Weak–Noise Modulation–Estimation of Vector Parameters

Neri Merhav

The Andrew & Erna Viterbi Faculty of Electrical Engineering

Technion – Israel Institute of Technology

Haifa 3200003, Israel

ISIT 2020, June 21–26, 2020.

Some Background

Consider the AWGN channel,

$$y_t = x_t + z_t, \qquad t = 1, \ldots, n, \qquad z_t \sim \mathcal{N}(0, \sigma^2),$$

x_t being the t–th coordinate of $\boldsymbol{x} = (x_1, \ldots, x_n) = f_n(u)$,

where $u \in [0, 1]$ is a parameter to be conveyed across the channel, and

$$\|f_n(u)\|^2 \leq nP.$$

The receiver estimates u according to $\hat{u} = g_n(\boldsymbol{y})$.

The modulation–estimation problem:

For a given error cost function, $\rho(u - \hat{u})$, find (f_n, g_n) with minimum

$$\mathbf{E}\{\rho(u - g_n[f_n(u) + \boldsymbol{Z}])\}, \qquad \boldsymbol{Z} \sim \mathcal{N}(0, \sigma^2 I).$$

Some Background (Cont'd)

- ♣ Waveform communication problem [Wozencraft & Jacobs, Chap. 8].
- ♣ Estimation–theoretic view: non/Bayesian upper & lower bounds.
- ♣ Info–theoretic view: JSC coding, Shannon-Kotel'nikov mappings.
- ♣ Linear modulation – $f_n(u) = u \cdot \boldsymbol{s}$: attains CRLB, but limited signaling.
- ♣ Non–linear modulation – better for high SNR, but $\exists$ threshold effect.
- ♣ Most literature – total MSE $=$ small + anomalous errors.
- ♣ [W& J, Chap. 8]: separated analysis, not formal.
- ♣ Köken *et al.* ('17): $P[\text{outage}]$/small MSE: no-outage data proc. bound.
- ♣ Merhav ('19): outage–exponent/weak–noise-MSE exponent.
- ♣ This work: scalar $\Rightarrow$ vector parameter – non–trivial converse.

Problem Setting

We consider the channel model

$$\boldsymbol{Y} = f_n(\boldsymbol{u}) + \boldsymbol{Z}, \qquad \boldsymbol{Z} \sim \mathcal{N}(0, \sigma^2 \cdot I)$$

where $\boldsymbol{u} \in [0,1]^d$, $\|f_n(\boldsymbol{u})\|^2 \leq nP$.

♠ SNR: $\gamma = P/\sigma^2$; capacity: $C(\gamma) = \frac{1}{2}\log(1+\gamma)$.

♠ Estimator: $\hat{u} = g_n[\boldsymbol{Y}]$.

♠ Outage event: $\mathcal{O}_n(\boldsymbol{u})$ = arbirtrary set of noise vectors, $\{\boldsymbol{z}\}$.

♠ Error cost:

$$\varepsilon(f_n, g_n, \mathcal{O}_n) = \sup_{\boldsymbol{u} \in [0,1]^d} \mathbf{E}\{\rho(\boldsymbol{u} - g_n[f_n(\boldsymbol{u}) + \boldsymbol{Z}]) | \boldsymbol{Z} \in \mathcal{O}_n^{\mathsf{c}}(\boldsymbol{u})\},$$

where

$$\rho(\boldsymbol{u} - \hat{\boldsymbol{u}}) = \sum_{i=1}^{d} W_i \cdot |u_i - \hat{u}_i|^q, \qquad q \geq 1.$$

Problem Setting (Cont'd)

We wish to minimize $\varepsilon(f_n, g_n, \mathcal{O}_n)$ s.t.

$$\|f_n(\boldsymbol{u})\|^2 \leq nP, \quad \text{Pr}\{\mathcal{O}_n\} \to 0,$$

or, more precisely, find

$$\mathcal{E}(\gamma) = \lim_{n\to\infty} \max_{f_n, g_n, \mathcal{O}_n} \left[-\frac{\log \varepsilon(f_n, g_n, \mathcal{O}_n)}{n} \right].$$

We take the weights $\{W_i\}$ of $\rho(\cdot)$ to be exponential, i.e.,

$$\rho(\boldsymbol{u} - \hat{\boldsymbol{u}}) = \sum_{i=1}^{d} e^{-na_i} \cdot |u_i - \hat{u}_i|^q, \qquad q \geq 1$$

Main Result

Let

$$
\begin{aligned}
E(\gamma) &= \min_{\{(R_1,\ldots,R_d):\ \sum_i R_i \le C(\gamma)\}} \min_i (a_i + qR_i) \\
&= \frac{qC(\gamma) + \sum_i a_i}{d},
\end{aligned}
$$

achieved by

$$
R_i^* = \frac{C(\gamma)}{d} + \frac{1}{q}\left(\frac{1}{d}\sum_{j=1}^{d} a_j - a_i\right).
$$

Theorem: Under certain assumptions,

$$
\mathcal{E}(\gamma) = E(\gamma).
$$

Achievability

♡ Quantize each u_i to $\tilde{u}_i$ with a uniform quantizer of $e^{n(R_i^* - \epsilon/d)}$ levels.

♡ Map $\tilde{\boldsymbol{u}}$ to a good channel code of rate $C(\gamma) - \epsilon$: $f_n(\boldsymbol{u}) = \boldsymbol{x}[\tilde{\boldsymbol{u}}]$.

♡ Decode $\boldsymbol{Y}$ and de–map the decoded codeword to $\tilde{\boldsymbol{u}}$: $g_n(\boldsymbol{Y}) = \tilde{\boldsymbol{u}}$.

♡ Outage event $=$ decoding error event.

$$
\begin{aligned}
\text{Weak–noise error} &= \text{quantization error} \\
&\leq \rho(e^{-nR_1^*}, \ldots, e^{-nR_d^*}) \\
&= e^{-nE(\gamma)}.
\end{aligned}
$$

Early (Fruitless) Thoughts About the Converse ...

Scalar case: weak–noise lower bound depends on f_n only via

$$L(f_n) = \int_0^1 \left\| \frac{\partial f_n(u)}{\partial u} \right\| \mathrm{d}u = \text{ length of signal locus curve}$$

and then, a universal upper bound on $L(f_n)$ yields the converse.

One expects: for $\boldsymbol{u} \in [0,1]^d$, $L(f_n) \to$ signal manifold area, e.g., $d = 2$:

$$S(f_n) = \int_{[0,1]^2} \left\| \frac{\partial f_n(u,v)}{\partial u} \right\| \cdot \left\| \frac{\partial f_n(u,v)}{\partial v} \right\| \cdot |\sin\theta(u,v)| \mathrm{d}u\mathrm{d}v,$$

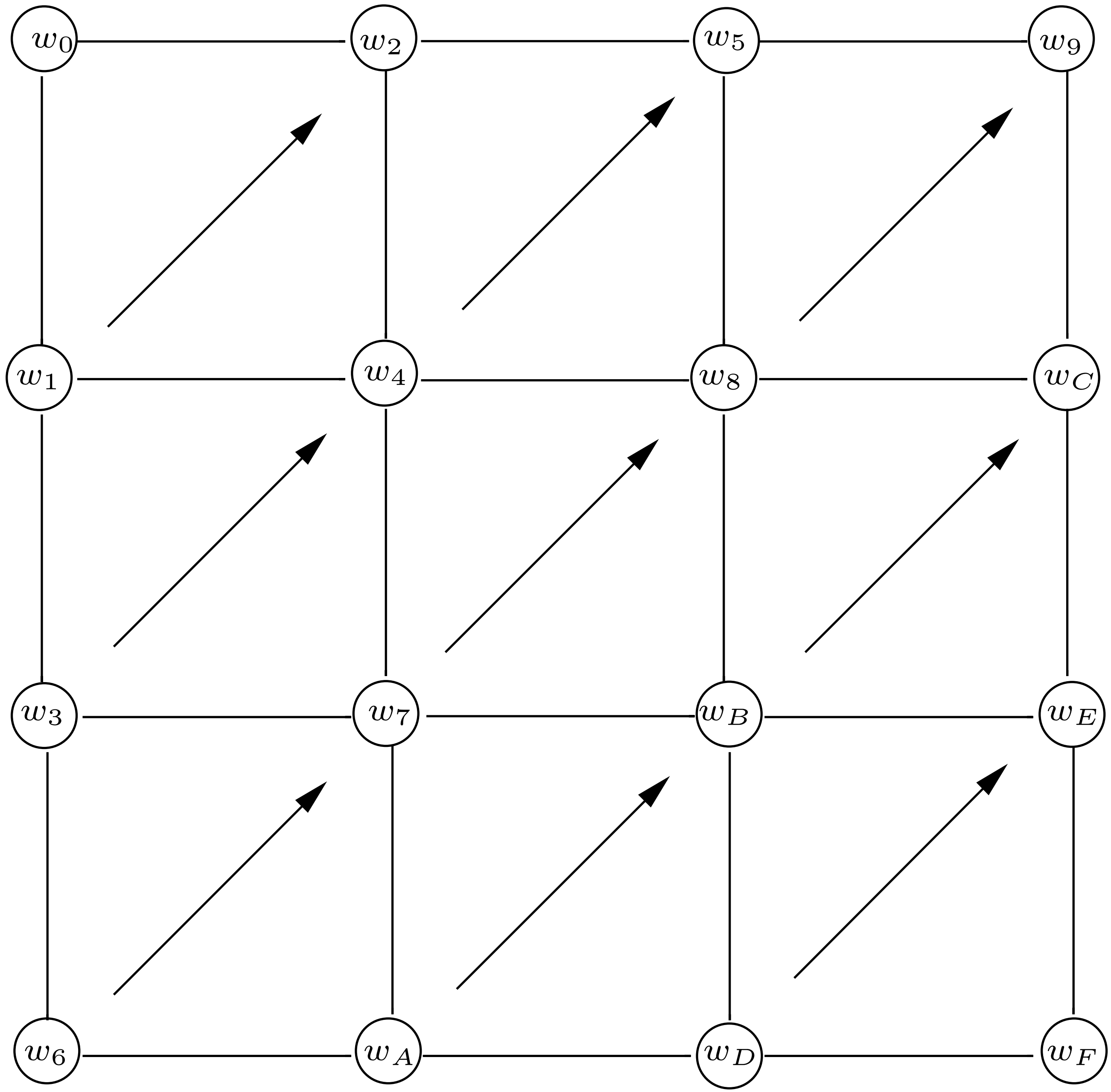
where

$$|\sin\theta(u,v)| = \sqrt{1 - \frac{\langle \partial f_n(u,v)/\partial u, \partial f_n(u,v)/\partial v \rangle^2}{\|\partial f_n(u,v)/\partial u\|^2 \|\partial f_n(u,v)/\partial v\|^2}}.$$

This line of thought turned out to be counterproductive ...

Converse

“Scan” the parameter space diagonally + apply the 1D converse.



The Dirty–Paper Channel

Suppose

$$y_t = f_n(\boldsymbol{u}, \boldsymbol{s}) + s_t + z_t, \quad t = 1, 2, \ldots.$$

Main result is left unaltered:

◇ Direct: scalar quantization + dirty–paper coding.

◇ Converse: genie–aided receiver with access to $\boldsymbol{s}$.

Universal Decoding Harnessed for Estimation

Suppose

$$y_t = f_n(\boldsymbol{u}) + s_t + z_t, \quad t = 1, 2, \ldots,$$

where

$$s_t = \sum_i \alpha_i \phi_{i,t}, \quad t = 1, 2, \ldots, \quad \{\alpha_i\} \text{ unknown}$$

Again, main result is the same using a universal decoder [Merhav ‘93].

Signal Structure Constraints

Suppose $\boldsymbol{u} = (u_1, u_2)$ and

$$f_n(u_1, u_2) = f_{n,1}(u_1) + f_{n,2}(u_2),$$

$$\|f_{n,1}(u_1)\|^2 \le nP_1, \qquad \|f_{n,2}(u_2)\|^2 \le nP_2, \qquad \langle f_{n,1}, f_{n,2} \rangle = 0.$$

Application: the Gaussian MAC.

1st lower bound – ignore the structural constraint:

$$E(\gamma_1 + \gamma_2) = \frac{qC(\gamma_1 + \gamma_2)}{2}.$$

2nd lower bound – treat individually each user:

$$\varepsilon \ge e^{-nqC(\gamma_1)} + e^{-nqC(\gamma_2)} \sim e^{-nq\min\{C(\gamma_1), C(\gamma_2)\}}.$$

If one SNR is small, the 2nd bound is tighter.

Thank You!