On the Capacity of the Carbon Copying onto Dirty Paper Channel

Stefano Rini * and Shlomo Shamai (Shitz) †
* National Chiao-Tung University, Hsinchu, Taiwan
E-mail: stefano@nctu.edu.tw
† Technion-Israel Institute of Technology, Haifa, Israel
E-mail: sshlomo@ee.technion.ac.il

Abstract—Costa’s “writing on dirty paper” capacity result establishes that full state pre-cancellation can be attained in the Gel’fand-Pinsker channel with additive state and additive Gaussian noise. The “carbon copy onto dirty paper” channel is the extension of Costa’s model to the compound setting: here a common message is to be decoded at $M$ receivers, each observing the sum of the channel input, Gaussian noise and one of $M$ possible state sequences. As in the Gel’fand-Pinsker channel, the state sequences are anti-causally known at the encoder, which can thus pre-code its transmissions so as to minimize the adverse effects of the channel states. In this paper we derive the capacity to within $2.25$ bits-per-channel-use of the $M$-user “carbon copying onto dirty paper” channel in which the state sequences are Gaussian distributed with the same variance and the same pairwise correlation.

Index Terms—Gel’fand-Pinsker Problem; Writing on Dirty Paper; Carbon Copying onto Dirty Paper; Costa Pre-Coding;

The Gel’fand-Pinsker (GP) channel [1] is the point-to-point channel in which the channel output is obtained as a random function of the input and a state sequence which is non-causally known at the encoder but is unknown at the decoder. Costa’s “Writing on Dirty Paper” (WDP) channel [2] is the Gaussian version of the GP channel in which the channel output is a linear combination of the input, the state sequence and iid Gaussian-distributed noise. Perhaps surprisingly, Costa showed that the capacity of the WDP channel is the same as the capacity of the point-to-point channel without state. In other words, the encoder can fully pre-code its transmissions against the channel state and the presence of the state does not affect capacity. The “Carbon Copying onto Dirty Paper” (CCDP) channel [3] is the extension of the GP channel to the compound setting: in this model the transmitter wishes to communicate a common message to $M$ receivers where each receiver observes the sum of the input, iid Gaussian noise and one of $M$ possible state sequences. As in the GP channel, all the state sequences are anti-causally known at the transmitter but unknown at the receivers. The CCDP channel can be used to model two scenarios: (i) a WDP channel in which the state sequence is randomly chosen among $M$ possible realizations, all known at the transmitter and (ii) an overlay multicast channel in which the transmitter wishes to communicate the same message to multiple receivers, while having knowledge of the underlying transmissions. In this correspondence we derive the approximate capacity of the CCDP channel with Gaussian-distributed state sequences with the same variance and the same pairwise correlation and any number of receivers.

While the GP channel is a well studied model, not many results are available in the literature for compound extensions of this channel, such as the CCDP. An achievable region for the 2-user compound GP channel is presented in [4] which employs a common codeword, simultaneously decoded at both receivers, and two private codewords, each decoded only at one receiver. In [3] the authors derive the first inner and outer bounds to capacity for the $M$-user CCDP. In [5] we focus on the 2-user CCDP channel and derive the approximate capacity for a certain set of correlations among the state sequences at each receiver.

A model related to the CCDP channel is the state-dependent broadcast channel with a common message: this model is obtained from the CCDP channel by introducing a private message to be transmitted from the encoder to each of the decoders. A first achievable region for this channel is obtained in [6] by combining coding strategies for the GP channel and the broadcast channel. In [7] Steinberg determines the capacity for this model when the channel output at one user is a degraded version of the output at the other user and, additionally, the non-degraded user is provided with the channel state sequences.

The remainder of the paper is organized as follows: in Sec. I we introduce the channel model while, in Sec. II, we present the relevant results available in the literature. In Sec. III we derive the approximate capacity for the 2-user CCDP with independent channel states. This result is generalized to any number of users in Sec. V. In Sec. IV we further extend the results of the previous sections by showing the approximate capacity for the $M$-users CCDP where states have equal variance and equal pairwise correlation. Finally, Sec. VII concludes the paper.

The work of S. Rini was funded by the Ministry Of Science and Technology (MOST) under the grant 103-2218-E-009-014-MY2. The work of S. Shamai has been supported by the European Union’s Horizon 2020 Research And Innovation Programme, grant agreement no. 694630.
The $M$-user “Carbon Copying onto Dirty Paper” (CCDP) channel, also depicted in Fig. 1, is the compound channel in which the channel outputs are obtained as

$$Y_m^N = X^N + cS_m^N + Z_m^N, \quad m \in [1 \ldots M],$$  \hspace{1cm} (1)

where $c \geq 0$, $Z_m^N$, $\forall m$ is an iid Gaussian sequence with zero mean and unitary variance and $\{S_m^N, m \in [1 \ldots M]\}$ is an iid vector of jointly Gaussian random variables with zero mean and covariance matrix $\Sigma_S$ with

$$1 = \text{Var}[S_1] \leq \text{Var}[S_2] \leq \ldots \leq \text{Var}[S_M].$$  \hspace{1cm} (2)

The channel input $X^N$ in (1) is additionally subject to the average power constraint $\sum_{i=1}^N \mathbb{E}[|X_i|^2] \leq NP$. The parameter $c$ is referred to as “state gain”.

Note that assumption of $c \geq 0$ and $\Sigma_S$ satisfying (2) are without loss of generality, given the symmetry of the state distribution and the fact that capacity is unchanged by a relabelling of the users.

In each of the three sections that follows we show the approximate capacity of three channel models of increasing generality:

- **Sec. III:** $M = 2$, $\text{Var}[S_1] = \text{Var}[S_2] = 1$ and $\mathbb{E}[S_1S_2] = 0$.
- **Sec. IV:** any $M$, $\text{Var}[S_m] = 1$ $\forall m$ and $\mathbb{E}[S_mS_q] = 0$, $m \neq q$.
- **Sec. V:** any $M$, $\text{Var}[S_m] = 1$ $\forall m$ and $\mathbb{E}[S_mS_q] = \rho$, $m \neq q$.

We term the first two models as “CCDP with Independent States” (CCDP-IS) channel and the latter model as “CCDP with Equivalent States” (CCDP-ES) channel.

The range of feasible pairwise correlation values $\rho$ for the CCPD-ES channel is shown in the next lemma.

**Lemma I.1. Feasible pairwise correlation.** The $M \times M$ matrix

$$\Sigma_S = \begin{bmatrix} 1 & \rho & \rho & \ldots \\ \rho & 1 & \ddots & \ddots \\ \rho & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & 1 \end{bmatrix},$$  \hspace{1cm} (3)

is positive defined for $-1/(M-1) \leq \rho \leq 1$.

A crucial observation that allows us to improve upon the available capacity outer bounds is stated by the next lemma.

**Lemma I.2. Capacity as a function of $c$.** The capacity of the CCDP channel is non-increasing in the state gain $c$.

This result is rather intuitive since capacity can only decrease if we increase the amplitude of the state in each channel output. We assume standard definitions for code, achievable rate, capacity and approximate capacity [8].

**II. RELATED RESULTS**

In this section we briefly review the results available in [3] for the CCDP channel with independent and equal-variance Gaussian states.

**Theorem II.1. Inner and outer bounds for the 2-CCDP-IS channel [3, Th. 3, Th. 4].** Consider the 2-CCDP channel in (1) for $\text{Var}[S_1] = \text{Var}[S_2] = 1$ and $\mathbb{E}[S_1S_2] = 0$, then the capacity is upper bounded as

$$C \leq R^{\text{OUT}} = \begin{cases} \frac{1}{4} \log \left( \frac{1+P}{c^2/4+P} \right) + \frac{1}{2} \log\left( 1 + \frac{c^2/4+2c\sqrt{P}}{c^2/4+P} \right) & c^2 < 4 \\ \frac{1}{4} \log(1 + P) - \frac{1}{2} \log(c^2) \end{cases}$$

and lower bounded as

$$C \geq R^{\text{IN}} = \begin{cases} \frac{1}{2} \log \left( \frac{1+c^2/2+P}{c^2} \right) & c^2/2 \leq 1 \\ \frac{1}{2} \log \left( \frac{P+c^2/2+1}{c^2} \right) + \frac{1}{2} \log \left( \frac{P}{2} \right) & 1 \leq c^2/2 < P + 1 \\ \frac{1}{4} \log(P+1) & c^2/2 \geq P + 1 \end{cases}$$  \hspace{1cm} (4)

A powerful bounding techniques is introduced in [3] to derive the outer bound in (4), although this bound is not tight in general. The inner bound in (5) is obtained by having the input be the superposition of two codewords, one treating the state as noise and one pre-coded against a linear combination of the channel states. The outer bounding technique developed in [3] for the case of $M = 2$ is also extended to the case of a general $M$.

**Theorem II.2. Outer bound for the $M$-user CCDP-IS channel [3, Eq. (31)].** Consider the $M$-user CCDP channel
in (1) with \( \text{Var}[S_m] = 1 \) and \( \mathbb{E}[S_m S_q] = 0, \ \forall \ m, q \neq m, q \), then capacity is upper bounded as

\[
C \leq R_{\text{OUT}} = \frac{1}{2} \log \left( P + c^2 + 2c\sqrt{P} \right) - \frac{M - 1}{2M} \log c^2 - \frac{1}{2M} \log M - \left[ \frac{1}{2M} \log \left( \frac{c^2}{M(P + 1)} \right) \right]^+.
\]

(6)

The inner bound in (5) can be generalized in to the case of any number of users \( M \) in a similar manner as in Th. II.2. Inner and outer bound for the case \( M = 2 \) are close for small values of \( P \) but otherwise no capacity characterization is possible using the bounds in Th. II.1. Additionally, the gap between inner and outer bound increases with the number of users, \( M \).

III. THE 2-USER CCDP-IS CHANNEL

We begin by deriving the approximate capacity for 2-user CCDP-IS with the aim of illustrating the main inner and outer bounding techniques. For the achievability part, we simplify the derivation of the inner bound in (5) in Th. II.1 by employing superposition coding and binning (while (5) is obtained by using lattice codes and linear combinations of the two channel states). More specifically, we consider the achievable scheme in Fig. 2: here the channel input is obtained as the superposition of two codewords: the bottom codeword, \( X_{\text{SAN}}^N \) (SAN for “State As Noise”) with power \( \alpha P \), carries the message \( W_{\text{SAN}} \) with rate \( R_{\text{SAN}} \) and treats the state sequences \( S_1^N \) and \( S_2^N \) as additional noise. Since the variance of \( S_1 \) is equal to that of \( S_2 \), the codeword \( X_{\text{SAN}}^N \) can be decoded at both receivers simultaneously. On top of the codeword \( X_{\text{SAN}}^N \) we superimpose two private codewords, \( X_{\text{PAS}}^{N-1} \) and \( X_{\text{PAS}}^{N-2} \) (PAS for “Pre-coded Against State”), with power \( \alpha^P \) for \( \alpha = 1 - \alpha \), each transmitted for half of the channel uses. The codeword \( X_{\text{PAS}}^{N-1} \) is pre-coded against the state sequence \( S_1^N \) as in the classical WDP channel and is decoded only at receiver 1. Similarly \( X_{\text{PAS}}^{N-2} \) is pre-coded against \( S_2^N \) and decoded only at receiver 2. Since each codeword is transmitted for half of the time, they can each have average power \( \alpha^P \) so that the overall power constraint on the channel input is met with equality. Both codewords carry the message \( W_{\text{PAS}} \) at rate \( R_{\text{PAS}} \), so that both receivers are able to decode both \( W_{\text{SAN}} \) and \( W_{\text{PAS}} \).

This strategy, accordingly, attains the rate

\[
R_{\text{IN}} = \frac{1}{2} \log \left( 1 + \frac{\alpha P}{c^2 + \alpha P} \right) + \frac{1}{4} \log (1 + \alpha P). \tag{7}
\]

The expression in (7) can be maximized over \( \alpha \), the ratio between the power of the bottom and top codewords. When \( P + 1 \geq c^2 \), the optimal value of \( \alpha \) is \((c^2 - 1)/P\), which corresponds to fixing the power of the top codewords to the same power as the state sequence. On the other hand, when \( c^2 > P + 1 \), the optimal strategy is for all the power to be allocated to the top codewords and the scheme reduces to pre-coding for receiver 1 half of the time and pre-coding for receiver 2 the remaining portion of the time.

For the outer bound, we are able to improve on the result of Th. II.1 using the observation in Lem. I.2. Note that the outer bound expression in (4) for \c^2 > 4\ is not monotonically increasing in \c \: for this reason it is possible to improve the outer bound by considering a channel with a parameter \( c' = \min\{\sqrt{P + 1}, c\} \leq c \). Such a channel has a larger capacity than the original channel but provides a tighter outer bound to the capacity of the original channel. By comparing these two expressions, we can show that the gap between inner and outer bound is at most 1 bit-per-channel-use (bpcu).

Theorem III.1. Approximate capacity for the 2-user CCDP-IS channel. Consider the 2-user CCDP-IS channel in Fig. 1, then an outer bound to capacity is

\[
C \leq R_{\text{OUT}} = \begin{cases} 
\frac{1}{2} \log(P + 1) & c^2 \leq 1 \\
\frac{1}{2} \log(P + c^2 + 1) & 1 < c^2 < P + 1 \\
\frac{1}{2} \log(P + 1) + 1 & c^2 \geq P + 1 
\end{cases}
\]

(8)

and the exact capacity \( C \) is to within a gap of 1 bpcu from the outer bound in (8).

The result in Th. III.1 is somewhat expected: when the states in a CCDP channel are independent, the best strategy is to send a common codeword at a power level higher than the state variance and a private, pre-coded codeword at the same power level of the state. In order for the two private codewords to communicate the same message at the two receiver, the two codewords must have equal time-sharing. The major difficulty in proving Th. III.1 is therefore in devising an outer bound which matches this very natural transmission strategy.

IV. THE M-USER CCDP-IS CHANNEL

In this section we extend the result in Sec. III from the case of \( M = 2 \) to the case of any \( M \). A generalization of the inner bound in Fig. 2 to the case of any \( M \) is rather straightforward: as shown in Fig. 3, we can employ \( M \) time-shared codewords \( X_{\text{PAS}}^{m-1}, m \in \{1, \ldots, M\} \), each having average power \( \alpha^P \) and each pre-coded against the state sequence \( S_m^N \) for a portion
1/M of the time. All the codewords $X_{SAN}^N$ convey the same message $W_{PAS}$ and receiver $m$ decodes both the codeword $X_{SAN}^N$ and $X_{PAS}^N$ so that, at the end of transmissions, all the decoders can correctly decode both $W_{SAN}$ and $W_{PAS}$.

The rate that we can attain with this strategy is

$$R^N = \frac{1}{2} \log \left( 1 + \frac{\alpha P}{c^2 + \alpha P + 1} \right) + \frac{1}{2M} \log (1 + \alpha P),$$

which can again be maximized over the power-sharing parameter $\alpha$. In this case the optimal value of $\alpha$ is

$$\alpha^* = \max \left\{ 0, \min \left\{ 1, \frac{c^2 + 1 - M}{P(M - 1)} \right\} \right\},$$

and the above scheme reduces to simple time-sharing and DPC when $c^2 > (M - 1)(P + 1)$.

The generalization of the outer bound in Th. III.1 can be obtained by establishing a recursive bound and using a very carefully chosen genie side information for each decoder. Again, the observation in Lem. I.2 is employed to tighten the outer bound expression by optimizing over the state gain $c$. This derivation is non trivial and we refer the interested reader to [9] for the complete proof.

**Theorem IV.1. Approximate capacity for the $M$-user CCDP-IS channel.** Consider the $M$-user CCDP-IS channel in Fig. 1, then an outer bound to capacity is

$$C \leq R^\text{OUT} = \left\{ \begin{array}{ll}
\frac{1}{2} \log \left( 1 + \frac{P}{c + \alpha} \right) + \frac{1}{2M} \log (1 + P) & M - 1 \geq c^2 \\
\frac{1}{2M} \log(1 + P) + \frac{M - 1}{c^2} \log \left( \frac{\alpha P}{c^2 + \alpha P + 1} \right) & M - 1 < c^2 \leq (M - 1)(P + 1) \\
\frac{1}{2M} \log(1 + P) + \frac{c^2}{2M} \log \left( \frac{\alpha P}{c^2 + \alpha P + 1} \right) & c^2 > (M - 1)(P + 1) \end{array} \right.$$  \hspace{1cm} (11)

and the exact capacity $C$ is to within 2.25 bpcu from the outer bound in (11).

**V. THE $M$-USER CCDP-ES CHANNEL**

In this section we consider the CCDP-ES channel which encompasses the models in the previous sections as special cases: we again begin by considering the case $M = 2$ and, successively, generalize our results to the case of any $M$.

In 2-user CCDP-ES, the channel output is obtained as

$$Y_1^N = X^N + cS_1^N + Z_1^N \quad (12a)$$
$$Y_2^N = X^N + cS_2^N + Z_2^N \quad (12b)$$

for $S_1^N$ and $S_2^N$ iid jointly Gaussian iid sequences with zero mean and covariance matrix

$$\Sigma_S = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$  \hspace{1cm} (13)

Note that the channel output (12) can be equivalently rewritten as

$$Y_1^N = X + c \left( aS_1^N + \sqrt{1 - a^2} Z_1^N \right) + Z_1^N \quad (14a)$$
$$Y_2^N = X + c \left( \frac{\rho}{a} S_1^N + \sqrt{1 - \frac{\rho^2}{a^2}} Z_2^N \right) + Z_2^N \quad (14b)$$

for $Z_1, Z_2 \sim N(0, 1)$ and any $a \in [\rho, 1]$.

The expression in (14) shows that, when $\rho$ is positive and by letting $a = \sqrt{\rho}$, we can treat the CCDP-ES channel as a combination of CCDP-IS channel and WDP channel: $S_1^N$ is a state sequence common to the two receivers as in a WDP channel while $S_1^N$ and $S_2^N$ are two independent state sequences as in the CCDP-IS channel. Note that when $\rho$ is negative this decomposition is no longer possible.

**Theorem V.1. Approximate capacity for the 2-user CCDP-ES channel.** Consider the general 2-user CCDP-ES channel in Fig. 1 in which the states have zero mean and state covariance matrix as in (13), then capacity can be upper bounded as

$$C \leq R^\text{OUT} = \left\{ \begin{array}{ll}
\frac{1}{2} \log (P + 1) & c^2 \rho \leq 1 \\
\frac{1}{2} \log (P + c^2 + 1) & 1 < c^2 \rho < P + 1 \\
-\frac{1}{4} \log(c^2) + \frac{P}{2} & c^2 \rho \geq P + 1 \end{array} \right.$$  \hspace{1cm} (16)

for $\rho = 1 - \max \{0, \rho\}$ and the exact capacity $C$ is to within 1 bpcu from the outer bound in (16).

The result in Th. V.1 is shown by exploiting the decomposition in (14) for $\rho$ positive, both in the inner bound and outer bound derivation. The achievability in Th. V.1 follows the achievability in Th. III.1 by additionally pre-coding the codeword $X_{SAN}^N$ against the common state sequence $S_1^N$. The converse is similarly obtained from the converse of Th. V.1 by additionally providing the common state sequence $S_1^N$ as a genie-aided side information to all the receivers. Note that, the result in Th. V.1 coincides with the results in Th. III.1 when $\rho$ is negative. This shows that the capacity of the channel with negative correlation $\rho$ is substantially the same as the capacity of the channel with independent channel states.

The result in Th. V.1 can be easily extended to the case of any number of users $M$.

**Theorem V.2. Approximate capacity for the $M$-user...**
CCDP-ES channel. Consider the $M$-user CCDP channel in Fig. 1 with $\forall \sigma[S_m] = 1$ and $E[S_m S_q] = \rho$, $\forall m, q \in [1 \ldots M], m \neq q$, then capacity can be upper bounded as

$$C \leq R^{\text{OUT}} =$$

$$\begin{cases}
\frac{1}{M} \log \left( 1 + \frac{P}{1+c^2 \rho} \right) + \frac{2}{3} (M - 1) \geq c^2 \rho \\
\frac{1}{2M} \log \left( 1 + P \right) + \frac{N}{2} \log \left( c^2 \right) + \frac{2}{3} (M - 1) \leq \left( M - 1 \right) (P + 1)
\end{cases}$$

$$\frac{1}{2M} \log \left( 1 + P \right) + \frac{N}{2} \log \left( c^2 \right) + \frac{2}{3} (M - 1) > \left( M - 1 \right) (P + 1)$$

(18)

for $\rho = 1 - \max \left\{ 0, \rho \right\}$ and the exact capacity $C$ is to within 2.25 bpcu from the outer bound in (18).

As in (14) and for $\rho > 0$, each channel output can be rewritten as

$$Y^N_m = X^N + c(\bar{S}^N_c + \bar{S}^N_m) + Z^N_i, \quad m \in [1 \ldots M].$$

(19)

for $S_i, \bar{S}_m \sim N(0, 1), \quad m \in [1 \ldots M]$. The capacity result in Th. V.2 is obtained by adapting the derivation in Th. IV.1 as follows: (i) for the achievability part, the common codeword in the achievable strategy in Fig. 3 is pre-coded against the common component of the state sequence $\bar{S}^N_c$. (ii) in the converse, $\bar{S}^N_i$ is provided as genie-aided side information to all the receivers.

When $\rho < 0$, the channel output can be equivalently expressed as

$$Y^N_m = X^N + Z^N_m + c \left( -\sqrt{|\rho|} \sum_{j=1}^{m-1} \bar{S}^N_{jm} \right)$$

$$+ \sqrt{|\rho|} \sum_{j=m+1}^{M} \bar{S}^N_{mj} + \sqrt{1 - (M - 1)|\rho|} \bar{S}^N_{mm} \right),$$

(20)

for $\bar{S}_{i,q} \sim N(0, 1), \quad i, q \in [1 \ldots M]^2$. Note that each term $\bar{S}_{mj}$ appears with negative sign in the expression of $Y_m$, and with a negative sign in the expression of $Y_j$, thus yielding the negative correlation among each two state terms $S_m$ and $S_j$. The expression in (20) intuitively shows why no common channel state term emerges from negatively correlated channel states. Note that the decomposition in (20) also ostensibly motivates why the minimum negative correlation $\rho$ is $-1/(M - 1)$ as in Lem I.1.

VI. DISCUSSION

In this section we would like to highlight some of the difficulties in determining the capacity for the general CCDP with any number of users and any state covariance matrix. It must be noted that the assumption of equal variance is crucial in obtaining a compact expression for the attainable region in Th. IV.1 and Th. V.2. When the states have the same variance, the attainable region only requires one common codeword: if the state sequences had different variances we could improve the achievable scheme by employing partially common codewords. Consider the case of $M = 3$ with independent state sequences of variance $\sigma^2_1 \leq \sigma^2_2 < \sigma^2_3$ respectively: the achievable scheme in Fig. 3 could be improved upon by additionally employing a partially common codeword which is decoded at receiver 1 and 2 but not at receiver 3. Since this codeword is decoded at two of the three receivers, it is necessary to adjust the time-sharing parameters and the power allocated to each private codeword to guarantee that all the decoders are able to decode the same set of sub-messages. Obtaining an explicit expressions of the attainable rate is therefore challenging, as it is challenging to determine a matching outer bound expression.

VII. CONCLUSION

In this paper we study the capacity of the “carbon copying onto dirty paper”; this model is a variation of the classic “writing on dirty paper” channel in which a common message is to be decoded at $M$ receivers, each observing a linear combination of the input, Gaussian noise and one of $M$ possible state sequences. The state sequences are non-causally known at the transmitter but are unknown at the receivers. We focus, in particular, on the case in which the state sequences are iid jointly Gaussian distributed with the same variance and the same pairwise correlation. For this model, we derive the capacity of to within 2.25 bits-per-channel-use for all channel parameters. For this model, capacity can be approached with a rather simple strategy in which the input is composed of the superposition of two codewords: (i) a bottom codeword which is pre-coded against the common component of the channel state and (ii) a top codeword which is pre-coded against the state observed at each receiver. While the top codeword is decoded by all receiver simultaneously, the top codeword is decoded only at one receiver at the time and for a portion $1/M$ of the time. During each of these phases, the top codeword is pre-codes against the state experienced at the intended receiver as in the “writing on dirty paper” channel.

REFERENCES

[1] S. Gel’fand and M. Pinsker, “Coding for channel with random parameters,” Problemy Peredachi Informatsii, vol. 6, no. 1, pp. 3–26, 1970.

[2] M. Costa, “Writing on dirty paper.” IEEE Trans. Inf. Theory, vol. 29, no. 3, pp. 439–441, 1983.

[3] A. Khisti, U. Erez, A. Lapidoth, and G. Wornell, “Carbon copying onto dirty paper,” IEEE Trans. Inf. Theory, vol. 53, no. 5, pp. 1814–1827, May 2007.

[4] P. Piantanida and S. Shamai (Shitz), “On the capacity of compound state-dependent channels with states known at the transmitter,” in Int. Symp. Inf. Theory, (ISIT). IEEE, 2010, pp. 624–628.

[5] S. Rini and S. Shamai, “On the approximate capacity of the carbon copy coding channel,” in Int. Symp. Inf. Theory, (ISIT). IEEE, 2014, pp. 2287–2291.

[6] Y. Steinberg and S. Shamai, “Achievable rates for the broadcast channel with states known at the transmitter,” in Int. Symp. Inf. Theory, (ISIT), 2005.

[7] Y. Steinberg, “Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information,” IEEE Trans. Inf. Theory, vol. 51, no. 8, pp. 2867–2877, 2005.

[8] T. Cover and J. Thomas, Elements of Information Theory. Wiley-Interscience, New York, 1991.

[9] S. Rini and S. Shamai, “On the approximate capacity of the carbon copy onto dirty paper channel,” arXiv preprint, arXiv:1602.02206, 2016.