} \) is distributed uniformly rather than normally.

\( C_1 < \frac{1}{2} \ln \gamma \) = loss associated with the fact that \( \{u_t\} \) has memory.

We could have increased \( C_1 \) up to \( C \) by decreasing \( q = \sum_x p^2(x) \geq \frac{1}{r} \), but recall that \( r \) is limited by \( \ln r \leq C \).
Feeding the Channel by the Itinerary Signal?

At first glance it seems counterintuitive that the itinerary sequence, \( \{x_t\} \) could do a better job than the state sequence, \( \{s_t\} \) (or \( \{u_t\} \)) since \( \{x_t\} \) is a quantized version of \( \{s_t\} \):

\[
x_t = \phi(s_{t-1}), \quad F(x_t) \leq s_{t-1} < F(x_t + 1).
\]

However, recall that \( \{s_t\} \) is generated from \( s_0 \) which in turn can be expressed in terms of \( (x_1, x_2, \ldots) \). Thus, \( (x_1, x_2, \ldots) \) and \( (s_1, s_2, \ldots) \) include exactly the same information about \( s_0 \).

- For large \( n \), \( (x_1, x_2, \ldots, x_n) \) and \( (s_1, s_2, \ldots, s_n) \) and include almost the same information about \( s_0 \).
- If \( S_0 \sim \text{Unif}[0, 1] \), \( \{X_t\} \) is an i.i.d. process governed by \( P \). No loss due to input memory.
- No limitation on \( r \): select \( P \) to approximate the capacity-achieving input distribution, \( \mathcal{N}(0, Q) \).
Proposed Modulation Scheme

- Given $\theta \in [0, 1]$, quantize it to $\theta_i$ using a fine grid of $M = e^{n(C-\epsilon)}$ points and a random mapping of the grid onto itself, $\eta_i = \psi(\theta_i)$.
- Let $s_0 = \eta_i$ be the initial state of the modulator.
- Transmit $x$ over the channel.
- Decode $\hat{\eta}_i$.
- $\hat{\theta}_i = \psi^{-1}(\hat{\eta}_i)$.

If the decoding is correct (which is the case w.h.p.), there is only a quantization error:

$$\rho(\theta_i - \theta) \leq \rho \left( \frac{1}{2M} \right) = \rho(e^{-n(C-\epsilon)}) = e^{-n\zeta(C-\epsilon)}.$$ 

Decoding error = anomaly.