On Sampling Continuous-Time Gaussian Channels

Guangyue Han
The University of Hong Kong
email: ghan@hku.hk

Shlomo Shamai
Technion-Israel Institute of Technology
email: sshlomo@ee.technion.ac.il

February 10, 2020

Abstract

For a continuous-time Gaussian channel, it has been shown that as sampling gets infinitesimally fine, the mutual information of the corresponding discrete-time counterparts converges to that of the original continuous-time channel. We give in this paper more quantitative strengthenings of this result, which, among other implications, characterize how over-sampling approaches the true mutual information of a continuous-time Gaussian channel with bandwidth limit. Compared to the Shannon-Nyquist sampling theorem, a widely used tool to connect continuous-time Gaussian channels to their discrete-time counterparts that requires the band-limitedness of the channel input, our results only require some integrability conditions on the power spectral density function of the input.

1 Introduction

Consider the following continuous-time Gaussian channel:

\[ Y(t) = \int_0^t X(s)ds + B(t), \quad t \in [0, T], \]  

where \( T > 0 \), \( \{B(t) : t \in [0, T]\} \) denotes the standard Brownian motion. Throughout the paper, we assume the channel input process \( \{X(t) : t \in [0, T]\} \) is a stationary stochastic process with power spectral density function \( f(\cdot) \). We say that \( t_0, t_1, \ldots, t_n \in \mathbb{R} \) are evenly spaced over \([0, T]\) if \( t_0 = 0, t_n = T \) and \( \delta_{T,n} \equiv t_i - t_{i-1} = T/n \) for all feasible \( i \). Sampling the continuous-time channel (1) with respect to evenly spaced \( t_0, t_1, \ldots, t_n \), we obtain the following discrete-time Gaussian channel

\[ Y(t_i) = \int_0^{t_i} X(s)ds + B(t_i), \quad i = 1, 2, \ldots, n. \]  

It turns out that if the sampling is “fine” enough, the mutual information of the continuous-time Gaussian channel (1) can be “well-approximated” by that of the discrete-time Gaussian channel (2). More precisely, it has been established in [3] that under some mild assumptions,

\[ \lim_{n \to \infty} I(X_0^T ; Y(\Delta T,n)) = I(X_0^T ; Y_0^T), \]  

\[ \overset{1}{\text{Here and hereafter, all } t_i \text{ depend on } T \text{ and } n, \text{ however we suppress this notational dependence for brevity.}} \]
where

\[ \Delta_{T,n} \triangleq \{ t_0, t_1, \ldots, t_n \}, \quad Y(\Delta_{T,n}) \triangleq \{ Y(t_0), Y(t_1), \ldots, Y(t_n) \}. \]

This result connects continuous-time Gaussian channels with their discrete-time counterparts, which has been used to recover a classical information-theoretic formula [1]. The main results of this paper are quantitative strengthenings of (3), which may be used for more quantitative information-theoretic analyses of the channel [1]; in particular, our results characterize how over-sampling approaches the true mutual information of a continuous-time Gaussian channel with bandwidth limit. We remark that the celebrated Shannon-Nyquist sampling theorem [6, 5] is a widely used tool to connect continuous-time Gaussian channels to their discrete-time counterparts, which requires the band-limitedness of the channel input. By comparison, our results only require some integrability conditions on the power spectral density function of the channel input.

## 2 Main Results and Proofs

The following theorem contains the main results of this paper and gives quantitative strengthenings of (3).

**Theorem 2.1.** For the continuous-time Gaussian channel (1), the following two statements hold:

(a) Suppose \( \int f(\lambda)|\lambda|d\lambda < \infty \). Then, for any \( T \) and \( n \),

\[
\sqrt{I(X_0^T; Y_0^T)} \leq \frac{\sqrt{2T \delta_{T,n} \int f(\lambda)|\lambda|d\lambda} + \sqrt{2T \delta_{T,n} \int f(\lambda)|\lambda|d\lambda} + 4I(X_0^T; Y(\Delta_{T,n}))}{2}.
\]

(b) Suppose \( \int f(\lambda)|\lambda|d\lambda < \infty \) and \( \int f(\lambda)d\lambda < \infty \). Then, for any \( T \) and \( n \),

\[
I(X_0^T; Y_0^T) - I(X_0^T; Y(\Delta_{T,n})) \leq T \sqrt{\delta_{T,n}} \left( \int f(\lambda)|\lambda|d\lambda \right)^{1/2} \left( \int f(\lambda)d\lambda \right)^{1/2}.
\]

Consequently, for any \( T \), choosing \( n = n(T) \) such that \( \lim_{T \to \infty} \delta_{T,n(T)} = 0 \), we have

\[
\frac{I(X_0^T; Y_0^T)}{T} - \frac{I(X_0^T; Y(\Delta_{T,n(T)}))}{T} = O \left( \sqrt{\delta_{T,n(T)}} \right),
\]

as \( T \) tends to infinity.

**Proof.** Consider the following parameterized version of the channel (1):

\[ Z(t) = \sqrt{\text{snr}} \int_0^t X(s)ds + B(t), \quad t \in [0, T], \] (4)

where the parameter \( \text{snr} > 0 \) can be regarded as the signal-to-noise ratio of the channel. Obviously, when \( \text{snr} \) is fixed to be 1, \( Z(t) = Y(t) \) for any \( t \in [0, T] \), and moreover, the channel (1) is exactly the same as the channel (1).
Sampling the channel \( \mathcal{H} \) with respect to sampling times \( t_0, t_1, \ldots, t_n \) that are evenly spaced over \([0, T]\), we obtain the following discrete-time Gaussian channel:

\[
Z(t_i) - Z(t_{i-1}) = \sqrt{\text{snr}} \int_{t_{i-1}}^{t_i} X(s) ds + B(t_i) - B(t_{i-1}), \quad i = 1, 2, \ldots, n, \tag{5}
\]

which can be “normalized” as follows:

\[
\frac{Z(t_i) - Z(t_{i-1})}{\sqrt{\delta T,n}} = \sqrt{\text{snr}} \frac{\int_{t_{i-1}}^{t_i} X(s) ds}{\sqrt{\delta T,n}} + \frac{B(t_i) - B(t_{i-1})}{\sqrt{\delta T,n}}, \quad i = 1, 2, \ldots, n, \tag{6}
\]

where, at each time \( i \), the channel noise \( \frac{B(t_i) - B(t_{i-1})}{\sqrt{\delta T,n}} \) is a standard Gaussian random variable, and \( \frac{\int_{t_{i-1}}^{t_i} X(s) ds}{\sqrt{\delta T,n}} \) and \( \frac{Z(t_i) - Z(t_{i-1})}{\sqrt{\delta T,n}} \) should be regarded as the channel input and output, respectively.

By the continuous-time I-MMSE relationship \( \mathcal{I} \) applied to the channel \( \mathcal{H} \), the mutual information of the channel \( \mathcal{H} \) can be computed as

\[
I(X_0^T; Y_0^T) = \frac{1}{2} \int_0^1 \int_0^T \mathbb{E}[(X(s) - \mathbb{E}[X(s)|Z_0^T])^2] ds \text{dnr}.
\]

And by the discrete-time I-MMSE relationship \( \mathcal{I} \) (or more precisely, its extension \( \mathcal{I} \) to Gaussian memory channels) applied to the channel \( \mathcal{H} \), we have

\[
I(X_0^T; Y(\Delta T,n)) = \frac{1}{2} \int_0^1 \sum_{i=1}^n \mathbb{E}\left[ \left( \frac{\int_{t_{i-1}}^{t_i} X(s) ds}{\sqrt{\delta T,n}} - \mathbb{E}\left[ \frac{\int_{t_{i-1}}^{t_i} X(s) ds}{\sqrt{\delta T,n}} \mid Z(\Delta T,n) \right] \right)^2 \right] d\text{snr}.
\]

It is obvious that \( I(X_0^T; Y_0^T) \geq I(X_0^T; Y(\Delta T,n)) \). In the following, we will give an upper bound on their difference \( I(X_0^T; Y_0^T) - I(X_0^T; Y(\Delta T,n)) \), thereby characterizing the closeness between the two quantities. Towards this goal, using the following easily verified that for each \( i \),

\[
\mathbb{E}\left[ \left( \int_{t_{i-1}}^{t_i} X(s) - \mathbb{E}[X(s)|Z_0^T] \right)^2 \right] \leq \mathbb{E}\left[ \left( \int_{t_{i-1}}^{t_i} X(s) - \mathbb{E}[X(s)|Z(\Delta T,n)] \right)^2 \right],
\]

we first note that

\[
I(X_0^T; Y_0^T) - I(X_0^T; Y(\Delta T,n))
\]

\[
= \frac{1}{2} \int_0^1 \int_0^T \mathbb{E}[(X(s) - \mathbb{E}[X(s)|Z_0^T])^2] ds - \sum_{i=1}^n \mathbb{E}\left[ \left( \int_{t_{i-1}}^{t_i} X(s) - \mathbb{E}[X(s)|Z(\Delta T,n)] \right)^2 \right] d\text{snr}
\]

\[
\leq \frac{1}{2} \int_0^1 \int_0^T \mathbb{E}[(X(s) - \mathbb{E}[X(s)|Z_0^T])^2] ds - \sum_{i=1}^n \mathbb{E}\left[ \left( \int_{t_{i-1}}^{t_i} X(s) - \mathbb{E}[X(s)|Z_0^T] \right)^2 \right] d\text{snr}
\]

\[
\leq \frac{1}{2} \int_0^1 \int_0^T \mathbb{E}[R^2(X(s); Z_0^T)] ds - \sum_{i=1}^n \mathbb{E}\left[ \left( \int_{t_{i-1}}^{t_i} R[X(s); Z_0^T] \right)^2 \right] d\text{snr}
\]

\[
= \frac{1}{2}(S_1 + S_2)
\]
where we have used the shorthand notation \( R[X(s); Z_0^T] \) for \( X(s) - \mathbb{E}[X(s)|Z_0^T] \) and

\[
S_1 \triangleq \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \mathbb{E}[R^2[X(s); Z_0^T]] ds - \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \mathbb{E}[R[X(s); Z_0^T]R[X(t_{i-1}); Z_0^T]] ds d\text{snr},
\]

\[
S_2 \triangleq \sum_{i=1}^{n} \int_{0}^{1} \mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i} R[X(s); Z_0^T] ds\right) R[X(t_{i-1}); Z_0^T]\right] - \sum_{i=1}^{n} \mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i} \frac{R[X(s); Z_0^T]}{\sqrt{\delta_{T,n}}} ds\right)^2\right] d\text{snr}.
\]

For the first term, we have

\[
S_1^2 = \left(\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s); Z_0^T]] ds d\text{snr} - \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R[X(s); Z_0^T]R[X(t_{i-1}); Z_0^T]] d\text{snr} ds\right)^2
\]

\[
= \left(\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R[X(s); Z_0^T]R[X(s) - X(t_{i-1}); Z_0^T]] d\text{snr} ds\right)^2
\]

\[
\leq n \sum_{i=1}^{n} \left(\int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R[X(s); Z_0^T]R[X(s) - X(t_{i-1}); Z_0^T]] d\text{snr} ds\right)^2
\]

\[
\leq n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s); Z_0^T]] ds d\text{snr} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s) - X(t_{i-1}); Z_0^T]] ds d\text{snr}, \quad (7)
\]

where we have used the Cauchy-Scharz inequality for the last inequality. Now, noticing the fact that

\[
\mathbb{E}[R^2[X(s) - X(t_{i-1}); Z_0^T]] \leq \mathbb{E}[(X(s) - X(t_{i-1}))^2],
\]

we continue as follows:

\[
S_1^2 \leq n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s); Z_0^T]] ds d\text{snr} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[(X(s) - X(t_{i-1}))^2] ds d\text{snr} \quad (8)
\]

\[
= n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s); Z_0^T]] ds d\text{snr} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[X^2(s) + X^2(t_{i-1}) - 2X(s)X(t_{i-1})] ds d\text{snr}. \quad (9)
\]

Now, using the fact that, for any \( u, v \in \mathbb{R} \),

\[
\mathbb{E}[X(u)X(u + v)] = \int f(\lambda)e^{iv\lambda} d\lambda,
\]
we have

\[
 S_1^2 \leq n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s); Z_0^T]] ds dr dsnr ds \int_{t_{i-1}}^{t_i} \int_{0}^{1} \int_{0}^{1} 2f(\lambda) |1 - e^{i\lambda(s-t_{i-1})}| d\lambda dsnr ds \tag{10}
\]

\[
 \leq n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E}[R^2[X(s); Z_0^T]] ds dr dsnr ds \int_{t_{i-1}}^{t_i} \int_{0}^{1} \int_{0}^{1} 2f(\lambda) |s-t_{i-1}| d\lambda dsnr ds \tag{11}
\]

\[
 \leq n \delta^2_{T,n} \int f(\lambda) |\lambda| d\lambda \int_{0}^{1} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \mathbb{E}[R^2[X(s); Z_0^T]] ds dr dsnr \tag{12}
\]

\[
 = T \delta_{T,n} \int f(\lambda) |\lambda| d\lambda \int_{0}^{1} \int_{0}^{T} \mathbb{E}[(X(s) - \mathbb{E}[X(s)|Z_0^T])^2] ds dr dsnr \tag{13}
\]

\[
 = 2T \delta_{T,n} I(X^T_0; Y^T_0) \int f(\lambda) |\lambda| d\lambda. \tag{14}
\]

For the second term, we have

\[
 S_2^2 = \left( \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \int_{0}^{1} \mathbb{E} \left[ R[X(s); Z_0^T] \int_{t_{i-1}}^{t_i} \frac{R[X(t_{i-1}) - X(u); Z_0^T]}{\delta_{T,n}} du \right] ds nr dsnr \right)^2.
\]

Starting from this and proceeding in a similar fashion as in (7)-(14), we derive

\[
 S_2^2 \leq 2T \delta_{T,n} I(X^T_0; Y^T_0) \int f(\lambda) |\lambda| d\lambda.
\]

Combining the bounds on $S_1$ and $S_2$, we have

\[
 |I(X^T_0; Y^T_0) - I(X^T_0; Y(\Delta_{T,n}))| \leq \sqrt{2T \delta_{T,n} I(X^T_0; Y^T_0) \int f(\lambda) |\lambda| d\lambda} \frac{1}{2}. \tag{15}
\]

It then immediately follows that

\[
 \sqrt{I(X^T_0; Y^T_0)} \leq \sqrt{2T \delta_{T,n} \int f(\lambda) |\lambda| d\lambda} + \sqrt{2T \delta_{T,n} \int f(\lambda) |\lambda| d\lambda} + 4I(X^T_0; Y^T_{t_0}) \frac{1}{2},
\]

establishing (a). Moreover, together with the fact that

\[
 I(X^T_0; Y^T_0) \leq \frac{1}{2} \int_{0}^{T} \mathbb{E}[X^2(s)] ds = \frac{T}{2} \int f(\lambda) d\lambda,
\]

the inequality \eqref{eq:15} implies that

\[
 |I(X^T_0; Y^T_0) - I(X^T_0; Y(\Delta_{T,n}))| \leq T \sqrt{\delta_{T,n} \left( \int f(\lambda) |\lambda| d\lambda \right)^{1/2} \left( \int f(\lambda) d\lambda \right)^{1/2}},
\]

establishing (b). \qed
The following corollary, which immediately follows from Theorem 2.2, characterizes, among others, how over-sampling approaches the true mutual information of the channel with bandwidth limit.

**Corollary 2.2.** For the continuous-time Gaussian channel, suppose that the channel input has bandwidth limit $W$ and average power $P$, more precisely, $f(\lambda) = 0$ for any $\lambda \in (-\infty, -W] \cup [W, \infty)$ and $E[X^2(s)] \leq P$ for any $s \geq 0$. Then, the following two statements hold:

(a) For any $T$ and $n$,

$$\sqrt{I(X_0^T;Y_0^T)} \leq \sqrt{2TWP\delta_{T,n}} + \sqrt{2TWP\delta_{T,n}} + 4I(X_0^T;Y(\Delta_{T,n}))$$

(b) For any $T$ and $n$,

$$I(X_0^T;Y_0^T) - I(X_0^T;Y(\Delta_{T,n})) \leq TP\sqrt{W\delta_{T,n}}.$$ 

Consequently, for each $T$, choosing $n = n(T)$ such that $\lim_{T \to \infty} \delta_{T,n(T)} = 0$, we have

$$\frac{I(X_0^T;Y_0^T)}{T} - \frac{I(X_0^T;Y(\Delta_{T,n}))}{T} = O\left(\sqrt{\delta_{T,n(T)}}\right),$$

as $T$ tends to infinity.

**References**

[1] D. Guo, S. Shamai, and S. Verdu. Mutual information and minimum mean-square error in Gaussian channels. *IEEE Trans. Info. Theory*, vol. 51, no. 4, pp. 1261-1282, 2005.

[2] G. Han and J. Song. Extensions of the I-MMSE relationship to Gaussian channels with feedback and memory. *IEEE Trans. Info. Theory*, vol. 62, no. 10, pp. 5422-5445, 2016.

[3] Liu and G. Han. On continuous-time Gaussian channels. *Entropy*, vol. 21, no. 1:67, 2019.

[4] X. Liu, R. Bustin, G. Han and S. Shamai. An elementary proof of a classical information-theoretic formula. *IEEE ISIT*, 2019.

[5] H. Nyquist. Certain factors affecting telegraph speed. *The Bell System Technical Journal*, pp. 324, 1924.

[6] C. Shannon. Communication in the presence of noise. *Proc. IRE*, vol. 37, pp. 10-21, 1949.