On Optimum Parameter Modulation–Estimation From a Large Deviations Perspective

Neri Merhav

Communications and Information Theory Seminar, May 3, 2012.

Special thanks to Yariv Ephraim for many useful discussions.

Thanks also to Tsachy Weissman and Yonina Eldar for interesting conversations.

Background

Consider the model

$$y(t) = x(t, u) + n(t), \quad 0 \leq t < T,$$

where:

$x(t, u) =$ a waveform parametrized by u;

$n(t) =$ AWGN with spectral density $N_0/2$.

Background

Consider the model

$$y(t) = x(t, u) + n(t), \quad 0 \leq t < T,$$

where:

$x(t, u) =$ a waveform parametrized by u;

$n(t) =$ AWGN with spectral density $N_0/2$.

Conveying info in a parameter u by modulating in $x(t, u)$:

Shannon–Kotel'nikov mappings (Floor '08, Floor & Ramstad '09, Hekland '07, Ramstad '02 + references).

Background

Consider the model

$$y(t) = x(t, u) + n(t), \quad 0 \leq t < T,$$

where:

$x(t, u) =$ a waveform parametrized by u;

$n(t) =$ AWGN with spectral density $N_0/2$.

Conveying info in a parameter u by modulating in $x(t, u)$:

Shannon–Kotel'nikov mappings (Floor '08, Floor & Ramstad '09, Hekland '07, Ramstad '02 + references).

Nonlinear modulation $\Longrightarrow$ threshold effect:

Below some critical SNR, anomalous errors dominate the MSE.

Background (Cont'd) - The Threshold Effect

- Not an artifact of a particular modulator–estimator pair.
- In the wideband regime, the threshold effect is abrupt: $\Pr\{\text{anomaly}\}$ jumps from ~ 0 to ~ 1.

Background (Cont'd) - The Threshold Effect

- Not an artifact of a particular modulator–estimator pair.
- In the wideband regime, the threshold effect is abrupt: Pr{anomaly} jumps from ~ 0 to ~ 1.

In the linear model

$$y(t) = u \cdot s(t) + n(t), \quad 0 \leq t < T$$

the ML estimator always achieves

$$\text{MSE} = \text{CRLB} = \frac{N_0}{2\mathcal{E}},$$

where $\mathcal{E} =$ energy of $\{s(t)\}$: $\Leftrightarrow$ No threshold effect.

Background (Cont'd) - The Threshold Effect

- Not an artifact of a particular modulator–estimator pair.
- In the wideband regime, the threshold effect is abrupt: Pr$\{$anomaly$\}$ jumps from ~ 0 to ~ 1.

In the linear model

$$y(t) = u \cdot s(t) + n(t), \quad 0 \leq t < T$$

the ML estimator always achieves

$$\text{MSE} = \text{CRLB} = \frac{N_0}{2\mathcal{E}},$$

where $\mathcal{E} =$ energy of $\{s(t)\}$: $\Leftrightarrow$ No threshold effect.

Only way to improve (at high SNR): non–linear modulation $x(t, u)$.

Background (Cont'd) – Nonlinear Modulation

Let

$$x(t, u) \approx x(t, u_0) + (u - u_0) \cdot \dot{x}(t, u_0).$$

like the linear case with $\dot{x}(t, u_0)$ in the role of $s(t)$.

Background (Cont'd) – Nonlinear Modulation

Let

$$x(t,u) \approx x(t,u_0) + (u-u_0)\cdot \dot{x}(t,u_0).$$

like the linear case with $\dot{x}(t,u_0)$ in the role of $s(t)$. Thus, at high SNR,

$$\text{MSE} \approx \text{CRLB} \approx \frac{N_0}{2\dot{\mathcal{E}}},$$

where $\dot{\mathcal{E}} =$ energy of $\dot{x}(t,u_0)$, which depends on more details.

Background (Cont'd) – Nonlinear Modulation

Let

$$x(t,u) \approx x(t,u_0) + (u-u_0)\cdot \dot{x}(t,u_0).$$

like the linear case with $\dot{x}(t,u_0)$ in the role of $s(t)$. Thus, at high SNR,

$$\text{MSE} \approx \text{CRLB} \approx \frac{N_0}{2\dot{\mathcal{E}}},$$

where $\dot{\mathcal{E}} =$ energy of $\dot{x}(t,u_0)$, which depends on more details.

For example, if $x(t,u) = s(t-u)$, $\dot{\mathcal{E}} = W^2\mathcal{E}$, where

$$W = \sqrt{\frac{1}{\mathcal{E}}\int_{-\infty}^{\infty} \mathrm{d}f \cdot (2\pi f)^2 S(f)} \qquad \text{Gabor bandwidth}$$

Background (Cont'd) – Nonlinear Modulation

Let

$$x(t,u) \approx x(t,u_0) + (u-u_0)\cdot \dot{x}(t,u_0).$$

like the linear case with $\dot{x}(t,u_0)$ in the role of $s(t)$. Thus, at high SNR,

$$\text{MSE} \approx \text{CRLB} \approx \frac{N_0}{2\dot{\mathcal{E}}},$$

where $\dot{\mathcal{E}} =$ energy of $\dot{x}(t,u_0)$, which depends on more details.

For example, if $x(t,u) = s(t-u)$, $\dot{\mathcal{E}} = W^2\mathcal{E}$, where

$$W = \sqrt{\frac{1}{\mathcal{E}}\int_{-\infty}^{\infty} \mathrm{d}f \cdot (2\pi f)^2 S(f)} \qquad \text{Gabor bandwidth}$$

Why not increase W without a limit?

Background (Cont'd) – Geometry of Anomalous Errors

Let $\bar{x}(u) = (x_1(u), \ldots, x_K(u)) =$ representation of $x(t, u)$ by K orthonormal basis functions. Consider the locus of $\{\bar{x}(u),\ a \le u \le b\}$ in $\mathbb{R}^K$.

Background (Cont'd) – Geometry of Anomalous Errors

Let $\bar{x}(u) = (x_1(u), \ldots, x_K(u))$ = representation of $x(t, u)$ by K orthonormal basis functions. Consider the locus of $\{\bar{x}(u),\ a \leq u \leq b\}$ in $\mathbb{R}^K$.

Assuming that $\mathcal{E}$ is independent of u, the locus lies on the hypersurface of the K–dimensional sphere of radius $\sqrt{\mathcal{E}}$.

Background (Cont'd) – Geometry of Anomalous Errors

Let $\bar{x}(u) = (x_1(u), \ldots, x_K(u))$ = representation of $x(t,u)$ by K orthonormal basis functions. Consider the locus of $\{\bar{x}(u),\ a \le u \le b\}$ in $\mathbb{R}^K$.

Assuming that $\mathcal{E}$ is independent of u, the locus lies on the hypersurface of the K–dimensional sphere of radius $\sqrt{\mathcal{E}}$.

The length of the curve

$$L = \int_a^b \mathrm{d}u \sqrt{\sum_i \dot{x}_i^2(u)} = (b-a)\sqrt{\dot{\mathcal{E}}}.$$

Background (Cont'd) – Geometry of Anomalous Errors

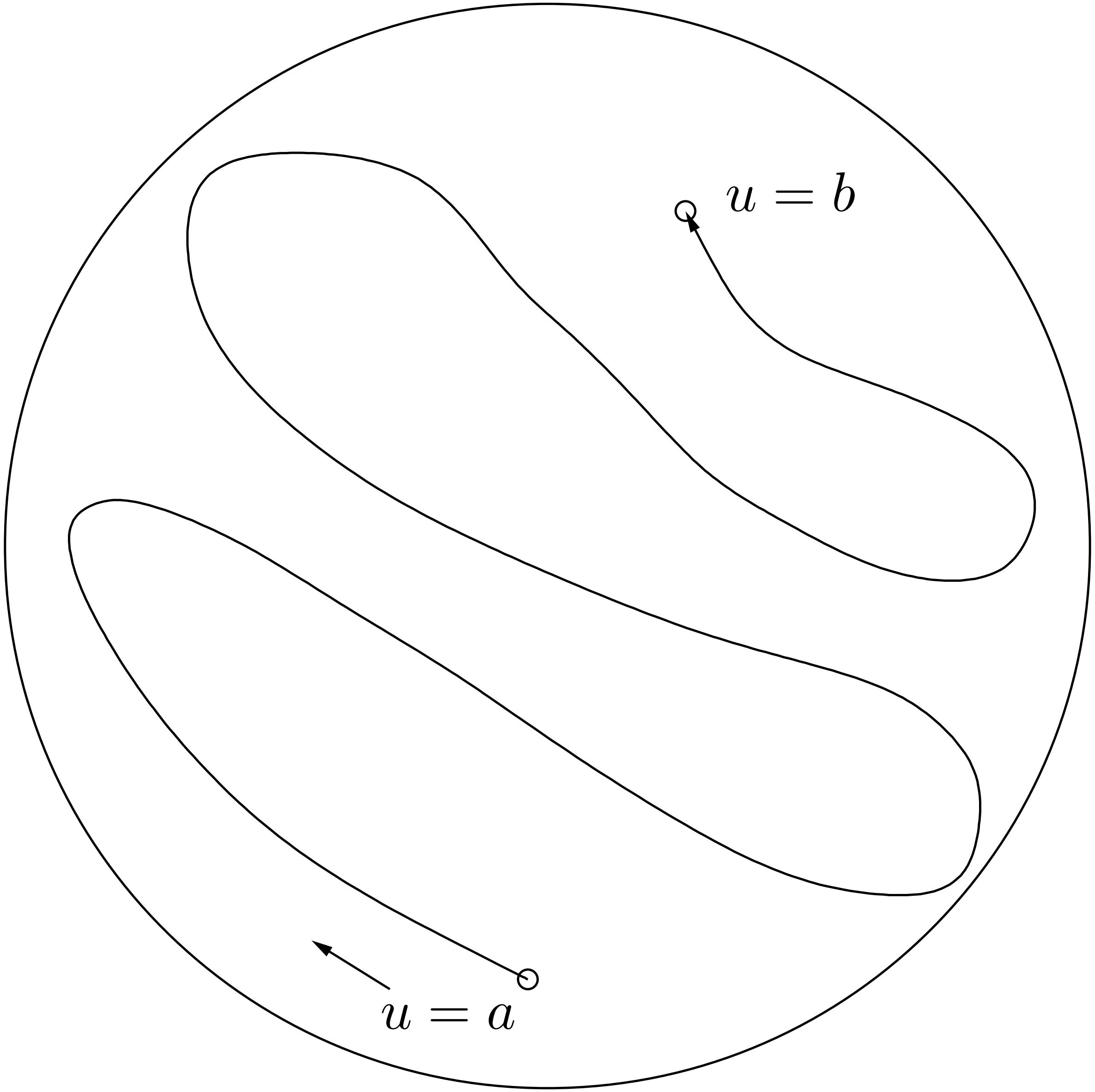
Let $\bar{x}(u) = (x_1(u), \ldots, x_K(u))$ = representation of $x(t,u)$ by K orthonormal basis functions. Consider the locus of $\{\bar{x}(u),\ a \le u \le b\}$ in $\mathbb{R}^K$.

Assuming that $\mathcal{E}$ is independent of u, the locus lies on the hypersurface of the K–dimensional sphere of radius $\sqrt{\mathcal{E}}$.

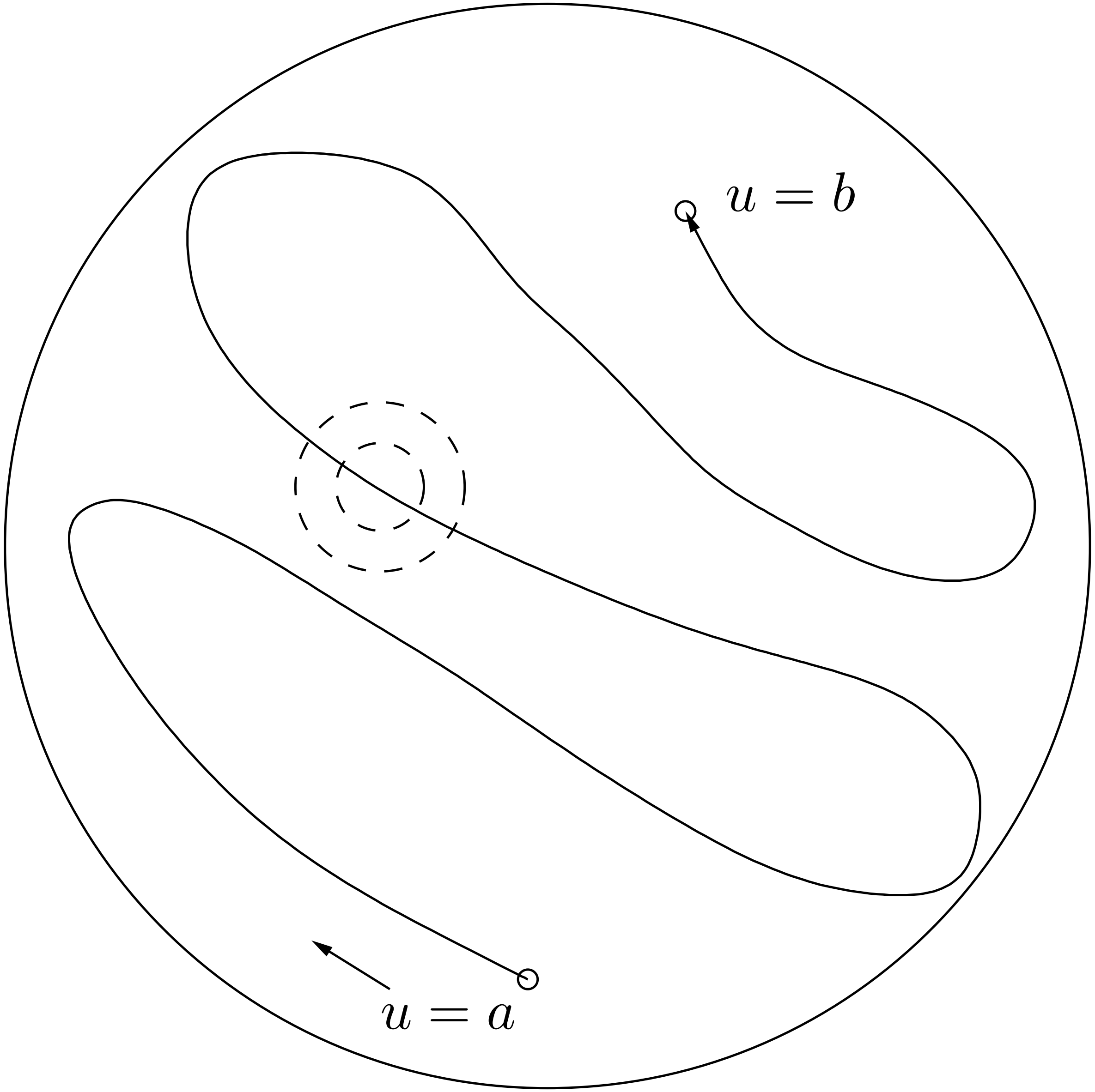
The length of the curve

$$L = \int_a^b \mathrm{d}u \sqrt{\sum_i \dot{x}_i^2(u)} = (b-a)\sqrt{\dot{\mathcal{E}}}.$$

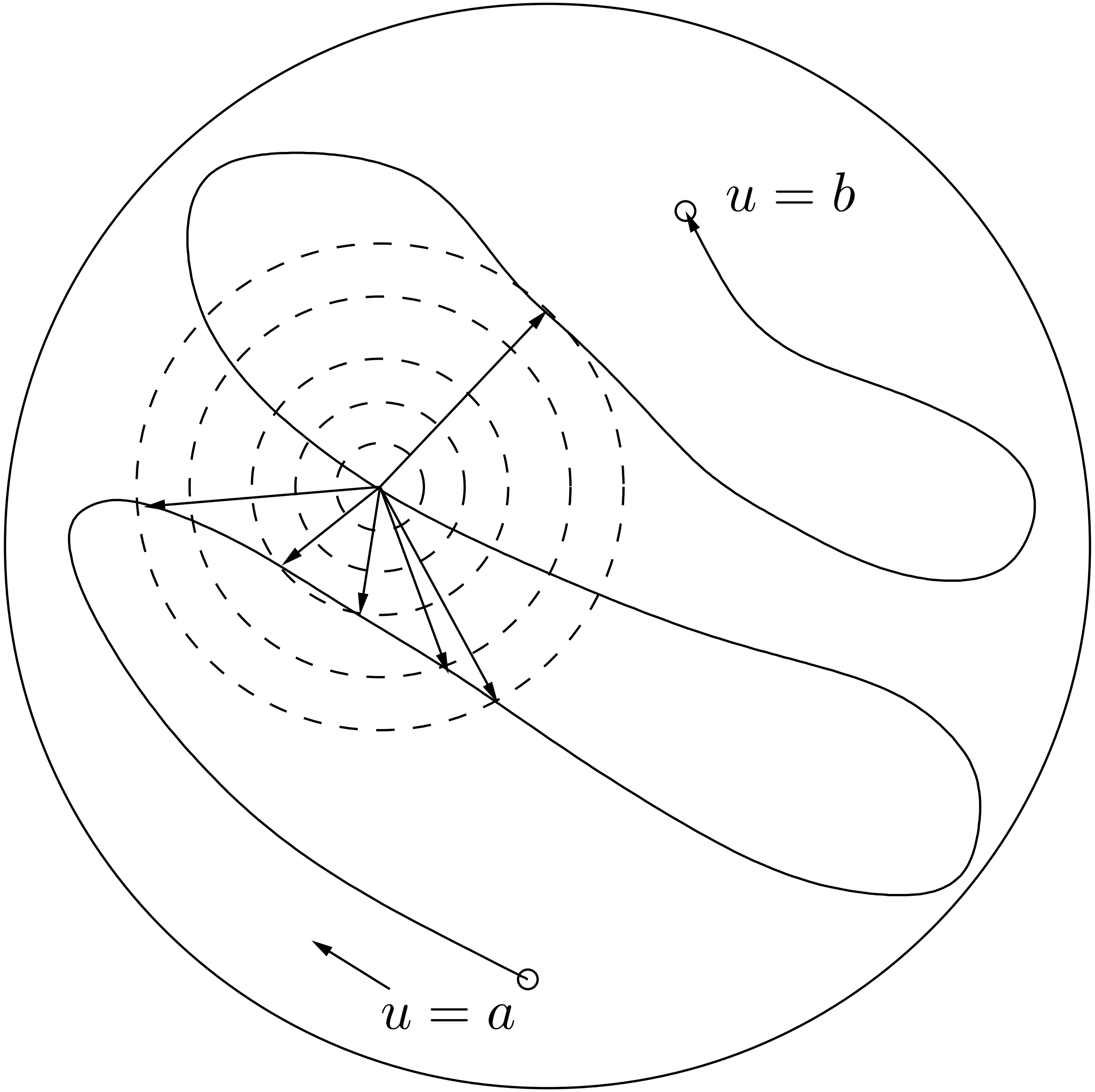
High–SNR MSE $\downarrow$ with $\dot{\mathcal{E}}$, we want $\dot{\mathcal{E}}\ \uparrow$, thus $L\ \uparrow$.



$u = b$
$u = a$



$u = b$
$u = a$


$u = b$
$u = a$

Anomalous Errors (Cont'd)

L – limited by the need of safe distances between folds – hot dog packing.

Anomalous Errors (Cont'd)

L – limited by the need of safe distances between folds – hot dog packing. Maximum achievable $L \sim e^{CT}$, $C = S/N_0$ (PPM). For PPM, $K \sim 2WT$,

$$\text{MSE} \approx \underbrace{\frac{N_0}{2W^2\mathcal{E}}}_{\text{small error}} + \underbrace{(b-a)^2 \cdot 2WT \cdot e^{-\mathcal{E}/(2N_0)}}_{\text{anomalous error}}$$

For fixed W, anomalous error term $\uparrow$ gracefully as $S/N_0 \downarrow$.

Anomalous Errors (Cont'd)

L – limited by the need of safe distances between folds – hot dog packing. Maximum achievable $L \sim e^{CT}$, $C = S/N_0$ (PPM). For PPM, $K \sim 2WT$,

$$\text{MSE} \approx \underbrace{\frac{N_0}{2W^2\mathcal{E}}}_{\text{small error}} + \underbrace{(b-a)^2 \cdot 2WT \cdot e^{-\mathcal{E}/(2N_0)}}_{\text{anomalous error}}$$

For fixed W, anomalous error term $\uparrow$ gracefully as $S/N_0 \downarrow$.
For a better balance between terms – let $W \sim e^{RT}$.

$$\text{MSE} \approx \frac{N_0}{2\mathcal{E}} e^{-2RT} + (b-a)^2 \cdot e^{-TE(R)} \qquad R < C$$

where $E(R) =$ reliability function of AWGN channel:

$$E(R) = \begin{cases} \frac{C}{2} - R & 0 \le R \le \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \le R \le C \end{cases}$$

Anomalous Errors (Cont'd)

L – limited by the need of safe distances between folds – hot dog packing.
Maximum achievable $L \sim e^{CT}$, $C = S/N_0$ (PPM). For PPM, $K \sim 2WT$,

$$\text{MSE} \approx \underbrace{\frac{N_0}{2W^2\mathcal{E}}}_{\text{small error}} + \underbrace{(b-a)^2 \cdot 2WT \cdot e^{-\mathcal{E}/(2N_0)}}_{\text{anomalous error}}$$

For fixed W, anomalous error term $\uparrow$ gracefully as $S/N_0 \downarrow$.
For a better balance between terms – let $W \sim e^{RT}$.

$$\text{MSE} \approx \frac{N_0}{2\mathcal{E}} e^{-2RT} + (b-a)^2 \cdot e^{-TE(R)} \qquad R < C$$

where $E(R) =$ reliability function of AWGN channel:

$$E(R) = \begin{cases} \frac{C}{2} - R & 0 \le R \le \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \le R \le C \end{cases}$$

Optimum compromise: $R = C/6 \quad \Longrightarrow \quad \text{MSE} \sim e^{-CT/3}$.

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).
- Most are useful for a given modulator (exception: ZZ-CZZ, DPT bounds).

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).
- Most are useful for a given modulator (exception: ZZ-CZZ, DPT bounds).
- Quest for universal bounds, independent of both modulator and estimator.

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).
- Most are useful for a given modulator (exception: ZZ-CZZ, DPT bounds).
- Quest for universal bounds, independent of both modulator and estimator.
- The modulator – limited by a power constraint only (power $\leq S$).

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).
- Most are useful for a given modulator (exception: ZZ-CZZ, DPT bounds).
- Quest for universal bounds, independent of both modulator and estimator.
- The modulator – limited by a power constraint only (power $\leq S$).
- Bounds that depend only on $\mathcal{E}/N_0 = CT$, $C = S/N_0$.

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).
- Most are useful for a given modulator (exception: ZZ-CZZ, DPT bounds).
- Quest for universal bounds, independent of both modulator and estimator.
- The modulator – limited by a power constraint only (power $\leq S$).
- Bounds that depend only on $\mathcal{E}/N_0 = CT$, $C = S/N_0$.
- We saw that MSE $\sim e^{-CT/3}$ is achievable for $T \to \infty$.

Lower Bounds on the MSE

- Plenty of lower bounds (CR, BZ, ZZ, CZZ, WW, ChRo, FG, BT, etc.)
- Some fail to capture the threshold effect (exception: WW).
- Most are useful for a given modulator (exception: ZZ-CZZ, DPT bounds).
- Quest for universal bounds, independent of both modulator and estimator.
- The modulator – limited by a power constraint only (power $\leq S$).
- Bounds that depend only on $\mathcal{E}/N_0 = CT$, $C = S/N_0$.
- We saw that MSE $\sim e^{-CT/3}$ is achievable for $T \to \infty$.
- Is there a compatible lower bound?

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$.

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D)$$

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D)$$

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

implies $\quad D = \mathbf{E}(\hat{U} - U)^2 \geq \dfrac{1}{2\pi e} \cdot e^{-2CT}.$

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

implies $\quad D = \mathbf{E}(\hat{U} - U)^2 \geq \frac{1}{2\pi e} \cdot e^{-2CT}.$

Zakai and Ziv (1975) – generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \overset{\sim}{\geq} e^{-CT}$.

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

$$\text{implies} \quad D = \mathbf{E}(\hat{U} - U)^2 \geq \frac{1}{2\pi e} \cdot e^{-2CT}.$$

Zakai and Ziv (1975) – generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \gtrsim e^{-CT}$.

Merhav (2011) – other generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \gtrsim e^{-2CT/3}$

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

$$\text{implies} \quad D = \mathbf{E}(\hat{U} - U)^2 \geq \frac{1}{2\pi e} \cdot e^{-2CT}.$$

Zakai and Ziv (1975) – generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT}$.

Merhav (2011) – other generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-2CT/3}$

$+$ bounds related to signal detection considerations: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT/2}$.

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi eD) = -\frac{1}{2}\ln(2\pi eD)$$

$$\text{implies} \quad D = \mathbf{E}(\hat{U} - U)^2 \geq \frac{1}{2\pi e} \cdot e^{-2CT}.$$

Zakai and Ziv (1975) – generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT}$.

Merhav (2011) – other generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-2CT/3}$

$+$ bounds related to signal detection considerations: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT/2}$.

Compare to the upper bound of $e^{-CT/3}$.

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

$$\text{implies} \quad D = \mathbf{E}(\hat{U} - U)^2 \geq \frac{1}{2\pi e} \cdot e^{-2CT}.$$

Zakai and Ziv (1975) – generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT}$.

Merhav (2011) – other generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-2CT/3}$

$+$ bounds related to signal detection considerations: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT/2}$.

Compare to the upper bound of $e^{-CT/3}$.

Attempts to close the gap have failed thus far...

Some Universal MSE Lower Bounds

Let u be a realization of $U \sim$ Unif$[-1/2, +1/2)$. Then,

$$CT \geq R(D) \geq h(U) - \frac{1}{2}\ln(2\pi e D) = -\frac{1}{2}\ln(2\pi e D)$$

$$\text{implies} \quad D = \mathbf{E}(\hat{U} - U)^2 \geq \frac{1}{2\pi e} \cdot e^{-2CT}.$$

Zakai and Ziv (1975) – generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT}$.

Merhav (2011) – other generalized DPI's: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-2CT/3}$

$+$ bounds related to signal detection considerations: $\mathbf{E}(\hat{U} - U)^2 \stackrel{\sim}{\geq} e^{-CT/2}$.

Compare to the upper bound of $e^{-CT/3}$.

Attempts to close the gap have failed thus far...
Conjecture: “Blame” the lower bound.

The Large Deviations Perspective

Instead of $\mathbf{E}(\hat{U} - U)^2$, consider minimizing

$$\mathbf{E}1\{|\hat{U} - U| \geq \Delta\} = \Pr\{|\hat{U} - U| > \Delta\}.$$

The Large Deviations Perspective

Instead of $\mathbf{E}(\hat{U} - U)^2$, consider minimizing

$$\mathbf{E}1\{|\hat{U} - U| \geq \Delta\} = \text{Pr}\{|\hat{U} - U| > \Delta\}.$$

Moreover, we allow $\Delta = e^{-RT}$.

The Large Deviations Perspective

Instead of $\mathbf{E}(\hat{U} - U)^2$, consider minimizing

$$\mathbf{E}1\{|\hat{U} - U| \geq \Delta\} = \text{Pr}\{|\hat{U} - U| > \Delta\}.$$

Moreover, we allow $\Delta = e^{-RT}$.

Assume $u =$ realization of $U \sim$ Unif$[-1/2, +1/2)$, allow any modulator $x(t, \cdot)$ with

$$\mathbf{E}\left\{\frac{1}{T}\int_0^T x^2(t, U)\text{d}t\right\} \leq S$$

and any estimator $\hat{U} = f\{y(t),\ 0 \leq t < T\}$.

The Large Deviations Perspective

Instead of $\mathbf{E}(\hat{U} - U)^2$, consider minimizing

$$\mathbf{E}1\{|\hat{U} - U| \geq \Delta\} = \mathsf{Pr}\{|\hat{U} - U| > \Delta\}.$$

Moreover, we allow $\Delta = e^{-RT}$.

Assume $u =$ realization of $U \sim$ Unif$[-1/2, +1/2)$, allow any modulator $x(t, \cdot)$ with

$$\mathbf{E}\left\{\frac{1}{T}\int_0^T x^2(t, U)\mathrm{d}t\right\} \leq S$$

and any estimator $\hat{U} = f\{y(t),\ 0 \leq t < T\}$.

We are interested in

$$E^*(R) = \limsup_{T\to\infty}\left[-\frac{1}{T}\log \inf \mathsf{Pr}\left\{|\hat{U} - U| > e^{-RT}\right\}\right].$$

Motivation

Separates between anomalous and non–anomalous error events:

Motivation

Separates between anomalous and non–anomalous error events:

- $|\hat{U} - U| \leq e^{-RT}$ – weak–noise (non–anomalous). Error $\sim e^{-RT}$.

Motivation

Separates between anomalous and non–anomalous error events:

- $|\hat{U} - U| \leq e^{-RT}$ – weak–noise (non–anomalous). Error $\sim e^{-RT}$.
- $|\hat{U} - U| > e^{-RT}$ – gross error (anomalous). Error $\sim e^{-TE^*(R)}$.

Motivation

Separates between anomalous and non–anomalous error events:

- $|\hat{U} - U| \leq e^{-RT}$ – weak–noise (non–anomalous). Error $\sim e^{-RT}$.
- $|\hat{U} - U| > e^{-RT}$ – gross error (anomalous). Error $\sim e^{-TE^*(R)}$.
- Typically, anomalous $\hat{U}$ falls at random away from $U \Rightarrow$ weigh all anomalous errors evenly.

Motivation

Separates between anomalous and non–anomalous error events:

- $|\hat{U} - U| \leq e^{-RT}$ – weak–noise (non–anomalous). Error $\sim e^{-RT}$.
- $|\hat{U} - U| > e^{-RT}$ – gross error (anomalous). Error $\sim e^{-TE^*(R)}$.
- Typically, anomalous $\hat{U}$ falls at random away from $U \Rightarrow$ weigh all anomalous errors evenly.

MSE does not distinguish between weak–noise errors and anomalous errors.

Basic Result

Theorem: For all $R>0$, the $\limsup$ of $E^*(R)$ is actually $\lim$ and

$$E^*(R) = E(R) = \begin{cases} \frac{C}{2} - R & 0 \le R \le \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \le R \le C \\ 0 & R \ge C \end{cases}$$

Basic Result

Theorem: For all $R>0$, the $\limsup$ of $E^*(R)$ is actually $\lim$ and

$$E^*(R) = E(R) = \begin{cases} \frac{C}{2} - R & 0 \le R \le \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \le R \le C \\ 0 & R \ge C \end{cases}$$

Achievability

Basic Result

Theorem: For all $R>0$, the $\limsup$ of $E^*(R)$ is actually $\lim$ and

$$E^*(R) = E(R) = \begin{cases} \frac{C}{2} - R & 0 \le R \le \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \le R \le C \\ 0 & R \ge C \end{cases}$$

Achievability

Modulator: Form a grid of $M = e^{RT}/2$ points in $[-1/2, +1/2)$:

$$\{-1/2 + 1 \cdot e^{-RT}, -1/2 + 3 \cdot e^{-RT}, -1/2 + 5 \cdot e^{-RT}, \ldots, 1/2 - e^{-RT}\}.$$

Map grid points to orthogonal signals $s_i(t)$ with power S: $x(t,u) = s_i(t)$, where $i =$ index of grid point NN to u.

Basic Result

Theorem: For all $R > 0$, the $\limsup$ of $E^*(R)$ is actually $\lim$ and

$$E^*(R) = E(R) = \begin{cases} \frac{C}{2} - R & 0 \leq R \leq \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \leq R \leq C \\ 0 & R \geq C \end{cases}$$

Achievability

Modulator: Form a grid of $M = e^{RT}/2$ points in $[-1/2, +1/2)$:

$$\{-1/2 + 1 \cdot e^{-RT}, -1/2 + 3 \cdot e^{-RT}, -1/2 + 5 \cdot e^{-RT}, \ldots, 1/2 - e^{-RT}\}.$$

Map grid points to orthogonal signals $s_i(t)$ with power S: $x(t, u) = s_i(t)$, where $i =$ index of grid point NN to u.

Estimator: Decode $\hat{i}$ and $\hat{u} = -1/2 + (2\hat{i} - 1)e^{-RT}$.

Basic Result

Theorem: For all $R>0$, the $\limsup$ of $E^*(R)$ is actually $\lim$ and

$$E^*(R) = E(R) = \begin{cases} \frac{C}{2} - R & 0 \le R \le \frac{C}{4} \\ (\sqrt{C} - \sqrt{R})^2 & \frac{C}{4} \le R \le C \\ 0 & R \ge C \end{cases}$$

Achievability

Modulator: Form a grid of $M = e^{RT}/2$ points in $[-1/2, +1/2)$:

$$\{-1/2 + 1 \cdot e^{-RT}, -1/2 + 3 \cdot e^{-RT}, -1/2 + 5 \cdot e^{-RT}, \ldots, 1/2 - e^{-RT}\}.$$

Map grid points to orthogonal signals $s_i(t)$ with power S: $x(t,u) = s_i(t)$, where $i =$ index of grid point NN to u.

Estimator: Decode $\hat{i}$ and $\hat{u} = -1/2 + (2\hat{i} - 1)e^{-RT}$.

Obviously,

$$\Pr\{|\hat{U} - U| > e^{-RT}\} \le \Pr\{\hat{i} \ne i\} \sim e^{-TE(R)}.$$

Converse Part

For a given u, consider the grid

$$\{u, u+2e^{-RT}, u+4e^{-RT}, \ldots, u+2(M-1)e^{-RT}\}, \qquad M = \frac{e^{(R-\epsilon)T}}{2}+1$$

Converse Part

For a given u, consider the grid

$$\{u, u + 2e^{-RT}, u + 4e^{-RT}, \ldots, u + 2(M-1)e^{-RT}\}, \qquad M = \frac{e^{(R-\epsilon)T}}{2} + 1$$

and define the hypothesis testing problem:

$$\mathcal{H}_i : \; y(t) = x\left(t, u + 2ie^{-RT}\right) + n(t) \qquad i = 0, 1, \ldots, M-1.$$

Converse Part

For a given u, consider the grid

$$\{u, u+2e^{-RT}, u+4e^{-RT}, \ldots, u+2(M-1)e^{-RT}\}, \qquad M = \frac{e^{(R-\epsilon)T}}{2}+1$$

and define the hypothesis testing problem:

$$\mathcal{H}_i : \; y(t) = x\left(t, u+2ie^{-RT}\right) + n(t) \qquad i = 0, 1, \ldots, M-1.$$

Consider a detector that chooses the grid point NN to $\hat{U}$.

Converse Part

For a given u, consider the grid

$$\{u, u+2e^{-RT}, u+4e^{-RT}, \ldots, u+2(M-1)e^{-RT}\}, \qquad M = \frac{e^{(R-\epsilon)T}}{2} + 1$$

and define the hypothesis testing problem:

$$\mathcal{H}_i : \; y(t) = x\left(t, u + 2ie^{-RT}\right) + n(t) \qquad i = 0, 1, \ldots, M-1.$$

Consider a detector that chooses the grid point NN to $\hat{U}$. Obviously,

$$P_e \geq e^{-T[E(R-\epsilon)+o(T)]}.$$

Converse Part

For a given u, consider the grid

$$\{u, u + 2e^{-RT}, u + 4e^{-RT}, \ldots, u + 2(M-1)e^{-RT}\}, \qquad M = \frac{e^{(R-\epsilon)T}}{2} + 1$$

and define the hypothesis testing problem:

$$\mathcal{H}_i : \; y(t) = x\left(t, u + 2ie^{-RT}\right) + n(t) \qquad i = 0, 1, \ldots, M-1.$$

Consider a detector that chooses the grid point NN to $\hat{U}$. Obviously,

$$\frac{1}{M}\sum_{i=0}^{M-1} \text{Pr}\{|\hat{U} - U| > e^{-RT} | U = u + 2ie^{-RT}) \geq P_e \geq e^{-T[E(R-\epsilon)+o(T)]}.$$

Converse Part

For a given u, consider the grid

$$\{u, u + 2e^{-RT}, u + 4e^{-RT}, \ldots, u + 2(M-1)e^{-RT}\}, \qquad M = \frac{e^{(R-\epsilon)T}}{2} + 1$$

and define the hypothesis testing problem:

$$\mathcal{H}_i : \; y(t) = x\left(t, u + 2ie^{-RT}\right) + n(t) \qquad i = 0, 1, \ldots, M-1.$$

Consider a detector that chooses the grid point NN to $\hat{U}$. Obviously,

$$\frac{1}{M}\sum_{i=0}^{M-1} \text{Pr}\{|\hat{U} - U| > e^{-RT} | U = u + 2ie^{-RT}) \geq P_e \geq e^{-T[E(R-\epsilon)+o(T)]}.$$

The result is obtained by integrating both sides over u.

The Case $R = 0$

The operational reliability – discontinuous at $R = 0$. For fixed M, P_e is dictated by $d_{\min} = \frac{2M\mathcal{E}}{M-1}$. In particular,

$$P_e \propto Q\left(\sqrt{\frac{\mathcal{E}}{N_0} \cdot \frac{M}{M-1}}\right) \sim \exp\left(-\frac{CT}{2} \cdot \frac{M}{M-1}\right).$$

The Case $R = 0$

The operational reliability – discontinuous at $R = 0$. For fixed M, P_e is dictated by $d_{\min} = \frac{2M\mathcal{E}}{M-1}$. In particular,

$$P_e \propto Q\left(\sqrt{\frac{\mathcal{E}}{N_0} \cdot \frac{M}{M-1}}\right) \sim \exp\left(-\frac{CT}{2} \cdot \frac{M}{M-1}\right).$$

In estimation, Pr$\{|\hat{U} - U| > \Delta\}$ for fixed Δ, is according to P_e but with $M \propto 1/\Delta$.

The Case $R = 0$

The operational reliability – discontinuous at $R = 0$. For fixed M, P_e is dictated by $d_{\min} = \frac{2M\mathcal{E}}{M-1}$. In particular,

$$P_e \propto Q\left(\sqrt{\frac{\mathcal{E}}{N_0} \cdot \frac{M}{M-1}}\right) \sim \exp\left(-\frac{CT}{2} \cdot \frac{M}{M-1}\right).$$

In estimation, Pr$\{|\hat{U} - U| > \Delta\}$ for fixed Δ, is according to P_e but with $M \propto 1/\Delta$.

Small gap between upper bound and the lower bound for every fixed Δ, but this gap $\to 0$ as $\Delta \to 0$. In particular,

$$\lim_{\Delta \to 0} \lim_{T \to \infty} \left[-\frac{\ln \text{Pr}\{|\hat{U} - U| > \Delta\}}{T}\right] = \frac{C}{2} = E(0).$$

The Case $R = 0$ (Cont'd)

Relation to the MSE:

$$\mathbf{E}(\hat{U} - U)^2 = 2 \int_0^1 \mathsf{d}\Delta \cdot \Delta \cdot \mathsf{Pr}\{|\hat{U} - U| \geq \Delta\}.$$

The Case $R = 0$ (Cont'd)

Relation to the MSE:

$$\mathbf{E}(\hat{U} - U)^2 = 2\int_0^1 \mathrm{d}\Delta \cdot \Delta \cdot \Pr\{|\hat{U} - U| \geq \Delta\}.$$

MSE can be lower bounded via a lower bound on $\Pr\{|\hat{U} - U| \geq \Delta\}$ (CZZ '75).

The Case $R = 0$ (Cont'd)

Relation to the MSE:

$$\mathbf{E}(\hat{U} - U)^2 = 2\int_0^1 \mathsf{d}\Delta \cdot \Delta \cdot \mathsf{Pr}\{|\hat{U} - U| \geq \Delta\}.$$

MSE can be lower bounded via a lower bound on Pr$\{|\hat{U} - U| \geq \Delta\}$ (CZZ '75).

Weakness: integration range of Δ – restricted to $[0, 1/(M-1)]$.

The Case $R = 0$ (Cont'd)

Relation to the MSE:

$$\mathbf{E}(\hat{U} - U)^2 = 2\int_0^1 \mathsf{d}\Delta \cdot \Delta \cdot \mathsf{Pr}\{|\hat{U} - U| \geq \Delta\}.$$

MSE can be lower bounded via a lower bound on Pr$\{|\hat{U} - U| \geq \Delta\}$ (CZZ '75).

Weakness: integration range of Δ – restricted to $[0, 1/(M-1)]$.

Large deviations performance metric avoids integration altogether.

The Case $R = 0$ (Cont'd)

Relation to the MSE:

$$\mathbf{E}(\hat{U} - U)^2 = 2\int_0^1 \mathrm{d}\Delta \cdot \Delta \cdot \mathrm{Pr}\{|\hat{U} - U| \geq \Delta\}.$$

MSE can be lower bounded via a lower bound on Pr$\{|\hat{U} - U| \geq \Delta\}$ (CZZ '75).

Weakness: integration range of Δ – restricted to $[0, 1/(M-1)]$.

Large deviations performance metric avoids integration altogether.

Open question: devise a system independent of Δ, yet minimizes Pr$\{|\hat{U} - U| \geq \Delta\}$ for every Δ.

Discussion

Strong Converse $\Leftrightarrow$ Sharp Threshold Effect

- Both achievability and converse rely on signal detection considerations.
- Strong converse: $\lim_{T \to \infty} P_e$ jumps from 0 to 1 as R crosses C.
- Equivalently, $E^*(R) = 0$ for $R > C$ in the strong sense.
- “Inheriting” strong converse — jump in Pr$\{|\hat{U} - U| > e^{-RT}\}$.
- For an optimum system, $|\hat{U} - U|$ “concentrates” around e^{-CT}.

Alternative Achievability Schemes

Achievability – quantization of U + orthogonal signaling.

Alternative Achievability Schemes

Achievability – quantization of U + orthogonal signaling.

Alternative modulators (+ ML estimation):

FPM: $x(t,u) = \sqrt{2S}\cos[(\omega_0 + u \cdot \Delta\omega)t + \phi] \quad \omega_0, \Delta\omega \propto e^{RT}$.

Alternative Achievability Schemes

Achievability – quantization of U + orthogonal signaling.

Alternative modulators (+ ML estimation):

FPM: $x(t,u) = \sqrt{2S}\cos[(\omega_0 + u \cdot \Delta\omega)t + \phi]$ $\quad \omega_0, \Delta\omega \propto e^{RT}$.

PPM: $x(t,u) = s[t - (u + 1/2)(T - \tau)]$ bandwidth $\propto e^{RT}$.

Alternative Achievability Schemes

Achievability – quantization of U + orthogonal signaling.

Alternative modulators (+ ML estimation):

FPM: $x(t,u) = \sqrt{2S}\cos[(\omega_0 + u \cdot \Delta\omega)t + \phi] \quad \omega_0, \Delta\omega \propto e^{RT}$.

PPM: $x(t,u) = s[t - (u + 1/2)(T - \tau)]$ bandwidth $\propto e^{RT}$.

In both, the event $\{|\hat{U} - U| > e^{-RT}\} =$ anomaly.

Alternative Achievability Schemes

Achievability – quantization of U + orthogonal signaling.

Alternative modulators (+ ML estimation):

FPM: $x(t,u) = \sqrt{2S}\cos[(\omega_0 + u \cdot \Delta\omega)t + \phi] \quad \omega_0, \Delta\omega \propto e^{RT}$.

PPM: $x(t,u) = s[t - (u + 1/2)(T - \tau)]$ bandwidth $\propto e^{RT}$.

In both, the event $\{|\hat{U} - U| > e^{-RT}\}$ = anomaly.

Common feature: correlation between $x(t,u)$ and $x(t,u')$ depends only on $|u - u'|$ with support $\sim e^{-RT}$.

Alternative Achievability Schemes

Achievability – quantization of U + orthogonal signaling.

Alternative modulators (+ ML estimation):

FPM: $x(t,u) = \sqrt{2S}\cos[(\omega_0 + u \cdot \Delta\omega)t + \phi] \quad \omega_0, \Delta\omega \propto e^{RT}$.

PPM: $x(t,u) = s[t - (u + 1/2)(T - \tau)]$ bandwidth $\propto e^{RT}$.

In both, the event $\{|\hat{U} - U| > e^{-RT}\} =$ anomaly.

Common feature: correlation between $x(t,u)$ and $x(t,u')$ depends only on $|u - u'|$ with support $\sim e^{-RT}$.

In AM:

$$\Pr\{|\hat{U} - U| > e^{-RT}\} = 2Q(e^{-RT}\sqrt{2CT}) \to 1 \quad \forall\, R > 0$$

Relation to Moments of the Estimation Error

By Chebyshev's inequality

$$e^{-T[E(R)+o(T)]} \leq \text{Pr}\{|\hat{U} - U| > e^{-RT}\} \leq \frac{\mathbf{E}(\hat{U} - U)^2}{e^{-2RT}}$$

Relation to Moments of the Estimation Error

By Chebyshev's inequality

$$e^{-T[E(R)+o(T)]} \leq \text{Pr}\{|\hat{U} - U| > e^{-RT}\} \leq \frac{\mathbf{E}(\hat{U} - U)^2}{e^{-2RT}}$$

implying that

$$\mathbf{E}(\hat{U} - U)^2 \geq e^{-T[E(R)+2R+o(T)]},$$

which is maximized for $R = 0$, yielding $\sim e^{-CT/2}$ as above.

Relation to Moments of the Estimation Error

By Chebyshev's inequality

$$e^{-T[E(R)+o(T)]} \leq \text{Pr}\{|\hat{U} - U| > e^{-RT}\} \leq \frac{\mathbf{E}(\hat{U} - U)^2}{e^{-2RT}}$$

implying that

$$\mathbf{E}(\hat{U} - U)^2 \geq e^{-T[E(R)+2R+o(T)]},$$

which is maximized for $R = 0$, yielding $\sim e^{-CT/2}$ as above.

For a general moment $\mathbf{E}|\hat{U} - U|^\alpha$ ($\alpha > 0$, arbitrary):

$$\mathbf{E}|\hat{U} - U|^\alpha \geq \begin{cases} e^{-CT/2} & \alpha \geq 1 \\ e^{-\alpha CT/(1+\alpha)} & 0 < \alpha < 1 \end{cases}$$

Relation to Joint Source–Channel Coding

Csiszár (1982): JSC problem under

$$\min \mathsf{Pr} \left\{ \sum_{i=1}^{N} d(U_i, \hat{U}_i) > ND \right\}.$$

Relation to Joint Source–Channel Coding

Csiszár (1982): JSC problem under

$$\min \Pr \left\{ \sum_{i=1}^{N} d(U_i, \hat{U}_i) > ND \right\}.$$

The exponential rate cannot exceed

$$e(D) = \min_R [F(D,R) + E(R)]$$

where

$$F(D,R) = \min_{Q':\ R(D,Q') \geq R} D(Q' \| Q)$$

is the source coding exponent of the source Q (Marton, 1974).

Relation to Joint Source–Channel Coding (Cont'd)

For separate source– and channel coding:

$$e_{sep}(D) = \sup_R \min\{F(D,R), E(R)\} \leq e(D)$$

Relation to Joint Source–Channel Coding (Cont'd)

For separate source– and channel coding:

$$e_{sep}(D) = \sup_R \min\{F(D,R), E(R)\} \leq e(D)$$

$\Rightarrow$ no separation theorem for error exponents.

Relation to Joint Source–Channel Coding (Cont'd)

For separate source– and channel coding:

$$e_{sep}(D) = \sup_R \min\{F(D,R), E(R)\} \leq e(D)$$

$\Rightarrow$ no separation theorem for error exponents.

But our achievability is based on separation:

- Quantize U – source coding
- Then map to $s_i(t)$ – channel coding.

Relation to Joint Source–Channel Coding (Cont'd)

For separate source– and channel coding:

$$e_{sep}(D) = \sup_R \min\{F(D,R), E(R)\} \leq e(D)$$

$\Rightarrow$ no separation theorem for error exponents.

But our achievability is based on separation:

- Quantize U – source coding
- Then map to $s_i(t)$ – channel coding.

Q: How does this settle?

Relation to Joint Source–Channel Coding (Cont'd)

Answer: Let Q^* maximize $R(D,Q)$ (often, uniform).

$$F(D,R) = \min_{Q:\ R(D,Q)\geq R} D(Q\|Q^*) = \begin{cases} 0 & R \leq R(D,Q^*) \\ \infty & R > R(D,Q^*) \end{cases}$$

Relation to Joint Source–Channel Coding (Cont'd)

Answer: Let Q^* maximize $R(D, Q)$ (often, uniform).

$$F(D, R) = \min_{Q:\ R(D,Q) \geq R} D(Q \| Q^*) = \begin{cases} 0 & R \leq R(D, Q^*) \\ \infty & R > R(D, Q^*) \end{cases}$$

Here, source–channel separation applies:

$$e_{sep}(D) = e(D) = E[R(D, Q^*)].$$

Relation to Joint Source–Channel Coding (Cont'd)

Answer: Let Q^* maximize $R(D,Q)$ (often, uniform).

$$F(D,R) = \min_{Q\colon\, R(D,Q)\geq R} D(Q\|Q^*) = \begin{cases} 0 & R \leq R(D,Q^*) \\ \infty & R > R(D,Q^*) \end{cases}$$

Here, source–channel separation applies:

$$e_{sep}(D) = e(D) = E[R(D,Q^*)].$$

Intuition:

- "Cover" source space by $e^{NR(D,Q^*)}$ D–spheres.
- Source encoder does not cause $\sum_i d(U_i,\hat{U}_i) > ND$.
- Excess distortion – only due to channel – w. p. $e^{-NE[R(D,Q^*)]}$.
- This is our case too.

Extensions

The Multidimensional Parameter Vector Case

Consider a d–dimensional vector $\boldsymbol{U} = (U_1, \ldots, U_d) \sim$ Unif$[-1/2, +1/2)^d$.

The Multidimensional Parameter Vector Case

Consider a d–dimensional vector $\boldsymbol{U} = (U_1, \ldots, U_d) \sim$ Unif$[-1/2, +1/2)^d$.

Minimize

$$\text{Pr}\left[\bigcup_{i=1}^{d}\left\{|\hat{U}_i - U_i| > e^{-R_i T}\right\}\right].$$

Let $E^*(R_1, \ldots, R_d) =$ best achievable exponent.

The Multidimensional Parameter Vector Case

Consider a d–dimensional vector $\boldsymbol{U} = (U_1, \ldots, U_d) \sim$ Unif$[-1/2, +1/2)^d$.

Minimize

$$\text{Pr}\left[\bigcup_{i=1}^{d} \left\{|\hat{U}_i - U_i| > e^{-R_i T}\right\}\right].$$

Let $E^*(R_1, \ldots, R_d) =$ best achievable exponent.

Thm above extends to

$$E^*(R_1, \ldots, R_d) = E(R_1 + R_2 + \ldots + R_d).$$

The Multidimensional Parameter Vector Case

Consider a d–dimensional vector $\boldsymbol{U} = (U_1, \ldots, U_d) \sim$ Unif$[-1/2, +1/2)^d$.

Minimize

$$\text{Pr}\left[\bigcup_{i=1}^{d}\left\{|\hat{U}_i - U_i| > e^{-R_i T}\right\}\right].$$

Let $E^*(R_1, \ldots, R_d) =$ best achievable exponent.

Thm above extends to

$$E^*(R_1, \ldots, R_d) = E(R_1 + R_2 + \ldots + R_d).$$

Think of a grid with $e^{R_i T}$ points in the i–th coordinate $\Rightarrow$ total $= e^{(R_1+\ldots+R_d)T}$.

The Vector Case (Cont'd)

Consider the case $R_1 = R_2 = \ldots = R_d \equiv R$:

$$E^*(R, R, \ldots, R) = E(R \cdot d).$$

The Vector Case (Cont'd)

Consider the case $R_1 = R_2 = \ldots = R_d \equiv R$:

$$E^*(R, R, \ldots, R) = E(R \cdot d).$$

For $R > 0$, due to the strong converse, $\exists$ dimensionality threshold effect:

$$\lim_{T \to \infty} \text{Pr}\left[\bigcup_{i=1}^{d} \left\{|\hat{U}_i - U_i| > e^{-RT}\right\}\right] = \begin{cases} 0 & d < d_c \triangleq \lfloor C/R \rfloor \\ 1 & d \geq d_c \end{cases}$$

The Vector Case (Cont'd)

Consider the case $R_1 = R_2 = \ldots = R_d \equiv R$:

$$E^*(R, R, \ldots, R) = E(R \cdot d).$$

For $R > 0$, due to the strong converse, $\exists$ dimensionality threshold effect:

$$\lim_{T\to\infty} \text{Pr}\left[\bigcup_{i=1}^{d}\left\{|\hat{U}_i - U_i| > e^{-RT}\right\}\right] = \begin{cases} 0 & d < d_c \triangleq \lfloor C/R \rfloor \\ 1 & d \geq d_c \end{cases}$$

For $R = 0$, $E(0) = C/2$ independently of d.

The Vector Case (Cont'd)

Consider the case $R_1 = R_2 = \ldots = R_d \equiv R$:

$$E^*(R, R, \ldots, R) = E(R \cdot d).$$

For $R > 0$, due to the strong converse, $\exists$ dimensionality threshold effect:

$$\lim_{T \to \infty} \text{Pr}\left[\bigcup_{i=1}^{d} \left\{|\hat{U}_i - U_i| > e^{-RT}\right\}\right] = \begin{cases} 0 & d < d_c \triangleq \lfloor C/R \rfloor \\ 1 & d \geq d_c \end{cases}$$

For $R = 0$, $E(0) = C/2$ independently of d.

Different from the common "curse of dimensionality", which is usually graceful in d.

Other Channels

- Gaussianity – not used very strongly.

Other Channels

- Gaussianity – not used very strongly.
- Main feature = known reliability func. (∞ bandwidth).

Other Channels

- Gaussianity – not used very strongly.
- Main feature = known reliability func. (∞ bandwidth).
- Reliability func. – known also for the Poisson channel (Wyner 1988) and others of ∞ bandwidth (Gallager 1987).

Other Channels

- Gaussianity – not used very strongly.
- Main feature = known reliability func. (∞ bandwidth).
- Reliability func. – known also for the Poisson channel (Wyner 1988) and others of ∞ bandwidth (Gallager 1987).
- For DMC's – known for $R \geq R_{crit}$.

Other Channels

- Gaussianity – not used very strongly.
- Main feature = known reliability func. (∞ bandwidth).
- Reliability func. – known also for the Poisson channel (Wyner 1988) and others of ∞ bandwidth (Gallager 1987).
- For DMC's – known for $R \geq R_{crit}$.
- For $R < R_{crit}$, not known, but separation still works.

Other Channels

- Gaussianity – not used very strongly.
- Main feature = known reliability func. (∞ bandwidth).
- Reliability func. – known also for the Poisson channel (Wyner 1988) and others of ∞ bandwidth (Gallager 1987).
- For DMC's – known for $R \geq R_{crit}$.
- For $R < R_{crit}$, not known, but separation still works.
- Applicable to bandlimited Gaussian channel with $N = 2WT$ channel uses.

Other Channels

- Gaussianity – not used very strongly.
- Main feature = known reliability func. (∞ bandwidth).
- Reliability func. – known also for the Poisson channel (Wyner 1988) and others of ∞ bandwidth (Gallager 1987).
- For DMC's – known for $R \geq R_{crit}$.
- For $R < R_{crit}$, not known, but separation still works.
- Applicable to bandlimited Gaussian channel with $N = 2WT$ channel uses.
- Unknown channels: universal decoding metrics – applicable for universal estimation.

Rayleigh Fading

Let

$$y(t) = a \cdot x(t, u) + n(t), \qquad 0 \leq t < T$$

where $a =$ realization of A, with density

$$f_A(a) = \frac{a}{\sigma^2} e^{-a^2/2\sigma^2} \qquad a \geq 0.$$

Rayleigh Fading

Let

$$y(t) = a \cdot x(t, u) + n(t), \qquad 0 \leq t < T$$

where $a =$ realization of A, with density

$$f_A(a) = \frac{a}{\sigma^2} e^{-a^2/2\sigma^2} \qquad a \geq 0.$$

For $R > 0$, the probability of excess error – dominated by channel outage

$$\text{Pr}\{A^2 C \leq R\} = 1 - e^{-R/2\bar{C}} \qquad \bar{C} = \sigma^2 C.$$

Rayleigh Fading

Let

$$y(t) = a \cdot x(t,u) + n(t), \qquad 0 \leq t < T$$

where $a =$ realization of A, with density

$$f_A(a) = \frac{a}{\sigma^2} e^{-a^2/2\sigma^2} \qquad a \geq 0.$$

For $R > 0$, the probability of excess error – dominated by channel outage

$$\Pr\{A^2 C \leq R\} = 1 - e^{-R/2\bar{C}} \qquad \bar{C} = \sigma^2 C.$$

For $R = 0$ – decays like $1/T$.

Summary and Conclusion

- Large deviations performance metric – natural for wideband communication.
- Precise characterization of the best achievable exponent.
- Intimately related to signal detection – reliability function.
- Simple considerations; simple to extend in many directions.
- Relation to JSCC: separate source– and channel coding is optimal.
- Open problem: close the gap between upper and lower bounds on the MMSE.

Thank You!