A Strong Converse for Joint Source-Channel Coding

Da Wang	Amir Ingber	Yuval Kochman
EECS, MIT	EE, Stanford	CSE, HUJI

ISIT 2012, Cambridge, MA

July 5, 2011

Results in information theory

Achievability

$$P_{\mathrm{e}} \to 0$$

Channel Coding:

- When $R < C$, reliable communication achievable.

Weak converse
When $R > C$, $P_{\mathrm{e}} \to 0$ is impossible.

Strong converse
When $R > C$, $P_{\mathrm{e}} \to 1$.

Exponentially-strong converse
When $R > C$,

- $P_{\mathrm{c}} \triangleq 1 - P_{\mathrm{e}} \to 0$ exponentially,
- The exponents are known for both source and channel coding.

Converse results
Channel coding (DMC)

Error exponent outer bound: when $R < C$,

$$P_{\mathrm{e}} \dot{\geq} e^{-nE_{\mathrm{sp}}(R,W)}$$

$$E_{\mathrm{sp}}(R,W) \triangleq \max_{\Phi} \min_{V:I(\Phi,V)\leq R} D(V \| W|\Phi).$$

- [C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, 1967], [E. A. Haroutunian, 1968]

Exponentially-strong converse: when $R > C$,

$$P_{\mathrm{c}} \dot{\leq} e^{-n\bar{E}_{\mathrm{sp}}(R,W)} \qquad \text{(tight)}$$

$$\bar{E}_{\mathrm{sp}}(R,W) \triangleq \max_{\Phi} \min_{V} \left[D(V \| W|\Phi) + |R - I(\Phi,V)|^{+} \right].$$

- [S. Arimoto, 1973], [G. Dueck and J. Körner, 1979]

Converse results

Lossy source coding

Error event: $\mathcal{E} \triangleq \{d(\mathbf{s}, \hat{\mathbf{s}}) \geq D\}$

Error exponent outer bound: when $R > R(P, D)$,

$$P_{\mathrm{e}} \dot{\geq} e^{-nE_{\mathrm{S}}(R,D,P)} \quad \text{(tight)}$$

$$E_{\mathrm{S}}(R, D, P) \triangleq \min_{Q: R(Q,D) \geq R} D(Q \| P).$$

- [K. Marton, 1974]

Exponentially-strong converse: when $R < R(P, D)$,

$$P_{\mathrm{c}} \dot{\leq} e^{-n\bar{E}_{\mathrm{S}}(R,D,P)} \quad \text{(tight)}$$

$$\bar{E}_{\mathrm{S}}(R, D, P) \triangleq \min_{Q} \left[D(Q \| P) + |R(Q, D) - R|^{+} \right].$$

- Problem 9.6, [I. Csiszár and J. Körner, 1981/2011]

Error exponent outer bound: when $R(D) < \rho C$,

$$P_{\mathrm{e}} \overset{\cdot}{\geq} e^{-nE_{\mathrm{JSCC}}(P,D,W,\rho)} \qquad \text{(tight)}$$

$$E_{\mathrm{JSCC}}\left(P,D,W,\rho\right) \triangleq \min_{R}\left[E_{\mathrm{S}}\left(R,D,P\right) + \rho E_{\mathrm{sp}}\left(R/\rho,W\right)\right].$$

- ρ: bandwidth expansion factor
- Shown in [I. Csiszár, 1980[1], 1982[2]].

How about the strong converse for the case $R(D) > \rho C$?

[1] Joint source-channel error exponent, 1980
[2] On the error exponent of source-channel transmission with a distortion threshold, 1982

Joint source-channel coding converse

When $R(D) > \rho C$,

Previous work: special cases of non-exponentially strong converse

- Lossless: [T. S. Han, 2002]
- Quadratic-Gaussian: [Y. Zhong, F. Alajaji, and L.L. Campbell, 2007]
- Modulo-additive DMC: [Y. Zhong, F. Alajaji, and L.L. Campbell, 2009]

Our contributions

- A general (non-exponentially) strong converse can be derived indirectly via:
  - the information spectrum method: [T. S. Han, 2002]
  - equivalence to channel coding: [M. Agarwal, A. Sahai, and S. Mitter, 2006]
  - JSCC dispersion: [D. Wang, A. Ingber, and Y. Kochman, 2011]
- Derived the exponentially strong converse for general DMS-DMC pairs:

$$P_{\mathrm{c}} \doteq e^{-n\bar{E}_{\mathrm{JSCC}}(P,D,W,\rho)}$$

$$\bar{E}_{\mathrm{JSCC}}(P,D,W,\rho) = \min_{R}[\bar{E}_{\mathrm{S}}(R,D,P) + \rho\bar{E}_{\mathrm{sp}}(R/\rho,W)].$$

Example: BSS + BSC

- BSS:

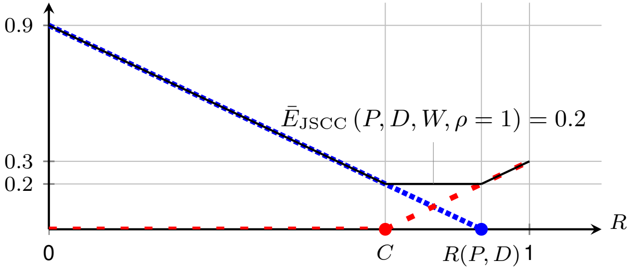
$$\bar{E}_{\mathrm{S}}\left(R, D, P\right) = \left|1 - H_b(D) - R\right|^+.$$

- BSC:

$$\bar{E}_{\mathrm{sp}}\left(R, W\right) = \left|R - (1 - H_b(\varepsilon))\right|^+.$$

- When $R(P, D) > C(W)$, i.e. $D < \varepsilon$,

$$\begin{aligned}\bar{E}_{\mathrm{JSCC}}\left(P, D, W, \rho = 1\right) &= \min_R[\bar{E}_{\mathrm{S}}\left(R, D, P\right) + \bar{E}_{\mathrm{sp}}\left(R, W\right)] \\ &= R(P, D) - C(W) = H_b\left(\varepsilon\right) - H_b\left(D\right).\end{aligned}$$



Example: BSS + BSC

- BSS:

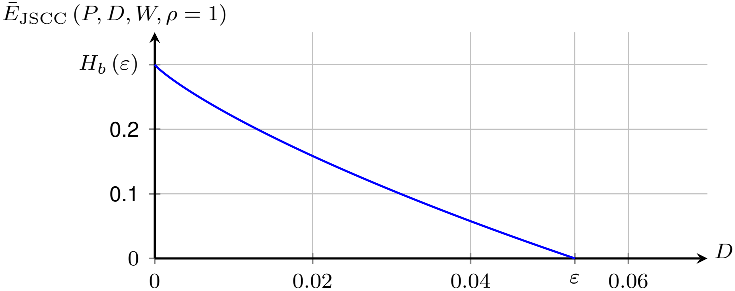
$$\bar{E}_{\mathrm{S}}(R, D, P) = |1 - H_b(D) - R|^+.$$

- BSC:

$$\bar{E}_{\mathrm{sp}}(R, W) = |R - (1 - H_b(\varepsilon))|^+.$$

- When $R(P, D) > C(W)$, i.e. $D < \varepsilon$,

$$\begin{aligned}\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho = 1) &= \min_R [\bar{E}_{\mathrm{S}}(R, D, P) + \bar{E}_{\mathrm{sp}}(R, W)] \\ &= R(P, D) - C(W) = H_b(\varepsilon) - H_b(D).\end{aligned}$$



Bigger picture

- Filled the missing piece:

	$P_\mathrm{e} \dot{\geq} e^{-nE}$	$P_\mathrm{c} \doteq e^{-n\bar{E}}$
Channel coding	$R < C$	$R > C$
	$E_\mathrm{sp}\left(R, W\right)$	$\bar{E}_\mathrm{sp}\left(R, W\right)$
Lossy source coding	$R > R(P, D)$	$R < R(P, D)$
	$E_\mathrm{S}\left(R, D, P\right)$	$\bar{E}_\mathrm{S}\left(R, D, P\right)$
JSCC	$R(D) < \rho C$	$R(D) > \rho C$
	$E_\mathrm{JSCC}\left(P, D, W, \rho\right)$	$\bar{E}_\mathrm{JSCC}\left(P, D, W, \rho\right)$

Technical details

Problem formulation

DMS $(\mathcal{S}, \hat{\mathcal{S}}, P, d)$

- source alphabet $\mathcal{S}$
- reproduction alphabet $\hat{\mathcal{S}}$
- source distribution P
- distortion $d : \mathcal{S} \times \hat{\mathcal{S}} \to \mathbb{R}_+$

DMC $W : \mathcal{X} \to \mathcal{Y}$

- input alphabet $\mathcal{X}$,
- output alphabet $\mathcal{Y}$
- conditional distribution $W(\cdot | \cdot)$

Discrete memoryless JSCC

- DMS $(\mathcal{S}, \hat{\mathcal{S}}, P, d)$
- DMC $W : \mathcal{X} \to \mathcal{Y}$
- bandwidth expansion factor $\rho \in \mathbb{R}_+$.

Joint source-channel coding problem formulation

JSCC scheme $\mathcal{C}^{(n)}_{\mathrm{JSCC}}$

- encoder $f_{J;n} : \mathcal{S}^n \to \mathcal{X}^{\lfloor \rho n \rfloor}$
- decoder $g_{J;n} : \mathcal{Y}^{\lfloor \rho n \rfloor} \to \hat{\mathcal{S}}^n$

$$\mathbf{s} \xrightarrow{f_{J;n}(\mathbf{s})} \mathbf{x} \to \boxed{W} \to \mathbf{y} \xrightarrow{g_{J;n}(\mathbf{y})} \hat{\mathbf{s}}$$

Error (and success) events

$$\mathcal{E}(D) \triangleq \{d(\mathbf{s}, \hat{\mathbf{s}}) > D\},$$

$$\bar{\mathcal{E}}(D) \triangleq \mathcal{E}(D)^c = \{d(\mathbf{s}, \hat{\mathbf{s}}) \leq D\}$$

JSCC exponents

$$\bar{\mathcal{E}}_n(D) \in \underset{\left\{\mathcal{C}^{(n)}_{\mathrm{JSCC}}\right\}}{\arg\min}\, \mathbb{P}\left[\bar{\mathcal{E}}(D)\right].$$

$$E_{\mathrm{JSCC}}(P, D, W, \rho) \triangleq \liminf_{n\to\infty} \sup_{\left\{\mathcal{C}^{(n)}_{\text{JSCC}}\right\}} -\frac{1}{n} \log \mathbb{P}\left[\mathcal{E}_n(D)\right]$$

$$\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho) \triangleq \liminf_{n\to\infty} \sup_{\left\{\mathcal{C}^{(n)}_{\text{JSCC}}\right\}} -\frac{1}{n} \log \mathbb{P}\left[\bar{\mathcal{E}}_n(D)\right]$$

Theorem

Let $\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho)$ be the exponent of the success probability for the best sequence of JSCC schemes:

$$\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho) \triangleq \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}\left[\bar{\mathcal{E}}_n(D)\right].$$

Then

$$\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho) = \min_{R}[\bar{E}_{\mathrm{S}}(R, D, P) + \rho \bar{E}_{\mathrm{sp}}(R/\rho, W)].$$

Actually a two-part result:

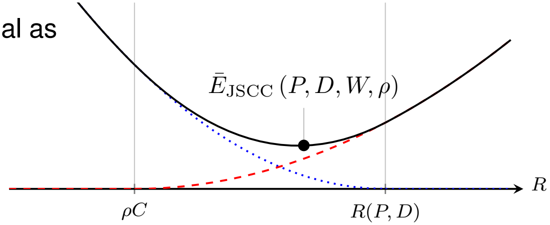
- Converse part: the exponent is upper bounded
- Direct part: the expoent is acheivable

Analogous to JSCC error exponent

$$\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho) = \min_R [\bar{E}_{\mathrm{S}}(R, D, P) + \rho \bar{E}_{\mathrm{sp}}(R/\rho, W)]$$
$$E_{\mathrm{JSCC}}(P, D, W, \rho) = \min_R [E_{\mathrm{S}}(R, D, P) + \rho E_{\mathrm{sp}}(R/\rho, W)]$$

Minimizing rate

- When $R(P,D) \leq \rho C$, trivial as $\bar{E}(P, D, W, \rho) = 0$.
- When $R(P,D) > \rho C$, the minimizing rate satisfies $\rho C < R < R(P,D)$.



Joint Source-Channel Coding

Separation is optimal

- $\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho)$ attainable by a separation scheme!
- For any chosen digital rate R,

$$\mathbb{P}[\text{no JSCC excess distortion}] \geq \mathbb{P}[\text{no source excess distortion}] \cdot \mathbb{P}[\text{no channel error}]$$

  where

$$\mathbb{P}[\text{no source excess distortion}] \doteq e^{-n\bar{E}_{\mathrm{S}}(R,D,P)}$$

$$\mathbb{P}[\text{no channel error}] \doteq e^{-n\rho\bar{E}_{\mathrm{sp}}(R/\rho,W)}$$

  Hence $\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho) \geq \bar{E}_{\mathrm{S}}(R, D, P) + \bar{E}_{\mathrm{sp}}(R, W)$.

- Surprising, because separation is suboptimal for
  - JSCC error exponent below capcity, and
  - JSCC dispersion.

Strong converse: alternative form

$$
\begin{aligned}
&\bar{E}_{\mathrm{JSCC}}\left(P, D, W, \rho\right) \\
&= \min_{Q \in \mathcal{P}(\mathcal{S})} \left[D\left(Q \,\|\, P\right) + \max_{\Phi \in \mathcal{P}(\mathcal{X})} \min_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} \left(\rho D\left(V \,\|\, W | \Phi\right) \right.\right. \\
&\qquad\qquad \left.\left. + \left| R(Q, D) - \rho I(\Phi, V) \right|^{+} \right) \right].
\end{aligned}
$$

- Minimize over distributions.
- Proof is based on this form.

Proof sketches

Condition on source type:

$$\mathbb{P}\left[\bar{\mathcal{E}}(D)\right] \leq (n+1)^{|\mathcal{S}|} \max_{Q \in \mathcal{P}_n(\mathcal{S})} \mathbb{P}\left[\bar{\mathcal{E}}(D) \,\middle|\, P_{\mathbf{S}} = Q\right] e^{-nD(Q \,\|\, P)}.$$

Condition on source type and channel input type:

$$\mathbb{P}\left[\bar{\mathcal{E}}(D) | P_{\mathbf{S}} = Q\right] = \sum_{\Phi \in \mathcal{A}_m} \mathbb{P}\left[P_{\mathbf{X}} = \Phi | P_{\mathbf{S}} = Q\right] \mathbb{P}\left[\bar{\mathcal{E}}(D) | P_{\mathbf{S}} = Q, P_{\mathbf{x}} = \Phi\right]$$

Condition on source type, channel input type, and channel conditional type:

$$\mathbb{P}\left[\bar{\mathcal{E}}(D) | P_{\mathbf{S}} = Q, P_{\mathbf{x}} = \Phi\right] = \Bigg(\sum_{V \in \mathcal{P}_m(\mathcal{Y}|\Phi)} \mathbb{P}\left[P_{\mathbf{y}|\mathbf{x}} = V | P_{\mathbf{x}} = \Phi\right] \cdot \mathbb{P}\left[\bar{\mathcal{E}}(D) | P_{\mathbf{S}} = Q, P_{\mathbf{x}} = \Phi, P_{\mathbf{y}|\mathbf{x}} = V\right] \Bigg).$$

Proof sketches

Key technical detail

Combine

- Refined results on source type covering
- Exponentially-strong channel coding converse: [G. Dueck and J. Körner, 1979]

$\Rightarrow$

$$\mathbb{P}\left[\bar{\mathcal{E}}(D) | P_{\mathbf{S}} = Q, P_{\mathbf{x}} = \Phi, P_{\mathbf{y}|\mathbf{x}} = V\right] \leq \mathrm{poly}(n) \frac{e^{-n|R(Q,D)-\rho I(\Phi,V)|^{+}}}{\mathbb{P}\left[P_{\mathbf{X}} = \Phi | P_{\mathbf{S}} = Q\right]}$$

Proof sketches

Putting things together…

DMS with distribution P has type Q

$$\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho) = \min_{Q \in \mathcal{P}(\mathcal{S})} \left[D(Q \,\|\, P) + \max_{\Phi \in \mathcal{P}(\mathcal{X})} \min_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} \left(\rho D(V \,\|\, W | \Phi) + |R(Q, D) - \rho I(\Phi, V)|^{+} \right) \right].$$

Channel conditional type V

JSCC exponent for source type Q, channel input type Φ and channel conditional type V

That's it!

	$P_\mathrm{e} \dot{\geq} e^{-nE}$	$P_\mathrm{c} \doteq e^{-n\bar{E}}$
Channel coding	$R < C$	$R > C$
	$E_{\mathrm{sp}}(R, W)$	$\bar{E}_{\mathrm{sp}}(R, W)$
Lossy source coding	$R > R(P, D)$	$R < R(P, D)$
	$E_{\mathrm{S}}(R, D, P)$	$\bar{E}_{\mathrm{S}}(R, D, P)$
JSCC	$R(D) < \rho C$	$R(D) > \rho C$
	$E_{\mathrm{JSCC}}(P, D, W, \rho)$	$\bar{E}_{\mathrm{JSCC}}(P, D, W, \rho)$

Backup slides

Proof sketches

Joint source channel coding converse with fixed types

$\hat{B}(\mathbf{s}, D)$: all the channel outputs that covers some source sequence $\mathbf{s}$ with distortion D

$$\hat{B}(\mathbf{s}, D) \triangleq \{\mathbf{y} \in \mathcal{Y}^m : d(\mathbf{s}, g_{J;n}(\mathbf{y})) \leq D\}$$

$$\frac{1}{|G(Q,\Phi)|} \sum_{\mathbf{s}_i \in G(Q,\Phi)} \frac{\left|\mathcal{T}_V^m\left(f(\mathbf{s}_i)\right) \cap \hat{B}(\mathbf{s}_i, D)\right|}{\left|\mathcal{T}_V^m\left(f(\mathbf{s}_i)\right)\right|} \leq \frac{p(n)}{\alpha(Q,\Phi)} e^{-n[R(Q,D)-\rho I(\Phi,V)]^+}$$

$\alpha(Q,\Phi) \triangleq \mathbb{P}\left[P_{\mathbf{X}} = \Phi | P_{\mathbf{S}} = Q\right]$

For $\mathbf{s}_i$, the faction of channel outputs that has distortion less than D.

$G(Q,\Phi) \triangleq \left\{\mathbf{s} \in \mathcal{T}_Q^n : \mathbf{x} = f_{J;n}(\mathbf{s}) \in \mathcal{T}_\Phi^m\right\}$, the set of source sequences in $\mathcal{T}_Q^n$ that are mapped (via JSCC encoder $f_{J;n}$) to channel codewords with type Φ.

Proof sketches

Restricted D-ball size

Given source type P and a reconstruction sequence $\hat{\mathrm{s}}$, define restricted D-ball as

$$B(\hat{\mathbf{s}}, P, D) \triangleq \{\mathbf{s} \in \mathcal{T}_P^n : d(\mathbf{s}, \hat{\mathbf{s}}) \leq D\}.$$

Then

$$|B(\hat{\mathbf{s}}, P, D)| \leq (n+1)^{|\mathcal{S}||\hat{\mathcal{S}}|} \exp\{n[H(P) - R(P, D)]\}.$$