The Gaussian State-Dependent Channel with Rate-Limited Decoder State-Information

Shraga I. Bross
Faculty of Engineering, Bar-Ilan University
Ramat-Gan 52900, Israel
brosss@biu.ac.il

Amos Lapidoth
ETH Zurich
CH-8092 Zurich, Switzerland
lapidoth@isi.ee.ethz.ch

***Abstract*—The capacity of a state-dependent channel with rate-limited decoder state-information is the highest rate at which data can be transmitted reliably over the channel when the decoder—in addition to observing the channel output—is also furnished with a rate-limited state-description that is provided by a state-encoder observing the state-sequence noncausally. A lower bound on this capacity is derived here for the Gaussian state-dependent channel, with the lower bound being tight in the absence of noise. In this regime every additional bit of state-description increases capacity by one bit. At the other extreme, when the signal power is much smaller than the noise and state powers, capacity behaves asymptotically as though the state were known to the decoder precisely.**

I. Introduction

Consider a state-dependent (SD) memoryless channel whose state sequence S^n is drawn IID p_S and then revealed to a state encoder, which provides a rate-limited description of the state sequence to the channel decoder. The capacity of this channel is the highest rate at which data can be transmitted reliably over the channel when the decoder has access to both the channel output and the state description.

In [1] Heegard and El Gamal derived a lower bound on the capacity of a memoryless SD channel, where the transmitter is furnished with a description of the state sequence at rate R_{e} (CSIT), and the receiver is furnished with such a description at rate R_{d} (CSIR). The special case when R_{e} exceeds the entropy of the state and the receiver is thus furnished with the state sequence, was solved by Steinberg who provided a single-letter expression for the capacity [2]. The general case of Heegard and El Gamal's problem is still open.

Here we consider the average-power limited Gaussian SD channel, whose time-i output $Y_i \in \mathcal{Y}$ is

$$Y_i = x_i + S_i + Z_i, \tag{1}$$

where $x_i \in \mathcal{X}$ is the time-i input, $S_i \in \mathcal{S}$ is the time-i state, and Z_i is the time-i noise. Here $\mathcal{X}$, $\mathcal{S}$, and $\mathcal{Y}$ all denote the reals. We assume that the state sequence and the noise sequence are independent with the former being IID $\mathcal{N}\left(0, \sigma_s^2\right)$ and the latter IID $\mathcal{N}\left(0, N\right)$. We only consider the case where R_{e} is zero, so the transmitter is not furnished with any information about the state, while the receiver is provided with a description of the state at rate R_{d}. We derive a lower bound on the capacity, which is tight in the absence of additive noise, i.e., when N is zero.

We also consider the noiseless average-power limited SD Gaussian multiple-access channel (MAC) whose time-i output $Y_i \in \mathcal{Y}$ is

$$Y_i = x_{1,i} + x_{2,i} + S_i, \tag{2}$$

where $x_{k,i} \in \mathcal{X}_k, k = 1, 2$ is the time-i input of user k, and $S_i \in \mathcal{S}$ is the time-i state. We derive inner and outer bounds on the capacity region and show that when the encoders exhibit symmetric average-power constraints these bounds determine the per-user capacity.

II. The Set-Up and Main Results

Consider the Gaussian channel (1), and let R denote the data rate and R_s the state-description rate. To convey the message W that is uniformly distributed over the message set $\mathcal{W} \triangleq \{1, 2, \ldots, 2^{nR}\}$, the transmitter produces the n-tuple $X^n(W)$, which must comply with the average power constraint

$$\sum_{k=1}^{n} X_k(w)^2 \leq nP, \quad w \in \mathcal{W}. \tag{3}$$

The state encoder—to which the state sequence S^n is provided noncausally—describes it as $T(S^n)$, where $T(S^n)$ takes values in the set $\mathcal{T} \triangleq \{1, 2, \ldots, 2^{nR_s}\}$. The description $T(S^n)$ is forwarded to the receiver noiselessly via some bit pipe of capacity R_s. Based on the channel outputs Y^n and the state-sequence description $T(S^n)$, the receiver guesses the message W. The decoder is thus a mapping of the form

$$g\colon \mathcal{Y}^n \times \mathcal{T} \to \mathcal{W}. \tag{4}$$

We shall sometimes use

$$f\colon \mathcal{W} \to \mathcal{X}^n, \tag{5}$$

to denote the transmitter's mapping of the message W to the transmitted sequence $X(W) = f(W)$, and use

$$f_s\colon \mathcal{S}^n \to \mathcal{T}, \tag{6}$$

to denote the description of the state sequence, so $T(S^n) = f_s(S^n)$.

A pair (R, R_s) is achievable if for every $\varepsilon > 0$ there exists some sufficiently large n_0 such that for every blocklength n

978-1-5090-2152-9/16/$31.00

exceeding n_0 there exists an message-encoding mapping f whose rate exceeds $R-\varepsilon$, a state-encoding mapping f_s, whose rate does not exceed $R_s+\varepsilon$, and a decoding mapping g whose average probability of error $P_{\mathrm{e}}^{(n)}$ is smaller than ε, where

$$P_{\mathrm{e}}^{(n)} = 2^{-nR} \sum_{w,s^n} P_s(s^n) P_{Y^n|X^n S^n}\left(\mathcal{A}_w^c(f_s(s^n)) \middle| f(w), s^n\right), \quad (7)$$

where

$$\mathcal{A}_w(t) = \{y^n \in \mathcal{Y}^n | g(y^n, t) = w\}$$

denotes the decoding set of the message w when the description of the state is $f_s(s^n) = t$, and $\mathcal{A}_w^c(t)$ is its complement $\mathcal{Y}^n \setminus \mathcal{A}_w(t)$.

We denote by $\mathscr{R}$ the set of achievable (R, R_s) pairs, and for every given maximal allowed description rate R_s we define the capacity $C(R_s)$ as the supremum of all the rates R for which (R, R_s) is achievable.

Proposition 1 (State-and-Noise): Consider the SD Gaussian channel (1) of noise variance N, state-variance σ_s^2 and maximal-allowed average power P. Let the state description rate be R_s. In this setting

$$C(R_s) \geq \frac{1}{2} \log \frac{P+N}{N + \frac{P d_{\mathrm{WZ}}}{P+N}} \quad (8)$$

where

$$d_{\mathrm{WZ}} \triangleq \sigma_s^2 \, 2^{-2R_s} \, \frac{P+N}{P+N+\sigma_s^2}. \quad (9)$$

Discussion: The achievable rate (8) may be interpreted as follows. The distortion of the state estimator $\hat{S}$ based on the state-encoder description and the channel output $Y = X + S + N$ is given by d_{WZ} of (9). The achievable rate (8) then equals the mutual information $I(X; Y\hat{S})$ for jointly Gaussian $(X, Y, \hat{S})$.

Furthermore, since the receiver decodes the message W, it can recover the transmitted codeword $X^n(W)$, and its least mean squared-error in estimating the state sequence is bounded by

$$\mathbb{E}\left[(\mathbf{S}-\hat{\mathbf{S}})^2\right] \leq \frac{N d_{\mathrm{WZ}}}{N + \frac{P d_{\mathrm{WZ}}}{P+N}}. \quad (10)$$

We next establish that asymptotically, as the power tends zero, capacity behaves as though the state were known to the receiver perfectly.

Proposition 2 (State-and-Noise—Low SNR*):* Consider the SD Gaussian channel (1) of noise variance N, state-variance σ_s^2, and maximal-allowed average power P. For every positive R_s

$$\lim_{P \downarrow 0} \frac{C(R_s)}{P/(2N)} = \log e. \quad (11)$$

In the absence of noise, i.e., when $N = 0$ we have a conclusive result:

Theorem 1 (State-without-Noise): In the setting of Proposition 1 with $N = 0$

$$C(R_s) = R_s + \frac{1}{2} \log \left(1 + \frac{P}{\sigma_s^2}\right). \quad (12)$$

Discussion: When R_s is zero, the receiver obtains no side information about the state sequence and it is forced to treat it as noise. The capacity is then the second term on the RHS of (12). Each additional bit of state description increases this capacity by one bit.

Furthermore, since the receiver decodes the message W, it can recover the transmitted codeword $X^n(W)$, and its least mean squared-error in estimating the state sequence can be made arbitrarily small

$$\mathbb{E}\left[(\mathbf{S}-\hat{\mathbf{S}})^2\right] \leq \varepsilon, \quad (13)$$

by choosing the blocklength sufficiently large.

Next we consider the Gaussian MAC model (2).

For a given $R_s > 0$ and a nonnegative tuple (σ_s^2, P_1, P_2) define $\mathcal{R}^{(\mathrm{in})}(R_s)$ as

$$\begin{aligned} \mathcal{R}^{(\mathrm{in})}(R_s) \triangleq \Big\{ & (R_1, R_2) \colon R_1 > 0, R_2 > 0, \\ & R_1 \leq \frac{1}{2} \log \left(\frac{P_1 + P_2 2^{-2R_s}}{P_1 + P_2} + \frac{P_1}{\sigma_s^2} \right) + R_s \\ & R_2 \leq \frac{1}{2} \log \left(\frac{P_1 2^{-2R_s} + P_2}{P_1 + P_2} + \frac{P_2}{\sigma_s^2} \right) + R_s \\ & R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + P_2}{\sigma_s^2} \right) + R_s \Big\}, \end{aligned} \quad (14)$$

and define $\mathcal{R}^{(\mathrm{out})}(R_s)$ as

$$\begin{aligned} \mathcal{R}^{(\mathrm{out})}(R_s) \triangleq \Big\{ & (R_1, R_2) \colon R_1 > 0, R_2 > 0, \\ & R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{\sigma_s^2} \right) + R_s \\ & R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_2}{\sigma_s^2} \right) + R_s \\ & R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + P_2}{\sigma_s^2} \right) + R_s \Big\}. \end{aligned} \quad (15)$$

Theorem 2 (MAC and State-without-Noise): Consider the SD Gaussian multiple-access channel $Y^n = X_1^n + X_2^n + S^n$ of state-variance σ_s^2 and maximal-allowed average powers P_1 and P_2, respectively. Let the state description rate be R_s. In this setting the capacity region is bounded as follows

$$\mathcal{R}^{(\mathrm{in})}(R_s) \subseteq \mathcal{C}(R_s) \subseteq \mathcal{R}^{(\mathrm{out})}(R_s), \quad (16)$$

hence the symmetric ($P_1 = P_2 = P$) per user capacity is

$$C(R_s) = R_s/2 + \frac{1}{4} \log \left(1 + \frac{2P}{\sigma_s^2}\right). \quad (17)$$

III. Techniques

A. Proof of Proposition 1

1) Coding scheme: Fix some $\varepsilon > 0$ and a rate pair (R, R_s).

Code Construction: Two codebooks $\mathcal{C}$ and $\mathcal{C}_{\rm w}$ are generated independently. Codebook $\mathcal{C}$ consists of $2^{n(R_s+\tilde{R})}$ codewords $\{\mathbf{U}(1), \mathbf{U}(2), \dots, \mathbf{U}(2^{n(R_s+\tilde{R})})\}$, drawn independently and uniformly over the surface of the centered $\mathbb{R}^n$-sphere $\mathcal{S}$ of radius $r = \sqrt{n\sigma_s^2(1-\delta_{\rm Q})}$, where

$$\delta_{\rm Q} \triangleq \frac{d_{\rm WZ}}{\sigma_s^2\left(1 - \frac{d_{\rm WZ}}{P+N}\right)} = 2^{-2(R_s+\tilde{R})}, \tag{18}$$

hence by (9)

$$\tilde{R} = \frac{1}{2}\log\left(\frac{\sigma_s^2 + P + N}{P+N}\frac{P+N-d_{\rm WZ}}{P+N}\right). \tag{19}$$

The codebook $\mathcal{C}$ is partitioned randomly into 2^{nR_s} equal-size bins $\mathcal{C}_{\rm b}(1), \mathcal{C}_{\rm b}(2), \dots, \mathcal{C}_{\rm b}(2^{nR_s})$. For $j \in \{1, 2, \dots, 2^{n(R_s+\tilde{R})}\}$ let $b(j) \in \{1, 2, \dots, 2^{nR_s}\}$ denote the bin to which $\mathbf{U}(j)$ belongs –i.e. $\mathbf{U}(j) \in \mathcal{C}_{\rm b}(b(j))$.

The codebook $\mathcal{C}_{\rm w}$ consists of a set of 2^{nR} codewords $\{\mathbf{X}(1), \mathbf{X}(2), \dots, \mathbf{X}(2^{nR})\}$. The codewords are drawn independently uniformly over the surface of the centered $\mathbb{R}^n$-sphere $\mathcal{S}_{\rm w}$ of radius $r_{\rm w} = \sqrt{nP}$.

For every $\mathbf{w}, \mathbf{v} \in \mathbb{R}^n$ where neither $\mathbf{w}$ nor $\mathbf{v}$ are the zero-sequence, denote the angle between $\mathbf{w}$ and $\mathbf{v}$ by $\sphericalangle(\mathbf{w}, \mathbf{v})$. i.e.,

$$\cos\sphericalangle(\mathbf{w}, \mathbf{v}) \triangleq \frac{\langle \mathbf{w}, \mathbf{v}\rangle}{\|\mathbf{w}\|\|\mathbf{v}\|}.$$

Encoding: Given the source sequence $\mathbf{s}$, let $\mathcal{F}(\mathbf{s}, \mathcal{C})$ be the set defined by

$$\mathcal{F}(\mathbf{s}, \mathcal{C}) \triangleq \left\{\mathbf{u} \in \mathcal{C} \colon \left|\cos\sphericalangle(\mathbf{s}, \mathbf{u}) - \sqrt{1-\delta_{\rm Q}}\right| \le \varepsilon\sqrt{1-\delta_{\rm Q}}\right\}.$$

The state encoder vector-quantizes $\mathbf{s}$ as follows.

If $\mathcal{F}(\mathbf{s}, \mathcal{C}) \neq \emptyset$ it chooses the vector $\mathbf{u}^*$ as the codeword $\mathbf{u}(j^*) \in \mathcal{F}(\mathbf{s}, \mathcal{C})$ where j^* minimizes $|\cos\sphericalangle(\mathbf{s}, \mathbf{u}(j)) - \sqrt{1-\delta_{\rm Q}}|$, while if $\mathcal{F}(\mathbf{s}, \mathcal{C}) = \emptyset$ then $\mathbf{u}^* = \mathbf{u}(1)$. Formally,

$$\mathbf{u}^* = \begin{cases} \underset{\mathbf{u}(j) \in \mathcal{F}(\mathbf{s},\mathcal{C})}{\arg\min} |\cos\sphericalangle(\mathbf{s}, \mathbf{u}(j)) - \sqrt{1-\delta_{\rm Q}}| & \mathcal{F}(\mathbf{s}, \mathcal{C}) \neq \emptyset \\ \mathbf{u}(1) & \mathcal{F}(\mathbf{s}, \mathcal{C}) = \emptyset \end{cases}$$

Given that $\mathbf{u}^* = \mathbf{u}(j^*)$, so that the state sequence can be expressed as

$$\mathbf{s} = \mathbf{u}^* + \mathbf{Z}_{\rm Q}$$

with $\mathbf{Z}_{\rm Q} \sim \mathcal{N}(0, n\sigma_s^2\delta_{\rm Q})$ independent of $\mathbf{u}^*$, the state encoder sends the bin index $b(j^*)$ to the decoder.

Given the message $W = w$, the channel encoder chooses the channel input as $\mathbf{X}(w)$.

Decoding: Given the channel output $\mathbf{y}$ and the state encoder input $b(j)$, the decoder forms its estimate $\hat{w}$ of the message W as follows.

Simultaneous decoder:

The decoder looks for a pair (j^*, w^*), where $\mathbf{u}(j^*) \in \mathcal{C}_{\rm b}(b(j^*))$, such that

$$(j^*, w^*) = \arg\min_{j,w}\|\mathbf{y} - \mathbf{u}(j) - \mathbf{x}(w)\|^2, \tag{20}$$

and chooses $\hat{w} = w^*$ as its estimate.

To analyze the error probability of the decoder (20) we view $\mathbf{y}$ as the output of an additive Gaussian noise MAC with two independent inputs; $\mathbf{x}(w)$ and $\mathbf{u}(j)$, where $\mathbf{X}(w) \sim \mathcal{N}(0, nP)$ and $\mathbf{U}(j) \sim \mathcal{N}(0, n\sigma_s^2(1-\delta_{\rm Q}))$. The codebook $\{\mathbf{X}(w)\}$ consists of 2^{nR} IID codewords while the "codebook" $\{\mathbf{U}(j)\}$ consists of the $2^{n\tilde{R}}$ source codewords within the bin $\mathcal{C}_{\rm b}(b(j)$. Consequently, the MAC can be modeled as follows

$$Y = X + U + Z_{\rm Q} + Z \tag{21}$$

where both $Z_{\rm Q} \sim \mathcal{N}(0, \sigma_s^2\delta_{\rm Q})$ and $Z \sim \mathcal{N}(0, N)$ are independent of each other and of the rest.

There are two relevant error events to be considered

1) The event $\{w^* \neq w, j^* \neq j\}$: Using [3], it can be verified that if

$$R + \tilde{R} \le \frac{1}{2}\log\left(\frac{\sigma_s^2 + P + N}{N + \sigma_s^2\delta_{\rm Q}}\right) \tag{22}$$

then $\Pr\{w^* \neq w, j^* \neq j\} \to 0$, for sufficiently large n.

2) The event $\{w^* \neq w, j^* = j\}$: Similarly, it can be verified that if

$$R \le \frac{1}{2}\log\left(\frac{\sigma_s^2\delta_{\rm Q} + P + N}{N + \sigma_s^2\delta_{\rm Q}}\right) \tag{23}$$

then $\Pr\{w^* \neq w, j^* = j\} \to 0$, for sufficiently large n.

It can now be verified that the substitution of $\delta_{\rm Q}$ as per (18) in (23) yields

$$R < \frac{1}{2}\log\frac{P+N}{N + \frac{Pd_{\rm WZ}}{P+N}}, \tag{24}$$

while the substitution of $\delta_{\rm Q}$ as per (18) and $\tilde{R}$ as per (19) in (22) yields also the constraint (24). This establishes the desired result.

Successive decoder:

The decoder looks first for a source codeword $\mathbf{u}(j^*)$, where $\mathbf{u}(j^*) \in \mathcal{C}_{\rm b}(b(j^*))$, such that

$$j^* = \arg\min_j\|\mathbf{y} - \mathbf{u}(j)\|^2. \tag{25}$$

Having decoded $\mathbf{u}$ the decoder looks for w^* such that

$$w^* = \arg\min_w\|\mathbf{y} - \mathbf{u}(j^*) - \mathbf{x}(w)\|^2, \tag{26}$$

and chooses $\hat{w} = w^*$ as its estimate.

To analyze the error probability of the successive decoder (25)–(26) we view $\mathbf{y}$ as the output of the MAC (21) and consider the following two error events:

1) The event $\{j^* \neq j\}$: It can be verified that if

$$\tilde{R} \le \frac{1}{2}\log\left(\frac{\sigma_s^2 + P + N}{P + N + \sigma_s^2\delta_{\rm Q}}\right) \tag{27}$$

then the probability that $j^* \neq j$ tends to zero for sufficiently large n.

2) The event $\{w^* \neq w, j^* = j\}$: It can be verified that if

$$R \leq \frac{1}{2}\log\left(\frac{\sigma_s^2\delta_{\rm Q} + P + N}{N + \sigma_s^2\delta_{\rm Q}}\right) \tag{28}$$

then the probability that $w^* \neq w$, conditioned on $j^* = j$, tends to zero for sufficiently large n.

It can be verified that the substitution of $\delta_{\rm Q}$ as per (18) in (27) yields

$$\tilde{R} < \frac{1}{2}\log\left(\frac{\sigma_s^2 + P + N}{P + N}\frac{P + N - d_{\rm WZ}}{P + N}\right), \tag{29}$$

which is in accordance with (19). Next, the substitution of $\delta_{\rm Q}$ as per (18) in (28) yields the constraint (24). This establishes the desired result.

State Estimation: Having decoded $(\mathbf{u}(j^*), \mathbf{x}(w^*))$ the receiver forms its estimate for $\mathbf{s}$ as follows

$$\begin{aligned}\hat{\mathbf{S}} &= \mathbf{u}(j^*) + \hat{\mathbf{Z}}_{\rm Q} \\ &= \mathbf{u}(j^*) + (\mathbf{y} - \mathbf{u}(j^*) - \mathbf{x}(w^*))\frac{\sigma_s^2\delta_{\rm Q}}{N + \sigma_s^2\delta_{\rm Q}}.\end{aligned} \tag{30}$$

Consequently

$$\frac{1}{n}\mathbb{E}\left[(\mathbf{S} - \hat{\mathbf{S}})^2\right] = \frac{N\sigma_s^2\delta_{\rm Q}}{N + \sigma_s^2\delta_{\rm Q}}. \tag{31}$$

The substitution of $\delta_{\rm Q}$ as per (18) into (31) yields the claimed bound (10).

B. Proof of Theorem 1

Converse : Consider a sequence of (e^{nR}, e^{nR_s}, n) codes with average block error probability ε_n, where $\varepsilon_n \to 0$ as $n \to \infty$, and a law on $\mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{S}^n$ given by

$$p_{WX^nS^nY^n} = p_W p_{S^n} I_{\{X^n = f(W)\}} \prod_{k=1}^{n} p_{Y_k|X_kS_k} \, . \tag{32}$$

Denote the state-encoder output by $T = f_s(S^n)$ then

$$p_{S^nT} = p_{S^n}p_{T|S^n} = p_{S^n}I_{\{T = f_s(S^n)\}}. \tag{33}$$

We consider first the state encoder

$$\begin{aligned} nR_s \geq H(T) &= H(T) - H(T|S^n) \\ &= I(T;S^n) \\ &= h(S^n) - h(S^n|T). \end{aligned} \tag{34}$$

So

$$\begin{aligned} h(S^n|T) &\geq h(S^n) - nR_s \\ &= \frac{n}{2}\log(2\pi e\sigma_s^2) - nR_s. \end{aligned} \tag{35}$$

Next we apply Fano's inequality at the decoder.

$$\begin{aligned} n(R - \eta_n) &\leq I(W;TY^n) \overset{(a)}{=} I(W;Y^n|T) \\ &= I(WX^n;Y^n|T) \overset{(b)}{=} I(X^n;Y^n|T), \end{aligned} \tag{36}$$

where (a) follows since T is independent of W and (b) follows since $W \multimap (X^n,T) \multimap Y^n$ is a Markov chain.

Next we expand the right-hand-side of (36) as follows

$$\begin{aligned} I(X^n;Y^n|T) &= h(Y^n|T) - h(Y^n|T,X^n) \\ &= h(X^n + Z^n + S^n|T) - h(X^n + Z^n + S^n|T,X^n) \\ &= h(X^n + Z^n + S^n|T) - h(Z^n + S^n|T,X^n) \\ &\overset{(c)}{=} h(X^n + Z^n + S^n|T) - h(Z^n + S^n|T) \\ &\leq h(X^n + Z^n + S^n) - h(Z^n + S^n|T), \end{aligned} \tag{37}$$

where (c) follows since X^n is independent of (S^n, Z^n, T).

When $Z^n = \mathbf{0}$ the combination of (36), (37) and (35) yields

$$\begin{aligned} n(R - \eta_n) &\leq h(X^n + S^n) - h(S^n|T) \\ &\leq h(X^n + S^n) - \frac{n}{2}\log(2\pi e\sigma_s^2) + nR_s \\ &\overset{(d)}{\leq} \frac{n}{2}\log\left(2\pi e(P + \sigma_s^2)\right) - \frac{n}{2}\log(2\pi e\sigma_s^2) + nR_s \\ &\leq \frac{n}{2}\log\left(1 + \frac{P}{\sigma_s^2}\right) + nR_s, \end{aligned} \tag{38}$$

where (d) follows since $\mathbb{E}[(X_k + S_k)^2] \leq P + \sigma_s^2$.

Direct : The substitution of $N = 0$ in (8) and (9) yields

$$C(R_s) \geq \frac{1}{2}\log\frac{P}{d_{\rm WZ}^{(0)}} \tag{39}$$

where

$$d_{\rm WZ}^{(0)} \triangleq \sigma_s^2 2^{-2R_s}\frac{P}{P + \sigma_s^2}. \tag{40}$$

So

$$C(R_s) \geq \frac{1}{2}\log\left(1 + \frac{P}{\sigma_s^2}\right) + R_s. \tag{41}$$

C. Proof of Proposition 2

First, note that

$$\frac{P + N}{N + \frac{Pd_{\rm WZ}}{P+N}} = \frac{P + N}{N + \frac{P}{P+N+\sigma_s^2}\sigma_s^2 2^{-2R_s}}\Bigg|_{\frac{P}{N} \ll 1} = \frac{1 + \frac{P}{N}}{1 + \frac{P}{N}\alpha}$$

where

$$\alpha \triangleq \frac{\sigma_s^2 2^{-2R_s}}{N + \sigma_s^2}.$$

Fix P and R_s and consider a time-sharing approach which at fraction $0 < \theta < 1$ of the time uses the code construction of Proposition 1 while at the rest of the time neither the data-encoder nor the state-encoder operate. Using the expression (8) the achievable rate with this strategy is

$$C(R_s) \geq \frac{1}{2}\log\frac{1 + \frac{P}{\theta N}}{1 + \frac{P}{\theta N}\frac{\sigma_s^2 2^{-2R_s/\theta}}{N + \sigma_s^2}}. \tag{42}$$

Next, fix $0 < \theta < 1$ such that

$$\frac{\sigma_s^2 2^{-2R_s/\theta}}{N + \sigma_s^2} \ll 1,$$

in which case the lower bound (42) becomes

$$C(R_s) \geq \frac{1}{2}\log\left(1+\frac{P}{\theta N}\right). \tag{43}$$

The claimed lower bound now follows by noting that

$$\lim_{P\to 0}\frac{N}{P}\log\left(1+\frac{P}{\theta N}\right)^{\theta} = \log e. \tag{44}$$

The upper bound follows by assuming that S^n is conveyed perfectly to the receiver.

Furthermore, since

$$\lim_{P\to 0}\frac{1+\frac{P}{N}}{1+\frac{P}{N}\alpha} \approx \left(1+\frac{P}{N}\right)\left(1-\alpha\frac{P}{N}\right) \approx 1+(1-\alpha)\frac{P}{N}$$

the non time-sharing strategy, that is used to prove Proposition 1, is sub-optimal.

D. Proof of Theorem 2

1) Coding scheme: Fix some $\varepsilon > 0$ and a rate pair (R, R_s). We propose a code construction for the average-power limited Gaussian SD-MAC, whose Time-i output $Y_i \in \mathcal{Y}$ is

$$Y_i = x_{1,i} + x_{2,i} + S_i + Z_i, \tag{45}$$

where $x_{k,i} \in \mathcal{X}_k,\ k = 1, 2$ is the Time-i input of Encoder k, $S_i \in \mathcal{S}$ is the Time-i state, and Z_i is the Time-i noise. The error probability analysis that follows refers to the model (45), and then in order to identify the error event which imposes the tightest constraint on the achievable symmetric rate we consider the special case where $Z_i = 0$.

Code Construction: Two codebooks $\mathcal{C}$, $\mathcal{C}_{\text{w}}^{(1)}$ and $\mathcal{C}_{\text{w}}^{(2)}$ are generated independently. Codebook $\mathcal{C}$ consists of $2^{n(R_s+\tilde{R})}$ codewords $\{\mathbf{U}(1), \mathbf{U}(2), \ldots, \mathbf{U}(2^{n(R_s+\tilde{R})})\}$, drawn independently and uniformly over the surface of the centered $\mathbb{R}^n$-sphere $\mathcal{S}$ of radius $r = \sqrt{n\sigma_s^2(1-\delta_{\text{Q}})}$, where

$$\delta_{\text{Q}} \triangleq \frac{d_{\text{WZ}}}{\sigma_s^2\left(1-\frac{d_{\text{WZ}}}{P_1+P_2+N}\right)} = 2^{-2(R_s+\tilde{R})} \tag{46}$$

$$d_{\text{WZ}} \triangleq \sigma_s^2 2^{-2R_s}\frac{P_1+P_2+N}{P_1+P_2+N+\sigma_s^2}, \tag{47}$$

hence

$$\tilde{R} = \frac{1}{2}\log\left(\frac{\sigma_s^2+P_1+P_2+N}{P_1+P_2+N}\frac{P_1+P_2+N-d_{\text{WZ}}}{P_1+P_2+N}\right). \tag{48}$$

The codebook $\mathcal{C}$ is partitioned randomly into 2^{nR_s} equal-size bins $\mathcal{C}_{\text{b}}(1), \mathcal{C}_{\text{b}}(2), \ldots, \mathcal{C}_{\text{b}}(2^{nR_s})$. For $j \in \{1, 2, \ldots, 2^{n(R_s+\tilde{R})}\}$ let $b(j) \in \{1, 2, \ldots, 2^{nR_s}\}$ denote the bin to which $\mathbf{U}(j)$ belongs –i.e. $\mathbf{U}(j) \in \mathcal{C}_{\text{b}}(b(j))$.

The codebook $\mathcal{C}_{\text{w}}^{(k)},\ k = 1, 2$ consists of a set of 2^{nR_k} codewords $\{\mathbf{X}_k(1), \mathbf{X}_k(2), \ldots, \mathbf{X}_k(2^{nR_k})\}$. The codewords are drawn independently uniformly over the surface of the centered $\mathbb{R}^n$-sphere $\mathcal{S}_{\text{w}}^{(k)}$ of radius $r_{\text{w}}^{(k)} = \sqrt{nP_k}$.

Encoding: Given the source sequence $\mathbf{s}$, the state encoder vector-quantizes $\mathbf{s}$ by choosing $\mathbf{u}^*$ as follows.

$$\mathbf{u}^* = \begin{cases} \arg\min_{\mathbf{u}(j)\in\mathcal{F}(\mathbf{s},\mathcal{C})} |\cos\sphericalangle(\mathbf{s},\mathbf{u}(j)) - \sqrt{1-\delta_{\text{Q}}}| & \mathcal{F}(\mathbf{s},\mathcal{C}) \neq \emptyset \\ \mathbf{u}(1) & \mathcal{F}(\mathbf{s},\mathcal{C}) = \emptyset \end{cases}$$

Given that $\mathbf{u}^* = \mathbf{u}(j^*)$, so that the state sequence can be expressed as

$$\mathbf{s} = \mathbf{u}^* + \mathbf{Z}_{\text{Q}}$$

with $\mathbf{Z}_{\text{Q}} \sim \mathcal{N}(0, n\sigma_s^2\delta_{\text{Q}})$ independent of $\mathbf{u}^*$, the state encoder sends the bin index $b(j^*)$ to the decoder.

Given the message pair $(W_1 = w_1, W_2 = w_2)$, the channel encoders choose their channel inputs as $\mathbf{X}_k(w_k),\ k = 1, 2$.

Decoding: Given the channel output $\mathbf{y}$ and the state encoder input $b(j)$, the decoder forms its estimate $(\hat{w}_1, \hat{w}_2)$ of the message pair (W_1, W_2) as follows.

Simultaneous decoder:

The decoder looks for a triple (j^*, w_1^*, w_2^*), where $\mathbf{u}(j^*) \in \mathcal{C}_{\text{b}}(b(j^*))$, such that

$$(j^*, w_1^*, w_2^*) = \arg\min_{j,w_1,w_2}\|\mathbf{y}-\mathbf{u}(j)-\mathbf{x}_1(w_1)-\mathbf{x}_2(w_2)\|^2, \tag{49}$$

and chooses $(\hat{w}_1 = w_1^*, \hat{w}_2 = w_2^*)$ as its estimate.

The error probability analysis for the decoder (49) establishes the achievability of any rate pair inside $\mathcal{R}^{(\text{in})}(R_s)$, while the proof of the outer bound $\mathcal{R}^{(\text{out})}(R_s)$ follows similarly to the proof of the converse of Theorem 1.

ACKNOWLEDGMENT

The work of S. Bross was supported by the Israel Science Foundation under Grant 455/14.

REFERENCES

[1] C. Heegard and A. El Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inform. Theory,* vol. IT-29, no. 5, pp. 731–739, Sep. 1983.

[2] Y. Steinberg, "Coding for channels with rate-limited side-information at the decoder, with applications," *IEEE Trans. Inform. Theory,* vol. IT-54, no. 9, pp. 4283–4295, Sep. 2008.

[3] A. Lapidoth,"Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inform. Theory,* vol. IT-42, no. 5, pp. 1520–1529, Sep. 1996.