On Covert Communication Over Infinite-Bandwidth Gaussian Channels

Ligong Wang

On Covert Communication Over Infinite-Bandwidth Gaussian Channels

Ligong Wang
ETIS—Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS
Cergy-Pontoise, France
ligong.wang@ensea.fr

***Abstract*—We consider a problem of communication over a continuous-time additive Gaussian noise channel. A covertness constraint is imposed on the communication protocol, which says that the channel output must statistically resemble pure noise. When there is no bandwidth constraint on the input, we argue that the covert communication capacity of this channel is positive, in contrast to the band-limited case where covertness requires that the amount of transmission grow at most like the square root of total communication time. This claim is formally proven in the case where the Gaussian noise is white with respect to the bandwidth used by the transmitter.**

***Index Terms*—Covert communication, low probability of detection, Gaussian channel, continuous time, waveform channel.**

I. Introduction

Covert communication, or communicaiton with low probability of detection [1]–[4], refers to scenarios where the transmitter and the receiver must keep a warden from discovering the fact that they are using the channel to communicate. On an additive white Gaussian noise (AWGN) channel, this means that the warden's observation should be statistically close to pure noise. It was first shown in [1] that the AWGN channel obeys the so-called "square-root law" for covert communication: the number of information nats that can be communicated covertly over the channel can only grow proportionally to the square root of the total number of channel uses. The exact scaling law, when covertness is measured in terms of relative entropy, was determined in [3]. Similar results have been obtained for the binary symmetric channel [2] and general discrete memoryless channels [3], [4]. A number of further works have extended these results in several directions.

A discrete-time AWGN channel is usually used to model a real-life continuous-time communication channel with a bandwidth constraint on its input waveform, corrupted by white Gaussian noise with respect to that bandwidth. Using the sampling theorem, such a continuous-time channel with bandwidth W Hz over the time interval $[0, T]$ is approximately equivalent to $2WT$ uses of a discrete-time AWGN channel [5]. Hence, one can roughly say (as in, e.g., a brief remark in [1]) that the number of nats that can be covertly communicated over this continuous-time channel is proportional to $\sqrt{WT}$. Thus, for any fixed finite bandwidth W, the number of covert nats can only grow like the square root of total communication time T.

In this paper, we investigate the regime where W is infinity or grows large with T. In most information-theoretic works, asymptotic results are obtained in the limit where $T \to \infty$ with W held fixed, or where W is set to infinity *after* one lets $T \to \infty$ (the latter is how one normally obtains the capacity of an "infinite-bandwidth" AWGN channel). We study a different regime with the intention to capture engineering insights to scenarios where information is transmitted over a large bandwidth and a relatively short time. This can happen, for example, in "spread-spectrum" communication [6].

We observe that, if W is infinity or grows large fast enough with T, then covert communication can have positive rates in nats per second. In the white-noise case, this might be somewhat obvious, since we already argued that information throughput grows like $\sqrt{WT}$. In fact, we can show that, under the same average-power constraint on the input, the covert communication capacity is the same as the capacity without covertness constraint. What is perhaps more interesting is the case where the noise is colored, or band-limited, or both. Our key observation there is that positive-rate covert communication is possible if and only if the noise itself spans an infinite bandwidth.

The above observations are related to our recent work [7], which shows that the continuous-time Poisson channel with neither bandwidth nor peak-power constraint permits transmission of infinitely many information nats per second.

The above observations are first made using heuristic approaches along the direction of [5]. We then try to formulate the problem in a mathematically rigorous fashion. Formal treatments of continuous-time Gaussian channels are often complicated because, in short, no nonzero signal can be both band-limited and time-limited (see, e.g., [8, Theorem 6.8.2]). Shannon's capacity formula for the band-limited Gaussian channel [5] called for several follow-up works to acquire a clear physical meaning; see [9], [10] and references therein. We start from these classic treatments, and observe some technicalities that arise in covert communication. In particular, we point out that, even when the transmitted signal is strictly time-limited, one should *not* assume that the warden can only observe the channel output within the same time limit. Instead, in our formulation, we let the warden observe the output waveform over the entire real line.

The main technical part of this paper is divided into two sections: Section II makes some heuristic observations, and

Section III provides a formal treatment. Our formal treatment, however, so far only covers the case where the noise is white with respect to the bandwidth of interest; colored noise is only discussed heuristically in Section II. Before proceeding, we first introduce our notation and give an overall description of the problem.

A. Notation and problem description

We use a boldface letter like $\mathbf{x}$ to denote a real function, either on the entire real line or on an interval. The value of $\mathbf{x}$ at a specific time t is denoted $x(t)$. When it is necessary to specify the domain of a function, we write it more explicitly as, for example, $x(t)$, $t \in \mathbb{R}$, or $x(t)$, $t \in [0, T]$; the latter is sometimes abbreviated as x_0^T. We use the upper case to denote random objects: $\mathbf{X}$ denotes a (real) random process, and $X(t)$ denotes the value of $\mathbf{X}$ at time t, hence $X(t)$ is a real random variable for any t in the domain of $\mathbf{X}$. We use $P_{\mathbf{X}}$ to denote the *finite-dimensional distribution* of the random process $\mathbf{X}$.

In this paper we are concerned with channels described by

$$Y(t) = X(t) + Z(t), \quad t \in \mathbb{R}, \tag{1}$$

where $\mathbf{X}$ is the channel input, $\mathbf{Y}$ is the channel output, an $\mathbf{Z}$ is the additive noise to the channel. Throughout the paper, we assume that $Z(t)$, $t \in \mathbb{R}$, is zero-mean stationary Gaussian.

As in [3], we let both the receiver and the warden observe $\mathbf{Y}$ (or $\mathbf{Z}$, if no communication is taking place), and provide the transmitter and the receiver with a sufficiently long secret key; we do not consider the resolvability aspect of covert communication as does [4]. Specifically, the transmitter maps every message, together with a sufficiently long secret key that it shares with the receiver, to some function $\mathbf{x}$. The receiver then maps the channel output $\mathbf{y}$ and the secret key to its decoded message. The warden's aim is to distinguish between the output waveform $\mathbf{Y}$ and a pure noise process $\mathbf{Z}$, so as to determine whether the channel is being used for communication or not. Our "covertness" constraint is that the Kullback-Leibler divergence $D(P_{\mathbf{Y}} \| P_{\mathbf{Z}})$ must be sufficiently small. Note that, by Pinsker's inequality [11], this would imply that the total variation distance between $P_{\mathbf{Y}}$ and $P_{\mathbf{Z}}$ also be small, but not reversely.

II. Heuristic Observations

In this section, we analyze our problem using some conventional arguments that are not mathematically rigorous. The "observations" below hence should not be considered as proven information-theoretic results. We shall formalize some of these observations in Section III.

A. White Gaussian noise over infinite bandwidth

In our first setting, we want the Gaussian noise $\mathbf{Z}$ to be white with infinite bandwidth; that is, we want its power-spectral density (PSD) to equal some positive constant $N_0/2$ for all frequency $f \in \mathbb{R}$. Such a noise model is however mathematically invalid, as the noise variance would equal infinity at any given time. To avoid this issue, we assume, instead, that the noise is white within a certain bandwidth W_T Hz: its PSD equals $N_0/2$ for $|f| \leq W_T$ and decays to zero sufficiently fast for $|f| > W_T$, where W_T depends on and grows to infinity with total communication time T. Then, for any finite T, the noise process is well-defined. The assumption that W_T grows with T is only for mathematical convenience, and has no engineering meaning.

Consider a constraint where $\mathbf{X}$ is power-limited to P per second. For our signaling strategy, we limit the bandwidth of $\mathbf{X}$ to W_T Hz. By a standard sampling argument [5], [12], the continuous-time channel can be reduced to $2W_T T$ uses of a discrete-time memoryless additive Gaussian noise channel with noise variance $N_0/2$. We choose the input for each sample to be independent and identically distributed (IID) Gaussian of mean zero and variance $P/2W_T$. Then $D(P_{\mathbf{Y}} \| P_{\mathbf{Z}})$ equals the sum of the Kullback-Leibler divergences corresponding to each use of the discrete-time channel, which can be upper-bounded as in [3, Section V] by

$$\frac{P^2}{2N_0^2} \cdot \frac{T}{W_T}.$$

This will vanish as $T \to \infty$ if W_T grows faster than T, for example, if $W_T \propto T^2$.

The input-output mutual information can be calculated as in normal (non-covert) communication. In particular, as W_T grows to infinity, the per-second mutual information tends to P/N_0. That this mutual information represents an achievable coding rate in nats per second can be shown using similar methods as in [3].

We thus see that, in the above setting, one can make $D(P_{\mathbf{Y}} \| P_{\mathbf{Z}})$ decay to zero as T grows large while still achieving its capacity P/N_0. To summarize, we have:

Observation 1: Under average-power constraint P, a channel of "infinite-bandwidth" with additive white Gaussian noise of double-sided PSD $N_0/2$ has communication capacity P/N_0 nats per second even under a covertness constraint that $D(P_{\mathbf{Y}} \| P_{\mathbf{Z}})$ must tend to zero as T grows large.

B. Noise with infinite bandwidth but finite variance

We have argued that positive-rate covert communication over a Gaussian channel is possible if the input signal can use infinite bandwidth and the noise is white over an infinite bandwidth. One might wonder whether this is an artifact resulting from the unbounded noise power, which allows one to hide a nonzero signal power in it. Thus, we now turn to scenarios where the additive noise has finite power. Let the stationary Gaussian noise $\mathbf{Z}$ in (1) have PSD $N(f)$, $f \in \mathbb{R}$, that is positive for all $f \in \mathbb{R}$, symmetric around $f = 0$, and satisfies

$$\int_{-\infty}^{\infty} N(f) \, \mathrm{d}f < \infty. \tag{2}$$

We choose the input signal $\mathbf{X}$ to be generated from a stationary Gaussian process with PSD

$$S(f) = \begin{cases} T^{-7/4} \cdot N(f), & f \in [-W_T, W_T] \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

where we choose $W_T = T^2$. (Note that W_T is a parameter of our choice and *not* given in the problem.) We then have [13, Theorem 10.5.1]

$$D(P_{\mathbf{Y}}\|P_{\mathbf{Z}}) = T\cdot\frac{1}{2}\int_{-W_T}^{W_T}\left(\frac{S(f)}{N(f)} - \log\left(1+\frac{S(f)}{N(f)}\right)\right)\mathrm{d}f \le T\cdot\frac{1}{4}\int_{-W_T}^{W_T}\left(\frac{S(f)}{N(f)}\right)^2\mathrm{d}f = \frac{T^{-1/2}}{2}, \tag{4}$$

which tends to zero as $T\to\infty$. We also have [13, Theorem 10.3.1]

$$I(\mathbf{X};\mathbf{Y}) = T\cdot\int_{-W_T}^{W_T}\frac{1}{2}\log\left(1+\frac{S(f)}{N(f)}\right)\mathrm{d}f = \frac{T^3}{2}\log\left(1+T^{-7/4}\right)\approx\frac{T^{5/4}}{2} \tag{5}$$

for large T. Hence the per-second mutual information scales like $T^{1/4}$ and grows to infinity with T. This suggests that infinite covert nats per second can be achieved. Further note that, by our choice, the average per-second input power, given by $\int_{-\infty}^{\infty}S(f)\,\mathrm{d}f$, tends to zero as T grows large. Summarizing the above we have the following.

Observation 2: If the Gaussian noise process has PSD $N(f)$ that is positive almost everywhere, then the covert communication capacity without bandwidth constraint on the input is infinity. Furthermore, this should hold irrespective of whether an average-power constraint is imposed on the input or not.

C. Band-limited noise

We next consider the case where the noise is band-limited:

$$N(f) = 0, \quad |f| > W. \tag{6}$$

Note that W is a constant that does not depend on T. We again restrict ourselves to using stationary Gaussian input processes.[1] Note that, if the input PSD $S(f)$ is positive on any interval where $N(f)=0$, then $D(P_{\mathbf{Y}}\|P_{\mathbf{Z}})$ will be infinity. Hence the input process must also be limited to the frequencies in $[-W,W]$. Let $\lambda(f)\triangleq S(f)/N(f)$ for $f\in[-W,W]$. If we require $D(P_{\mathbf{Y}}\|P_{\mathbf{Z}})\le\delta$ for some positive constant δ, then, again by [13, Theorem 10.5.1],

$$\delta \ge D(P_{\mathbf{Y}}\|P_{\mathbf{Z}}) = T\cdot\frac{1}{2}\int_{-W}^{W}\big(\lambda(f)-\log(1+\lambda(f))\big)\,\mathrm{d}f. \tag{7}$$

The integrand is convex in $\lambda(f)$, so we obtain

$$\bar{\lambda} - \log(1+\bar{\lambda}) \le \frac{\delta}{WT} \tag{8}$$

where

$$\bar{\lambda}\triangleq\frac{1}{2W}\int_{-W}^{W}\lambda(f)\,\mathrm{d}f. \tag{9}$$

From (8) we obtain that, for large T,

$$\bar{\lambda}\lesssim\sqrt{\frac{2\delta}{WT}}. \tag{10}$$

[1]Gaussian inputs are known to be optimal for covert communication over discrete-time Gaussian channels [3]. We expect them to be optimal for the current problem as well, though we do not prove it in this paper.

On the other hand,

$$I(\mathbf{X};\mathbf{Y}) = T\cdot\int_{-W}^{W}\frac{1}{2}\log(1+\lambda(f))\,\mathrm{d}f \le WT\log(1+\bar{\lambda})\lesssim\sqrt{2\delta WT}, \tag{11}$$

where the first inequality follows because the integrand is concave in $\lambda(f)$. We thus see that $I(\mathbf{X};\mathbf{Y})$ can only grow like $\sqrt{T}$.

Observation 3: For the channel (1) where $\mathbf{Z}$ is band-limited, the number of nats that can be covertly communicated can at most grow proportionally to $\sqrt{T}$ as T becomes large, even if no explicit constraint is imposed on the bandwidth of the input signal.

III. Formal Treatment

In this section we provide a mathematically rigorous treatment of the problem. We only consider white Gaussian noise, over a bandwidth that is either finite or unbounded.

Consider the channel (1), where $Z(t)$ is a zero-mean stationary Gaussian process. Let $D(t_1,t_2)$ denote the divergence $D(P_{\mathbf{Y}}\|P_{\mathbf{Z}})$ with both $\mathbf{Y}$ and $\mathbf{Z}$ restricted to the interval $[t_1,t_2]$. Ideally, we would like to have a model where $\mathbf{X}$ is strictly time-limited, while the covertness constraint is on the entire real line, i.e., on $D(-\infty,\infty)$, but so far we have not been able to solve the capacity for such a model. We discuss two relaxed models. As we shall see, the first model is meaningful, whereas the second model is flawed and gives pathological results. Before presenting these models and their results, we first provide some preliminary.

A. Preliminary: prolate spheroidal wave functions

The prolate spheroidal wave functions (PSWFs) are useful tools for analyzing band-limited continuous-time Gaussian channels. We give a brief introduction to them below; for more details, we refer the reader to [9], [10], [14] and references therein.

Given any $W,T>0$ there exists a countably infinite set of real positive numbers

$$1>\lambda_1>\lambda_2>\cdots \tag{12}$$

and a corresponding set of real functions $\{\boldsymbol{\psi}_i\colon\mathbb{R}\to\mathbb{R}\}_{i=1}^{\infty}$ such that the following properties are satisfied.

1) Each $\boldsymbol{\psi}_i$ is band-limited to W Hz. Further, the functions $\{\boldsymbol{\psi}_i\}$ are orthonormal on $\mathbb{R}$, and complete in the space of functions that are band-limited to W Hz.
2) The restrictions of $\{\boldsymbol{\psi}_i\}$ to the interval $[0,T]$ are orthogonal:
$$\int_0^T\psi_i(t)\psi_j(t)\,\mathrm{d}t = \begin{cases}\lambda_i, & i=j,\\ 0, & i\ne j.\end{cases} \tag{13}$$
Further, the restrictions of $\{\psi_i\}$ to $[0,T]$ are complete in the space of square integrable functions on $[0,T]$.

Note it is clear from the above properties that

$$\sum_{i=1}^{\infty}\lambda_i = 2WT. \tag{14}$$

It was shown by Slepian [15] that the coefficients $\{\lambda_i\}$ above satisfy the following: for any $\epsilon \in (0,1)$, as $WT \to \infty$,

$$\lambda_{2(1-\epsilon)WT} \to 1 \tag{15}$$

$$\lambda_{2(1+\epsilon)WT} \to 0. \tag{16}$$

Further, let $\mathbf{Z}$ be stationary Gaussian noise with PSD

$$N(f) = \begin{cases} \frac{N_0}{2}, & |f| \le W, \\ 0, & |f| > W \end{cases} \tag{17}$$

restricted to the interval $[0,T]$, then $\mathbf{Z}$ can be written in the Karhunen-Loève expansion using the above PSWFs:

$$Z(t) = \sum_{i=1}^{\infty} Z_i \psi_i(t), \quad t \in [0,T], \tag{18}$$

where $\{Z_i\}$ are IID Gaussian random variables of mean zero and variance $N_0/2$.

B. Covertness constraint on the entire real line

Model 1: We require the input signal to be "approximately time-limited," and impose the covertness constraint on the entire real line. Specifically, for every $T > 0$,

- the transmitter maps a message to $x(t)$, $t \in \mathbb{R}$, subject to the condition that the ratio
$$\frac{\int_0^T |x(t)|^2 \, \mathrm{d}t}{\int_{-\infty}^{\infty} |x(t)|^2 \, \mathrm{d}t}$$
must tend to one as T grows large;
- the receiver maps $y(t)$, $t \in [0,T]$, to a decoded message; and
- the covertness constraint is that $D(-\infty, \infty)$ must tend to zero as T grows large.

The first two conditions are taken from a classic way of treating band-limited Gaussian channel; see Wyner [9] and Gallager [10].

Proposition 1: Assume, for every T, the noise process $\mathbf{Z}$ has PSD $N_0/2$ over $[-W_T, W_T]$, where $W_T = T^2$. Under Model 1, and under power constraint

$$\mathsf{E}\left[\int_{-\infty}^{\infty} |X(t)|^2 \, \mathrm{d}t\right] \le PT, \tag{19}$$

the covert communication capacity of the channel is P/N_0 nats per second.

Proof sketch: The proof is a slight generalization of the classic approach in [9], [10]. Fix $\epsilon \in (0,1)$. Our coding scheme is to generate $2(1-\epsilon)T^3$ IID Gaussian random variables $\{X_i\}$ each of mean zero and variance $PT^{-2}/2$, and transmit the signal

$$X(t) = \sum_{i=1}^{(1-\epsilon)T^3} X_i \psi_i(t), \quad t \in \mathbb{R}, \tag{20}$$

where $\{\psi_i\}$ are PSWFs for the frequency band $[-T^2, T^2]$ and time interval $[0,T]$. The power constraint is satisfied because each $\boldsymbol{\psi}_i$ has unit energy. Further, it follows from (15) that, as $T \to \infty$, the power in $\mathbf{X}$ becomes concentrated in $[0,T]$.

For covertness, since the $\{\boldsymbol{\psi}_i\}$ are orthonormal on $\mathbb{R}$,

$$\begin{aligned} D(-\infty,\infty) &= \sum_{i=1}^{2(1-\epsilon)T^3} D(P_{X_i+Z_i} \| P_{Z_i}) \\ &= 2(1-\epsilon)T^3 \cdot \left(\frac{P}{2N_0T^2} - \log\left(1 + \frac{P}{2N_0T^2}\right)\right) \\ &\le 2(1-\epsilon)T^3 \cdot \frac{P^2}{4N_0^2T^4} = \frac{(1-\epsilon)P^2}{2N_0^2T}, \end{aligned} \tag{21}$$

which indeed tends to zero as $T \to \infty$.

Finally, we analyze the achievable rate with this scheme. For brevity, we only compute the per-second input-output mutual information; that this mutual information represents an achievable communication rate can be proven using similar methods as in [3]. To compute the mutual information between $\mathbf{X}$ and $\mathbf{Y}$ on the interval $[0,T]$, we use (13) and (18) to see that the receiver can recover every $X_i + Z_i$, $i \in \{1, \ldots, 2(1-\epsilon)T^3\}$, by computing the inner product of $\mathbf{Y}$ and $\boldsymbol{\psi}_i$ on $[0,T]$ and then dividing by λ_i. The continuous-time channel can then be reduced to $2(1-\epsilon)T^3$ parallel discrete-time Gaussian channels, with IID Gaussian noise. It can be easily computed that

$$\begin{aligned} \lim_{T\to\infty} \frac{I\left(X_0^T; Y_0^T\right)}{T} &= \lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{2(1-\epsilon)T^3} I(X_i; X_i + Z_i) \\ &= \lim_{T\to\infty} 2(1-\epsilon)T^2 \cdot \frac{1}{2} \log\left(1 + \frac{P}{N_0T^2}\right) \\ &= (1-\epsilon)\frac{P}{N_0}. \end{aligned} \tag{22}$$

By choosing ϵ arbitrarily close to zero, we approach the claimed capacity P/N_0. ∎

Proposition 2: Let $Z(t)$ be band-limited such that its PSD equals $N_0/2$ over $[-W, W]$ and equals zero elsewhere. Under Model 1, positive-rate covert communication is not possible, irrespectively of whether a power constraint is imposed on the input or not.

Proof sketch: Since the noise is band-limited to W Hz, covertness requires that the input signal $x(t)$, $t \in \mathbb{R}$ be also band-limited to W Hz. Then, since the $\{\boldsymbol{\psi}_i\}$ are complete in the space of functions band-limited to W Hz, we can write $\mathbf{X}$ as

$$X(t) = \sum_{i=1}^{\infty} X_i \psi_i(t), \quad t \in \mathbb{R} \tag{23}$$

for some infinite sequence of random variables $X_1, X_2, \ldots$. By (16) and by the requirement that $\mathbf{X}$ must be "almost time-limited" to $[0,T]$, we must have, for any $\epsilon > 0$,

$$\lim_{T\to\infty} \sum_{i=2(1+\epsilon)WT}^{\infty} \mathsf{E}[X_i] = 0. \tag{24}$$

By an argument similar to, e.g., [9], we know that $\{X_i, i \ge 2(1+\epsilon)WT\}$ cannot contribute to any positive communication rate. The remaining X_is can be thought of as inputs to $2(1+\epsilon)WT$ parallel channels with IID Gaussian noise. By [1], [3], the amount of information that can be covertly communicated over these channels can at most grow proportionally

to $\sqrt{(1+\epsilon)WT}$ as $T \to \infty$, hence cannot contribute to a positive rate, either. ∎

C. Time-limited model

Model 2: The input signal is strictly time-limited, and the covertness constraint is imposed on the duration of the input signal. Specifically, for $T > 0$,

- the transmitter maps a message to $x(t)$, $t \in [0, T]$;
- the receiver maps $y(t)$, $t \in [0, T]$, to a decoded message; and
- the covertness constraint is that $D(0, T)$ must tend to zero as T grows large.

Proposition 3: Let $Z(t)$ have PSD that equals $N_0/2$ on $[-W, W]$ and zero elsewhere. Under Model 2, the covert communication capacity of the channel (1) is infinity.

Proof sketch: Our signaling scheme is again to modulate on the PSWFs. The idea is simply to use the fact that, for any $W, T > 0$, the number of nonzero PSWFs is infinity.

For any positive integer k, we generate a sequence of k^3 IID Gaussian random variables $\{X_i\}$ of mean zero and variance k^{-2}. Let

$$X(t) = \begin{cases} \sum_i X_i \psi_i(t), & t \in [0, T], \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

Clearly, $X(t)$ is strictly time-limited to $[0, T]$. By the orthogonality of the PSWFs on $[0, T]$, the channel can be reduced, for both the warden and the receiver, to a set of k^3 parallel Gaussian channels

$$Y_i = X_i + Z_i. \tag{26}$$

The total divergence is

$$D(P_{\mathbf{Y}} \| P_{\mathbf{Z}}) = k^3 \left(\frac{2k^{-2}}{N_0} - \log\left(1 + \frac{2k^{-2}}{N_0}\right) \right) \leq \frac{2}{k N_0^2}. \tag{27}$$

The input-output mutual information is

$$I(\mathbf{X}; \mathbf{Y}) = \frac{k^3}{2} \log\left(1 + \frac{2k^{-2}}{N_0}\right). \tag{28}$$

Clearly, as we let k grow large, $D(P_{\mathbf{Y}} \| P_{\mathbf{Z}})$ tends to zero, while $I(\mathbf{X}; \mathbf{Y})$ tends to infinity. (We again omit the proof that $I(\mathbf{X}; \mathbf{Y})$ represents an achievable amount of communication.) Thus, even for a fixed T, the amount of information that can be covertly communicated is unbounded. ∎

Proposition 3 contradicts Observation 3, which we made earlier. This should be seen as an artifact of Model 2, in particular, of restricting the covertness criterion to the interval $[0, T]$. Because the additive noise $\mathbf{Z}$ has memory, its value on $(-\infty, 0)$ and (T, ∞) can provide information about its values on $[0, T]$, helping the warden detect communication. For example, consider a communication scheme where $X(0) \neq 0$ with a nonzero probability. If the warden observes the entire real line, then it will see a discontinuity in $Y(t)$ at $t = 0$, from which it can immediately determine that communication is taking place. This would not be possible if the warden only had access to $Y(t)$, $t \in [0, T]$. Hence the message of this sub-section is that Model 2 should *not* be adopted.

IV. Discussion

We showed that, over an AWGN channel where the transmitter can employ unbounded bandwidth, covert communication has the same per-second capacity as standard, non-covert communication. This is proven in a continuous-time setting, in the spirit of [9], [10]. For the case where the noise is colored, we provided calculations to suggest that infinitely many nats per second can be communicated covertly, but we have not proven this formally as in the white-noise case, because, in short, we are not aware of properties similar to (15) and (16) for the Karhunen-Loève expansion of colored Gaussian noise.

As we pointed out, one must be careful when formulating the above continuous-time model. In particular, the model should allow the warden to observe not only the time window when communication might take place, but also before and after that time window. This is because our channel has *memory*. The same issue would also arise in discrete-time channels with memory, unless one assumes, for example, that the channel behaves independently before, during, and after the communication window, as in [16], [17].

References

[1] B. A. Bash, D. Goekel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Select. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sept. 2013.

[2] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Proc. IEEE Int. Symp. Inform. Theory*, Istanbul, Turkey, July 10–15 2013.

[3] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Trans. Inform. Theory*, vol. 62, no. 6, pp. 3493–3503, June 2016.

[4] M. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inform. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.

[5] C. E. Shannon, "A mathematical theory of communication," *Bell System Techn. J.*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[6] M. Simon, J. Omura, R. Scholtz, and B. Levitt, *Spread Spectrum Communications Handbook*. McGraw-Hill, 1994.

[7] L. Wang, "The continuous-time Poisson channel has infinite covert communication capacity," in *Proc. IEEE Int. Symp. Inform. Theory*, Vail, CO, USA, June 17–22 2018.

[8] A. Lapidoth, *A Fundation in Digital Communication*, 2nd ed. Cambridge University Press, 2017.

[9] A. D. Wyner, "Capacity of the band-limited Gaussian channel," *Bell System Techn. J.*, vol. 45, pp. 359–395, 1966.

[10] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[11] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, 1981.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: John Wiley & Sons, 2006.

[13] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco: Holden-Day, 1964.

[14] D. Slepian, H. Landau, and H. Pollak, "Prolate spheroidal wave functions, fourier analysis and uncertainty – i & ii," *Bell System Techn. J.*, vol. 40, pp. 43–84, 1961.

[15] D. Slepian, "Some asymptotic expansions of prolate spheroidal wave functions," *J. Math. and Phys.*, vol. 44, pp. 99–140, 1965.

[16] P. H. Che, M. Bakshi, C. Chan, and S. Jaggi, "Reliable deniable communication with channel uncertainty," in *Proc. Inform. Theory Workshop (ITW)*, Hobart, Australia, Nov. 2–5, 2014.

[17] T. Sobers, B. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Trans. Wireless Comm.*, vol. 16, no. 9, pp. 6193–6206, 2017.