On Dispersion and Mismatched Decoding

Vincent Tan

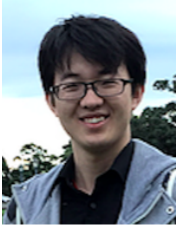
Department of Electrical and Computer Engineering,
Department of Mathematics,
National University of Singapore

JSWCIT, NTU
March 2018

Collaborators

J. Scarlett (NUS)

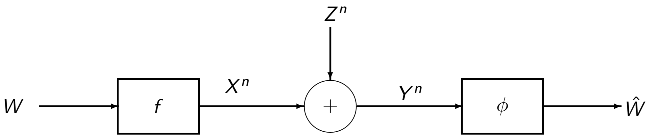
G. Durisi (Chalmers)



L. Zhou⋆ (NUS)

M. Motani (NUS)

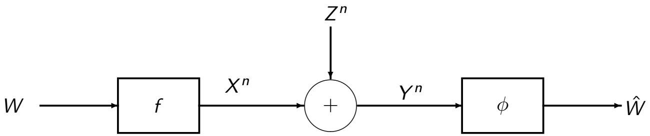
Channel Coding for AWGN Channel: Model and Definitions



- $W \sim \mathrm{Unif}(\mathcal{M})$ where $\mathcal{M} := [M]$;
- Z^n: $Z_i \sim \mathcal{N}(0, 1)$ for all $i \in [n]$;
- Power constraint:

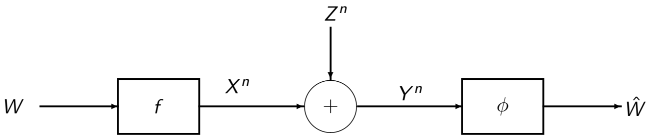
$$\frac{1}{n} \sum_{i=1}^{n} X_i^2 \leq P$$

Channel Coding for AWGN Channel: Model and Definitions



- $W \sim \mathrm{Unif}(\mathcal{M})$ where $\mathcal{M} := [M]$;
- Z^n: $Z_i \sim \mathcal{N}(0,1)$ for all $i \in [n]$;
- Power constraint:
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \le P$$
- (n, M, P)-code consists of $f : \mathcal{M} \to \mathcal{X}^n$ and $\phi : \mathcal{Y}^n \to \mathcal{M}$

Channel Coding for AWGN Channel: Model and Definitions



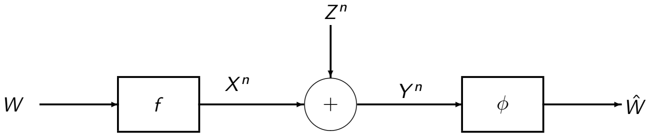
- $W \sim \mathrm{Unif}(\mathcal{M})$ where $\mathcal{M} := [M]$;
- Z^n: $Z_i \sim \mathcal{N}(0,1)$ for all $i \in [n]$;
- Power constraint:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \le P$$

- (n, M, P)-code consists of $f : \mathcal{M} \to \mathcal{X}^n$ and $\phi : \mathcal{Y}^n \to \mathcal{M}$
- Non-asymptotic fundamental limit:

$$M^*(n, P, \varepsilon) := \sup\{M : \exists \text{ an } (n, M, P)\text{-code s.t. } \Pr\{\hat{W} \neq W\} \le \varepsilon\}.$$

Channel Coding for AWGN Channel: Existing Results



Tomamichel-Tan showed that for any $\varepsilon \in [0, 1)$,

$$\log M^*(n, P, \varepsilon) := n\mathrm{C}(P) - \sqrt{n\mathrm{V}(P)}\mathrm{Q}^{-1}(\varepsilon) + \frac{1}{2}\log n + O(1),$$

where the Gaussian capacity (Shannon 1948) and Gaussian dispersion function (Hayashi 2009, PPV 2010) are resp.

$$\mathrm{C}(P) = \frac{1}{2}\log(1 + P), \qquad \mathrm{V}(P) = \frac{P(P+2)}{2(P+1)^2}.$$

Channel Coding: Mismatched Setting

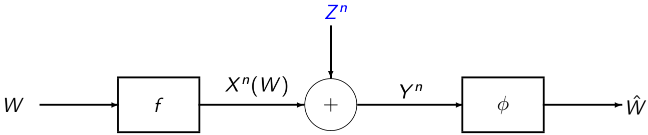
TRANSACTIONS ON INFORMATION THEORY, VOL. 42, NO. 5, SEPTEMBER 1996

Nearest Neighbor Decoding for Additive Non-Gaussian Noise Channels

Amos Lapidoth, *Member, IEEE*

***Abstract*—We study the performance of a transmission scheme employing random Gaussian codebooks and nearest neighbor decoding over a power limited additive non-Gaussian noise channel. We show that the achievable rates depend on the noise distribution only via its power and thus coincide with the capacity region of a white Gaussian noise channel with signal and noise power equal to those of the original channel. The results are presented for single-user channels as well as multiple-access channels, and are extended to fading channels with side information at the receiver.**

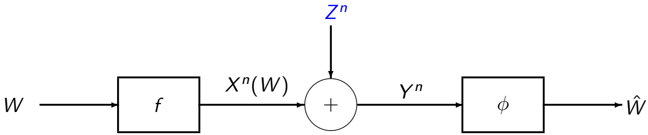
Channel Coding: Mismatched Setting[1]



- Non-Gaussian noise Z^n: $Z_i \sim P_Z$ such that $\mathbb{E}[Z^2] = 1$;

[1] A. Lapidoth. "Nearest neighbor decoding for additive non-Gaussian noise channels," IEEE Trans. on Inform. Th., 42(5):1520-1529, 1996.

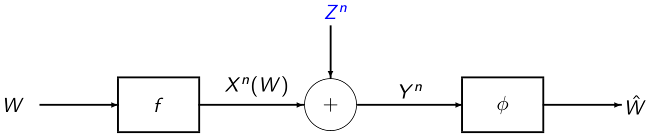
Channel Coding: Mismatched Setting[1]



- Non-Gaussian noise Z^n: $Z_i \sim P_Z$ such that $\mathbb{E}[Z^2] = 1$;
- Random Gaussian codebook $\{X^n(1), \ldots, X^n(M)\}$;

[1] A. Lapidoth. "Nearest neighbor decoding for additive non-Gaussian noise channels," IEEE Trans. on Inform. Th., 42(5):1520-1529, 1996.

Channel Coding: Mismatched Setting[1]

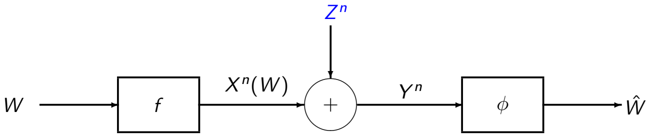


- Fixed coding scheme
  - Encoding: Given W, the encoder outputs channel input $X^n(W)$;
  - Minimum distance decoding: Given Y^n, the decoder outputs $\hat{W}$ if

$$\hat{W} = \arg\min_{\tilde{w} \in \mathcal{M}} \|X^n(\tilde{w}) - Y^n\|_2^2.$$

[1] A. Lapidoth. "Nearest neighbor decoding for additive non-Gaussian noise channels," IEEE Trans. on Inform. Th., 42(5):1520-1529, 1996.

Channel Coding: Mismatched Setting[1]



- Fixed coding scheme
  - Encoding: Given W, the encoder outputs channel input $X^n(W)$;
  - Minimum distance decoding: Given Y^n, the decoder outputs $\hat{W}$ if

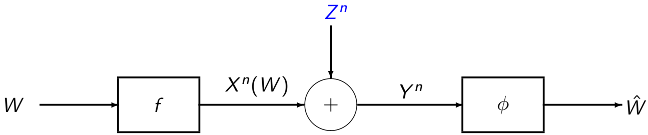
$$\hat{W} = \arg\min_{\tilde{w} \in \mathcal{M}} \|X^n(\tilde{w}) - Y^n\|_2^2.$$

- Ensemble error probability

$$\overline{\mathrm{P}}_{\mathrm{e},n}(M) := \Pr\{\hat{W} \neq W\}$$

[1] A. Lapidoth. "Nearest neighbor decoding for additive non-Gaussian noise channels," IEEE Trans. on Inform. Th., 42(5):1520-1529, 1996.

Channel Coding: Mismatched Setting[1]

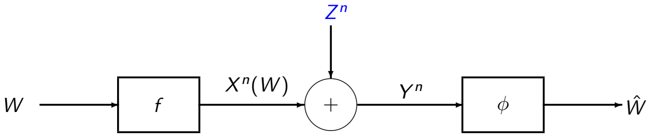


- Non-asymptotic fundamental limit: for any $\varepsilon \in [0, 1)$, let

$$M^*(n, P, \varepsilon) := \sup\{M : \overline{\mathrm{P}}_{\mathrm{e},n}(M) \leq \varepsilon\}.$$

[1] A. Lapidoth. "Nearest neighbor decoding for additive non-Gaussian noise channels," IEEE Trans. on Inform. Th., 42(5):1520-1529, 1996.

Channel Coding: Mismatched Setting[1]



- Non-asymptotic fundamental limit: for any $\varepsilon \in [0,1)$, let

$$M^*(n, P, \varepsilon) := \sup\{M : \overline{\mathrm{P}}_{\mathrm{e},n}(M) \le \varepsilon\}.$$

- Lapidoth showed that

$$\lim_{n\to\infty} \frac{1}{n} \log M^*(n, P, \varepsilon) = \mathrm{C}(P) = \frac{1}{2}\log(1+P).$$

[1] A. Lapidoth. "Nearest neighbor decoding for additive non-Gaussian noise channels," IEEE Trans. on Inform. Th., 42(5):1520-1529, 1996.

Channel Coding: Dispersion for Mismatched Setting[2]

- Consider P_Z s.t. $\mathbb{E}[Z^2] = 1$ and $\zeta_{\mathrm{c}} := \mathbb{E}[Z^4] < \infty$

[2] "The dispersion of nearest-neighbor decoding for additive Non-Gaussian channels", J. Scarlett, V. Y. F. Tan, and G. Durisi, IEEE Trans. Inf. Theory, vol. 63, no. 1, pp. 8192, 2017.

Channel Coding: Dispersion for Mismatched Setting[2]

- Consider P_Z s.t. $\mathbb{E}[Z^2] = 1$ and $\zeta_{\mathrm{c}} := \mathbb{E}[Z^4] < \infty$
- Mismatched dispersions:

$$\mathrm{V}_{\mathrm{sp}}(P, \zeta_{\mathrm{c}}) := \frac{P^2(\zeta_{\mathrm{c}} - 1) + 4P}{4(P+1)^2}, \qquad \text{Spherical Gaussian codebook}$$

$$\mathrm{V}_{\mathrm{iid}}(P, \zeta_{\mathrm{c}}) := \mathrm{V}_{\mathrm{sp}}(P, \zeta_{\mathrm{c}}) + \frac{P^2}{2(P+1)^2} \qquad \text{i.i.d. Gaussian codebook;}$$

[2] "The dispersion of nearest-neighbor decoding for additive Non-Gaussian channels", J. Scarlett, V. Y. F. Tan, and G. Durisi, IEEE Trans. Inf. Theory, vol. 63, no. 1, pp. 8192, 2017.

Channel Coding: Dispersion for Mismatched Setting[2]

- Consider P_Z s.t. $\mathbb{E}[Z^2] = 1$ and $\zeta_{\rm c} := \mathbb{E}[Z^4] < \infty$
- Mismatched dispersions:

$$\mathrm{V}_{\rm sp}(P, \zeta_{\rm c}) := \frac{P^2(\zeta_{\rm c} - 1) + 4P}{4(P+1)^2}, \qquad \text{Spherical Gaussian codebook}$$

$$\mathrm{V}_{\rm iid}(P, \zeta_{\rm c}) := \mathrm{V}_{\rm sp}(P, \zeta_{\rm c}) + \frac{P^2}{2(P+1)^2} \qquad \text{i.i.d. Gaussian codebook;}$$

Theorem 1 (Scarlett-Tan-Durisi 2017)

For any $\dagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M_\dagger^*(n, P, \varepsilon) = n\mathrm{C}(P) - \sqrt{n\mathrm{V}_\dagger(P, \zeta_{\rm c})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

[2] "The dispersion of nearest-neighbor decoding for additive Non-Gaussian channels", J. Scarlett, V. Y. F. Tan, and G. Durisi, IEEE Trans. Inf. Theory, vol. 63, no. 1, pp. 8192, 2017.

Channel Coding: Remarks for Mismatched Setting

Theorem 2 (Scarlett-Tan-Durisi 2017)

For any $\dagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M_\dagger^*(n, P, \varepsilon) = n\mathrm{C}(P) - \sqrt{n\mathrm{V}_\dagger(P, \zeta_\mathrm{c})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

- $\mathrm{V}_\dagger$ depend on P_Z only through the second and fourth moment;

Channel Coding: Remarks for Mismatched Setting

Theorem 2 (Scarlett-Tan-Durisi 2017)

For any $\dagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M_\dagger^*(n, P, \varepsilon) = n\mathrm{C}(P) - \sqrt{n\mathrm{V}_\dagger(P, \zeta_\mathrm{c})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

- $\mathrm{V}_\dagger$ depend on P_Z only through the second and fourth moment;
- For Gaussian noise

$$\mathrm{V}_\mathrm{sp}(n, P, \varepsilon) = \mathrm{V}(P)$$

  recovers the second-order coding rate for AWGN channels (cf. Hayashi 2009 and PPV 2010);

Channel Coding: Remarks for Mismatched Setting

Theorem 2 (Scarlett-Tan-Durisi 2017)

For any $\dagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M^*_\dagger(n, P, \varepsilon) = n\mathrm{C}(P) - \sqrt{n\mathrm{V}_\dagger(P, \zeta_\mathrm{c})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

- $\mathrm{V}_\dagger$ depend on P_Z only through the second and fourth moment;
- For Gaussian noise
$$\mathrm{V}_{\mathrm{sp}}(n, P, \varepsilon) = \mathrm{V}(P)$$
recovers the second-order coding rate for AWGN channels (cf. Hayashi 2009 and PPV 2010);
- Spherical Gaussian codebook outperforms the i.i.d. Gaussian codebook for $\varepsilon \in [0, 0.5)$.

Channel Coding: Intuition for Mismatched Dispersions

- The "mismatched" information density

$$\tilde{\imath}(x^n; y^n) := \mathrm{C}(P) + \frac{\|y^n\|_2^2}{2(P+1)} - \frac{\|y^n - x^n\|_2^2}{2};$$

where x^n obeys an i.i.d. or spherical Gaussian distribution.

Channel Coding: Intuition for Mismatched Dispersions

- The "mismatched" information density

$$\tilde{\imath}(x^n; y^n) := \mathrm{C}(P) + \frac{\|y^n\|_2^2}{2(P+1)} - \frac{\|y^n - x^n\|_2^2}{2};$$

  where x^n obeys an i.i.d. or spherical Gaussian distribution.
- Dominant error event governed by the cumulative distribution function of $\tilde{\imath}(X^n; Y^n)$, i.e.,

$$\Pr\{\tilde{\imath}(X^n; Y^n) \le \gamma\},$$

  which depends on the codebook distribution, i.e., P_{X^n};

Channel Coding: Intuition for Mismatched Dispersions

- The "mismatched" information density

$$\tilde{\imath}(x^n; y^n) := \mathrm{C}(P) + \frac{\|y^n\|_2^2}{2(P+1)} - \frac{\|y^n - x^n\|_2^2}{2};$$

  where x^n obeys an i.i.d. or spherical Gaussian distribution.

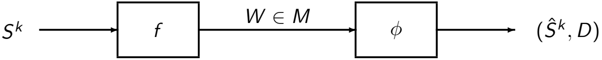
- Dominant error event governed by the cumulative distribution function of $\tilde{\imath}(X^n; Y^n)$, i.e.,

$$\Pr\{\tilde{\imath}(X^n; Y^n) \le \gamma\},$$

  which depends on the codebook distribution, i.e., P_{X^n};

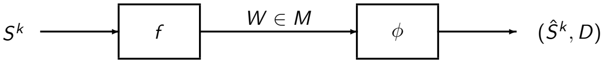
- Spherical codebook implies exact power P while i.i.d. Gaussian codebook corresponds to average power P

Rate-Distortion Problem: Model and Definitions



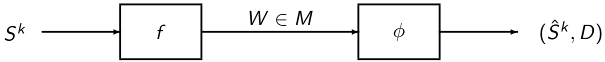
- Memoryless Source: S^k is i.i.d. according to P_S on $\mathcal{S}$;

Rate-Distortion Problem: Model and Definitions



- Memoryless Source: S^k is i.i.d. according to P_S on $\mathcal{S}$;
- Reproduced Sequence $\hat{S}^k \in \hat{\mathcal{S}}^k$;

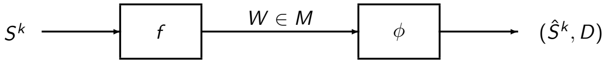
Rate-Distortion Problem: Model and Definitions



- Memoryless Source: S^k is i.i.d. according to P_S on $\mathcal{S}$;
- Reproduced Sequence $\hat{S}^k \in \hat{\mathcal{S}}^k$;
- Distortion measure:

$$d(s^k, \hat{s}^k) := \frac{1}{k} \sum_{i=1}^{k} d(s_i, \hat{s}_i);$$

Rate-Distortion Problem: Model and Definitions

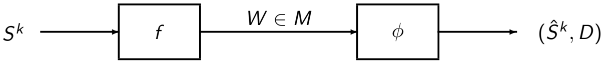


- Memoryless Source: S^k is i.i.d. according to P_S on $\mathcal{S}$;
- Reproduced Sequence $\hat{S}^k \in \hat{\mathcal{S}}^k$;
- Distortion measure:

$$d(s^k, \hat{s}^k) := \frac{1}{k} \sum_{i=1}^{k} d(s_i, \hat{s}_i);$$

- (k, M)-code consists of $f : \mathcal{S}^k \to [M]$ and $\phi : [M] \to \hat{\mathcal{S}}^k$;

Rate-Distortion Problem: Model and Definitions



- Memoryless Source: S^k is i.i.d. according to P_S on $\mathcal{S}$;
- Reproduced Sequence $\hat{S}^k \in \hat{\mathcal{S}}^k$;
- Distortion measure:

$$d(s^k, \hat{s}^k) := \frac{1}{k} \sum_{i=1}^{k} d(s_i, \hat{s}_i);$$

- (k, M)-code consists of $f : \mathcal{S}^k \to [M]$ and $\phi : [M] \to \hat{\mathcal{S}}^k$;
- Nonasymptotic fundamental limit:

$$M^*(k, D, \varepsilon) := \inf\{M : \exists\, (k, M)\text{−code s.t. } \Pr\{d(S^k, \hat{S}^k) > D\} \leq \varepsilon\}.$$

Rate-Distortion Problem: Existing Results for GMS

$$S^k \longrightarrow \boxed{f} \xrightarrow{W \in [M]} \boxed{\phi} \longrightarrow (\hat{S}^k, D)$$

Kostina and Verdú[3] showed that for any Gaussian memoryless sources (i.e., $P_S = \mathcal{N}(0, \sigma^2)$) under the quadratic distortion measure, for any $\varepsilon \in (0, 1)$,

$$\log M^*(k, D, \varepsilon) = k\mathrm{R}(\sigma^2, D) + \sqrt{\frac{k}{2}}\mathrm{Q}^{-1}(\varepsilon) + O(\log k),$$

where the Gaussian rate-distortion function (cf. Shannon 1948) is

$$\mathrm{R}(\sigma^2, D) = \max\left\{\frac{1}{2}\log\frac{\sigma^2}{D}, 0\right\}.$$

[3] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime," IEEE Trans. Inf. Theory, vol. 58, no. 6, pp. 3309–3338, 2012

Rate-Distortion: Mismatched Setting

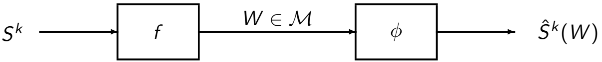
On the Role of Mismatch in Rate Distortion Theory

Amos Lapidoth, *Member, IEEE*

***Abstract*—Using a codebook $\mathcal{C}$, a source sequence is described by the codeword that is closest to it according to the distortion measure $d_0(x, \hat{x}_0)$. Based on this description, the source sequence is reconstructed to minimize the reconstruction distortion as measured by $d_1(x, \hat{x}_1)$, where, in general, $d_1(x, \hat{x}_1) \neq d_0(x, \hat{x}_0)$. We study the minimum resulting $d_1(x, \hat{x}_1)$-distortion between the reconstructed sequence and the source sequence as we optimize over the codebook subject to a rate constraint. Using a random coding argument we derive an upper bound on the resulting distortion. Applying this bound to blocks of source symbols we construct a sequence of bounds which are shown to converge to the least distortion achievable in this setup. This solves the rate distortion dual of an open problem related to the capacity of channels with a given decoding rule—the mismatch capacity. Addressing a different kind of mismatch, we also study the mean-squared error description of non-Gaussian sources with random Gaussian codebooks. It is shown that the use of a Gaussian codebook to compress any ergodic source results in an average distortion which depends on the source via its second moment only. The source with a given second moment that is most difficult to describe is the memoryless zero-mean Gaussian source, and it is best described using a Gaussian codebook. Once a Gaussian codebook is used, we show that all sources of a given second moment become equally hard to describe.**

Our interest in this paper is in a situation where the distortion measure[1] $d_1(x, \hat{x}_1)$ that best describes the sensitivities of the end user is different from the distortion measure $d_0(x, \hat{x}_0)$ according to which the source is encoded. Such a situation can arise if encoding to minimize $d_0(x, \hat{x}_0)$ is easier to implement than encoding to minimize $d_1(x, \hat{x}_1)$, or when an individual who has his own special sensitivities that are best described by $d_1(x, \hat{x}_1)$, attempts to reconstruct a source sequence that was compressed using a lossy compression algorithm that minimizes $d_0(x, \hat{x}_0)$ and over which he has no control. It should be noted, however, that in our setup we allow the choice of the codebook to depend on the two distortion measures (as well as on the source law) and it need not be optimal for the distortion measure $d_0(x, \hat{x}_0)$. In this respect our problem arises more naturally when the mismatch in distortion measures is introduced to reduce complexity rather than due to a change in the distortion criteria after the source was compressed. We do, however, briefly address the latter case as well by analyzing the performance attained with a random codebook that is drawn according to the optimal distribution for the distortion

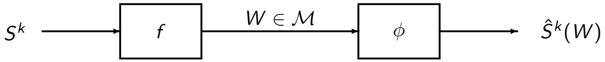
Rate-Distortion: Mismatched Setting[4]



- Consider arbitrary source distribution P_S s.t. $\mathbb{E}[S^2] = \sigma^2 > D$;

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]



- Consider arbitrary source distribution P_S s.t. $\mathbb{E}[S^2] = \sigma^2 > D$;
- Consider quadratic distortion measure, i.e., $d(s^k, \hat{s}^k) = \|s^k - \hat{s}^k\|_2^2$;

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]

$$S^k \longrightarrow \boxed{f} \xrightarrow{W \in \mathcal{M}} \boxed{\phi} \longrightarrow \hat{S}^k(W)$$

- Consider arbitrary source distribution P_S s.t. $\mathbb{E}[S^2] = \sigma^2 > D$;
- Consider quadratic distortion measure, i.e., $d(s^k, \hat{s}^k) = \|s^k - \hat{s}^k\|_2^2$;
- Fixed Gaussian codebook with M codewords $\hat{S}^k(1), \ldots, \hat{S}^k(M)$

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]

$$S^k \longrightarrow \boxed{\;f\;} \xrightarrow{\;W \in \mathcal{M}\;} \boxed{\;\phi\;} \longrightarrow \hat{S}^k(W)$$

- Fixed coding scheme
  - Minimum distance encoding: Given S^k, the encoder f outputs W if

$$W = \operatorname*{arg\,min}_{\tilde{w} \in \mathcal{M}} \|S^k - \hat{S}^k(\tilde{w})\|^2;$$

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

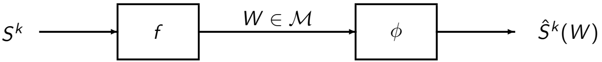
Rate-Distortion: Mismatched Setting[4]

- Fixed coding scheme
  - Minimum distance encoding: Given S^k, the encoder f outputs W if

$$W = \arg\min_{\tilde{w}\in\mathcal{M}} \|S^k - \hat{S}^k(\tilde{w})\|^2;$$

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]



- Fixed coding scheme
  - Minimum distance encoding: Given S^k, the encoder f outputs W if

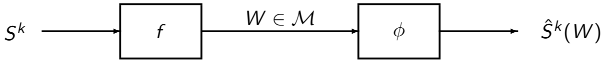
$$W = \arg\min_{\tilde{w} \in \mathcal{M}} \|S^k - \hat{S}^k(\tilde{w})\|^2;$$

  - Decoding: Given W, the decoder ϕ declares $\hat{S}^k(W)$ as the estimate.
- Ensemble excess-distortion probability

$$\overline{\mathrm{P}}_{\mathrm{e},k}(M, D) := \Pr\{d(S^k, \hat{S}^k(W)) > D\}$$

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]

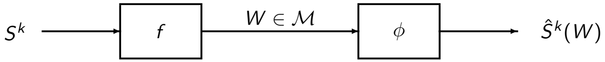


- Non-asymptotic fundamental limit:

$$M^*_{\mathrm{sp}}(k, \varepsilon, \sigma^2, D) := \inf\{M : \overline{\mathrm{P}}_{\mathrm{e},k}(M, D) \le \varepsilon\}.$$

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]



- Non-asymptotic fundamental limit:

$$M^*_{\mathrm{sp}}(k, \varepsilon, \sigma^2, D) := \inf\{M : \overline{\mathrm{P}}_{\mathrm{e},k}(M, D) \le \varepsilon\}.$$

Similarly, we define $M^*_{\mathrm{iid}}(k, \varepsilon, \sigma^2, D)$.

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Rate-Distortion: Mismatched Setting[4]

$$S^k \longrightarrow \boxed{f} \xrightarrow{W \in \mathcal{M}} \boxed{\phi} \longrightarrow \hat{S}^k(W)$$

- Non-asymptotic fundamental limit:

$$M_{\mathrm{sp}}^*(k, \varepsilon, \sigma^2, D) := \inf\{M : \overline{\mathrm{P}}_{\mathrm{e},k}(M, D) \le \varepsilon\}.$$

  Similarly, we define $M_{\mathrm{iid}}^*(k, \varepsilon, \sigma^2, D)$.

- Lapidoth showed that for any $\varepsilon \in [0, 1)$,

$$\lim_{n \to \infty} \frac{1}{n} \log M_{\mathrm{sp}}^*(k, \varepsilon, \sigma^2, D) = \mathrm{R}(\sigma^2, D) = \frac{1}{2} \log \frac{\sigma^2}{D}.$$

[4] "On the role of mismatch in rate distortion theory", A. Lapidoth, IEEE Trans. Inf. Theory, vol. 43, no. 1, pp. 38-47, 1997.

Dispersion for Mismatched Rate-Distortion Problem

- Any source distribution P_S s.t. $\mathbb{E}[S^2] = \sigma^2 > D$ and $\zeta_{\mathrm{s}} := \mathbb{E}[S^4] < \infty$

Dispersion for Mismatched Rate-Distortion Problem

- Any source distribution P_S s.t. $\mathbb{E}[S^2] = \sigma^2 > D$ and $\zeta_{\mathrm{s}} := \mathbb{E}[S^4] < \infty$
- Mismatched distortion-dispersion function

$$\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}}) := \frac{\zeta_{\mathrm{s}} - \sigma^4}{4\sigma^4} = \frac{\mathrm{Var}[S^2]}{4(\mathbb{E}[S^2])^2}.$$

Dispersion for Mismatched Rate-Distortion Problem[5]

- Any source distribution P_S s.t. $\mathbb{E}[S^2] = \sigma^2 > D$ and $\zeta_{\mathrm{s}} := \mathbb{E}[S^4] < \infty$
- Mismatched distortion-dispersion function

$$\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}}) := \frac{\zeta_{\mathrm{s}} - \sigma^4}{4\sigma^4} = \frac{\mathrm{Var}[S^2]}{4(\mathbb{E}[S^2])^2}.$$

Theorem 3 (Zhou-Tan-Motani 2017)

For any $\ddagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M_\ddagger^*(k, \varepsilon, \sigma^2, D) = k\mathrm{R}(\sigma^2, D) + \sqrt{k\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log k).$$

[5] L. Zhou, V. Y. F. Tan and M. Motani, "Refined asymptotics for rate-distortion using Gaussian codebooks for arbitrary sources," arXiv:1708.04778.

Remarks for Mismatched Dispersion

Theorem 4 (Zhou-Tan-Motani 2017)

For any $\ddagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M_{\ddagger}^{*}(k, \varepsilon, \sigma^2, D) = k\mathrm{R}(\sigma^2, D) + \sqrt{k\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log k).$$

- For a Gaussian source we recover Kostina and Verdú (TIT 2012);

Remarks for Mismatched Dispersion

Theorem 4 (Zhou-Tan-Motani 2017)

For any $\ddagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0,1)$,

$$\log M_{\ddagger}^{*}(k,\varepsilon,\sigma^2,D) = k\mathrm{R}(\sigma^2,D) + \sqrt{k\mathrm{V}(\sigma^2,\zeta_{\mathrm{s}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log k).$$

- For a Gaussian source we recover Kostina and Verdú (TIT 2012);
- Strengthen Lapidoth's result

Remarks for Mismatched Dispersion

Theorem 4 (Zhou-Tan-Motani 2017)

For any $\ddagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

$$\log M_{\ddagger}^{*}(k, \varepsilon, \sigma^2, D) = k\mathrm{R}(\sigma^2, D) + \sqrt{k\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log k).$$

- For a Gaussian source we recover Kostina and Verdú (TIT 2012);
- Strengthen Lapidoth's result
- Spherical and i.i.d. Gaussian codebooks achieve the same second-order asymptotics/dispersion

Remarks for Mismatched Dispersion

Theorem 4 (Zhou-Tan-Motani 2017)

For any $\ddagger \in \{\mathrm{sp}, \mathrm{iid}\}$ *and any* $\varepsilon \in [0, 1)$,

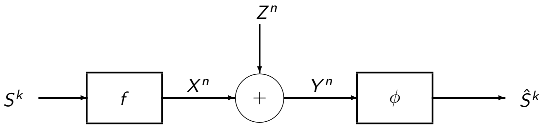
$$\log M_{\ddagger}^{*}(k, \varepsilon, \sigma^2, D) = k\mathrm{R}(\sigma^2, D) + \sqrt{k\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log k).$$

- For a Gaussian source we recover Kostina and Verdú (TIT 2012);
- Strengthen Lapidoth's result
- Spherical and i.i.d. Gaussian codebooks achieve the same second-order asymptotics/dispersion
  - Intuition: dominant error event is the atypicality of X^n:

$$\Pr\left\{\frac{1}{n}\sum_{i=1}^{n} X_i^2 > a\right\}$$

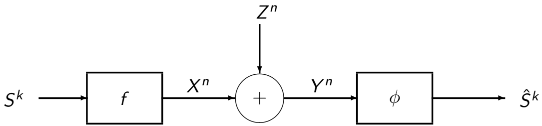
and the exact nature of the Gaussian codebooks play a diminished role

JSCC of Transmitting a GMS over an AWGN Channel



- S^k i.i.d. $\sim \mathcal{N}(0, \sigma^2)$ and Z^n i.i.d. $\sim \mathcal{N}(0, 1)$;

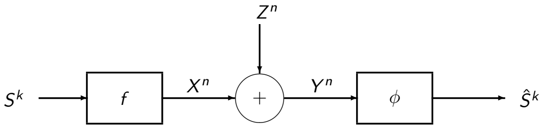
JSCC of Transmitting a GMS over an AWGN Channel



- S^k i.i.d. $\sim \mathcal{N}(0, \sigma^2)$ and Z^n i.i.d. $\sim \mathcal{N}(0, 1)$;
- An (k, n, P)-code consists of one encoder $f : \mathcal{S}^k \to \mathcal{X}^n$ and one decoder $\phi : \mathcal{Y}^n \to \hat{\mathcal{S}}^k$ s.t.

$$\frac{1}{n} \sum_{i=1}^{n} X_i^2 \leq P;$$

JSCC of Transmitting a GMS over an AWGN Channel



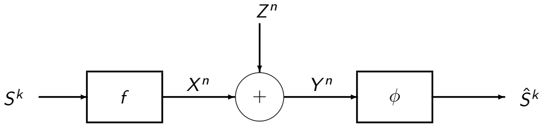
- S^k i.i.d. $\sim \mathcal{N}(0, \sigma^2)$ and Z^n i.i.d. $\sim \mathcal{N}(0, 1)$;
- An (k, n, P)-code consists of one encoder $f : \mathcal{S}^k \to \mathcal{X}^n$ and one decoder $\phi : \mathcal{Y}^n \to \hat{\mathcal{S}}^k$ s.t.

$$\frac{1}{n} \sum_{i=1}^{n} X_i^2 \leq P;$$

- Non-asymptotic fundamental limit

$$k^*(n, \varepsilon, D) := \sup\{k : \exists \text{ an } (k, n)\text{−code s.t. } \Pr\{d(S^k, \hat{S}^k) > D\} \leq \varepsilon\}.$$

JSCC of Transmitting a GMS over an AWGN Channel[6]

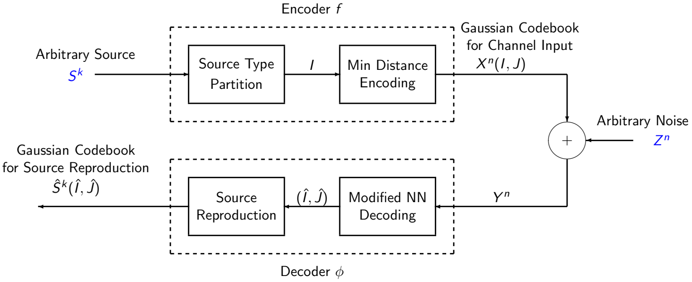


- Kostina and Verdú showed that for any $\varepsilon \in [0, 1)$,

$$k^*(n, \varepsilon, D) = \frac{n\mathrm{C}(P)}{\mathrm{R}(\sigma^2, D)} - \sqrt{\frac{n\mathrm{V}(P) + \frac{k}{2}}{(\mathrm{R}(\sigma^2, D))^2}}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

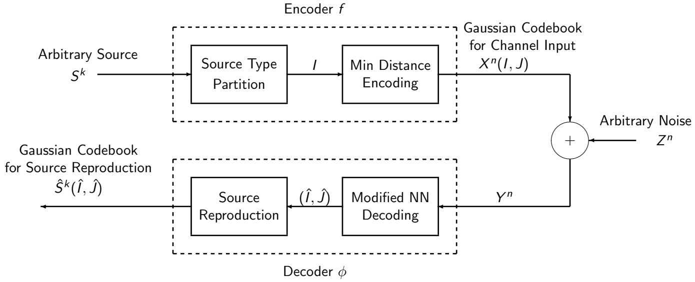
[6] V. Kostina and S. Verdú, "Lossy joint source-channel coding in the finite blocklength regime," IEEE Trans. Inf. Theory, vol. 59, no. 5, pp. 2545–2575, 2013.

Mismatched JSCC: Problem Setting and Definitions



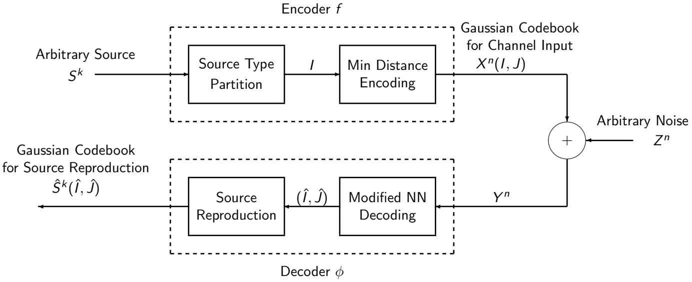
- Transmit any memoryless source over an additive arbitrary noise channel using coding scheme optimized for GMS over AWGN
  - Arbitrary memoryless source: S^k i.i.d. according to any P_S s.t. $\mathbb{E}[S^2]=\sigma^2$ and $\zeta_{\mathrm{s}}=\mathbb{E}[S^4]<\infty$
  - Arbitrary i.i.d. noise: Z^n i.i.d. according to any distribution P_Z s.t. $\mathbb{E}[Z^2]=1$ and $\zeta_{\mathrm{c}}=\mathbb{E}[Z^4]<\infty$

Mismatched JSCC: Problem Setting and Definitions



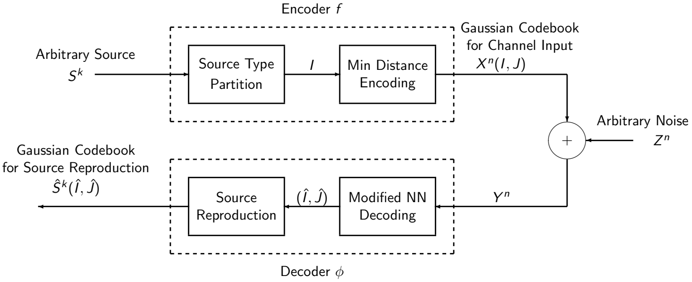
- Random Gaussian codebooks
  - Channel codebook for channel input
    - spherical
    - i.i.d.
  - Source codebook for source reproduction
    - spherical
    - i.i.d.

Mismatched JSCC: Problem Setting and Definitions



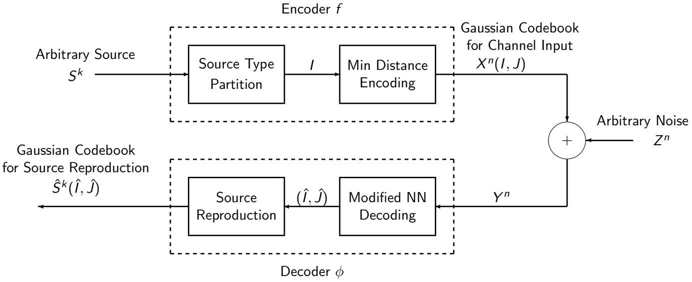
- Random Gaussian codebooks
  - Channel codebook for channel input
    - spherical
    - i.i.d.
  - Source codebook for source reproduction
    - spherical
    - i.i.d.

Mismatched JSCC: Problem Setting and Definitions



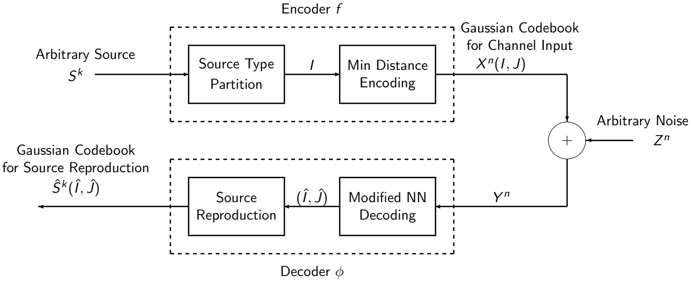
- A (k, n)-code consists of an encoder f and a decoder ϕ

Mismatched JSCC: Problem Setting and Definitions



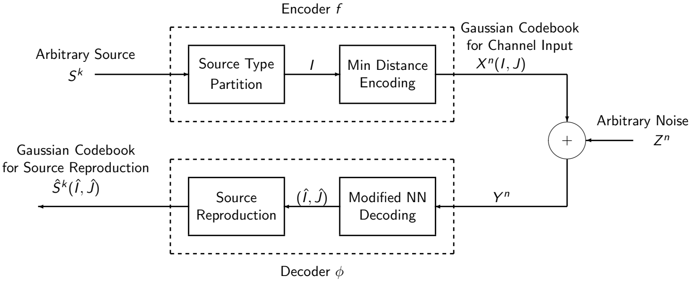
- A (k, n)-code consists of an encoder f and a decoder ϕ
- Encoder f uses modified minimum distance encoding
  - Source partition: Fix integer N and let $\{\mathcal{T}_i\}_{i\in[N]}$ be partition $\mathcal{S}^k$;

Mismatched JSCC: Problem Setting and Definitions



- A (k, n)-code consists of an encoder f and a decoder ϕ
- Encoder f uses modified minimum distance encoding
  - Source partition: Fix integer N and let $\{\mathcal{T}_i\}_{i \in [N]}$ be partition $\mathcal{S}^k$;
  - Source Codebook: Let $\{M_i\}_{i \in [N]}$ be a sequence of integers. For each $i \in [N]$, let $\{\hat{S}^k(i, \tilde{j})\}_{\tilde{j} \in [M_i]}$ be a sub-codebook

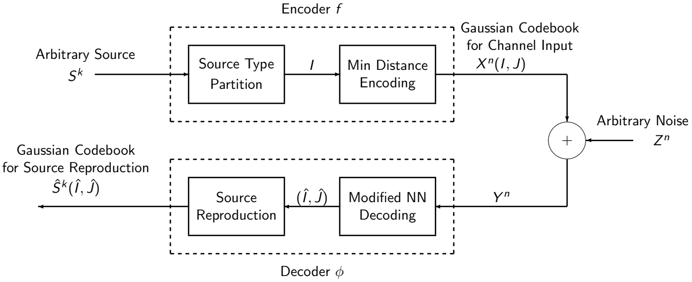
Mismatched JSCC: Problem Setting and Definitions



- A (k, n)-code consists of an encoder f and a decoder ϕ
- Encoder f uses modified minimum distance encoding
  - Given S^k, if $S^k \notin \mathcal{T}_i$ for any $i \in [N]$, the encoder f declares an error
  - Otherwise, if $S^k \in \mathcal{T}_i$ for some $i \in [N]$, the encoder f transmits $X^n(I, J)$ using modified minimum distance encoding, i.e.,

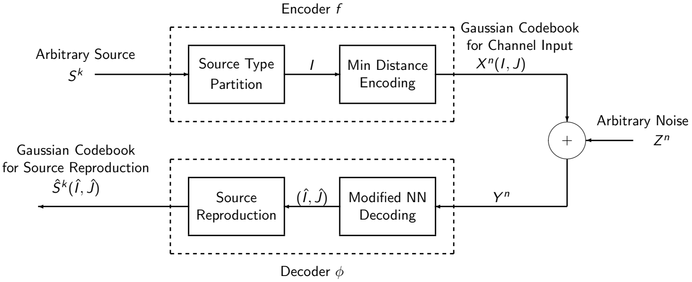
$$J = \arg\min_{\tilde{j} \in [M_I]} \|S^k, \hat{S}^k(I, \tilde{j})\|^2.$$

Mismatched JSCC: Problem Setting and Definitions



- Decoder ϕ employs the modified nearest neighbor decoding
  - For simplicity, let $\mathcal{D} := \{(r,s) \in \mathbb{N}^2 : \; r \in [N], \; s \in [M_r]\}$;

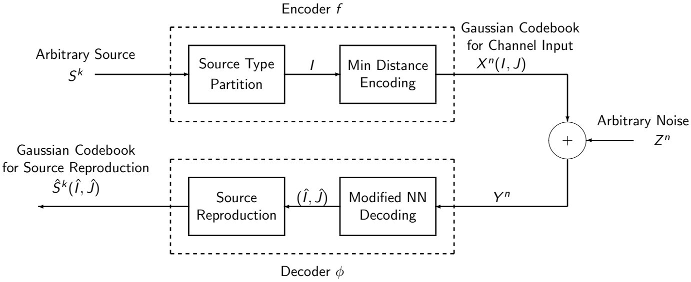
Mismatched JSCC: Problem Setting and Definitions



- Decoder ϕ employs the modified nearest neighbor decoding
  - For simplicity, let $\mathcal{D} := \{(r,s) \in \mathbb{N}^2 : \; r \in [N], \; s \in [M_r]\}$;
  - Given channel output Y^n, the decoder ϕ declares $\hat{S}^k(\hat{I},\hat{J})$ as the source estimate if

$$(\hat{I},\hat{J}) = \arg\min_{(\tilde{i},\tilde{j}) \in \mathcal{D}} \|X^n(\tilde{i},\tilde{j}) - Y^n\|_2^2 + 2\log M_{\tilde{i}}.$$

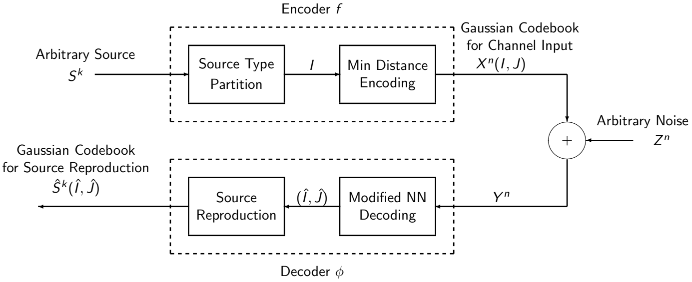
Mismatched JSCC: Problem Setting and Definitions



- Ensemble excess-distortion probability

$$\overline{\mathrm{P}}_{\mathrm{e},k,n}(D) := \Pr\{d(S^k, \phi(f(S^k)) > D\}$$

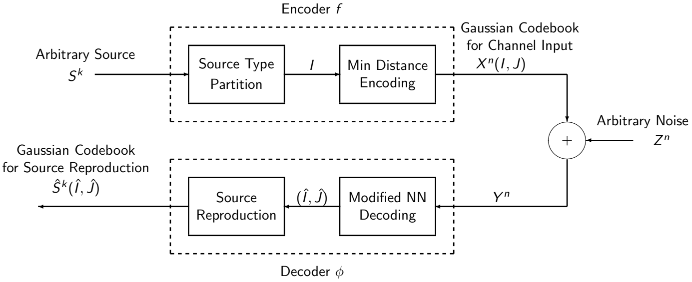
Mismatched JSCC: Problem Setting and Definitions



- Fundamental limit: for any $\varepsilon \in [0,1)$,

$$k^*_{\mathrm{sp,sp}}(n,\varepsilon,D) := \sup\{k : \overline{P}_{\mathrm{e},k,n}(D) \le \varepsilon\}.$$

Mismatched JSCC: Problem Setting and Definitions



- Fundamental limit: for any $\varepsilon \in [0,1)$,

$$k^*_{\mathrm{sp,sp}}(n,\varepsilon,D) := \sup\{k : \overline{P}_{\mathrm{e},k,n}(D) \le \varepsilon\}.$$

Similarly, we define $k^*_{\mathrm{sp,iid}}(n,\varepsilon,D)$, $k^*_{\mathrm{iid,sp}}(n,\varepsilon,D)$ and $k^*_{\mathrm{iid,iid}}(n,\varepsilon,D)$.

Dispersion for Mismatched JSCC: Preliminaries

- Optimal bandwidth expansion ratio

$$\rho^*(P, \sigma^2, D) = \frac{\mathrm{C}(P)}{\mathrm{R}(\sigma^2, D)} = \frac{\frac{1}{2}\log(1+P)}{\frac{1}{2}\log\frac{\sigma^2}{D}};$$

Dispersion for Mismatched JSCC: Preliminaries

- Optimal bandwidth expansion ratio

$$\rho^*(P, \sigma^2, D) = \frac{\mathrm{C}(P)}{\mathrm{R}(\sigma^2, D)} = \frac{\frac{1}{2}\log(1+P)}{\frac{1}{2}\log\frac{\sigma^2}{D}};$$

- Joint source-channel mismatched dispersions: for any $\dagger \in \{\mathrm{sp}, \mathrm{iid}\}$,

$$\mathrm{V}_\dagger(\sigma^2, \zeta_{\mathrm{s}}, P, \zeta_{\mathrm{c}}) := \frac{\rho^*(P, \sigma^2, D)\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}}) + \mathrm{V}_\dagger(P, \zeta_{\mathrm{c}})}{(\mathrm{R}(\sigma^2, D))^2},$$

where $\mathrm{V}(\sigma^2, \zeta_{\mathrm{s}})$ (resp. $\mathrm{V}_\dagger(P, \zeta_{\mathrm{c}})$) is the source (resp. channel) dispersion.

Dispersion for Mismatched JSCC: Main Result[7]

Theorem 5 (Zhou-Tan-Motani 2017)

With proper choices of $\mathcal{T}_i$ *and* M_i *for* $i \in [N]$, *for any* $\varepsilon \in [0,1)$ *and any* $(\ddagger, \dagger) \in \{\mathrm{sp}, \mathrm{iid}\}^2$,

$$k^*_{\ddagger,\dagger}(n, \varepsilon, D) = n\rho^*(P, \sigma^2, D) - \sqrt{n\mathrm{V}_\dagger(\sigma^2, \zeta_\mathrm{s}, P, \zeta_\mathrm{c})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

[7] L. Zhou, V. Y. F. Tan and M. Motani, "The Dispersion of Universal Joint Source-Channel Coding for Arbitrary Sources and Additive Channels", arXiv:1711.11206.

Dispersion for Mismatched JSCC: Main Result[7]

Theorem 5 (Zhou-Tan-Motani 2017)

With proper choices of $\mathcal{T}_i$ *and* M_i *for* $i \in [N]$, *for any* $\varepsilon \in [0,1)$ *and any* $(\ddagger, \dagger) \in \{\mathrm{sp}, \mathrm{iid}\}^2$,

$$k^*_{\ddagger,\dagger}(n, \varepsilon, D) = n\rho^*(P, \sigma^2, D) - \sqrt{n\mathrm{V}_\dagger(\sigma^2, \zeta_{\mathrm{s}}, P, \zeta_{\mathrm{c}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

- The second-order asymptotics depend only on the type of the channel codebook;

[7] L. Zhou, V. Y. F. Tan and M. Motani, "The Dispersion of Universal Joint Source-Channel Coding for Arbitrary Sources and Additive Channels", arXiv:1711.11206.

Dispersion for Mismatched JSCC: Main Result[7]

Theorem 5 (Zhou-Tan-Motani 2017)

With proper choices of $\mathcal{T}_i$ and M_i for $i \in [N]$, for any $\varepsilon \in [0,1)$ and any $(\ddagger, \dagger) \in \{\mathrm{sp}, \mathrm{iid}\}^2$,

$$k^*_{\ddagger,\dagger}(n,\varepsilon,D) = n\rho^*(P,\sigma^2,D) - \sqrt{n\mathrm{V}_\dagger(\sigma^2,\zeta_{\mathrm{s}},P,\zeta_{\mathrm{c}})}\mathrm{Q}^{-1}(\varepsilon) + O(\log n).$$

- The second-order asymptotics depend only on the type of the channel codebook;
- Recover the dispersion of transmitting GMS over an AWGN channel by Kostina and Verdú (TIT 2013) when setting $P_S = \mathcal{N}(0,\sigma^2)$ and $P_Z = \mathcal{N}(0,1)$;

[7] L. Zhou, V. Y. F. Tan and M. Motani, "The Dispersion of Universal Joint Source-Channel Coding for Arbitrary Sources and Additive Channels", arXiv:1711.11206.

Summary

- Mismatched channel coding
  - Spherical Gaussian codebooks outperform i.i.d. Gaussian codebooks in second-order asymptotics;

Summary

- Mismatched channel coding
  - Spherical Gaussian codebooks outperform i.i.d. Gaussian codebooks in second-order asymptotics;
- Mismatched rate-distortion problem
  - Spherical and i.i.d. Gaussian codebooks achieve the same second-order and moderate deviations asymptotics;

Summary

- Mismatched channel coding
  - Spherical Gaussian codebooks outperform i.i.d. Gaussian codebooks in second-order asymptotics;
- Mismatched rate-distortion problem
  - Spherical and i.i.d. Gaussian codebooks achieve the same second-order and moderate deviations asymptotics;
- Mismatched joint source-channel coding
  - Fixed coding scheme using UEP, modified minimum distance encoding and modified nearest neighbor decoding;
  - Second-order asymptotics depend only on the type of the channel codebook;