PERGAMON

Applied Mathematics Letters 12 (1999) 1–3

Applied Mathematics Letters

www.elsevier.nl/locate/aml

A Counterexample in Rate-Distortion Theory for Correlated Sources

R. AHLSWEDE AND N. CAI
Fakultät für Mathematik, Universität Bielefeld
D-33501 Bielefeld, Germany

(Received March 1999; accepted April 1999)

Keywords—Correlated sources, Fidelity criteria, Rate region, Two helper problem.

In the forthcoming paper "Multi-terminal source coding—achievable rates and reliability" Haroutunian claims the solution of an outstanding problem in source coding, namely, a characterisation of the rate region for discrete memoryless correlated sources with two separate encoders and one decoder under two fidelity criteria.

Such a source model is specified by a sequence $(X^n, Y^n)_{n=1}^{\infty}$ with generic random variables (X, Y) taking values in $\mathcal{X} \times \mathcal{Y}$ and having joint distribution $P_{XY} = P^* \times W^*$ and (sum-type) distortion measures with per letter distortions $d_{\mathcal{X}} : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^+$ and $d_{\mathcal{Y}} : \mathcal{Y} \times \mathcal{V} \to \mathbb{R}^+$.

For a given pair of nonnegative numbers $\Delta = (\Delta_{\mathcal{X}}, \Delta_{\mathcal{Y}})$ and $E > 0$ denote by $\mathcal{R}(E, \Delta)$ the set of nonnegative pairs of numbers $(R_{\mathcal{X}}, R_{\mathcal{Y}})$ such that for all $\varepsilon > 0$ and sufficiently large n there exists (encoding) functions $f_{\mathcal{X}} : \mathcal{X}^n \to \mathbb{N}$, $f_{\mathcal{Y}} : \mathcal{X}^n \to \mathbb{N}$, and a (decoding) function $F : \mathbb{N} \times \mathbb{N} \to \mathcal{U}^n \times \mathcal{V}^n$ with rate $(f_{\mathcal{X}}) \le R_{\mathcal{X}} + \varepsilon$, rate $(f_{\mathcal{Y}}) \le R_{\mathcal{Y}} + \varepsilon$ such that for $(U^n, V^n) \triangleq F(f_{\mathcal{X}}(X^n), f_{\mathcal{Y}}(Y^n))$

$$1 - \Pr\left(\left\{\frac{1}{n} d_{\mathcal{X}}(X^n, Y^n) \le \Delta_{\mathcal{X}}, \frac{1}{n} d_{\mathcal{Y}}(Y^n, V^n) \le \Delta_{\mathcal{Y}}\right\}\right) \le \exp\{-nE\}.$$

Now, the paper presents an inner bound on $\mathcal{R}(E, \Delta)$ and an outer bound, called $\mathcal{R}_{sp}(E, \Delta)$. By passing with E to 0 those bounds coincide. Unfortunately the outer bound $\mathcal{R}_{sp}(E, \Delta)$ is incorrect.

We recall first its definition and then we give our counterexample.

For any $E > 0$ define

$$\alpha(E) = \{P \times W \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : D(P \times W \,\|\, P^* \times W^*) \le E\}.$$

Denote by $\varphi = (\varphi_{\mathcal{X}}, \varphi_{\mathcal{Y}})$ a function which associates pairs of PDs (P, PW) with pairs of conditional PDs (Q_P, G_{PW}), i.e., $\varphi(P, PW) = (\varphi_{\mathcal{X}}(P), \varphi_{\mathcal{Y}}(PW)) = (Q_P, G_{PW})$, such that

$$\mathbb{E}_{P,Q_P} d_{\mathcal{X}}(X, U) \triangleq \sum_{x,u} P(x) Q_P(u \mid x)\, d_{\mathcal{X}}(x, u) \le \Delta_{\mathcal{X}} \tag{1}$$

0893-9659/99/$ - see front matter. Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TeX
PII:S0893-9659(99)00093-2

and

$$\mathbb{E}_{PW,G_{PW}}\, d_{\mathcal{Y}}(Y,V) \triangleq \sum_{y,v} PW(y)G_{PW}(v \mid y)\, d_{\mathcal{Y}}(y,v) \leq \Delta_{\mathcal{Y}}. \tag{2}$$

Here the RVs (X,Y,U,V) have the joint distribution

$$P_{XYUV}(x,y,u,v) = P(x)W(y \mid x)Q_P(u \mid x)G_{PW}(v \mid y) \quad \text{for } x \in \mathcal{X},\quad y \in \mathcal{Y},\quad u \in \mathcal{U},\quad \text{and} \quad v \in \mathcal{V}. \tag{3}$$

To indicate the dependence on φ we write $I_{P,W,\varphi}(X \wedge U \mid V)$ for $I(X \wedge U \mid V)$, $I_{P,W,\varphi}(X \wedge U \mid V)$ for $I(XY \wedge UV)$, and so on.

Now we are ready to define the outer region in terms of the three inequalities

(i) $R_{\mathcal{X}} \geq \max_{P\times W\in\alpha(E)} I_{P,W,\varphi}(X \wedge U \mid V)$,
(ii) $R_{\mathcal{Y}} \geq \max_{P\times W\in\alpha(E)} I_{P,W,\varphi}(Y \wedge V \mid U)$, and
(iii) $R_{\mathcal{X}} + R_{\mathcal{Y}} \geq \max_{P\times W\in\alpha(E)} I_{P,W,\varphi}(XY \wedge UV)$,

as follows:

$$\mathcal{R}_{sp}(E,\Delta) = \bigcup_{\varphi\in\Phi(\Delta)} \mathcal{R}_{sp}(E,\Delta,\varphi), \tag{4}$$

where

$$\mathcal{R}_{sp}(E,\Delta,\varphi) = \{(R_{\mathcal{X}},R_{\mathcal{Y}}) : R_{\mathcal{X}} \text{ and } R_{\mathcal{Y}} \text{ satisfies (i), (ii), and (iii)}\} \tag{5}$$

and $\Phi(\Delta)$ denotes the set of all functions φ, for which (1) and (2) hold.

This description invokes equation (3), which is equivalent to the Markovity

$$U \ominus X \ominus Y \ominus V.$$

The "proof" for $\mathcal{R}(E,\Delta) \subset R_{sp}(E,\Delta)$ has a gap; namely, this Markovity does not appear in it. Moreover, the gap cannot be closed, because the statement itself is false.

EXAMPLE. $\mathcal{R}(E,\Delta) \not\subset \mathcal{R}_{sp}(E,\Delta)$.

Choose $\mathcal{X} = \mathcal{Y} = \mathcal{U} = \mathcal{V} = \{0,1\}$, the source distribution $P^* \times W^*$ as $P^*(0) = P^*(1) = 1/2$, $W^*(x \mid x) = 1-p$ for $x \in \mathcal{X}$ and any $p \in (0,1/2)$, and the distortion measures $d_{\mathcal{X}}, d_{\mathcal{Y}}$ as Hamming distance.

It is easy to see that for $\Delta = (0,\delta)$ with $\delta > p$ and some $E_\delta \triangleq -\delta\log p-(1-\delta)\log(1-p)-h(\delta) > 0$

$$R = (R_{\mathcal{X}},R_{\mathcal{Y}}) = (1,0) \in \mathcal{R}(E_\delta,\Delta), \tag{6}$$

but

$$R = (1,0) \notin \mathcal{R}_{sp}(E_\delta,\Delta). \tag{7}$$

Indeed, to verify (6), consider the code $(f_{\mathcal{X}}, f_{\mathcal{Y}}, F)$ defined by an injective $f_{\mathcal{X}}$, a constant $f_{\mathcal{Y}}$, and for all $x^n \in \mathcal{X}^n$, $y^n \in \mathcal{Y}^n$

$$F\left(f_{\mathcal{X}}\left(x^n\right), f_{\mathcal{Y}}\left(y^n\right)\right) = \left(x^n, x^n\right). \tag{8}$$

Thus, $R_{\mathcal{X}} = \text{rate } (f_{\mathcal{X}}) = 1$ and $R_{\mathcal{Y}} = \text{rate } (f_{\mathcal{Y}}) = 0$.

For $(U^n,V^n) \triangleq F(f_{\mathcal{X}}(X^n),\ f_{\mathcal{Y}}(Y^n)) = (X^n,X^n)$, clearly

$$1-\Pr\left(d_H(X^n,U^n)=0,\ d_H(Y^n,V^n)\leq\delta\right) = \Pr\left(d_H(X^n,Y^n)>\delta\right) = \sum_{k>n\delta}\binom{n}{k}p^k(1-p)^{n-k} = 2^{-n(-\delta\log p-(1-\delta)\log(1-p)-h(\delta)+o(1))}$$

(since $\delta > p$) $= 2^{-nE_\delta}$, and (6) holds.

It remains to show equation (7). Obviously, for all $E > 0$, $P^* \times W^* \in \alpha(E)$, because $D(P^* \times W^* \mid P^* \times W^*) = 0 \leq E$. For any $\varphi \in \Phi(\Delta)$, $\Delta = (0,\delta)$, we have for $(Q,W) =$

$\varphi(P^*, P^*W^*)\sum_{x,u} P^*(x)Q(u \mid x)d_H(x,u) = 0$ and therefore $Q(x \mid x) = 1$ for $x \in \mathcal{X}$. This implies the first equality in

$$I_{P^*,W^*,\varphi}(Y \wedge V \mid U) = I_{P^*,W^*,\varphi}(Y \wedge V \mid X) = 0,$$

and the second equality holds, because $R_{\mathcal{Y}} = 0$ and (ii) should hold. Therefore, we have the Markovity

$$Y \ominus X \ominus V. \tag{9}$$

This and (3) yield

$$P_{XYV}(x,y,v) = P^*(x)W^*(x \mid x)G(v \mid y) = P_{XY}(x,y)P_{v|x}(v \mid x), \qquad \text{for all } x, y, v. \tag{10}$$

Since for all x, y $P_{XY}(x,y) = P^*(x)W^*(y \mid x) > 0$, the second equality in (10) implies that

$$P_{V|X}(v \mid x) = G(v \mid y), \qquad \text{for all } x, y.$$

This implies in particular that Y and V are independent and that we can write $G(v \mid y)$ as $\tilde{G}(v)$. In this notation

$$\begin{aligned}\Delta_{\mathcal{Y}} &\geq \sum_{y,v} P^*W^*(y)\tilde{G}(y)d_H(y,v)\\ &= \frac{1}{2}\sum_{y,v}\tilde{G}(y)d_H(y,v) = \frac{1}{2}.\end{aligned}$$

Consequently, for every $E > 0$, $\delta < 1/2$, $\Delta = (0,\delta)$, and every $(\mathcal{R}_{\mathcal{X}}, 0)$ necessarily $(\mathcal{R}_{\mathcal{X}}, 0) \notin \mathcal{R}_{sp}(E,\Delta)$. In particular for E_δ, (7) holds.

REMARKS.

1. We have chosen the extremal points $R = (1,0)$, $\Delta = (0,\delta)$ only to get a simple example. By continuity there are also counterexamples of the form $R = (1-\eta_1, \eta_2)$, $\Delta = (\eta_3, \eta_4)$ with small η_1, η_2, and η_3.
2. Unfortunately it cannot be excluded that the same kind of mistake has entered other papers in this area.