Article

New Result on the Feedback Capacity of the Action-Dependent Dirty Paper Wiretap Channel

Guangfen Xie ^{1,2,†} and Bin Dai ^{1,2,*,†}
¹ School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China; gfxie@my.swjtu.edu.cn

² Peng Cheng Laboratory, Shenzhen 518055, China

* Correspondence: daibin@home.swjtu.edu.cn; Tel.: +86-135-480-53724

† These authors contributed equally to this work.

Abstract: The Gaussian wiretap channel with noncausal state interference available at the transmitter, which is also called the dirty paper wiretap channel (DP-WTC), has been extensively studied in the literature. Recently, it has been shown that taking actions on the corrupted state interference of the DP-WTC (also called the action-dependent DP-WTC) helps to increase the secrecy capacity of the DP-WTC. Subsequently, it has been shown that channel feedback further increases the secrecy capacity of the action-dependent DP-WTC (AD-DP-WTC), and a sub-optimal feedback scheme is proposed for this feedback model. In this paper, a two-step hybrid scheme and a corresponding new lower bound on the secrecy capacity of the AD-DP-WTC with noiseless feedback are proposed. The proposed new lower bound is shown to be optimal (achieving the secrecy capacity) and tighter than the existing one in the literature for some cases, and the results of this paper are further explained via numerical examples.

Keywords: action encoder; channel feedback; dirty paper channel; intelligent reflecting surfaces; wiretap channel

Citation: Xie, G.; Dai, B. New Result on the Feedback Capacity of the Action-Dependent Dirty Paper Wiretap Channel. *Entropy* **2021**, *23*, 929. <https://doi.org/10.3390/e23080929>

Academic Editor: Song-Nam Hong

Received: 30 May 2021

Accepted: 20 July 2021

Published: 21 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Dirty paper coding is one of the most important pre-coding schemes in wireless communications and has a wide range of applications in information hiding [1]. Dirty paper coding was first investigated by Costa in his well-known paper on the dirty paper channel (DPC) [2], where a corrupted Gaussian state interference of a white Gaussian channel is noncausally known at the transmitter and not available at the receiver. Costa showed that the capacity of the DPC equals the capacity of the same model without state interference, which indicates that though the receiver does not know the state interference, it can still be perfectly removed by using dirty paper coding.

Note that in [2], the channel state is assumed to be generated by nature. However, in some practical scenarios, the state is affected or controlled by the communication systems, e.g., in intelligent reflecting surface-aided communication systems, the state is formed in part by the reflecting phase shift which is actually controlled by the transceiver [3]. Such a case was first investigated by Weissman in his paper on the action-dependent dirty paper channel (AD-DPC) [4], where the corrupted state interference of the DPC is affected and noncausally known by the transmitter. In [4], a lower bound on the capacity of AD-DPC was proposed, and the capacity was fully determined in [5].

In recent years, the study of the above dirty paper channels under additional secrecy constraints has received a lot of attention. Specifically, [6] studied the discrete memoryless wiretap channel with state interference noncausally known by the transmitter, and obtained upper and lower bounds on the secrecy capacity. The authors of [7] extended the model studied in [6] to the broadcast situation, and also provided bounds on the secrecy capacity region of this extended model. The authors of [8,9] studied the Gaussian case of [6],

namely, the dirty paper wiretap channel, and proposed bounds on the secrecy capacity. Furthermore, [8,9] pointed out that the proposed secret dirty paper coding increases the secrecy capacity of the Gaussian wiretap channel [10]. The above works mainly adopted the tools in [11,12] for establishing the secrecy rate/capacity. Very recently, the authors of [13] studied the AD-DPC with an additional eavesdropper, which is also called the action-dependent dirty paper wiretap channel (AD-DP-WTC), and proposed lower and upper bounds on the secrecy capacity. Note that the capacity results given in [9,13] indicate that:

- There is a penalty term between secrecy capacity and capacity of the same model without secrecy constraints.
- If the eavesdropper's channel is less noisy than the legitimate receiver's, the secrecy capacity may equal zero, i.e., no positive rate can be guaranteed for secure communications.

Very recently, it has been shown that channel feedback is an effective way to enhance the secrecy capacities of the DP-WTC [14,15] and the multi-input multi-output (MIMO) X-channels [16]. In [16], it has been shown that feedback increases the secure degrees of freedom (SDoF) region of the MIMO X-channel with secrecy constraints, which indicates that feedback also enlarges the secrecy capacity of the same model, even if the eavesdropper's channel is less noisy than the legitimate receiver's. The authors of [14–16] mainly adopted the idea of a generating secret key from the channel feedback and using this key to encrypt the transmitted message. Subsequently, [13] showed that a variation of the classical Schalkwijk–Kailath (SK) feedback scheme [17] for the point-to-point white Gaussian channel achieves the secrecy capacity of the DP-WTC with noiseless feedback, and the secrecy capacity equals the capacity of the same model without secrecy constraints, i.e., to achieve perfect secrecy, no rate needs to be sacrificed even if the eavesdropper gains advantage over the legitimate receiver. In addition, the authors of [13] also proposed a sub-optimal SK type feedback scheme for the AD-DP-WTC with noiseless feedback (see Figure 1), i.e., this scheme achieves a lower bound on the secrecy capacity of the AD-DP-WTC with noiseless feedback, and the secrecy capacity remains open.

In this paper, we derive a new lower bound on the secrecy capacity of the AD-DP-WTC with noiseless feedback, which is based on a hybrid two-step feedback scheme. The proposed new lower bound is shown to be optimal and tighter than the existing one in the literature for some cases, and the results of this paper are further explained via numerical examples. The remainder of this paper is organized as follows. A formal definition of the model studied in this paper and previous results are introduced in Section 2. The main result and the corresponding proof are given in Sections 3 and 4, respectively. Section 5 includes the summary of all results in this paper and discusses future work.

Figure 1. The AD-DP-WTC with noiseless feedback.

2. Model Formulation and Previous Results

2.1. Model Formulation

For the AD-DP-WTC with noiseless feedback (see Figure 1), the i -th ($i \in \{1, 2, \dots, N\}$) channel inputs and outputs are given by

$$S_i = A_i + W_i, \quad Y_i = X_i + S_i + \eta_{1,i}, \quad Z_i = Y_i + \eta_{2,i} \quad (1)$$

where X_i is the output of the channel encoder subject to an average power constraint P ($\frac{1}{N} \sum_{i=1}^N E[X_i^2] \leq P$), A_i is the output of the action encoder subject to an average power constraint P_A ($\frac{1}{N} \sum_{i=1}^N E[A_i^2] \leq P_A$), Y_i and Z_i are channel outputs, respectively, at the legitimate receiver and the eavesdropper, and $W_i, \eta_{1,i}, \eta_{2,i}$ are independent Gaussian noises and are independently identically distributed (i.i.d.) across the time index i . Note that $W_i \sim \mathcal{N}(0, \sigma_w^2)$, $\eta_{1,i} \sim \mathcal{N}(0, \sigma_1^2)$, and $\eta_{2,i} \sim \mathcal{N}(0, \sigma_2^2)$. The transmitted message M is uniformly drawn in its alphabet $\mathcal{M} = \{1, 2, \dots, |\mathcal{M}|\}$.

At time instant i ($i \in \{1, 2, \dots, N\}$), a corrupted state interference S_i is generated through a white Gaussian channel with i.i.d. noise $W_i \sim \mathcal{N}(0, \sigma_w^2)$ and channel input A_i , where A_i is a (stochastic) function of the message M . Since the corrupted state interference $S^N = (S_1, \dots, S_N)$ is noncausally known by the channel encoder, the i -th channel input X_i is a (stochastic) function of the message M , S^N , and the feedback $Y^{i-1} = (Y_1, Y_2, \dots, Y_{i-1})$.

The legitimate receiver generates an estimation $\hat{M} = \psi(Y^N)$, where ψ is the legitimate receiver's decoding function, and the average decoding error probability equals

$$P_e = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}_1} \Pr\{\psi(Y^N) \neq m | m \text{ sent}\}. \quad (2)$$

The eavesdropper's equivocation rate of the message M is denoted by

$$\Delta = \frac{1}{N} H(M | Z^N). \quad (3)$$

A rate R is said to be achievable with perfect weak secrecy if for any ϵ and sufficiently large N , there exists a channel encoder–decoder such that

$$\frac{\log |\mathcal{M}|}{N} = R, \quad \Delta \geq R - \epsilon, \quad P_e \leq \epsilon. \quad (4)$$

The *secrecy capacity* of the AD-DP-WTC with feedback, which is the maximum achievable secrecy rate defined in (4), is denoted by $\mathcal{C}_{sag}^f$. A new lower bound, R_{sag}^f on $\mathcal{C}_{sag}^f$, will be given in the next section.

2.2. SK Feedback Scheme for the Point-to-Point White Gaussian Channel

For the point-to-point white Gaussian channel with feedback, at time i ($i \in \{1, 2, \dots, N\}$), channel input and output are given by

$$Y_i = X_i + \eta_i, \quad (5)$$

where X_i is the channel input subject to an average power constraint P ($\frac{1}{N} \sum_{i=1}^N E[X_i^2] \leq P$), and $\eta_i \sim \mathcal{N}(0, \sigma^2)$ is the i.i.d. channel noise. The channel input X_i is a function of the message M and the feedback Y^{i-1} . It is well known that the capacity $\mathcal{C}_g^f$ of the white Gaussian channel with feedback equals the capacity $\mathcal{C}_g$ of the same model without feedback, i.e.,

$$\mathcal{C}_g^f = \mathcal{C}_g = \frac{1}{2} \log\left(1 + \frac{P}{\sigma^2}\right). \quad (6)$$

The authors of [17] showed that the classical SK scheme achieves $\mathcal{C}_g^f$, and this scheme is described below.

Since M takes the values in $\mathcal{M} = \{1, 2, \dots, 2^{NR}\}$, we divide the interval $[-0.5, 0.5]$ into 2^{NR} equally spaced sub-intervals, and the center of each sub-interval is mapped to a message value in $\mathcal{M}$. Let θ be the center of the sub-interval with respect to (w.r.t.) the message M .

At time 1, the transmitter transmits

$$X_1 = \sqrt{12P}\theta. \quad (7)$$

The receiver receives $Y_1 = X_1 + \eta_1$, and obtains an estimation of θ by computing

$$\hat{\theta}_1 = \frac{Y_1}{\sqrt{12P}} = \theta + \frac{\eta_1}{\sqrt{12P}} = \theta + \epsilon_1, \quad (8)$$

where $\epsilon_1 = \hat{\theta}_1 - \theta = \frac{\eta_1}{\sqrt{12P}}$, and $\alpha_1 \triangleq \frac{\sigma^2}{12P}$.

At time k ($2 \leq k \leq N$), the receiver obtains $Y_k = X_k + \eta_k$, and obtains an estimation of θ_k by computing

$$\hat{\theta}_k = \hat{\theta}_{k-1} - \frac{E[Y_k \epsilon_{k-1}]}{E[Y_k^2]} Y_k, \quad (9)$$

where

$$\epsilon_k = \hat{\theta}_k - \theta, \quad (10)$$

(10) yields that

$$\epsilon_k = \epsilon_{k-1} - \frac{E[Y_k \epsilon_{k-1}]}{E[Y_k^2]} Y_k. \quad (11)$$

At time k ($k \in \{2, 3, \dots, N\}$), the transmitter sends

$$X_k = \sqrt{\frac{P}{\alpha_{k-1}}} \epsilon_{k-1}, \quad (12)$$

where $\alpha_{k-1} \triangleq \text{Var}(\epsilon_{k-1})$.

In [17], it has been shown that the decoding error P_e of the above coding scheme is upper bounded by

$$P_e \leq \Pr\{|\epsilon_N| > \frac{1}{2|\mathcal{M}| - 1}\} \leq 2Q\left(\frac{1}{2 \cdot 2^{NR}} \frac{1}{\sqrt{\alpha_N}}\right), \quad (13)$$

where $Q(x)$ is the tail of the unit Gaussian distribution evaluated at x , and

$$\alpha_N = \frac{\sigma^2}{12P} \left(\frac{\sigma^2}{P + \sigma^2}\right)^{N-1}. \quad (14)$$

From (13) and (14), we conclude that if

$$R < \frac{1}{2} \log\left(1 + \frac{P}{\sigma^2}\right) = \mathcal{C}_g^f, \quad (15)$$

$P_e \rightarrow 0$ as $N \rightarrow \infty$. Recently, [18] showed that the above classical SK scheme, which is not designed with the consideration of secrecy, already achieves the secrecy capacity of the Gaussian wiretap channel with noiseless feedback, i.e., the corresponding secrecy capacity equals the capacity $\mathcal{C}_g^f$ of the same model without secrecy constraints.

2.3. Previous Results on the AD-DP-WTC with Feedback

For the capacity $\mathcal{C}_{sag}^f$ of the AD-DP-WTC with feedback, [13] showed that it is bounded by

$$L \leq \mathcal{C}_{sag}^f \leq \mathcal{C}_{ag}, \quad (16)$$

where

$$\begin{aligned} \mathcal{C}_{ag} = & \max_{(\rho_1, \rho_2): \rho_1^2 + \rho_2^2 \leq 1} \frac{1}{2} \log \left(1 + \frac{P(1 - \rho_1^2 - \rho_2^2)}{\sigma_1^2} \right) \\ & + \frac{1}{2} \log \left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2} \right), \end{aligned} \quad (17)$$

$$L = \max_{(\rho_1, \rho_2): \rho_1^2 + \rho_2^2 \leq 1} \frac{1}{2} \log \left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2} \right). \quad (18)$$

Note that $\mathcal{C}_{ag}$ is the capacity of the AD-DPC without feedback, and is given in [5].

Proof Sketch of (16). Since the capacity $\mathcal{C}_{sag}^f$ of the AD-DP-WTC with feedback is no larger than that of the same model without secrecy constraints, and feedback does not increase the capacity of the AD-DPC [4], and $\mathcal{C}_{ag}$ serves as a trivial upper bound on $\mathcal{C}_{sag}^f$. The lower bound $\mathcal{C}_{sag}^f \geq L$ is derived by constructing

$$X = \frac{\rho_2^* \sqrt{P}}{\sqrt{P_A}} A + \frac{\rho_1^* \sqrt{P}}{\sigma_w} W, \quad (19)$$

for $\rho_1^{*2} + \rho_2^{*2} = 1$, and

$$X = \frac{\rho_2^* \sqrt{P}}{\sqrt{P_A}} A + \frac{\rho_1^* \sqrt{P}}{\sigma_w} W + G, \quad (20)$$

for $\rho_1^{*2} + \rho_2^{*2} < 1$, where G is randomly generated according to $G \sim \mathcal{N}(0, P(1 - \rho_1^{*2} - \rho_2^{*2}))$ and it is independent of A and W . Substituting (19) and (20) into (1), the AD-DP-WTC with feedback is equivalent to the Gaussian wiretap channel with feedback with input A , channel noise $W + G + \eta_1 + \eta_2$, channel output Y of the legitimate receiver, and the channel output Z of the eavesdropper. Directly applying the SK scheme introduced in the preceding subsection, we conclude that L is achievable. Moreover, from [18], we know that the SK scheme also achieves the secrecy capacity of the Gaussian wiretap channel with noiseless feedback, which indicates that L is achievable with perfect weak secrecy, and the proof is completed. $\square$

Note that the above lower and upper bounds on $\mathcal{C}_{sag}^f$ do not meet in general, and exploring a tighter lower bound on $\mathcal{C}_{sag}^f$ is the motivation of this paper.

3. Main Results of this Paper

3.1. A New Lower Bound on the Secrecy Capacity of the AD-DP-WTC with Feedback

Theorem 1. A new lower bound on the secrecy capacity C_{sag}^f of the AD-DP-WTC with feedback is given by

$$R_{sag}^f = \max_{(\rho_1, \rho_2): \rho_1^2 + \rho_2^2 \leq 1} \left\{ \frac{1}{2} \log\left(1 + \frac{(1 - \rho_1^2 - \rho_2^2)P}{\sigma_1^2}\right) + \frac{1}{2} \log\left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2}\right) - \frac{1}{2} \log\left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2 + \sigma_2^2}\right) \right\}. \quad (21)$$

Remark 1. Comparing the new lower bound in (21) with the upper bound C_{ag} in (16), it is easy to see that there still exists a gap between the two bounds due to the penalty term $\frac{1}{2} \log\left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2 + \sigma_2^2}\right)$.

The following Corollary 1 shows that the proposed new lower bound in (21) is optimal for a special case. Moreover, the following Corollary 2 shows that the new lower bound in (21) is tighter than the existing lower bound in (16) when σ_w^2 tends to infinity.

Corollary 1. For $\sigma_2^2 \rightarrow \infty$,

$$\lim_{\sigma_2^2 \rightarrow \infty} R_{sag}^f = C_{ag}, \quad (22)$$

which indicates that the secrecy capacity is determined for this case.

Proof. Formula (22) is directly obtained by using (21) and letting $\sigma_2^2 \rightarrow \infty$. Hence the proof is completed. $\square$

Corollary 2.

$$\lim_{\sigma_w^2 \rightarrow \infty} R_{sag}^f \geq \lim_{\sigma_w^2 \rightarrow \infty} L, \quad (23)$$

where L is the existing lower bound defined in (18).

Proof. Define the gap R_G between the new lower bound in (21) and the existing lower bound in (18) by

$$R_G = R_{sag}^f - L = \max_{(\rho_1, \rho_2): \rho_1^2 + \rho_2^2 \leq 1} \left\{ \frac{1}{2} \log\left(1 + \frac{(1 - \rho_1^2 - \rho_2^2)P}{\sigma_1^2}\right) - \frac{1}{2} \log\left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2 + \sigma_2^2}\right) \right\}. \quad (24)$$

- For the case that the maximum is achieved when $\rho_1^2 + \rho_2^2 = 1$,

$$\lim_{\sigma_w^2 \rightarrow \infty} R_G = 0. \quad (25)$$

- For the case that the maximum is achieved when $\rho_1^2 + \rho_2^2 < 1$,

$$\lim_{\sigma_w^2 \rightarrow \infty} R_G > 0. \quad (26)$$

Combining (25) and (26), we conclude that $\lim_{\sigma_w^2 \rightarrow \infty} R_G \geq 0$, which indicates that $\lim_{\sigma_w^2 \rightarrow \infty} R_{\text{Sag}}^f \geq \lim_{\sigma_w^2 \rightarrow \infty} L$. The proof of Corollary 2 is completed. $\square$

Proof sketch of Theorem 1. The main idea of the achievable scheme is briefly illustrated by Figures 2 and 3. In Figure 2, we split the message M into two parts $M = (M_1, M_2)$, where the sub-message M_j ($j = 1, 2$) is uniformly distributed in $\mathcal{M}_j = \{1, \dots, 2^{NR_j}\}$. The sub-message M_1 is available at both the action encoder and channel encoder, and the sub-message M_2 is only available at the channel encoder. Let $X_i = U_i + V_i + G_i$, where M_1 is encoded as U_i with power $\rho_2^2 P$, M_2 is encoded as V_i with power $P' = (1 - \rho_1^2 - \rho_2^2)P$, $G_i = \sqrt{\frac{\rho_1^2 P}{\sigma_w^2}} W_i$, and $-1 \leq \rho_j \leq 1$ for $j = 1, 2$. Moreover, M_1 is also encoded as A_i with power P_A , and

$$U_i = \sqrt{\frac{\rho_2^2 P}{P_A}} A_i, \quad (27)$$

which indicates that U_i is a deterministic function of A_i . $\square$

Figure 2. The new coding scheme of AD-DP-WTC with noiseless feedback.

Figure 3. Equivalent coding scheme. (a) The equivalent channel model of the sub-message M_1 . (b) The equivalent channel model of the sub-message M_2 .

The i -th ($i \in \{1, 2, \dots, N\}$) channel inputs and outputs are rewritten as

$$\begin{aligned} Y_i &= X_i + S_i + \eta_{1,i} = U_i + V_i + G_i + A_i + W_i + \eta_{1,i} = A_i^* + V_i + W'_i + \eta_{1,i} \\ Z_i &= Y_i + \eta_{2,i}, \end{aligned} \quad (28)$$

where $W'_i = G_i + W_i$ is an i.i.d Gaussian noise process with zero mean and variance $\sigma_w'^2 = (\sigma_w + \sqrt{\rho_1^2 P})^2$, and $A_i^* = A_i + U_i$ is subject to an average power constraint $P^* = P_A + \rho_2^2 P + 2\sqrt{\rho_2^2 P P_A}$.

In Figure 3, since M_1 is known by the channel encoder, the codeword $A_i^* = A_i + U_i$ can be subtracted when applying an SK type feedback scheme to M_2 , i.e., the transmission of M_2 is through an equivalent channel with input V_i , output

$$Y'_i = Y_i - A_i^* = V_i + W'_i + \eta_{1,i}. \quad (29)$$

Moreover, since M_1 and S^N are known by the channel encoder, $W^N = S^N - A^N$ and $G^N = \sqrt{\frac{\rho_1^2 P}{\sigma_w'^2}} W^N$, $W'^N = W^N + G^N$ can be viewed as state interference which is noncausally known at the encoder of the equivalent channel of M_2 (namely, the equivalent encoder of M_2).

In addition, the transmission of M_1 is through an equivalent channel with inputs $A_i^* = A_i + U_i$, output $Y_i = A_i^* + V_i + W'_i + \eta_{1,i}$, and channel noise $\eta'_{1,i} = V_i + W'_i + \eta_{1,i}$, which is nonwhite Gaussian noise since V_i is not i.i.d generated.

Then, applying an SK type scheme to M_2 , and Wyner's random binning scheme [11] together with Feinstein's greedy coding scheme [19] to M_1 , the lower bound in Theorem 1 is obtained. The detail of the proof is given in the next section.

3.2. Numerical Results

Figure 4 plots the upper bound, the existing lower bound and the new lower bound on the secrecy capacity of the AD-DP-WTC with feedback for $P_A = 10$, $\sigma_w^2 = 2000$, $\sigma_1^2 = 30$, $\sigma_2^2 = 100$, and P taking values in $[0, 30]$. From Figure 4, we see that our new scheme performs better than the existing one when the noise variance σ_w^2 is large enough.

Figure 4. Comparison of the bounds on the secrecy capacity of AD-DP-WTC with feedback for $P_A = 10$, $\sigma_w^2 = 2000$, $\sigma_1^2 = 30$, $\sigma_2^2 = 100$, and P taking values in $[0, 30]$.

Figure 5 plots the bounds for $P_A = 5$, $\sigma_w^2 = 240$, $\sigma_1^2 = 30$, $\sigma_2^2 = 2500$, and P taking values in $[0, 30]$. From Figure 5, we see that our new lower bound almost meets the upper bound when the eavesdropper's channel noise variance σ_2^2 is large enough.

Figure 5. Comparison of the bounds on the secrecy capacity of AD-DP-WTC with feedback for $P_A = 5$, $\sigma_w^2 = 240$, $\sigma_1^2 = 30$, $\sigma_2^2 = 2500$, and P taking values in $[0, 30]$.

4. Proof of Theorem 1

The encoding and decoding procedure of Figure 3 is described below. Since M_2 takes values in $\mathcal{M}_2 \in \{1, 2, \dots, 2^{NR_2}\}$, the interval $[-0.5, 0.5]$ is divided into 2^{NR_2} equally spaced sub-intervals, and the center of each sub-interval is mapped to a message value. Let θ be the center of the sub-interval w.r.t. the message M_2 (the variance of θ approximately equals $\frac{1}{12}$). Let $A^{*N} = (A_1^* = 0, A_2^*, \dots, A_N^*) = (0, A_2^{*N})$, where $A_2^{*N} = (A_2^*, \dots, A_N^*)$. Here, A_2^{*N} is the codeword of the sub-message M_1 and a dummy message M_1^* , and it is generated by Feinstein's greedy construction for the non-i.i.d. Gaussian channel [19]. Moreover, M_1 and M_1^* are uniformly distributed in $\{1, \dots, 2^{(N-1)R_1}\}$ and $\{1, \dots, 2^{(N-1)R^*}\}$, respectively.

Encoding procedure: Before the transmission, for a given sub-message M_1 , first, a dummy message M_1^* is randomly chosen from its alphabet $\{1, \dots, 2^{(N-1)R^*}\}$. Then, Feinstein's greedy construction [19] is applied to encode the message pair (M_1, M_1^*) as the codeword $A_2^{*N} = (A_2^*, \dots, A_N^*)$.

At time 1, the equivalent encoders of M_1 and M_2 , respectively, send

$$A_1^* = 0, \quad V_1 = \sqrt{12P'}\left(\theta - \frac{W_1'}{\sqrt{12P'}} + B\right), \quad (30)$$

where

$$B = \sum_{i=2}^N \frac{E(Y_i'' \varepsilon_{i-1})}{E(Y_i'')^2} W_i', \quad P' = (1 - \rho_1^2 - \rho_2^2)P, \quad Y_i'' = V_i + \eta_{1,i} \quad (31)$$

and ε_{i-1} will be defined later.

At time 2, once receiving the feedback $Y_1 = V_1 + W_1' + \eta_{1,1}$, the equivalent encoder of M_2 computes

$$\frac{Y_1}{\sqrt{12P'}} = \theta + B + \frac{\eta_{1,1}}{\sqrt{12P'}} = \theta + B + \varepsilon_1, \quad (32)$$

where $\varepsilon_1 = \frac{\eta_{1,1}}{\sqrt{12P'}}$ and $\alpha_1 = \text{Var}(\varepsilon_1)$. The equivalent encoders of M_1 and M_2 , respectively, send A_2^* (the first component of A_2^{*N}) and

$$V_2 = \sqrt{\frac{P'}{\alpha_1}} \varepsilon_1. \quad (33)$$

At time k ($3 \leq k \leq N$), once receiving the feedback $Y_{k-1} = A_{k-1}^* + V_{k-1} + \eta_{1,k-1} + W_{k-1}'$, the equivalent encoder of M_2 computes

$$\varepsilon_{k-1} = \varepsilon_{k-2} - \frac{E(Y_{k-1}' \varepsilon_{k-2})}{E(Y_{k-1}'')^2} Y_{k-1}'', \quad (34)$$

where $Y_{k-1}'' = Y_{k-1} - A_{k-1}^* - W_{k-1}' = V_{k-1} + \eta_{1,k-1}$ and $\alpha_{k-1} = \text{Var}(\varepsilon_{k-1})$. The equivalent encoders of M_1 and M_2 , respectively, send A_k^* (the $k-1$ -th component of A_2^{*N}) and

$$V_k = \sqrt{\frac{P'}{\alpha_{k-1}}} \varepsilon_{k-1}. \quad (35)$$

Lemma 1. For $3 \leq k \leq N$, the general terms of α_{k-1} and V_k can be expressed as

$$\alpha_{k-1} = \text{Var}(\varepsilon_{k-1}) = \alpha_1 \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k-2}, \quad (36)$$

$$V_k = \frac{1}{\sqrt{12\alpha_{k-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k-2} \eta_{1,1} - \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{k-1}}} \frac{P'(\sigma_1^2)^{k-2-i}}{(P' + \sigma_1^2)^{k-(i+1)}} \eta_{1,i+1}. \quad (37)$$

Proof. The proof of Lemma 1 is in Appendix A. $\square$

Decoding procedure: The legitimate receiver does a two-step decoding scheme. First, by applying Feinstein's decoding rule [19] to the decoding of A_2^{*N} . According to Feinstein's lemma [19], the decoding error probability of M_1 and M_1^* , denoted by P_{e1} , can be arbitrarily small if

$$R_1 + R^* \leq \frac{I(A_2^{*N}; Y_2^N)}{N-1}, \quad (38)$$

where $Y_2^N = (Y_2, Y_3, \dots, Y_N)$.

Second, after decoding M_1 , the legitimate receiver obtains A_k^* for all $2 \leq k \leq N$, and subtracts A_k^* from Y_k , then the legitimate receiver obtains $Y_k' = Y_k - A_k^* = V_k + \eta_{1,k} + W_k'$.

At time 1, the legitimate receiver obtains $Y_1 = V_1 + W_1' + \eta_{1,1}$ and computes

$$\hat{\theta}_1 = \frac{Y_1}{\sqrt{12P'}} = \theta + B + \varepsilon_1. \quad (39)$$

At time k ($2 \leq k \leq N$), the legitimate receiver's estimation $\hat{\theta}_k$ of θ is given by

$$\begin{aligned}\hat{\theta}_k &= \hat{\theta}_{k-1} - \frac{E(Y_k'' \varepsilon_{k-1})}{E(Y_k'')^2} Y_k' \\ &\stackrel{(a)}{=} \hat{\theta}_{k-1} + \varepsilon_k - \varepsilon_{k-1} - \frac{E(Y_k'' \varepsilon_{k-1})}{E(Y_k'')^2} W_k' \\ &= \hat{\theta}_1 + \varepsilon_k - \varepsilon_1 - \sum_{j=2}^k \frac{E(Y_j'' \varepsilon_{j-1})}{E(Y_j'')^2} W_j' \\ &\stackrel{(b)}{=} \theta + \varepsilon_k + B - \sum_{j=2}^k \frac{E(Y_j'' \varepsilon_{j-1})}{E(Y_j'')^2} W_j',\end{aligned}\quad (40)$$

where $Y_k' = V_k + \eta_{1,k} + W_k'$, $Y_k'' = V_k + \eta_{1,k}$, (a) follows from (34), and (b) follows from (39). From (40), we can conclude that for $k = N$,

$$\begin{aligned}\hat{\theta}_N &= \theta + \varepsilon_N + B - \sum_{j=2}^N \frac{E(Y_j'' \varepsilon_{j-1})}{E(Y_j'')^2} W_j' \\ &\stackrel{(c)}{=} \theta + \varepsilon_N,\end{aligned}\quad (41)$$

where (c) follows from (31).

Now we bound the decoding error probability P_{e2} of M_2 as follows. From $\hat{\theta}_N = \theta + \varepsilon_N$ and the definition of θ , we have

$$\begin{aligned}P_{e2} &\leq \Pr[|\varepsilon_N| \geq \frac{1}{2(|\mathcal{M}_2| - 1)}] \\ &\stackrel{(d)}{\leq} 2Q\left(\frac{1}{2 \times 2^{NR_2} \sqrt{\alpha_N}}\right) \\ &\stackrel{(e)}{=} 2Q\left(\frac{1}{2\sqrt{\frac{\sigma_1^2 + P'}{12P'}}} 2^{-NR_2} 2^{\frac{N}{2} \log(1 + \frac{P'}{\sigma_1^2})}\right) \\ &= 2Q\left(\frac{1}{2\sqrt{\frac{\sigma_1^2 + P'}{12P'}}} 2^{N(\frac{1}{2} \log(1 + \frac{P'}{\sigma_1^2}) - R_2)}\right),\end{aligned}\quad (42)$$

where (d) following from $Q(x)$ is the tail of the unit Gaussian distribution evaluated at x , and (e) is from Lemma 1. Since $Q(x)$ is decreasing while x is increasing, from (42), we can conclude that if

$$R_2 \leq \frac{1}{2} \log\left(1 + \frac{P'}{\sigma_1^2}\right) = \frac{1}{2} \log\left(1 + \frac{(1 - \rho_1^2 - \rho_2^2)P}{\sigma_1^2}\right), \quad (43)$$

$P_{e2} \rightarrow 0$ as N is large enough.

Note that the total decoding error probability P_e of $M = (M_1, M_2)$ is upper bounded by $P_e \leq P_{e1} + P_{e2}$. From the above analysis, we conclude that $P_e \rightarrow 0$ as N tends to infinity if (38) and (43) are guaranteed.

Equivocation analysis: We bound the equivocation rate $\frac{H(M|Z^N)}{N}$ as follows:

$$\frac{H(M|Z^N)}{N} = \frac{H(M_1, M_2|Z^N)}{N} = \frac{H(M_1|Z^N) + H(M_2|Z^N, M_1)}{N}. \quad (44)$$

The first part $\frac{H(M_1|Z^N)}{N}$ of (44) is bounded by

$$\begin{aligned}
 \frac{H(M_1|Z^N)}{N} &= \frac{H(M_1|Z_1, Z_2^N)}{N} \stackrel{(f)}{=} \frac{H(M_1|Z_2^N)}{N} \\
 &= \frac{1}{N}(h(M_1, Z_2^N) - h(Z_2^N)) \\
 &= \frac{1}{N}(h(M_1, Z_2^N, A_2^{*N}) - h(A_2^{*N}|M_1, Z_2^N) - h(Z_2^N)) \\
 &= \frac{1}{N}(h(M_1, A_2^{*N}) + h(Z_2^N|M_1, A_2^{*N}) - h(A_2^{*N}|M_1, Z_2^N) - h(Z_2^N)) \\
 &= \frac{1}{N}(h(M_1, A_2^{*N}) - I(Z_2^N; M_1, A_2^{*N}) - h(A_2^{*N}|M_1, Z_2^N)) \\
 &= \frac{1}{N}(h(M_1, A_2^{*N}|Z_2^N) - h(A_2^{*N}|M_1, Z_2^N)) \\
 &= \frac{1}{N}(h(A_2^{*N}|Z_2^N) + H(M_1|A_2^{*N}, Z_2^N) - h(A_2^{*N}|M_1, Z_2^N)) \\
 &= \frac{1}{N}(h(A_2^{*N}|Z_2^N) - h(A_2^{*N}|M_1, Z_2^N) + h(A_2^{*N}) - h(A_2^{*N})) \\
 &= \frac{1}{N}(h(A_2^{*N}) - I(Z_2^N; A_2^{*N}) - h(A_2^{*N}|M_1, Z_2^N)) \\
 &\stackrel{(g)}{=} \frac{1}{N}(H(M_1, M_1^*) - I(Z_2^N; A_2^{*N}) - h(A_2^{*N}|M_1, Z_2^N)) \\
 &= \frac{1}{N}((N-1)(R_1 + R^*) - I(Z_2^N; A_2^{*N}) - h(A_2^{*N}|M_1, Z_2^N)) \\
 &\stackrel{(h)}{\geq} \frac{N-1}{N}(R_1 + R^*) - \frac{I(Z_2^N; A_2^{*N})}{N} - \frac{\delta(\epsilon)}{N} \\
 &= \frac{N-1}{N}(R_1 + R^*) - \frac{I(Z_2^N; A_2^{*N})}{N-1} \frac{N-1}{N} - \frac{\delta(\epsilon)}{N}, \tag{45}
 \end{aligned}$$

where (f) follows from the Markov chain $M_1 \rightarrow Z_2^N \rightarrow Z_1$ and $Z_2^N = (Z_2, Z_3, \dots, Z_N)$, (g) is due to the fact that A_2^{*N} is a deterministic function of M_1, M_1^* , and (h) follows from Fano's inequality when the codeword A_2^{*N} is generated by Feinstein's greedy construction [19], i.e., if

$$R^* \leq \frac{I(Z_2^N; A_2^{*N})}{N-1}, \tag{46}$$

given M_1 , the eavesdropper's decoding error probability P_{ew} of M_1^* is arbitrarily small ($P_{ew} \leq \epsilon$) as N tends to infinity, then using Fano's inequality, we have

$$h(A_2^{*N}|M_1, Z_2^N) \leq \delta(P_{ew}) \leq \delta(\epsilon). \tag{47}$$

Formula (45) indicates that N is sufficiently large, if

$$R^* \geq \frac{I(Z_2^N; A_2^{*N})}{N-1}, \tag{48}$$

we have

$$\frac{H(M_1|Z^N)}{N} \geq R_1 - \epsilon', \tag{49}$$

where $\epsilon' \rightarrow 0$ as $N \rightarrow \infty$.

From (46) and (48), we conclude that

$$R^* = \frac{I(Z_2^N; A_2^{*N})}{N-1}. \tag{50}$$

Substituting (50) into (38) and noting that the maximum of R_1 is achieved when $N \rightarrow \infty$, then we have

$$R_1 \leq \lim_{N \rightarrow \infty} \left\{ \frac{I(A_2^{*N}; Y_2^N) - I(A_2^{*N}; Z_2^N)}{N-1} \right\} \quad (51)$$

Lemma 2. The fundamental limit of R_1 in (51) is upper bounded by

$$\begin{aligned} R_1 &\leq \lim_{N \rightarrow \infty} \left\{ \frac{I(A_2^{*N}; Y_2^N) - I(A_2^{*N}; Z_2^N)}{N-1} \right\} \\ &\leq \frac{1}{2} \log \left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2} \right) \\ &\quad - \frac{1}{2} \log \left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2 + \sigma_2^2} \right) \end{aligned} \quad (52)$$

Proof. The Proof of Lemma 2 is in Appendix B. $\square$

The second part $\frac{H(M_2|Z^N, M_1)}{N}$ of (44) is bounded by

$$\begin{aligned} \frac{H(M_2|Z^N, M_1)}{N} &\stackrel{(a)}{=} \frac{H(\theta|Z^N, M_1)}{N} \geq \frac{1}{N} H(\theta | Z^N, \eta_{1,1}, \dots, \eta_{1,N}, \eta_{2,2}, \dots, \eta_{2,N}, W'_1, \dots, W'_N, M_1, M_1^*) \\ &\stackrel{(b)}{=} \frac{1}{N} H(\theta | \sqrt{12P'}(\theta - \frac{W'_1}{\sqrt{12P'}} + B) + W'_1 + \eta_{1,1} + \eta_{2,1}, \\ &\quad A_2^* + \sqrt{\frac{P'}{\alpha_1}} \varepsilon_1 + W'_2 + \eta_{1,2} + \eta_{2,2}, A_3^* + f_{2,3}(\eta_{1,1}, \eta_{1,2}) + W'_3 + \eta_{1,3} + \eta_{2,3}, \dots, \\ &\quad A_N^* + f_{2,N}(\eta_{1,1}, \dots, \eta_{1,N-1}) + W'_N + \eta_{1,N} + \eta_{2,N}, \eta_{1,1}, \dots, \eta_{1,N}, \eta_{2,2}, \dots, \eta_{2,N}, W'_1, \dots, W'_N, M_1, M_1^*) \\ &= \frac{1}{N} H(\theta | \sqrt{12P'}\theta + \eta_{2,1}, \eta_{1,1}, \dots, \eta_{1,N}, \eta_{2,2}, \dots, \eta_{2,N}, W'_1, \dots, W'_N, M_1, M_1^*) \\ &\stackrel{(c)}{=} \frac{1}{N} H(\theta | \sqrt{12P'}\theta + \eta_{2,1}) \\ &\stackrel{(d)}{=} \frac{1}{N} \{ H(\theta) - h(\sqrt{12P'}\theta + \eta_{2,1}) + h(\eta_{2,1}) \} \\ &\stackrel{(e)}{\geq} R_2 - \frac{1}{2N} \log(1 + \frac{P'}{\sigma_2^2}) \end{aligned} \quad (53)$$

where (a) follows from the fact that there is a one-to-one mapping between M_2 and θ , (b) follows from (28), (30), (33) and (A5), (c) follows from the fact that θ and $\eta_{2,1}$ are independent of $\eta_{1,1}, \dots, \eta_{1,N}, \eta_{2,2}, \dots, \eta_{2,N}, W'_1, \dots, W'_N, M_1, M_1^*$, (d) follows from the fact that θ and $\eta_{2,1}$ are independent of each other, and (e) follows from $H(\theta) = NR_2$, and the variance of θ equals $\frac{1}{12}$ as N tends to infinity.

Substituting (49) and (53) into (44), we have

$$\begin{aligned} \frac{H(M|Z^N)}{N} &= \frac{H(M_1|Z^N) + H(M_2|Z^N, M_1)}{N} \\ &\geq R_1 - \epsilon' + R_2 - \frac{1}{2N} \log(1 + \frac{P'}{\sigma_2^2}) \\ &= R - (\epsilon' + \frac{1}{2N} \log(1 + \frac{P'}{\sigma_2^2})) \end{aligned} \quad (54)$$

where $R = R_1 + R_2$. Choosing sufficiently large N , $\frac{H(M|Z^N)}{N} \geq R - \epsilon$ is satisfied. Finally, combining (43) and (52), the lower bound in *Theorem 1* is derived, and the proof is completed.

5. Conclusions and Future Work

This paper shows a new lower bound on the secrecy capacity of the AD-DP-WTC with noiseless feedback by proposing a novel two-step hybrid feedback scheme. The numerical result shows that when the noise variance σ_w^2 is large enough, our new feedback scheme performs better than the existing one in the literature. Moreover, when the eavesdropper's channel noise variance is fairly large, our new feedback scheme is almost optimal (the new lower bound almost equals the upper bound). Possible future work could be to extend the proposed feedback scheme to the multiple-access situation.

Author Contributions: G.X. did the theoretical work, performed the experiments, analyzed the data, and drafting the work; B.D. designed the work, did the theoretical work, analyzed the data, interpreted the data for the work, and revised the work. Both of the authors approve the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1806405; in part by the National Natural Science Foundation of China under Grant 62071392; and in part by the 111 Project No.111-2-14.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Lemma 1

From (32), we know that $\varepsilon_1 = \frac{\eta_{1,1}}{\sqrt{12P'}}$ and $\alpha_1 = \text{Var}(\varepsilon_1)$, and hence $\alpha_1 = E(\frac{\eta_{1,1}}{\sqrt{12P'}})^2 = \frac{\sigma_1^2}{12P'}$.

In the same way, since $\varepsilon_2 = \varepsilon_1 - \frac{E(Y_2''\varepsilon_1)}{E(Y_2'')^2}Y_2''$ and $\alpha_2 = \text{Var}(\varepsilon_2)$, we have

$$\begin{aligned}\alpha_2 &= \text{Var}(\varepsilon_2) = E\left(\varepsilon_1 - \frac{E(Y_2''\varepsilon_1)}{E(Y_2'')^2}Y_2''\right)^2 = E(\varepsilon_1)^2 - \frac{E(Y_2''\varepsilon_1)^2}{E(Y_2'')^2} \\ &= \alpha_1 - \frac{\alpha_1 P'}{P' + \sigma_1^2} = \alpha_1 \frac{\sigma_1^2}{P' + \sigma_1^2},\end{aligned}\quad (\text{A1})$$

where $Y_2'' = V_2 + \eta_{1,2} = \sqrt{\frac{P'}{\alpha_1}}\varepsilon_1 + \eta_{1,2}$.

For $3 \leq k \leq N$,

$$\begin{aligned}\alpha_{k-1} &= \text{Var}(\varepsilon_{k-1}) \stackrel{(a)}{=} E\left(\varepsilon_{k-2} - \frac{E(Y_{k-1}''\varepsilon_{k-2})}{E(Y_{k-1}'')^2}Y_{k-1}''\right)^2 = E(\varepsilon_{k-2})^2 - \frac{E(Y_{k-1}''\varepsilon_{k-2})^2}{E(Y_{k-1}'')^2} \\ &= \alpha_{k-2} - \frac{\alpha_{k-2}P'}{P' + \sigma_1^2} = \alpha_{k-2} \frac{\sigma_1^2}{P' + \sigma_1^2} = \alpha_1 \left(\frac{\sigma_1^2}{P' + \sigma_1^2}\right)^{k-2},\end{aligned}\quad (\text{A2})$$

where (a) follows from (34), and $Y_{k-1}'' = V_{k-1} + \eta_{1,k-1} = \sqrt{\frac{P'}{\alpha_{k-2}}}\varepsilon_{k-2} + \eta_{1,k-1}$.

Now, it remains to further compute the second part of Lemma 1. According to (35), we have $V_k = \sqrt{\frac{P'}{\alpha_{k-1}}} \varepsilon_{k-1}$, and it can be further expressed as

$$\begin{aligned} V_k &= \sqrt{\frac{P'}{\alpha_{k-1}}} \varepsilon_{k-1} \stackrel{(b)}{=} \sqrt{\frac{P'}{\alpha_{k-1}}} \left\{ \varepsilon_{k-2} - \frac{E(V_{k-1} + \eta_{1,k-1}) \varepsilon_{k-2}}{E(V_{k-1} + \eta_{1,k-1})^2} (V_{k-1} + \eta_{1,k-1}) \right\} \\ &= \sqrt{\frac{P'}{\alpha_{k-2}}} \sqrt{\frac{\alpha_{k-2}}{\alpha_{k-1}}} \left\{ \varepsilon_{k-2} - \frac{E(V_{k-1} + \eta_{1,k-1}) \varepsilon_{k-2}}{E(V_{k-1} + \eta_{1,k-1})^2} (V_{k-1} + \eta_{1,k-1}) \right\} \\ &= \sqrt{\frac{\alpha_{k-2}}{\alpha_{k-1}}} V_{k-1} - \sqrt{\frac{P'}{\alpha_{k-1}}} \frac{\sqrt{\alpha_{k-2} P'}}{P' + \sigma_1^2} (V_{k-1} + \eta_{1,k-1}) \\ &= \sqrt{\frac{\alpha_{k-2}}{\alpha_{k-1}}} \frac{\sigma_1^2}{P' + \sigma_1^2} V_{k-1} - \sqrt{\frac{\alpha_{k-2}}{\alpha_{k-1}}} \frac{P'}{P' + \sigma_1^2} \eta_{1,k-1}, \end{aligned} \quad (A3)$$

where (b) follows from (34), and $Y''_{k-1} = V_{k-1} + \eta_{1,k-1} = \sqrt{\frac{P'}{\alpha_{k-2}}} \varepsilon_{k-2} + \eta_{1,k-1}$.

Therefore, we conclude that

$$\begin{aligned} V_3 &= \sqrt{\frac{\alpha_1}{\alpha_2}} \frac{\sigma_1^2}{P' + \sigma_1^2} V_2 - \sqrt{\frac{\alpha_1}{\alpha_2}} \frac{P'}{P' + \sigma_1^2} \eta_{1,2}, \\ V_4 &= \sqrt{\frac{\alpha_2}{\alpha_3}} \frac{\sigma_1^2}{P' + \sigma_1^2} V_3 - \sqrt{\frac{\alpha_2}{\alpha_3}} \frac{P'}{P' + \sigma_1^2} \eta_{1,3} \\ &= \sqrt{\frac{\alpha_1}{\alpha_3}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^2 V_2 - \sum_{i=1}^2 \sqrt{\frac{\alpha_i}{\alpha_3}} \frac{P' (\sigma_1^2)^{2-i}}{(P' + \sigma_1^2)^{4-(i+1)}} \eta_{1,i+1} \\ &\vdots \\ V_k &= \sqrt{\frac{\alpha_1}{\alpha_{k-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k-2} V_2 - \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{k-1}}} \frac{P' (\sigma_1^2)^{k-2-i}}{(P' + \sigma_1^2)^{k-(i+1)}} \eta_{1,i+1} \\ &\stackrel{(c)}{=} \frac{1}{\sqrt{12\alpha_{k-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k-2} \eta_{1,1} - \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{k-1}}} \frac{P' (\sigma_1^2)^{k-2-i}}{(P' + \sigma_1^2)^{k-(i+1)}} \eta_{1,i+1} \end{aligned} \quad (A4)$$

where (c) follows from $V_2 = \sqrt{\frac{P'}{\alpha_1}} \varepsilon_1 = \sqrt{\frac{P'}{\alpha_1}} \frac{\eta_{1,1}}{\sqrt{12P'}}$. From (A4) and (33), we see that V_k is a function of $(\eta_{1,1}, \eta_{1,2}, \dots, \eta_{1,k-1})$, i.e.,

$$V_k = f_{2,k}(\eta_{1,1}, \eta_{1,2}, \dots, \eta_{1,k-1}), \quad 2 \leq k \leq N. \quad (A5)$$

Finally, according to (A2) and (A4), we complete the proof of Lemma 1.

Appendix B. Proof of Lemma 2

Since

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{I(A_2^{*N}; Y_2^N) - I(A_2^{*N}; Z_2^N)}{N-1} &= \lim_{N \rightarrow \infty} \frac{h(Y_2^N) - h(Y_2^N | A_2^{*N}) - h(Z_2^N) + h(Z_2^N | A_2^{*N})}{N-1} \\ &\stackrel{(a)}{=} \lim_{N \rightarrow \infty} \left\{ \frac{h(Y_2^N)}{N-1} - \frac{h(Z_2^N)}{N-1} - \frac{h(\eta_{1,2}^N)}{N-1} + \frac{h(\eta_{1,2}^N + \eta_{2,2}^N)}{N-1} \right\}, \end{aligned} \quad (A6)$$

where (a) is due to the fact that $Y_2^N = A_2^{*N} + \eta_{1,2}^N$, $Z_2^N = Y_2^N + \eta_{2,2}^N$, $\eta_{1,2}^N = V_2^N + W_2^N + \eta_{1,2}^N$ and A_2^{*N} is independent of $\eta_{1,2}^N, \eta_{2,2}^N$, and $\eta_{1,2}^N = (\eta_{1,2}^N, \eta_{1,3}^N, \dots, \eta_{1,N}^N)$, $\eta_{2,2}^N = (\eta_{2,2}^N, \eta_{2,3}^N, \dots, \eta_{2,N}^N)$.

For the part $\frac{h(Y_2^N)}{N-1} - \frac{h(Z_2^N)}{N-1}$ of (A6), we have

$$\begin{aligned} \lim_{N \rightarrow \infty} \left\{ \frac{h(Y_2^N)}{N-1} - \frac{h(Z_2^N)}{N-1} \right\} &= \lim_{N \rightarrow \infty} \left\{ \frac{h(Y_2^N)}{N-1} - \frac{h(Y_2^N + \eta_{2,2}^N)}{N-1} \right\} \\ &\stackrel{(b)}{\leq} \lim_{N \rightarrow \infty} \left\{ \frac{h(Y_2^N)}{N-1} - \frac{1}{2} \log(2^{\frac{2h(Y_2^N)}{N-1}} + 2\pi e \sigma_2^2) \right\} \\ &\stackrel{(c)}{\leq} \frac{1}{2} \log 2\pi e(P^* + P' + \sigma_w'^2 + \sigma_1^2) - \frac{1}{2} \log 2\pi e(P^* + P' + \sigma_w'^2 + \sigma_1^2 + \sigma_2^2), \quad (\text{A7}) \end{aligned}$$

where (b) follows from the entropy power inequality, (c) follows from $\frac{h(Y_2^N)}{N-1} - \frac{1}{2} \log(2^{\frac{2h(Y_2^N)}{N-1}} + 2\pi e \sigma_2^2)$ increasing while $\frac{h(Y_2^N)}{N-1}$ is increasing and

$$\begin{aligned} \frac{h(Y_2^N)}{N-1} &\leq \frac{1}{N-1} \sum_{i=2}^N h(Y_i) \leq \frac{1}{N-1} \sum_{i=2}^N \frac{1}{2} \log 2\pi e E(Y_i^2) \\ &= \frac{N}{N-1} \frac{1}{N} \sum_{i=2}^N \frac{1}{2} \log 2\pi e E(Y_i^2) \\ &\leq \frac{1}{2} \frac{N}{N-1} \log 2\pi e(P^* + P' + \sigma_w'^2 + \sigma_1^2), \quad (\text{A8}) \end{aligned}$$

where $P^* = P_A + \rho_2^2 P + 2\rho_2 \sqrt{P P_A}$, $P' = P(1 - \rho_1^2 - \rho_2^2)$, $\sigma_w'^2 = (\sigma_w + \rho_1 \sqrt{P})^2$, and the last inequality of (A8) is from Jensen's inequality.

Based on the joint differential entropy, for the part $\frac{h(\eta_{1,2}^N)}{N-1}$ of (A6), we have

$$\lim_{N \rightarrow \infty} h(\eta_{1,2}^N) = \lim_{N \rightarrow \infty} \frac{1}{2} \log(2\pi e)^{N-1} |K|, \quad (\text{A9})$$

where K is a covariance matrix, and

$$|K| = \begin{vmatrix} E(\eta'_{1,2}\eta'_{1,2}) & E(\eta'_{1,2}\eta'_{1,3}) & \cdots & E(\eta'_{1,2}\eta'_{1,N}) \\ E(\eta'_{1,3}\eta'_{1,2}) & E(\eta'_{1,3}\eta'_{1,3}) & \cdots & E(\eta'_{1,3}\eta'_{1,N}) \\ \vdots & \vdots & \cdots & \vdots \\ E(\eta'_{1,N}\eta'_{1,2}) & E(\eta'_{1,N}\eta'_{1,3}) & \cdots & E(\eta'_{1,N}\eta'_{1,N}) \end{vmatrix}_{(N-1) \times (N-1)} \quad (\text{A10})$$

with

$$E(\eta'_{1,k}\eta'_{1,k}) = P' + \sigma_w'^2 + \sigma_1^2, \quad 2 \leq k \leq N \quad (\text{A11})$$

$$\begin{aligned} E(\eta'_{1,k}\eta'_{1,j}) &= E(V_k + W'_k + \eta_{1,k})(V_j + W'_j + \eta_{1,j}) \\ &= E(V_k V_j) + E(\eta_{1,k} V_j), \quad 2 \leq k \leq N, \quad k < j \leq N. \end{aligned} \quad (\text{A12})$$

According to (A4), we have

$$\begin{aligned}
 V_k &= \sqrt{\frac{\alpha_1}{\alpha_{k-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k-2} V_2 - \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{k-1}}} \frac{P'(\sigma_1^2)^{k-2-i}}{(P' + \sigma_1^2)^{k-(i+1)}} \eta_{1,i+1}, \\
 V_j &= \sqrt{\frac{\alpha_1}{\alpha_{j-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{j-2} V_2 - \sum_{i=1}^{j-2} \sqrt{\frac{\alpha_i}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-2-i}}{(P' + \sigma_1^2)^{j-(i+1)}} \eta_{1,i+1} \\
 &= \sqrt{\frac{\alpha_1}{\alpha_{j-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{j-2} V_2 - \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-2-i}}{(P' + \sigma_1^2)^{j-(i+1)}} \eta_{1,i+1} \\
 &\quad - \sqrt{\frac{\alpha_{k-1}}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-k-1}}{(P' + \sigma_1^2)^{j-k}} \eta_{1,k} - \sum_{i=k}^{j-2} \sqrt{\frac{\alpha_i}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-2-i}}{(P' + \sigma_1^2)^{j-(i+1)}} \eta_{1,i+1},
 \end{aligned} \tag{A13}$$

Based on (A13), we can conclude that

$$\begin{aligned}
 E(\eta_{1,k} V_j) &= -E \left(\eta_{1,k} \sqrt{\frac{\alpha_{k-1}}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-k-1}}{(P' + \sigma_1^2)^{j-k}} \eta_{1,k} \right) \\
 &= -\sqrt{\frac{\alpha_{k-1}}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-k}}{(P' + \sigma_1^2)^{j-k}} \stackrel{(d)}{=} -P' \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{\frac{j-k}{2}},
 \end{aligned} \tag{A14}$$

where (d) follows from (A2). For $E(V_k V_j)$, we have

$$\begin{aligned}
 E(V_k V_j) &\stackrel{(e)}{=} E \left\{ \left[\sqrt{\frac{\alpha_1}{\alpha_{k-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k-2} V_2 - \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{k-1}}} \frac{P'(\sigma_1^2)^{k-2-i}}{(P' + \sigma_1^2)^{k-(i+1)}} \eta_{1,i+1} \right] \right. \\
 &\quad \times \left. \left[\sqrt{\frac{\alpha_1}{\alpha_{j-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{j-2} V_2 - \sum_{i=1}^{j-2} \sqrt{\frac{\alpha_i}{\alpha_{j-1}}} \frac{P'(\sigma_1^2)^{j-2-i}}{(P' + \sigma_1^2)^{j-(i+1)}} \eta_{1,i+1} \right] \right\} \\
 &= P' \sqrt{\frac{\alpha_1}{\alpha_{k-1}}} \sqrt{\frac{\alpha_1}{\alpha_{j-1}}} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{k+j-4} + \sum_{i=1}^{k-2} \sqrt{\frac{\alpha_i}{\alpha_{k-1}}} \sqrt{\frac{\alpha_i}{\alpha_{j-1}}} \frac{(P')^2 (\sigma_1^2)^{k+j-3-2i}}{(P' + \sigma_1^2)^{k+j-2(i+1)}} \\
 &\stackrel{(f)}{=} P' \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{\frac{k+j-4}{2}} + \sum_{i=1}^{k-2} \frac{(P')^2}{\sigma_1^2} \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{\frac{k+j-2(i+1)}{2}} \\
 &= P' \left(\frac{\sigma_1^2}{P' + \sigma_1^2} \right)^{\frac{j-k}{2}},
 \end{aligned} \tag{A15}$$

where (e) follows from (A13), and (f) follows from (A2).

Hence, (A12) can be further calculated by

$$E(\eta'_{1,k} \eta'_{1,j}) = E(V_k V_j) + E(\eta_{1,k} V_j) \stackrel{(g)}{=} 0, \quad 2 \leq k \leq N, \quad k < j \leq N, \tag{A16}$$

where (g) follows from (A15) and (A14).

According to (A16) and (A11),

$$|K| = \begin{vmatrix} P' + \sigma_w^2 + \sigma_1^2 & 0 & \cdots & 0 \\ 0 & P' + \sigma_w^2 + \sigma_1^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & P' + \sigma_w^2 + \sigma_1^2 \end{vmatrix}_{(N-1) \times (N-1)} \tag{A17}$$

Therefore, we can obtain that

$$\lim_{N \rightarrow \infty} \frac{1}{N-1} h(\eta'_{1,2}) = \lim_{N \rightarrow \infty} \frac{1}{2(N-1)} \log(2\pi e)^{N-1} |K| = \frac{1}{2} \log 2\pi e (P' + \sigma_w'^2 + \sigma_1^2). \quad (\text{A18})$$

Analogously, we have

$$\lim_{N \rightarrow \infty} \frac{1}{N-1} h(\eta'_{1,2} + \eta'_{2,2}) = \frac{1}{2} \log 2\pi e (P' + \sigma_w'^2 + \sigma_1^2 + \sigma_2^2). \quad (\text{A19})$$

Substituting (A7), (A18) and (A19) into (A6), we prove that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{I(A_2^{*N}; Y_2^N) - I(A_2^{*N}; Z_2^N)}{N-1} \\ & \leq \frac{1}{2} \log \left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2} \right) \\ & \quad - \frac{1}{2} \log \left(1 + \frac{(\sqrt{P_A} + \rho_2 \sqrt{P})^2}{P(1 - \rho_1^2 - \rho_2^2) + (\sigma_w + \rho_1 \sqrt{P})^2 + \sigma_1^2 + \sigma_2^2} \right), \end{aligned} \quad (\text{A20})$$

which completes the proof.

References

1. Zaidi, A.; Vandendorpe, L. Coding schemes for relay-assisted information embedding. *IEEE Trans. Inf. Forensics Secur.* **2009**, *4*, 70–85. [\[CrossRef\]](#)
2. Costa, M.H.M. Writing on dirty paper. *IEEE Trans. Inf. Theory* **1983**, *29*, 439–441. [\[CrossRef\]](#)
3. Yan, W.; Yuan, X.; He, Z.; Kuai, X. Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 1793–1808. [\[CrossRef\]](#)
4. Weissman, T. Capacity of channels with action-dependent states. *IEEE Trans. Inf. Theory* **2010**, *56*, 5396–5411. [\[CrossRef\]](#)
5. Dikstein, L.; Permuter, H.H.; Shamai, S. MAC with action-dependent state information at one encoder. *IEEE Trans. Inf. Theory* **2015**, *61*, 173–188. [\[CrossRef\]](#)
6. Chen, Y.; Vinck, A.J.H. Wiretap channel with side information. *IEEE Trans. Inf. Theory* **2008**, *54*, 395–402. [\[CrossRef\]](#)
7. Treust, M.L.; Zaidi, A.; Lasaulce, S. An achievable rate region for the broadcast channel with asymmetric side information. In Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, Monticello, IL, USA, 28–30 September 2011; pp. 68–75. [\[CrossRef\]](#)
8. El-Halabi, M.; Liu, T.; Georgiades, C.N.; Shamai (Shitz), S. Secret writing on dirty paper: A deterministic view. *IEEE Trans. Inf. Theory* **2012**, *58*, 3419–3429. [\[CrossRef\]](#)
9. Mitrpan, C.; Vinck, A.; Luo, Y. An achievable region for the Gaussian wiretap channel with side information. *IEEE Trans. Inf. Theory* **2006**, *52*, 2181–2190. [\[CrossRef\]](#)
10. Leung-Yan-Cheong, S.; Hellman, M. The Gaussian wire-tap channel. *IEEE Trans. Inf. Theory* **1978**, *24*, 451–456. [\[CrossRef\]](#)
11. Wyner, A.D. The wire-tap channel. *Bell Syst. Tech. J.* **1975**, *54*, 1355–1387. [\[CrossRef\]](#)
12. Csiszar, I.; Körner, J. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory* **1978**, *24*, 339–348.
13. Dai, B.; Li, C.; Liang, Y.; Ma, Z.; Shamai, S. Impact of action-dependent state and channel feedback on Gaussian wiretap channels. *IEEE Trans. Inf. Theory* **2020**, *66*, 3435–3455. [\[CrossRef\]](#)
14. Zhang, H.; Yu, L.; Dai, B. Feedback schemes for the action dependent wiretap channel with noncausal state at the transmitter. *Entropy* **2019**, *21*, 278. [\[CrossRef\]](#)
15. Zhang, H.; Yu, L.; Wei, C.; Dai, B. A new feedback scheme for the state-dependent wiretap channel with noncausal state at the transmitter. *IEEE Access* **2019**, *7*, 45594–45604. [\[CrossRef\]](#)
16. Zaidi, A.; Awan, Z.H.; Shamai, S.; Vandendorpe, L. Secure degrees of Freedom of MIMO-X channels with output feedback and delayed CSIT. *IEEE Trans. Inf. Forensics Secur.* **2013**, *8*, 1760–1774. [\[CrossRef\]](#)
17. Schalkwijk, J.P.M.; Kailath, T. A coding scheme for additive noise channels with feedback. part I: No bandwidth constraint. *IEEE Trans. Inf. Theory* **1966**, *12*, 172–182. [\[CrossRef\]](#)
18. Gunduz, D.; Brown, D.R.; Poor, H.V. Secret communication with feedback. In Proceedings of the International Symposium on Information Theory and Its Applications, ISITA, Auckland, New Zealand, 7–10 December 2008; pp. 1–6. [\[CrossRef\]](#)
19. Feinstein, A. A new basic theorem of information theory. *IEEE Trans. Inf. Theory* **1954**, *4*, 2–22. [\[CrossRef\]](#)