How Compressible are Innovation Processes?

Hamid Ghourchian, Arash Amini, *Senior Member, IEEE*, and Amin Gohari, *Senior Member, IEEE*

***Abstract*—The sparsity and compressibility of finite-dimensional signals are of great interest in fields such as compressed sensing. The notion of compressibility is also extended to infinite sequences of i.i.d. or ergodic random variables based on the observed error in their nonlinear k-term approximation. In this work, we use the entropy measure to study the compressibility of continuous-domain innovation processes (alternatively known as white noise). Specifically, we define such a measure as the entropy limit of the doubly quantized (time and amplitude) process. This provides a tool to compare the compressibility of various innovation processes. It also allows us to identify an analogue of the concept of "entropy dimension" which was originally defined by Rényi for random variables. Particular attention is given to stable and impulsive Poisson innovation processes. Here, our results recognize Poisson innovations as the more compressible ones with an entropy measure far below that of stable innovations. While this result departs from the previous knowledge regarding the compressibility of impulsive Poisson laws compared to continuous fat-tailed distributions, our entropy measure ranks α-stable innovations according to their tail.**

***Index Terms*—Compressibility, entropy, impulsive Poisson process, stable innovation, white Lévy noise.**

I. Introduction

The compressible signal models have been extensively used to represent or approximate various types of data such as audio, image and video signals. The concept of compressibility has been separately studied in information theory and signal processing societies. In information theory, this concept is usually studied via the well-known entropy measure and its variants. For instance, the notion of *entropy dimension* was introduced in [1] for continuous random variables based on the concept of differential entropy. The entropy dimension was later studied for discrete-domain random processes in [2] and [3] in the context of compressed sensing. In signal processing, a signal is intuitively called compressible if in its representation using a known dictionary only a few atoms contribute significantly and the rest amount to negligible contribution. Sparse signals are among the special cases for which the mentioned representation consists of a few non-zero (instead of insignificant) contributions. Compressible signals in general, and sparse signals in particular, are of fundamental importance in fields such as compressed sensing [4], [5], dimensionality reduction [6], and nonlinear approximation theory [7].

In this work, we consider the compressibility of continuous-domain innovation processes which were originally studied in the context of signal processing [8], [9]. However, we try to apply information theoretic tools to measure the compressibility. To cover the existing literature in both parts, we begin by the signal processing perspective of compressibility.

Traditionally, compressible signals are defined as infinite sequences within the Besov spaces [7], where the decay rate of the k-term approximation error could be efficiently controlled. A more recent deterministic approach towards modeling compressibility is via weak-ℓ_p spaces [5], [10]. The latter approach is useful in compressed sensing, where ℓ_1 (or ℓ_p) regularization techniques are used.

The study of stochastic models for compressibility started with identifying compressible priors. For this purpose, independent and identically distributed (i.i.d.) sequences of random variables with a given probability law are examined. Cevher in [11] defined the compressibility criterion based on the decay rate of the mean values of order statistics. A more precise definition in [12] revealed the connection between compressibility and the decay rate of the tail probability. In particular, heavy-tailed priors with infinite p-order moments were identified as ℓ_p-compressible probability laws. It was later shown in [13] that this sufficient condition is indeed, necessary as well. A similar identification of heavy-tailed priors (with infinite variance) was obtained in [14] with a different definition of compressibility[1].

The first non-i.i.d. result appeared in [15]. By extending the techniques used in [13], and based on the notion of ℓ_p-compressibility of [12], it is shown in [15] that discrete-domain stationary and ergodic processes are ℓ_p-compressible if and only if the invariant distribution of the process is an ℓ_p-compressible prior.

The recent framework of sparse stochastic processes introduced in [8], [9], [16] extends the discrete-domain models to continuous-domain. In practice, most of the compressible discrete-domain signals arise from discretized versions of continuous-domain physical phenomena. Thus, it might be beneficial to have continuous-domain models that result in compressible/sparse discrete-domain models for a general class of sampling strategies. Indeed, this goal is achieved in [8], [9] by considering non-Gaussian stochastic processes. The building block of these models are the innovation processes (widely known as white noise) that mimic i.i.d. sequences in continuous-domain. Unlike sequences, the probability laws of innovation processes are bound to a specific family known as *infinitely divisible* that includes α-stable distributions ($\alpha = 2$ corresponds to Gaussians).

The authors are with the department of Electrical Engineering and Advanced Communication Research Institute (ACRI), Sharif University of Technology, Tehran, Iran. (e-mails: h_ghourchian@ee.sharif.edu, {aamini, aminzadeh}@sharif.ir)

[1]Based on the definition in [14], a prior is called compressible if the compressed measurements (by applying a Gaussian ensemble) of a high-dimensional vector of i.i.d. values following this law could be better recovered using ℓ_1 regularization techniques than the classical ℓ_2 minimization approaches.

The discretization of innovation processes are known to form stationary and ergodic sequences of random variables with infinitely divisible distributions. As the tail probability of all non-Gaussian infinitely divisible laws are slower than Gaussians [17], [18], they exhibit more compressible behavior than Gaussians according to [12], [15].

In this paper, we investigate the compressibility of continuous-domain stochastic processes using the quantization entropy. As a starting point, we restrict our attention in the present work to innovation processes as the building blocks of more general stochastic processes. We postpone the evaluation of the quantization entropy for more general processes to future works.

In information theory, entropy naturally arises as the proper measure of compression for Shannon's lossless source coding problem. It also finds a geometrical interpretation as the volume of the typical sets. As the definition of entropy ignores the amplitude distribution of the involved random variables and only takes into account the distribution (or concentration) of the probability measure, it provides a fundamentally different perspective of compressibility compared to the previously studied k-term approximation. More precisely, the entropy reveals a universal compressibility measure that is not limited to a specific measurement technique, while the k-term approximation is tightly linked with the linear sampling strategy of compressed sensing. Hence, it is not surprising that our results based on the quantization entropy show that impulsive Poisson innovation processes are by far more compressible than heavy-tailed α-stable innovation processes; the previous studies on their k-term approximation sort them in the opposite order when the jump distribution in the Poisson innovation is not heavy-tailed. It is interesting to mention that the same ordering of impulsive Poisson and heavy-tailed -stable innovation processes is observed in [19].

The two main challenges that are addressed in this paper are 1) defining a quantization entropy for continuous-domain innovation processes that translates into operational lossless source coding, and 2) evaluating such a measure for particular instances to allow for their comparison. We recall that the *differential entropy* of a random variable X with continuous range is defined by finely quantizing X with resolution $1/m$, followed by canceling a diverging term $\log(m)$ from the discrete entropy of its quantized version. Obviously, we shall expect more elaborate diverging terms when dealing with continuous-domain processes. More specifically, after appropriate quantization in time and amplitude with resolutions $1/n$ and $1/m$ respectively, we propose the one of the following two expressions to cancel out the diverging terms:

$$\frac{\mathcal{H}_{m,n}(X)}{\kappa(n)} - \log(m) - \zeta(n), \tag{1}$$

or

$$\mathcal{H}_{m,n}(X) - \zeta(n), \tag{2}$$

where $\mathcal{H}_{m,n}(X)$ is the discrete entropy of the time/amplitude quantized process, and $\kappa(\cdot)$ and $\zeta(\cdot)$ are univariate functions. We prove that depending on the white noise process, (1) or (2) give the correct way to cancel out the diverging terms for a wide class of white Lévy noises with a suitable choice of $\kappa(n)$ and $\zeta(n)$. We may view $\kappa(n)/n$ as the analogue of entropy dimension for a white noise process. A general expression for $\kappa(n)$ is given. However, while we prove existence of a function $\zeta(n)$, we are able to provide its explicit expression only for special cases of stable and Poisson white noise processes. While the term $\log(m)$ is reminiscent of the amplitude quantization effect, functions $\kappa(\cdot)$ and $\zeta(\cdot)$ quantify the compressibility of a given Lévy process: the higher the growth rate of $\kappa(n)$, the less compressible the process. If two processes have the same growth rate of $\kappa(n)$, then, the $\zeta(n)$ with the smaller growth rate is the more compressible process.

Finally, ϵ-metric entropy of [20] and ϵ-entropy of [21] could be considered as other alternatives for quantifying compressibility of stochastic processes. The interested reader may refer to [22] for a general discussion of defining relative entropy for stochastic processes.

The organization of the paper is as follows. We begin by reviewing the preliminaries, including some of the basic definitions and results regarding differential entropy and white Lévy noises in Section II. Next, we present our main contributions in Section III wherein we study the quantization entropy of a wide class of white Lévy noise processes. Furthermore, special attention is given to the stable and impulsive Poisson innovation processes. To facilitate reading of the paper, we have separated the results from their proofs. The main body of proofs are postponed to Section IV. Some of the key propositions and lemmas for obtaining the final claims are stated in Appendix A, while their proofs are provided in Appendix B.

II. Preliminaries

The goal of this paper is to define a quantization entropy for certain random processes. Hence, we first review the concept of entropy for random variables. For this purpose, we provide the definition of entropy for three main types of probability distributions. This is followed by the definition of innovation processes (white Lévy noises) and in particular, the stable and Poisson white noise processes.

All the logarithms in this paper are in base e. In Table I, we summarize the notation used in this paper.

A. Types of Random variables

The main types of random variables considered in this paper are discrete, continuous, and discrete-continuous, which are defined below.

Definition 1 (Continuous Random Variables). *[23] Let $\mathcal{B}$ be the Borel σ-field of $\mathbb{R}$ and let X be a real-valued random variable with distribution (cdf) $F(x)$ that is measurable with respect to $\mathcal{B}$. We call X a* continuous *random variable if its probability measure μ, induced on $(\mathbb{R}, \mathcal{B})$, is absolutely continuous with respect to the Lebesgue measure for $\mathcal{B}$ (i.e., $\mu(\mathcal{A}) = 0$ for all $\mathcal{A} \in \mathcal{B}$ with zero Lebesgue measure). We denote the set of all absolutely continuous distributions by $\mathcal{AC}$.*

It is a well-known result that X being a continuous random variable is equivalent to the fact that the cdf $F(x)$ is an

Symbol	Definition
Sets	Caligraphic letters like $\mathcal{A}, \mathcal{C}, \mathcal{D}, \ldots$
Real and natural numbers	$\mathbb{R}$, $\mathbb{N}$
Borel sets in $\mathbb{R}$	$\mathcal{B}(\mathbb{R})$ or just $\mathcal{B}$
Random variables	Capital letters like: $A, X, Y, Z, \ldots$
Probability density function (pdf) of (continous) X	p_X or q_X (lowercase letter p)
Probability mass function (pmf) of (discrete) X	P_X (uppercase letter P)
Cumulative distribution function (cdf) of X	F_X
X_0 for a given white noise $X(t)$	$\int_0^1 X(t)\,\mathrm{d}t$; a random variable
$\frac{1}{n}$	step-size for time quantization (time sampling)
$\frac{1}{m}$	step-size for amplitude quantization
$X_i^{(n)}$ for a given white noise $X(t)$	$\int_{\frac{i-1}{n}}^{\frac{i}{n}} X(t)\,\mathrm{d}t$; a random variable
$X_1, X_2, \cdots$	A sequence of random variables Distinguished from $X_i^{(n)}$ by lack of (n) superscript.
$[x]_m$ $\left[X_i^{(n)}\right]_m$	Quantization of the value of x with step size $1/m$; Quantization of $X_i^{(n)}$ with step size $1/m$
Discrete random variable associated to a discrete-continuous random variable X	X_D
Continuous random variable associated to a discrete-continuous random variable X	X_c
Entropy	H
Differential entropy	h
$\mathcal{H}_{m,n}(X)$	Quantization entropy associated to a white noise $X(t)$ in (6) and (7)
Characteristic function	$\widehat{p}$

TABLE I: Summary of the notation.

absolutely continuous function. The Radon-Nikodym theorem implies that for each $X \sim F \in \mathcal{AC}$ there exists a $\mathcal{B}$-measurable function $p : \mathbb{R} \to [0, \infty)$, such that for all $\mathcal{A} \in \mathcal{B}$ we have that

$$\Pr\{X \in \mathcal{A}\} = \int_{\mathcal{A}} p(x)\,\mathrm{d}x.$$

The function p is called the probability density function (pdf) of X. The property $F \in \mathcal{AC}$ is alternatively written as $p \in \mathcal{AC}$. [22, p. 21]

Definition 2 (Discrete Random Variable)**.** *[23] A random variable X is called discrete if it takes values in a countable alphabet set $\mathcal{X} \subset \mathbb{R}$. We can assign a probability mass function to the discrete random variable X. Probability mass functions are denoted in uppercase letters as $P_X[x]$ to distinguish them from probability density functions, which are shown in lowercase letters as $p_X(x)$.*

Definition 3 (Discrete-Continuous Random Variable)**.** *[23] A random variable X is called discrete-continuous with parameters $(p_c, P_D, \Pr\{X \in \mathcal{D}\})$ if there exists a countable set $\mathcal{D}$, a discrete probability mass function P_D, whose support is $\mathcal{D}$, and a pdf $p_c \in \mathcal{AC}$ such that*

$$0 < \Pr\{X \in \mathcal{D}\} < 1,$$

as well as for every Borel set $\mathcal{A}$ in $\mathbb{R}$ we have that

$$\Pr\{X \in \mathcal{A} | X \notin \mathcal{D}\} = \int_{\mathcal{A}} p_c(x)\,\mathrm{d}x,$$
$$\Pr\{X \in \mathcal{A} | X \in \mathcal{D}\} = \sum_{x \in \mathcal{D} \cap \mathcal{A}} P_D[x].$$

It is clear that we can write the pdf of a discrete-continuous random variable X, p_X, as follows:

$$p_X(x) = \Pr\{X \in \mathcal{D}\} P_d(x) + (1 - \Pr\{X \in \mathcal{D}\})\, p_c(x),$$

where $p_c \in \mathcal{AC}$, and P_d is the probability density function of the discrete part, which is a convex combination of Dirac's delta functions.

In this paper, the probability mass function of discrete random variables is denoted by capital letters like P and Q, while the probability density function of continuous or discrete-continuous random variables is denoted by lowercase letters like p and q.

B. *Definition of Entropy*

We first define the entropy and differential entropy for discrete and continuous random variables, respectively. Next, we define entropy dimension for discrete-continuous random variables via amplitude quantization.

Definition 4 (Entropy and Differential Entropy)**.** *[24, Chapter 2] We define entropy $\mathrm{H}(X)$, or $\mathrm{H}(P_X)$ for a discrete random variable X with probability mass function (pmf) $P_X[x]$ as*

$$\mathrm{H}(P_X) = \mathrm{H}(X) := \sum_x P_X[x] \log \tfrac{1}{P_X[x]},$$

if the summation converges. For a continuous random variable X with pdf $p_X(x) \in \mathcal{AC}$, we define differential entropy $\mathrm{h}(X)$, or $\mathrm{h}(p_X)$ as

$$\mathrm{h}(p_X) = \mathrm{h}(X) := \int_{\mathbb{R}} p_X(x) \log \tfrac{1}{p_X(x)}\,\mathrm{d}x,$$

if

$$\int_{\mathbb{R}} p_X(x) \left|\log \tfrac{1}{p_X(x)}\right| \,\mathrm{d}x < \infty.$$

Similarly, for a discrete or continuous random vector $\mathbf{X}$*, the entropy or the differential entropy is defined as*

$$\mathrm{H}(\mathbf{X}) = \mathbb{E}\left[\log \frac{1}{P_{\mathbf{X}}[\mathbf{X}]}\right], \quad \textit{or} \quad \mathrm{h}(\mathbf{X}) = \mathbb{E}\left[\log \frac{1}{p_{\mathbf{X}}(\mathbf{X})}\right],$$

where $P_{\mathbf{X}}[\mathbf{x}]$ *(*$p_{\mathbf{X}}(\mathbf{x})$*) is the pmf (pdf) of the random vector* $\mathbf{X}$*, respectively.*

In brief, we say that the (differential) entropy is well-defined if the corresponding (integral) summation is convergent to a finite value.

Next, we identify a class of absolutely continuous probability distributions, and show that differential entropy is uniformly convergent over this space under the total variation distance metric.

Definition 5. *[25] Given* $\alpha, \ell, v \in (0, \infty)$*, we define* (α, ℓ, v)*–*$\mathcal{AC}$ *to be the class of all* $p \in \mathcal{AC}$ *such that the corresponding density function* $p : \mathbb{R} \mapsto [0, \infty)$ *satisfies*

$$\int_{\mathbb{R}} |x|^{\alpha} p(x) \,\mathrm{d}x \leq v,$$
$$\operatorname*{ess\,sup}_{x \in \mathbb{R}} \; p(x) \leq \ell.$$

Lemma 1. *[25] The differential entropy of any distribution in* (α, ℓ, v)*–*$\mathcal{AC}$ *is well-defined, and for all* $p_X, p_Y \in (\alpha, \ell, v)$*–*$\mathcal{AC}$ *satisfying* $\|p_X - p_Y\|_1 \leq m$*, we have that*

$$|\mathrm{h}(p_X) - \mathrm{h}(p_Y)| \leq c_1 D_{X,Y} + c_2 D_{X,Y} \log \tfrac{1}{D_{X,Y}},$$

where

$$D_{X,Y} = \|p_X - p_Y\|_1 := \int_{\mathbb{R}} |p_X(x) - p_Y(x)| \,\mathrm{d}x,$$

and

$$\begin{aligned} c_1 &= \tfrac{1}{\alpha} |\log(2\alpha v)| + |\log(\ell \mathrm{e})| + \log \tfrac{\mathrm{e}}{2} \\ &\quad + \log\left[2\Gamma\left(1 + \tfrac{1}{\alpha}\right)\right] + \tfrac{1}{\alpha} + 1, \\ c_2 &= \tfrac{1}{\alpha} + 2. \end{aligned}$$

Now, we define the quantization of a random variable in amplitude domain.

Definition 6 (Quantization of Random Variables)**.** *The quantized version of a random variable* X *with the step size* $1/m$ *(for* $m > 0$*) is defined as*

$$[X]_m = \tfrac{1}{m} \left\lfloor \tfrac{1}{2} + mX \right\rfloor.$$

Thus, $[X]_m$ *has the pmf* $P_{X;m}$ *given by*

$$P_{X;m}[i] := \Pr\left\{[X]_m = \tfrac{i}{m}\right\} = \int_{\left[\frac{i-0.5}{m}, \frac{i+0.5}{m}\right)} p_X(x) \,\mathrm{d}x.$$

Also, we define a continuous random variable $\widetilde{X}_m$ *with pdf* $q_{X;m} \in \mathcal{AC}$ *as follows*

$$q_{X;m}(x) = m\, P_{X;m}[i], \qquad x \in \left[\tfrac{i-0.5}{m}, \tfrac{i+0.5}{m}\right).$$

We state a useful lemma about the entropy of quantized random variables here:

Lemma 2. *[1] Let* $X \sim p_X(x)$ *be a continuous random variable. Then,*

$$\mathrm{H}\left([X]_m\right) - \log m = \mathrm{h}\left(\widetilde{X}_m\right),$$

where $[X]_m$ *is the quantization of* X *with step size* $1/m$*, and* $\widetilde{X}_m \sim q_{X;m}$ *is the random variable defined in Definition 6.*

The following lemma, proved in [1], measures the entropy of quantized discrete-continuous random variables by defining *entropy dimension*.

Lemma 3. *[1] Let* X *be a discrete-continuous random variable defined by the triplet* $(p_c, P_D, \Pr\{X \in \mathcal{D}\})$*. Let* X_D *be a discrete random variable with pmf* P_D *and* X_c *be a continuous random variable with pdf* p_c*. If* $\mathrm{H}([X]_1) < \infty$*,* $\mathrm{H}(X_D) < \infty$*, and* $\mathrm{h}(X_c)$ *is well-defined and finite (i.e.,* $\mathbb{E}[|\log(p_{X_c})|] < \infty$*), where* $[X]_1$ *is the quantized version of* X *with step size* 1*, then the entropy of quantized* X *with step size* $1/m$ *can be expressed as*

$$\begin{aligned} \mathrm{H}\left([X]_m\right) = d \log m + d\mathrm{h}\left(X_c\right) + (1-d)\mathrm{H}\left(X_D\right) \\ + \mathrm{H}_2(d) + o_m(1), \end{aligned}$$

where $o_m(1)$ *vanishes as* m *tends to* ∞*, and*

$$\begin{aligned} d &= \Pr\{X \notin \mathcal{D}\}, \\ h &= d\mathrm{h}(X_c) + (1-d)\mathrm{H}(X_D) + \mathrm{H}_2(d), \\ \mathrm{H}_2(d) &\triangleq d \log 1/d + (1-d) \log[1/(1-d)]. \end{aligned}$$

The variable d *is called the* Entropy Dimension *of* X*. The lemma is true for the discrete and continuous case with* $d = \mathrm{h}(X_c) = \mathrm{H}_2(d) = 0$ *and* $d = 1, \mathrm{H}(X_D) = \mathrm{H}_2(d) = 0$*, respectively.*

Corollary 1. *[1] According to Lemma 3, if a random variable* X *is continuous with pdf* p_c*, then, the entropy of quantized* X *with step size* $1/m$ *is*

$$\mathrm{H}\left([X]_m\right) = \log m + \mathrm{h}(p_c) + o_m(1),$$

provided that

$$\mathrm{H}\left([X]_1\right) < \infty, \qquad \int_{\mathbb{R}} p_c(x) \left|\log \tfrac{1}{p_c(x)}\right| \,\mathrm{d}x < \infty,$$

where $\mathrm{h}(p_c)$ *is the differential entropy of* X*.*

C. White Lévy Noises

The family of white Lévy noises were originally defined by Gelfand and Vilenkin [26, Chapter 4] (with the name *Generalized Random Processes with Independent Values at Every Point*). To introduce this family, we first define the concept of a *generalized function* that generalizes ordinary functions to include distributions such as Dirac's delta functional and its derivatives [16].

Definition 7 (Schwartz Space)**.** *[16, p. 30] The* Schwartz *space, denoted as* $\mathcal{S}(\mathbb{R})$*, consists of infinitely differentiable functions* $\phi : \mathbb{R} \mapsto \mathbb{R}$*, for which*

$$\sup_{t \in \mathbb{R}} \left| t^m \frac{d^n}{dt^n} \phi(t) \right| < \infty, \qquad \forall m, n \in \mathbb{N}.$$

In other words, $\mathcal{S}(\mathbb{R})$ is the class of smooth functions that, together with all of their derivatives, decay faster than the inverse of any polynomial at infinity.

The space of *tempered distributions* or alternatively, the continuous dual of the Schwartz space denoted by $\mathcal{S}'(\mathbb{R})$, is the set of all continuous linear mappings from $\mathcal{S}(\mathbb{R})$ into $\mathbb{R}$ (also known as generalized functions). In other words, for all $x \in \mathcal{S}'(\mathbb{R})$ and $\varphi \in \mathcal{S}(\mathbb{R})$, $x(\varphi)$ is a well-defined real number. Due to the linearity of the mapping with respect to φ, the following notations are interchangeably used:

$$x(\varphi) = \langle x, \varphi \rangle = \int_{\mathbb{R}} x(t)\varphi(t)\,\mathrm{d}t,$$

where $\langle x, \varphi \rangle$, $x(t)$ and the integral on the right-hand side are merely notations. This formalism is useful because it allows for a precise mathematical definition of generalized functions, such as impulse function that are common in engineering textbooks.

Just as generalized functions extend ordinary functions, generalized stochastic processes extend ordinary stochastic processes. In particular, a generalized stochastic process is a probability measure on $\mathcal{S}'(\mathbb{R})$. Further, observing a generalized stochastic process $X(\cdot)$ is done by applying its realizations to Schwartz functions; *i.e.*, for a given $\varphi \in \mathcal{S}(\mathbb{R})$, $X_\varphi = \langle X, \varphi \rangle = \int_{\mathbb{R}} X(t)\varphi(t)\,\mathrm{d}t$ represents a real-valued random variable with

$$\underset{\mathcal{I}\in\mathcal{B}(\mathbb{R})}{\Pr}\{X_\varphi \in \mathcal{I}\} = \mu\big(\{x \in \mathcal{S}'(\mathbb{R}) \,\big|\, \langle x, \varphi \rangle \in \mathcal{I}\}\big),$$

where $\mu(\cdot)$ stands for the probability measure on $\mathcal{S}'(\mathbb{R})$ that defines the generalized stochastic process.

The white Lévy noises are a subclass of generalized stochastic processes with certain properties. Before we introduce them, we define Lévy exponents:

Definition 8 (Lévy Exponent)**.** *[16, p. 59] A function $f : \mathbb{R} \to \mathbb{C}$ is called a Lévy exponent if*

1) *$f(0) = 0$,*
2) *f is continuous at 0,*
3) *$\forall n \in \mathbb{N}, \forall \boldsymbol{\omega} \in \mathbb{R}^n$, and $\forall \mathbf{a} \in \mathbb{C}^n$ satisfying $\sum_{i=1}^n a_i = 0$, we have that*

$$\sum_{i,j=1}^{n} a_i a_j^* f(\omega_i - \omega_j) \geq 0.$$

The following lemma provides the algebraic characterization of Lévy exponents.

Lemma 4 (Lévy-Khintchin)**.** *[16, p. 61] A function $f(\omega)$ is a Lévy exponent if and only if it can be written as*

$$-\frac{\sigma^2}{2}\omega^2 + \mathrm{j}\mu\omega + \int_{\mathbb{R}\setminus\{0\}} \left(\mathrm{e}^{\mathrm{j}\omega a} - 1 - \mathrm{j}\omega a \mathbb{1}_{(-1,1)}(a)\right)\,\mathrm{d}V(a),$$

where $\sigma, \mu \in \mathbb{R}$ are arbitrary constants. The function $\mathbb{1}_{(-1,1)}(a)$ is an indicator function which is 1 when $|a| < 1$ and is 0 otherwise. The function $V(x)$ is a non-negative non-decreasing function that is continuous at $a = 0$ and satisfies

$$\lim_{a\to-\infty} V(a) = 0,$$
$$\int_{\mathbb{R}\setminus\{0\}} \min\{1, a^2\}\,\mathrm{d}V(a) < \infty.$$

Definition 9 (White Lévy Noises)**.** *[16, Sec. 4.4] A generalized stochastic process X is called a white Lévy noise, if*

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{j}\langle X,\varphi\rangle}\right] = \exp\left(\int_{\mathbb{R}} f\big(\varphi(t)\big)\,\mathrm{d}t\right), \quad \forall\varphi \in \mathcal{S}(\mathbb{R}),$$

where $\langle X, \varphi \rangle$ is the output of the linear operator φ under X, and f is a valid Lévy exponent for which

$$\int_{\mathbb{R}\setminus[-1,1]} |a|^\theta\,\mathrm{d}V(a) < \infty,$$

for some $\theta > 0$.

The desired properties of a white noise could be inferred from Definition 9:

Lemma 5. *[18] A white Lévy noise X is a stationary process in the sense that $\langle X, \varphi_1 \rangle$ and $\langle X, \varphi_2 \rangle$ have the same probability law when $\varphi_2(t) = \varphi_1(t - t_0)$. In addition, the independent atom property of white noise could be expressed as the statistical independence of $\langle X, \varphi_1 \rangle$ and $\langle X, \varphi_2 \rangle$ when φ_1 and φ_2 have disjoint supports ($\varphi_1(t)\varphi_2(t) \equiv 0$).*

Next, we explain two important types of white Lévy noises, namely stable and impulsive Poisson, that are studied in this paper.

Definition 10 (Stable random variables)**.** *[27, p. 5] A random variable X is stable with parameters $(\alpha, \beta, \sigma, \mu)$ if and only if its characteristic function $\widehat{p}(\omega)$ is given by*

$$\widehat{p}(\omega) := \mathbb{E}\left[\mathrm{e}^{\mathrm{j}\omega X}\right] = \mathrm{e}^{f(\omega)},$$

where $f : \mathbb{R} \mapsto \mathbb{C}$ is

$$f(\omega) = \mathrm{j}\omega\mu - \sigma^\alpha |\omega|^\alpha \left(1 - \mathrm{j}\beta \,\mathrm{sgn}\,(\omega)\,\Phi(\omega)\right), \tag{3}$$

with

$$\Phi(\omega) = \begin{cases} \tan\frac{\pi\alpha}{2} & \alpha \neq 1, \\ -\frac{2}{\pi}\log|\omega| & \alpha = 1, \end{cases}$$

such that $\alpha \in (0, 2]$ is the stability coefficient, $\beta \in [-1, 1]$ is the skewness coefficient, $\sigma \in (0, \infty)$ is the scale coefficient, and $\mu \in \mathbb{R}$ is the shift coefficient. In addition, function $\mathrm{sgn}\,(x) : \mathbb{R} \mapsto \{-1, 0, 1\}$ is the sign *function defined as*

$$\mathrm{sgn}\,(x) = \begin{cases} 1 & x > 0, \\ 0 & x = 0, \\ -1 & x < 0. \end{cases}$$

Definition 11 (Stable white noise)**.** *[28, p. 87] The random process X is a stable innovation process with parameters $(\alpha, \beta, \sigma, \mu)$ if X is a white Lévy noise with the Lévy exponent $f(\omega)$ defined in* (3).

Definition 12 (Impulsive Poisson white noise)**.** *[16, p. 64] When the Lévy exponent $f(\omega)$ of a white Lévy noise satisfies*

$$f(\omega) = \lambda \int_{\mathbb{R}\setminus 0} \left(\mathrm{e}^{\mathrm{j}a\omega} - 1\right)\,\mathrm{d}F_A(a)$$

for some scalar $\lambda > 0$ (known as the rate of impulses) and cumulative distribution function F_A over $\mathbb{R}$ (called the amplitude cdf), then, it is called an impulsive Poisson white noise.

Lemma 6. *Let $X(t)$ be a white Lévy noise with parameters $\sigma = 0$, μ and $V(a)$ such that*

$$\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a) < \infty.$$

Then, $X(t)$ can be decomposed as $X(t) = Y(t) + \mu'$, where $Y(t)$ is an impulsive Poisson white noise with impulse rate $\lambda = \int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a) < \infty$ and impulse amplitude cdf $F_A(a) = \frac{1}{\lambda}V(a)$. The constant μ' is also given by

$$\mu' = \mu - \lambda \int_{[-1,1]\setminus\{0\}} a\,\mathrm{d}F_A(a).$$

The proof of the lemma can be found in Section IV-H.

As can be seen from Lemma 6, the function V has a great influence on the type of the random variables derived linearly from an innovation process. To clarify this fact, we consider the following decomposition on a generic V to three non-negative increasing functions as [1], [29]:

$$V(a) = V_{ac}(a) + V_d(a) + V_{cs}(a), \tag{4}$$

where

- the function V_d known as the discrete part of V, consists of countable finite jumps at $a \in \mathbb{R}$ such that $0 < V_d(a^+) - V_d(a^-) < \infty$, while $V_d(a)$ is constant elsewhere,
- the function V_{ac} known as the absolutely continuous part of V, *i.e.* for any Borel set $\mathcal{A}$ with Lebesgue measure of 0, $\int_{\mathcal{A}} \mathrm{d}V_{ac}(a) = 0$, and
- the function V_{cs} known as the continuous singular part of V, is a continuous function, but not absolutely continuous.

Then, we have the following lemma:

Lemma 7. *Let $X(t)$ be a white noise with the triplet $(\mu, \sigma, V = V_d + V_{ac} + V_{cs})$, and define the random variable X_0 as*

$$X_0 = \int_0^1 X(t)\,\mathrm{d}t.$$

Then,

1) *[30] X_0 is discrete if and only if*
   - $\sigma = 0$,
   - $\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a) < \infty$,
   - $V_{ac}(a) \equiv V_{cs}(a) \equiv 0$.
2) *[29] X_0 is continuous if at least one of the following conditions is satisfied:*
   - $\sigma > 0$,
   - $\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V_{ac}(a) = \infty$.
3) *[30] X_0 is discrete-continuous if and only if*
   - $\sigma = 0$,
   - $\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a) < \infty$,
   - $V_{cs}(a) \equiv 0$ *and* $V_{ac}(a) \not\equiv 0$.

Note that the conditions for the discrete and the discrete-continuous cases are necessary and sufficient, while the conditions for the continuous case are only sufficient.

III. Main Results

We first propose a criterion for the compressibility of stochastic processes, and study its operational meaning from the viewpoint of source coding. This criterion and our general results regarding its evaluation are described in Section III-A. In Section III-B, we evaluate the compressibility criterion for some special cases namely stable white noise, impulsive Poisson white noise and their sum. Finally, we present a qualitative comparison between the compressibility of the considered innovation processes in Section III-C.

A. Compressibility via quantization

A generic continuous-domain stochastic process is spread over the continuum of time and amplitude. Hence, we doubly discretize the process by applying time and amplitude quantization. This enables us to utilize the conventional definition of the entropy measure. Then, we monitor the entropy trends as the quantization becomes finer.

The amplitude quantization was previously defined in Definition 6. The time quantization, or equivalently, the sampling in time shall be defined in a similar fashion:

Definition 13 (Time Quantization)**.** *The time quantization with step size $1/n$ of a white Lévy noise (an innovation process) $X(t)$ is defined as the sequence $\left\{X_i^{(n)}\right\}_{i\in\mathbb{Z}}$ of random variables*

$$X_i^{(n)} = \langle X, \phi_{i,n} \rangle,$$

where $\phi_{i,n}(t) = \phi(nt - i + 1)$ and

$$\phi(t) = \begin{cases} 1 & t \in [0,1), \\ 0 & t \notin [0,1). \end{cases} \tag{5}$$

Remark 1. *Observe that $\phi(t)$ is not a member of $\mathcal{S}(\mathbb{R})$ as defined in Definition 7. Hence, strictly speaking, for a white Lévy noise $X(t)$ with sample space $\mathcal{S}'(\mathbb{R})$, we cannot automatically define $\langle X, \phi \rangle$ based on Definition 9. However, the random variables $X_i^{(n)}$ could be easily interpreted as the increments of the Lévy process corresponding to this white noise. Alternatively, one can define $\langle X, \phi \rangle$ as the limit of $\langle X, \psi_k \rangle$ as $k \to \infty$, when $\{\psi_k(t)\}_{k=1}^{\infty} \subset \mathcal{S}(\mathbb{R})$ satisfy $\lim_{k\to\infty} \int_{\mathbb{R}} |\psi_k(t) - \phi(t)|\,\mathrm{d}t = 0$. For definitions and arguments in this paper, convergence in probability is sufficient for $\langle X, \psi_k \rangle \underset{k\to\infty}{\to} \langle X, \phi \rangle$ to hold.*

In some cases, instead of just one function $\phi(t)$, we have multiple step functions ϕ_i for $i = 0, 1, \ldots, n$ which are not members of $\mathcal{S}(\mathbb{R})$ and we want to simultaneously define the random variables $\langle X, \phi_0 \rangle, \ldots, \langle X, \phi_n \rangle$. To account for the simultaneous definition of $\{\langle X, \phi_i \rangle\}_{i=0}^{n}$ that captures the joint distribution, we need the convergence in probability of the multivariate random variable $[\langle X, \psi_{k,0} \rangle, \ldots, \langle X, \psi_{k,n} \rangle]$ to $[\langle X, \phi_0 \rangle, \ldots, \langle X, \phi_n \rangle]$, when

$$\lim_{k\to\infty} \sum_{i=0}^{n} \int_{\mathbb{R}} |\psi_{k,i}(t) - \phi_i(t)|\,\mathrm{d}t = 0.$$

Such convergence results could be achieved via the approach of [31].

Our next step, is to find the entropy rate of a (doubly) quantized random process. Let $X(t)$ be a white Lévy noise and define the random vectors of size n

$$\widetilde{\mathbf{X}}_i^{m,n} := \left(\left[X_{n(i-1)+1}^{(n)}\right]_m, \cdots, \left[X_{ni}^{(n)}\right]_m \right),$$

where $X_i^{(n)}$ refers to the time quantization of the process (Definition 13) followed by amplitude quantization in the form $\left[X_i^{(n)}\right]_m$ as shown in Definition 6. The time quantization in $\widetilde{\mathbf{X}}_i^{m,n}$ spans the interval $t \in [i-1, i)$ of the process $X(t)$ via n random variables $\left[X_{n(i-1)+j}^{(n)}\right]_m$, $j = 1, \dots, n$. Thus, the sequence $\{\widetilde{\mathbf{X}}_i^{m,n}\}_{i\in\mathbb{Z}}$ represents the innovation process over the whole real axis $t \in \mathbb{R}$ in a quantized way. We evaluate the quantization entropy rate (entropy per unit interval of time) for $\{\widetilde{\mathbf{X}}_i^{m,n}\}_{i\in\mathbb{Z}}$

$$\mathcal{H}_{m,n}(X) := \lim_{T\to\infty} \frac{\mathrm{H}\left(\widetilde{\mathbf{X}}_{-T+1}^{m,n}, \cdots, \widetilde{\mathbf{X}}_T^{m,n}\right)}{2T}, \qquad (6)$$

where $\mathrm{H}(\cdot)$ stands for the discrete entropy. The above definition has an operational meaning in terms of the number of bits required for asymptotic lossless compression of the source as T tends to ∞. For fixed m, n and varying i, since $\left[X_i^{(n)}\right]_m$'s depend on equilength and non-overlapping time intervals of the white noise, they are independent and identically distributed (Lemma 5 in Sec. II-C). Therefore,

$$\mathcal{H}_{m,n}(X) = \mathrm{H}\left(\widetilde{\mathbf{X}}_1^{m,n}\right) = n\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right). \qquad (7)$$

To compensate for the quantization effect, we shall study the behavior of $\mathcal{H}_{m,n}(X)$ as $m, n \to \infty$.

The following theorems consider the behavior of $\mathcal{H}_{m,n}(X)$ by showing that for a wide class of white noises we have one of the following cases

$$\lim_{n\to\infty} \sup_{m\geq m(n)} \left| \frac{\mathcal{H}_{m,n}(X)}{\kappa(n)} - \log(m) - \zeta(n) \right| = 0, \qquad (8)$$

$$\lim_{n\to\infty} \sup_{m\geq m(n)} \left| \mathcal{H}_{m,n}(X) - \zeta(n) \right| = 0, \qquad (9)$$

for appropriate functions $m(n)$, $\kappa(n)$, and $\zeta(n)$. Intuitively, the asymptotic value of $\kappa(n)/n$ in the case of (8) generalizes the concept of Rényi entropy dimension to random processes; however, for the asymptotic results, we need m to be larger than a function of n. The case of (9) identifies random processes that are so discrete that the extended Rényi entropy dimension becomes zero. The following two theorems prove the existence and uniqueness of $\kappa(n)$ and $\zeta(n)$ such that (8) or (9) hold. Let us begin with the asymptotic uniqueness of $\kappa(n)$ and $\zeta(n)$ first.

Theorem 1 (Asymptotic Uniqueness of $\kappa(n)$ and $\zeta(n)$)**.** *Let $X(t)$ be a white Lévy noise. If one can find functions $\kappa_i(n)$, $\zeta_i(n)$, and $m_i(n)$ for $i = 1, 2$ such that*

$$\lim_{n\to\infty} \sup_{m\geq m_i(n)} \left| \frac{\mathcal{H}_{m,n}(X)}{\kappa_i(n)} - \log m - \zeta_i(n) \right| = 0,$$

then,

$$\exists n_0 : \kappa_1(n) = \kappa_2(n), \qquad \forall n \geq n_0 \qquad (10)$$

$$\lim_{n\to\infty} |\zeta_1(n) - \zeta_2(n)| = 0. \qquad (11)$$

Likewise, if

$$\lim_{n\to\infty} \sup_{m\geq m(n)} |\mathcal{H}_{m,n}(X) - \zeta_i(n)| = 0,$$

then,

$$\lim_{n\to\infty} |\zeta_1(n) - \zeta_2(n)| = 0.$$

The theorem is proved in Section IV-A.

The above theorem shows that $\kappa(n)$ and $\zeta(n)$ are essentially unique, if they exist. In the following theorem, we show the existence of $\kappa(n)$ and $\zeta(n)$ under certain conditions. We further express $\kappa(n)$ in terms of the parameters of the white noise. Next, we state some facts about $\zeta(n)$ for a class of white noise processes.

In the following theorem, we first identify the discrete and continuous parts of $X_i^{(n)}$ that are inherited from the random process, and eventually derive the entropy of the quantized random variables based on the concept of entropy dimension.

Theorem 2. *Let $X(t)$ be a white noise with the triplet $(\mu, \sigma, V = V_d + V_{ac} + V_{cs})$. Assume that X_0 defined by*

$$X_0 = \int_0^1 X(t)\,\mathrm{d}t,$$

is either a discrete, continuous, or discrete-continuous random variable (as discussed in Lemma 7). Then, the following statements hold:

- *If X_0 is discrete:*
  There exist functions $m(n)$ and $\zeta(n)$ such that
  $$\lim_{n\to\infty} \sup_{m:\ m\geq m(n)} |\mathcal{H}_{m,n}(X) - \zeta(n)| = 0,$$
  and
  $$c_1 \log n \leq \zeta(n) \leq c_2 \log n,$$
  for some constants $0 < c_1 \leq c_2 < \infty$ (not depending on n) provided that $\mathrm{H}(X_0)$ and $\mathrm{H}(A)$ are finite, where
  $$P_A[x] := \frac{1}{\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a)} \tfrac{\mathrm{d}}{\mathrm{d}x} V(x).$$
  Since X_0 is discrete $V(a)$ is discrete and bounded (Lemma 7), so the above definition of $P_A(x)$ corresponds to a discrete random variable A and the discrete entropy $\mathrm{H}(A)$ is meaningful.
- *If X_0 is discrete-continuous:*
  Set
  $$\kappa(n) = n\left(1 - \exp\left[-\frac{\lambda}{n}(1-\alpha)\right]\right),$$
  where
  $$\lambda := \int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a), \qquad \alpha := \frac{1}{\lambda}\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V_d(a).$$
  Then, there exist functions $m(n)$ and $\zeta(n)$ such that
  $$\lim_{n\to\infty} \sup_{m:\ m\geq m(n)} \left| \frac{\mathcal{H}_{m,n}(X)}{\kappa(n)} - \log(m) - \zeta(n) \right| = 0, \qquad (12)$$
  and
  $$c_1 \log n \leq \zeta(n) \leq c_2 \log n,$$

for some constants $0 < c_1 \leq c_2 < \infty$ (not depending on n) if random variables $X_{0,c}$, $X_{0,D}$, $X_{1,c}^{(n)}$, A_c, and A_D defined below satisfy the technical assumptions that $\mathrm{H}([X_0]_1)$, $\mathrm{h}\left(X_{1,c}^{(n)}\right)$, $\mathrm{h}(A_c)$, $\mathrm{H}(X_{0,D})$ and $\mathrm{H}(A_D)$ are well-defined for all n: random variables $X_{0,D}$ and $X_{0,c}$ are the discrete and continuous parts of X_0, respectively. Similarly, $X_{1,c}^{(n)}$ stands for the continuous part of $X_1^{(n)}$ (Lemma 7, shows that $X_1^{(n)}$ is also discrete-continuous if X_0 is discrete-continuous). The random variables A_c and A_D correspond to the distributions

$$p_{A_c}(x) := \frac{1}{\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V_{ac}(a)} \tfrac{\mathrm{d}}{\mathrm{d}x} V_{ac}(x),$$
$$P_{A_D}[x] := \frac{1}{\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V_{d}(a)} \tfrac{\mathrm{d}}{\mathrm{d}x} V_{d}(x),$$

respectively.

- *If X_0 is continuous:*
  Set $\kappa(n) = n$ and $\zeta(n) = \mathrm{h}\left(X_1^{(n)}\right)$ where $X_1^{(n)} = \int_0^{\frac{1}{n}} X(t)\,\mathrm{d}t$. If at least one of the assumptions in part (2) of Lemma 7 is satisfied, $\mathrm{H}([X_0]_1)$ is finite and $\mathrm{h}\left(X_1^{(n)}\right)$ is well-defined for all n, then, we have that
  $$\lim_{n\to\infty} \sup_{m:\ m\geq m(n)} \left| \frac{\mathcal{H}_{m,n}(X)}{\kappa(n)} - \log(m) - \zeta(n) \right| = 0. \tag{13}$$
  Furthermore, $\zeta(n) = \mathrm{h}\left(X_1^{(n)}\right)$ is a non-increasing function in n.

The proof can be found in Section IV-B.

Remark 2. *In order to define A_c and A_D, which are used in the case that X_0 is a discrete-continuous random variable, note that Lemma 7 in conjunction with Lemma 6 shows that X_0 is discrete-continuous only if the white noise is an impulsive Poisson white noise, plus a constant. Utilizing Lemma 7, we can define continuous random variable A_c with pdf p_{A_c} such that:*

$$p_{A_c}(x) := \frac{1}{\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V_{ac}(a)} \tfrac{\mathrm{d}}{\mathrm{d}x} V_{ac},$$

where V_{ac} was defined in (4). *Note that, according to Lemma 7, even if $V(a)$ does not have any discrete part, X_0 can still be discrete-continuous. Hence, if the discrete part of $V(a)$ exists, one can define A_D with pmf P_{A_D} such that*

$$P_{A_D}[x] := \frac{1}{\int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V_{d}(a)} \tfrac{\mathrm{d}}{\mathrm{d}x} V_{d}.$$

B. Special Cases

In this section, we evaluate functions $\zeta(n)$ and $m(n)$ in cases where white noise is impulsive Poisson, stable, or sum of impulsive Poisson and stable.

Proposition 1. *Let $X(t)$ be a stable white noise with parameters $(\alpha, \beta, \sigma, \mu)$. We define*

$$X_0 := \langle X, \phi\rangle = \int_0^1 X(t)\,\mathrm{d}t,$$

where ϕ is the function defined in (5). *Then, for all $m(\cdot) : \mathbb{N} \to \mathbb{N}$ satisfying $\lim_{n\to\infty} \frac{m(n)}{\sqrt[\alpha]{n}} = \infty$ we have that*

$$\lim_{n\to\infty} \sup_{m: m\geq m(n)} \left| \frac{\mathcal{H}_{m,n}(X)}{n} - \log \tfrac{m}{\sqrt[\alpha]{n}} - \mathrm{h}(X_0) \right| = 0,$$

where $\mathcal{H}_{m,n}(X)$ is defined in (7), *and $\mathrm{h}(X_0)$ is the differential entropy of continuous random variable X_0.*

The proof is given in Section IV-C.

Proposition 2. *Let $X(t)$ be an impulsive Poisson white noise with rate λ and continuous amplitude pdf p_A (Definition 1) such that*

$$p_A \in (\alpha, \ell, v)\text{–}\mathcal{AC}, \tag{14}$$

for some positive constants α, ℓ, v (defined in Definition 5) and

$$\int_{\mathbb{R}} p_A(x) \left|\log \tfrac{1}{p_A(x)}\,\mathrm{d}x\right| < \infty, \qquad \mathrm{H}([A]_1) < \infty,$$

where $[A]_1$ is the quantization with unit step size of a random variable A with pdf p_A. Then, for any function $m(\cdot) : \mathbb{N} \to \mathbb{N}$ satisfying $\lim_{n\to\infty} \frac{m(n)}{\log n} = \infty$ we have that

$$\lim_{n\to\infty} \sup_{m:\ m\geq m(n)} \left| \frac{\mathcal{H}_{m,n}(X)}{\kappa(n)} - \log(mn) - \mathrm{h}(A) + \log\lambda - 1 \right| = 0,$$

where $\kappa(n) = n(1 - \mathrm{e}^{-\lambda/n})$.

The proof is given in Section IV-D.

Corollary 2. *Let $X(t)$ be a white Lévy noise with parameters $\sigma = 0$, μ, and absolutely continuous and bounded function $V(a) = V_{ac}(a)$, which was defined in* (4). *Then, utilizing Lemma 6, the compressibility of $X(t)$ is similar to the compressibility of the impulsive Poisson case with*

$$\lambda := \int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a), \qquad A \sim p_A(x) := \tfrac{1}{\lambda} \tfrac{\mathrm{d}}{\mathrm{d}x} V(x),$$

provided that p_A satisfies the conditions in Proposition 2.

Proposition 3. *Let $X(t)$ be a stable white noise with parameters $(\alpha, \beta, \sigma, \mu)$ and $Y(t)$ be an impulsive Poisson white noise, independent of $X(t)$, with rate λ and amplitude random variable A, such that there exists $\theta > 0$ where*

$$\mathbb{E}\left[|A|^\theta\right] < \infty.$$

We define

$$X_0 := \langle X, \phi\rangle = \int_0^1 X(t)\,\mathrm{d}t,$$

where ϕ is the function defined in (5). *Then, for $Z(t) = X(t) + Y(t)$ and all $m(\cdot) : \mathbb{N} \to \mathbb{N}$ satisfying $\lim_{n\to\infty} \frac{m(n)}{\sqrt[\alpha]{n}} = \infty$ we have that*

$$\lim_{n\to\infty} \sup_{m: m\geq m(n)} \left| \frac{\mathcal{H}_{m,n}(Z)}{n} - \log \tfrac{m}{\sqrt[\alpha]{n}} - \mathrm{h}(X_0) \right| = 0,$$

where $\mathcal{H}_{m,n}(Z)$ is defined in (7).

The proposition is proved in Section IV-E.

Remark 3. *Note that Proposition 3 assumes that the amplitude measure of the impulsive Poisson only has a finite moment*

θ for an arbitrary $\theta > 0$, while Proposition 2 has the stronger assumption that amplitude measure is in (θ, ℓ, v)–$\mathcal{AC}$.

Corollary 3. *Let $Z(t)$ be a white Lévy noise with parameters $\sigma > 0$, μ, and* bounded *function $V(a)$. Then, $Z(t)$ can be decomposed as $Z(t) = X_1(t) + Y_1(t)$ where $X_1(t)$ is a Gaussian white noise with parameters $(0, \sigma)$ and $Y_1(t)$ is a white Lévy noise with parameters $\sigma = 0$, μ, and the bounded function $V(a)$. According to Lemma 6, we have that $Y_1(t) = \mu' + Y(t)$ where $Y(t)$ is an impulsive Poisson white noise with parameters λ and amplitude cdf F_A where*

$$\lambda := \int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a) < \infty,$$
$$A \sim F_A(x) := \frac{1}{\lambda} V(x),$$
$$\mu' := \mu - \lambda \int_{[-1,1]\setminus\{0\}} a \, \mathrm{d}V(a).$$

Then, the compressibility of $Z(t)$ is similar to the above case with

$$\alpha = 2, \qquad \mathrm{h}\,(X_0) = \frac{1}{2}\log(2\pi \mathrm{e}\sigma^2).$$

provided that $\mathbb{E}\left[|A|^\theta\right] < \infty$ for some $\theta > 0$.

C. Comparison and discussion

In this section, we compare different white Lévy noises based on Theorem 2. Take two white Lévy noises $X(t)$ and $Y(t)$ belonging to the general class of processes considered in Theorem 2. Let $m_X(n)$ and $m_Y(n)$ be the functions associated with these two white noises in Theorem 2. Observe that if Theorem 2 holds for a choice of $m(n)$, it will also hold for any function $m'(n) \geq m(n)$. As a result, we can look at the sequence of pairs (n, m) where $m \geq \max\{m_X(n), m_Y(n)\}$ when we compare the two white noise processes. Therefore to make the comparison we only need to compare the values of $\kappa(n)$ and $\zeta(n)$ for the two white noise processes asymptotically.

Theorem 3. *Let $X(t)$ and $Y(t)$ be two white Lévy noises and define $X_0 := \int_0^1 X(t)\,\mathrm{d}t$ and $Y_0 := \int_0^1 Y(t)\,\mathrm{d}t$. Assume that conditions in Theorem 2 hold for $X(t)$ and $Y(t)$ based on their type (X_0 and Y_0 might have different types). Let $\kappa_X(n), \kappa_Y(n)$, $\zeta_X(n), \zeta_Y(n)$, and $m_X(n), m_Y(n)$ be the functions associated in Theorem 2 to X and Y, respectively.*

1) *If X_0 is discrete and Y_0 is discrete-continuous, then*

$$\lim_{n\to\infty} \sup_{m \geq m(n)} \frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)} = 0,$$

where $m(n)$ is a function such that

$$m(n) \geq \max\{m_X(n), m_Y(n)\},$$
$$\lim_{n\to\infty} \frac{\log m(n)}{\log n} = \infty.$$

2) *If X_0 is discrete-continuous and Y_0 is continuous, then*

$$\lim_{n\to\infty} \sup_{m \geq m(n)} \frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)} = 0,$$

where $m(n)$ is a function such that

$$m(n) \geq \max\{m_X(n), m_Y(n)\},$$
$$\lim_{n\to\infty} \frac{\log m(n)}{\zeta_Y(n)} = \infty.$$

3) *If X_0 is continuous and $Y(t)$ is a Gaussian white noise with parameters $\sigma > 0$ and $V(a) = 0$ for all $a \in \mathbb{R}$, then*

$$\limsup_{n\to\infty} \sup_{m: m \geq m(n)} \frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)} \leq 1,$$

where $m(n)$ is a function such that

$$m(n) \geq \max\{m_X(n), m_Y(n)\},$$
$$\lim_{n\to\infty} \frac{m(n)}{\sqrt{n}} = \infty.$$

The proof of the theorem is presented in Section IV-F.

The theorem, intuitively states that when m grows sufficiently faster than n, the discrete white noises are asymptotically more compressible than discrete-continuous noises, and the discrete-continuous white noises are more compressible than continuous noises. In addition, Gaussian white noise is the least compressible white noise.

In the following theorem, we compare impulsive Poisson and stable white noises amongst themselves.

Theorem 4. *The following statements apply to impulsive Poisson and stable white noises.*

i) *A stable white noise becomes less compressible as its stability parameter α increases. Let X_1 and X_2 be stable white noises with stability parameters $0 < \alpha_1 < \alpha_2 \leq 2$, and functions $m_1(n)$ and $m_2(n)$ defined in Theorem 2, respectively. Then, for $m(n) \geq \max\{m_1(n), m_2(n)\}$, we have that*

$$\limsup_{n\to\infty} \sup_{m \geq m(n)} \frac{\mathcal{H}_{m,n}(X_1)}{\mathcal{H}_{m,n}(X_2)} \leq 1, \tag{15}$$
$$\lim_{n\to\infty} \sup_{m \geq m(n)} \mathcal{H}_{m,n}(X_1) - \mathcal{H}_{m,n}(X_2) = -\infty. \tag{16}$$

ii) *An impulsive Poisson white noise becomes less compressible as its rate parameter λ increases. Let X_1 and X_2 be impulsive Poisson white noises with rate parameters $0 < \lambda_1 < \lambda_2$, absolutely continuous amplitude distributions, and functions $m_1(n)$ and $m_2(n)$ defined in Theorem 2, respectively. Then, for $m(n) \geq \max\{m_1(n), m_2(n)\}$, we have that*

$$\lim_{n\to\infty} \sup_{m \geq m(n)} \frac{\mathcal{H}_{m,n}(X_1)}{\mathcal{H}_{m,n}(X_2)} = \frac{\lambda_1}{\lambda_2} < 1, \tag{17}$$
$$\lim_{n\to\infty} \sup_{m \geq m(n)} \mathcal{H}_{m,n}(X_1) - \mathcal{H}_{m,n}(X_2) = -\infty. \tag{18}$$

The proof can be found in Section IV-G.

We should mention that Theorem 4 confirms with previously studied notions of compressibility in [12], [14], [15]. However, Part 2 of Theorem 3 deviates from the available literature by identifying the impulsive Poisson white noises as more compressible than heavy-tailed stable white noises. This difference is fundamental and is caused by basing the compressibility definition on the probability concentration properties of processes rather than their amplitude distribution. For instance, by shifting a part of the probability law to higher amplitude values, the tail behavior changes, however, the differential entropy remains unchanged. To illustrate this fact with an example, let X be a non-negative-valued random variable and define

$$Y=\sum_{i=0}^{\infty}\mathbb{1}_{X\in[i,i+1]}(X+2^i).$$

Obviously, if the zero intervals in the pdf of Y are omitted, we obtain the pdf of X. Note that, compared to the pdf of X, the pdf of Y contains arbitrarily large zero intervals. Therefore, Y exhibits a different tail behavior from X (Y is more fat-tailed than X). Hence, the classical ℓ_p-compressibility identifies Y as more compressible than X [15]. However, the differential entropy of the two random variables are the same, and with our new entropy-based definition they are equally compressible.

IV. Proofs

A. Proof of Theorem 1

Proof of the first case:

Proof of (10): There is $n_0\in\mathbb{N}$ such that for any $n\geq n_0$ we have

$$\sup_{m\geq m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_i(n)}-\log m-\zeta_i(n)\right| \leq 1,\qquad i=1,2,\dots, \tag{19}$$

where $m(n)$ can be any function larger than $\max\{m_1(n),m_2(n)\}$. We can write

$$\begin{aligned}&\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_1(n)}-\log m-\zeta_1(n)\right|\\ &=\left|\frac{\kappa_2(n)}{\kappa_1(n)}\left[\frac{\mathcal{H}_{m,n}(X)}{\kappa_2(n)}-\log m-\zeta_2(n)\right]\right.\\ &\quad\left.-\left(1-\frac{\kappa_2(n)}{\kappa_1(n)}\right)\log m-\zeta_1(n)+\zeta_2(n)\frac{\kappa_2(n)}{\kappa_1(n)}\right|\\ &\geq\left|1-\frac{\kappa_2(n)}{\kappa_1(n)}\right|\log m\\ &\quad-\frac{\kappa_2(n)}{\kappa_1(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_2(n)}-\log m-\zeta_2(n)\right|\\ &\quad-\left|\zeta_1(n)-\zeta_2(n)\frac{\kappa_2(n)}{\kappa_1(n)}\right|.\end{aligned}$$

Hence, utilizing (19) for $i=2$, we have that for all $m\geq m(n)$, we have that

$$\begin{aligned}&\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_1(n)}-\log m-\zeta_1(n)\right|\\ &\quad\geq\left|1-\frac{\kappa_2(n)}{\kappa_1(n)}\right|\log m-\frac{\kappa_2(n)}{\kappa_1(n)}-\left|\zeta_1(n)-\zeta_2(n)\frac{\kappa_2(n)}{\kappa_1(n)}\right|.\end{aligned}$$

If $\kappa_1(n)\neq\kappa_2(n)$, we can select $m>m(n)$ large enough such that the right-hand side of the above equation becomes larger than 1. Therefore, (19) cannot be satisfied for $i=1$. Thus $\kappa_1(n)=\kappa_2(n)$ for any $n\geq n_0$.

Proof of (11): From (10) we obtain that for large enough n we have that $\kappa_1(n)=\kappa_2(n)$. Hence, for large enough n, we can write

$$\begin{aligned}&\sup_{m\geq m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_1(n)}-\log m-\zeta_1(n)\right|\\ &\qquad+\sup_{m\geq m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_2(n)}-\log m-\zeta_2(n)\right|\\ &=\sup_{m\geq m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_1(n)}-\log m-\zeta_1(n)\right|\\ &\quad+\sup_{m\geq m(n)}\left|-\frac{\mathcal{H}_{m,n}(X)}{\kappa_2(n)}+\log m+\zeta_2(n)\right|\\ &\geq\sup_{m\geq m(n)}|\zeta_1(n)-\zeta_2(n)|\\ &=|\zeta_1(n)-\zeta_2(n)|,\end{aligned}$$

where $m(n):=\max\{m_1(n),m_2(n)\}$. Therefore, according to the assumption, we obtain that

$$\begin{aligned}&\lim_{n\to\infty}|\zeta_1(n)-\zeta_2(n)|\\ &\leq\lim_{n\to\infty}\sup_{m\geq m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_1(n)}-\log m-\zeta_1(n)\right|\\ &\quad+\lim_{n\to\infty}\sup_{m\geq m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_2(n)}-\log m-\zeta_2(n)\right|\\ &=0.\end{aligned}$$

Thus, the statement was proved.

Proof of the second case: For this case, similarly, we can write

$$\begin{aligned}&\lim_{n\to\infty}|\zeta_1(n)-\zeta_2(n)|\\ &\leq\lim_{n\to\infty}\sup_{m\geq m(n)}|\mathcal{H}_{m,n}(X)-\zeta_1(n)|\\ &\quad+\lim_{n\to\infty}\sup_{m\geq m(n)}|\mathcal{H}_{m,n}(X)-\zeta_2(n)|\\ &=0.\end{aligned}$$

Therefore the statement is proved. □

B. Proof of Theorem 2

From (7) we obtain that

$$\mathcal{H}_{m,n}(X)=n\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right).$$

Now, we are going to use Lemma 3 in order to show that

$$n\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right)=nd_n\log m+nh_n+ne_{m,n},$$

where d_n is the entropy dimension of $X_1^{(n)}$ and $e_{m,n}$ vanishes for every fixed n when m tends to ∞. To this end, first, we prove the conditions of Lemma 3 which are the same for all types of discrete, continuous and discrete-continuous:

1) $X_1^{(n)}$ is a discrete, continuous, or discrete-continuous random variable, and has the same type as X_0,
2) $\mathrm{H}\left(\left[X_1^{(n)}\right]_1\right)$ is finite.

In the next step, we prove the exclusive conditions for each case. By choosing $\kappa(n)=1$ in the discrete or $\kappa(n)=nd_n$ in other two cases, and $\zeta(n) = nh_n/\kappa(n)$, we only need to find $m(n)$ such that

$$\lim_{n\to\infty} \sup_{m:m\geq m(n)} \left|\frac{e_{m,n}}{\kappa(n)}\right| = 0.$$

Since $e_{m,n}$ vanishes as m tends to ∞ for any fixed n, there is some $m(n)$ such that for any $m \geq m(n)$ we have

$$\left|\frac{e_{m,n}}{\kappa(n)}\right| \leq \frac{1}{n}.$$

This proves the existence of the function $m(n)$.

It remains to prove items 1 and 2, and then, find $\kappa(n)$ and $\zeta(n)$ for cases of X_0 being discrete, continuous, or discrete-continuous.

Proof of item 1: Random variable $X_1^{(n)}$ is the time quantization with step size $1/n$ of a white Lévy noise with parameters σ, μ and $V(a)$. It can be alternatively expressed as the time quantization with unit step size of a new white Lévy noise with scaled parameters $\sigma/\sqrt{n}$, μ/n and $V(a)/n$. Lemma 7 gives conditions for X_0 (time quantization with unit step size) being discrete, continuous, or discrete-continuous random variable in terms of σ, μ, and $V(a)$. These conditions for σ, μ, and $V(a)$ are equivalent for the conditions for $\sigma/\sqrt{n}$, μ/n, and $V(a)/n$. Therefore, random variable $X_1^{(n)}$ has the same type as X_0.

Proof of item 2: From Definition 9, observe that X_0 can be written as the sum of n i.i.d. random variables $X_i^{(n)}$ for $i \in \{1, \cdots, n\}$. Hence, we have that

$$\left[\sum_{i=1}^{n} X_i^{(n)}\right]_1 = [X_0]_1.$$

Observe that for any quantization of sum of variables we can write

$$\left[\sum_{i=1}^{n} X_i^{(n)}\right]_1 = \sum_{i=1}^{n} \left[X_i^{(n)}\right]_1 + E_n,$$

where E_n is a random variable taking values from $\{0, \cdots, n-1\}$. Observe that

$$\begin{aligned}
\mathrm{H}\left(\left[X_1^{(n)}\right]_1\right) &= \mathrm{H}\left(\sum_{i=1}^{n}\left[X_i^{(n)}\right]_1 \middle| \left[X_2^{(n)}\right]_1, \cdots, \left[X_n^{(n)}\right]_1\right) \\
&\leq \mathrm{H}\left(\sum_{i=1}^{n}\left[X_i^{(n)}\right]_1\right) \\
&= \mathrm{H}\left([X_0]_1 - E_n\right) \\
&\leq \mathrm{H}\left([X_0]_1, E_n\right) \\
&\leq \mathrm{H}\left([X_0]_1\right) + \log n
\end{aligned}$$

Therefore, $\mathrm{H}\left(\left[X_1^{(n)}\right]_1\right) < \infty$.

Now, we aim to check the exclusive constraints and find $\kappa(n)$ and $\zeta(n)$ for each case.

Case 1: X_0 is continuous and at least one of the assumptions in part (2) of Lemma 7 is satisfied: For the first claim, as it was proved in item 1 at the beginning of the proof, observe that $X_1^{(n)}$ is continuous. Therefore, according to Lemma 3, d_n is 1 for every n, and

$$\zeta(n) := h_n = \mathrm{h}\left(X_1^{(n)}\right),$$

provided that $\mathrm{h}\left(X_1^{(n)}\right)$ is well-defined, which is true based on the assumption.

Based on Definitions 9 and 13, we can write

$$X_1^{(n-1)} = X_1^{(n)} + E,$$

where E and $X_1^{(n)}$ are independent and

$$\mathrm{e}^{\mathrm{j}\omega E} = \exp\left(\omega \int_{\frac{1}{n}}^{\frac{1}{n-1}} X(t)\,\mathrm{d}t\right).$$

Now observe that for any two independent continuous random variables X, Y we have that

$$\mathrm{h}\left(X+Y\right) \geq \mathrm{h}\left(X+Y|Y\right) = \mathrm{h}\left(X\right).$$

Therefore,

$$\zeta(n) \leq \zeta(n-1).$$

As a result, $\zeta(n)$ is a non-increasing function.

Case 2: X_0 is discrete-continuous: From item 1 from the beginning of the proof, $X_1^{(n)}$ is discrete-continuous if X_0 is discrete-continuous. From Lemma 6 we have that $X_1^{(n)}$ is sum of a constant $\mu'/n := (\mu/n) - (1/n)\int_{[-1,1]\setminus\{0\}} a\,\mathrm{d}V(a)$ and the integral over $[0, \frac{1}{n}]$ of an impulsive Poisson white noise with rate λ and discrete-continuous amplitude density $p_A(x)$ where

$$\begin{aligned}
\lambda &:= \int_{\mathbb{R}\setminus\{0\}} \mathrm{d}V(a) < \infty, \\
A &\sim p_A(x) := \frac{1}{\lambda}\tfrac{\mathrm{d}}{\mathrm{d}x}V(x).
\end{aligned} \tag{20}$$

The reason that A is discrete-continuous is that, according to Lemma 7, when X_0 is discrete-continuous, the function $V(a)$ must be bounded and discrete-continuous *i.e.* $V_{cs}(a) \equiv 0$. Note that, according to Remark 2, the discrete and the continuous parts of A are A_D and A_c with pmf P_{A_D} and p_{A_c}, respectively.

The distribution of the integral of an impulsive Poisson white noise over $[0, \frac{1}{n}]$ is given in Proposition 4. Consider a Poisson random variable Q_1 with rate λ/n. Then, utilizing Proposition 4, we can define the following conditional distribution:

$$\begin{aligned}
p_{X_1^{(n)}|Q_1}(x|k) &= p_A^{*k}(x) \\
&:= \begin{cases} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x) & k \in \mathbb{N}, \\ \delta(x) & k = 0. \end{cases}
\end{aligned} \tag{21}$$

According to Lemma 3, we have that

$$d_n = 1 - \Pr\left\{X_1^{(n)} \text{ is discrete}\right\}.$$

Therefore, by the definition of Q we can write

$$\begin{aligned}
&\Pr\left\{X_1^{(n)} \text{ is discrete}\right\} \\
&\quad= \sum_{k=0}^{\infty} P_{Q_1}[k]\Pr\left\{X_1^{(n)} \text{ is discrete}\middle| Q_1 = k\right\} \\
&\quad= \mathrm{e}^{-\frac{\lambda}{n}} \times 1 + \sum_{k=1}^{\infty} P_{Q_1}[k]\Pr\left\{\sum_{i=1}^{k} A_i \text{ is discrete}\right\},
\end{aligned}$$

where A_i are i.i.d. random variable with probability density $p_A(x)$. Utilizing the result of Lemma 11 that

$$\begin{aligned}&\Pr\{X+Y \text{ is discrete}\}\\&\quad=\Pr\{X \text{ is discrete}\}\Pr\{Y \text{ is discrete}\},\end{aligned}$$

where X, Y are independent discrete-continuous random variables, we have that

$$\begin{aligned}d_n &= 1-\Pr\left\{X_1^{(n)} \text{ is discrete}\right\}\\&=1-\mathrm{e}^{-\frac{\lambda}{n}}-\mathrm{e}^{-\frac{\lambda}{n}}\sum_{k=1}^{\infty}\frac{(\lambda/n)^k}{k!}\Pr\{p_A \text{ is discrete}\}^k\\&=1-\mathrm{e}^{-\frac{\lambda}{n}}-\mathrm{e}^{-\frac{\lambda}{n}}(\mathrm{e}^{\frac{\lambda}{n}\Pr\{A \text{ is discrete}\}}-1)\\&=1-\exp\left[-\frac{\lambda}{n}\left(1-\Pr\{A \text{ is discrete}\}\right)\right].\end{aligned}\tag{22}$$

Therefore, the problem of $\kappa(n)$ is settled.

In order to find $\zeta(n)$, we apply Lemma 3 to $X_1^{(n)}$. Let $X_c^{(n)}$ and $X_D^{(n)}$ be the continuous and the discrete part of $X_1^{(n)}$, respectively. Then, to apply Lemma 3, we need to show that $\left|\mathrm{h}\left(X_c^{(n)}\right)\right|$ and $\mathrm{H}\left(X_D^{(n)}\right)$ are well-defined. Then, Lemma 3 implies that

$$\begin{aligned}\zeta(n)=\mathrm{h}\left(X_c^{(n)}\right)+\frac{1-d_n}{d_n}\mathrm{H}\left(X_D^{(n)}\right)\\+\frac{d_n\log\frac{1}{d_n}+(1-d_n)\log\frac{1}{1-d_n}}{d_n}.\end{aligned}$$

Since d_n given in (22) vanishes as n goes to infinity, to show the claim of the theorem for this part, it suffices to prove that

$$|\mathrm{h}\left(X_c^{(n)}\right)|\le c_3<\infty, \tag{23}$$

$$\frac{\mathrm{H}\left(X_D^{(n)}\right)}{d_n}\le c_4\log n, \tag{24}$$

$$c_5\log n\le\log\frac{1}{d_n}+\frac{1}{d_n}\log\frac{1}{1-d_n}\le c_6\log n. \tag{25}$$

for some constants $0\le c_3, c_4<\infty$ and $0<c_5\le c_6<\infty$ not depending on n. Observe that (23) and (24) imply that $\left|\mathrm{h}\left(X_c^{(n)}\right)\right|$ and $\mathrm{H}\left(X_D^{(n)}\right)$ are well-defined. Thus, the conditions of Lemma 3 are satisfied if we can show (23)-(25). It remains to show (23)-(25).

Proof of (23): From Definition 9, for $X_0 := \int_0^1 X(t)\,\mathrm{d}t$, we have that $X_0=\sum_{i=1}^n X_i^{(n)}$, where $X_i^{(n)}$ are *i.i.d.*. We denote the discrete and continuous part of X_0 and $X_1^{(n)}$ with $X_{0,D}, X_{0,c}$ and $X_D^{(n)}, X_c^{(n)}$, respectively. Therefore, according to Corollary 5 we have that

$$\mathrm{h}\left(X_{0,c}\right)\ge\mathrm{h}\left(X_c^{(n)}\right).$$

Hence, if $\mathrm{h}\left(X_{0,c}\right)<\infty$, then $\mathrm{h}\left(X_c^{(n)}\right)<\infty$ uniformly on n. Note that based on the assumption, $\mathrm{h}\left(X_c^{(n)}\right)$ well-defined.

Now, it only remains to prove $\mathrm{h}\left(X_c^{(n)}\right)>-\infty$. To this end, similar to what we did in (21). Let

$$S_k=\sum_{i=1}^k A_i, \tag{26}$$

where A_i are i.i.d. random variable with probability density $p_A(x)$. Let $(S_k)_c$ be the continuous part of S_k. Utilizing the definitions given in (20), Proposition 4 and Lemma 11, one can define random variable Q_2 as follows:

$$\begin{aligned}P_{Q_2}[k]&=\frac{1}{\mathrm{e}^{\lambda/n}-\mathrm{e}^{\Pr\{A \text{ is discrete}\}\lambda/n}}\frac{(\lambda/n)^k}{k!}\\&\quad\times(1-\Pr\{A \text{ is discrete}\}^k),\qquad\forall k\in\mathbb{N},\\p_{X_c^{(n)}|Q_2}(x|k)&=p_{(S_k)_c}(x),\qquad\forall k\in\mathbb{N}.\end{aligned}$$

Hence, we can write

$$\mathrm{h}\left(X_c^{(n)}\right)\ge\mathrm{h}\left(X_c^{(n)}\middle|Q_2\right)=\sum_{k=1}^{\infty}P_{Q_2}[k]\mathrm{h}\left((S_k)_c\right).$$

Now, from Corollary 5, we obtain that

$$\mathrm{h}\left(X_c^{(n)}\right)\ge\sum_{k=1}^{\infty}P_{Q_2}[k]\mathrm{h}\left(A_c\right)=\mathrm{h}\left(A_c\right),$$

where A_c is the continuous part of A. Therefore, if $\mathrm{h}\left(A_c\right)>-\infty$, the statement is proved.

Proof of (24): From Proposition 4, we can write

$$p_{X_1^{(n)}}(x)=\mathrm{e}^{-\frac{\lambda}{n}}\delta(x)+\mathrm{e}^{-\frac{\lambda}{n}}\sum_{k=1}^{\infty}\frac{\left(\frac{\lambda}{n}\right)^k}{k!}\big(\overbrace{p_A*\cdots*p_A}^{k}\big)(x).$$

Let us define $S_k=\sum_{i=1}^k A_k$ where A_i's are i.i.d. random variables with probability density function $p_A(x)$. We use $p_{S_k}(x)$ to denote the probability density function of S_k. We have

$$p_{X_1^{(n)}}(x)=\mathrm{e}^{-\frac{\lambda}{n}}\delta(x)+\mathrm{e}^{-\frac{\lambda}{n}}\sum_{k=1}^{\infty}\frac{\left(\frac{\lambda}{n}\right)^k}{k!}p_{S_k}(x).$$

Using Lemma 11, the discrete part of S_k, denoted by $(S_k)_D$, has probability $\Pr\{A \text{ is discrete}\}^k$, and equals

$$(S_k)_D=\sum_{i=1}^k (A_i)_D, \tag{27}$$

where $(A_i)_D$ is the discrete part of A_i. Hence, the distribution of the discrete part of $X_1^{(n)}$ without normalization is

$$\mathrm{e}^{-\frac{\lambda}{n}}\delta[x]+\mathrm{e}^{-\frac{\lambda}{n}}\sum_{k=1}^{\infty}\frac{\left(\frac{\lambda}{n}\right)^k}{k!}\Pr\{A \text{ is discrete}\}^k P_{(S_k)_D}[x],$$

where $P_{(S_k)_D}[x]$ is the pmf of the discrete part of S_k. Since

$$\begin{aligned}&\mathrm{e}^{-\frac{\lambda}{n}}+\mathrm{e}^{-\frac{\lambda}{n}}\sum_{k=1}^{\infty}\frac{\left(\frac{\lambda}{n}\right)^k}{k!}\Pr\{A \text{ is discrete}\}^k\\&\quad=\mathrm{e}^{-\frac{\lambda}{n}}+\mathrm{e}^{-\frac{\lambda}{n}}\left(\mathrm{e}^{\frac{\lambda\Pr\{A \text{ is discrete}\}}{n}}-1\right)\\&\quad=\mathrm{e}^{-\frac{\lambda}{n}+\frac{\lambda\Pr\{A \text{ is discrete}\}}{n}},\end{aligned}$$

after normalization, we obtain that

$$\begin{aligned}&P_{(X_1^{(n)})_D}[x]\\&\quad=\mathrm{e}^{-\frac{\lambda}{n}\Pr\{A \text{ is discrete}\}}\\&\qquad\times\left(\delta[x]+\sum_{k=1}^{\infty}\frac{\left(\frac{\lambda}{n}\Pr\{A \text{ is discrete}\}\right)^k}{k!}P_{(S_k)_D}[x]\right).\end{aligned}$$

Hence, one can define a Poisson random variable random variable Q_3 with rate $\Pr\{A \text{ is discrete}\}\lambda/n$ such that:

$$P_{X_D^{(n)}|Q_3}[x|k] = \begin{cases} P_{(S_k)_D}[x] & k \in \mathbb{N}, \\ \delta[x] & k = 0, \end{cases}$$

We can write that

$$\begin{aligned}\mathrm{H}\left(X_D^{(n)}\right) &\leq \mathrm{H}\left(X_D^{(n)}, Q_3\right) \\ &= \mathrm{H}\left(X_D^{(n)}\middle|Q_3\right) + \mathrm{H}\left(Q_3\right).\end{aligned}$$

Now, in order to prove (24), it suffices to show that

$$\begin{aligned}&n\mathrm{H}\left(X_D^{(n)}\middle|Q_3\right) \leq c_7 < \infty, \\ &n\mathrm{H}\left(Q_3\right) \leq c_8 \log n,\end{aligned}$$

for some constants c_7 and c_8.

For the first inequality, since

$$\mathrm{H}\left(X+Y\right) \leq \mathrm{H}\left(X, Y\right) \leq \mathrm{H}\left(X\right) + \mathrm{H}\left(Y\right),$$

using (27) we have

$$\begin{aligned}\mathrm{H}\left(X_D^{(n)}|Q\right) &= \sum_{k=1}^{\infty} P_{Q_3}[k]\mathrm{H}\left(\sum_{i=1}^{k}(A_i)_D\right) \\ &\leq \sum_{k=1}^{\infty} P_{Q_3}[k]k\mathrm{H}\left(A_D\right) \\ &= \mathbb{E}[Q_3]\mathrm{H}\left(A_D\right) \\ &= \frac{\lambda}{n}\Pr\{A \text{ is discrete}\}\,\mathrm{H}\left(A_D\right),\end{aligned}$$

where the last equality is true since Q_3 is a Poisson random variable with rate $\lambda\Pr\{A \text{ is discrete}\}/n$. Thus, $n\mathrm{H}\left(X_D^{(n)}|Q\right)$ is bounded uniformly on n.

For the second inequality, note that among all the discrete distributions defined on $\{0, 1, \cdots\}$ with a specified mean μ, we have that [24, Theorem 13.5.4]:

$$\mathrm{H}\left(X\right) \leq \log(1+\mu) + \mu\log\left(1 + \frac{1}{\mu}\right).$$

Therefore,

$$\mathrm{H}\left(Q_3\right) \leq \log(1+\mathbb{E}\left[Q_3\right]) + \mathbb{E}\left[Q_3\right]\log\left(1 + \frac{1}{\mathbb{E}\left[Q_3\right]}\right).$$

Note that $\mathbb{E}\left[Q_3\right] = \lambda\Pr\{A \text{ is discrete}\}/n$. Hence,

$$\frac{c_9}{n} \leq \mathbb{E}\left[Q_3\right] \leq \frac{c_{10}}{n}, \tag{28}$$

for some positive constants c_9 and c_{10}. This shows the existence of n_0 such that for $n > n_0$,

$$1 + \frac{1}{\mathbb{E}\left[Q_3\right]} \leq \left(\frac{n}{c_9}\right)^2.$$

Then, we have that

$$\begin{aligned}n\mathrm{H}\left(Q_3\right) &\leq n\mathbb{E}\left[Q_3\right] + n\mathbb{E}\left[Q_3\right]\log\left(1 + \frac{1}{\mathbb{E}\left[Q_3\right]}\right) \\ &\leq c_{10} + 2c_{10}\log(n) - 2c_{10}\log(c_9).\end{aligned}$$

Therefore, (24) is proved.

Proof of (25): It can be proved that

$$\begin{aligned}&\lim_{n\to\infty} \log\frac{1}{d_n} + \frac{1}{d_n}\log\frac{1}{1-d_n} \\ &\quad - \log\frac{n}{\lambda(1-\Pr\{A \text{ is discrete}\})} \\ &= 1.\end{aligned}$$

Therefore, (25) is proved.

Case 3: X_0 is discrete: As it was proved in item 1, observe that $X_1^{(n)}$ is discrete if X_0 is discrete. Therefore, from Lemma 3, d_n is 0 for every n, and there is no term of $\log m$ in $\mathcal{H}_{m,n}(X)$. However, regardless of the results of Lemma 3, we choose $\kappa(n) = 1$, and,

$$\zeta(n) = n\mathrm{H}\left(X_1^{(n)}\right).$$

From Lemma 7, we obtain that if X_0 is discrete, then $V(a) = V_d(a)$ and is bounded. Hence, according to Lemma 6, $X_1^{(n)}$ is sum of the integral of a Poisson impulsive white noise, plus a constant. Therefore, the calculation of finding $\zeta(n)$ is completely similar to (24) in the discrete-continuous case. Thus, the theorem is proved. □

C. *Proof of Proposition 1*

From Lemma 9 we obtain that

$$X_1^{(n)} \stackrel{d}{=} \frac{X_0 - b_n}{\sqrt[\alpha]{n}},$$

where b_n is defined in Lemma 9. Hence,

$$\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right) = \mathrm{H}\left(\left[\frac{X_0 - b_n}{\sqrt[\alpha]{n}}\right]_m\right) \tag{29}$$

Therefore, from Lemma 10 we obtain that

$$\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right) = \mathrm{H}\left(\left[X_0 - b_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right). \tag{30}$$

Lemma 8 shows that the distribution of X_0 satisfies the properties of Proposition 6. Thus, from Proposition 6 we obtain that when $m \geq m(n)$ and $\frac{m(n)}{\sqrt[\alpha]{n}}$ tends to ∞, it can be shown that

$$\begin{aligned}&\lim_{n\to\infty} \sup_{m:m\geq m(n)} \left|\mathrm{H}\left(\left[X_0 - b_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right) - \log\frac{m}{\sqrt[\alpha]{n}} - \mathrm{h}\left(X_0\right)\right| \\ &\quad = 0.\end{aligned} \tag{31}$$

Statement of the proposition follows from (29)-(31). □

D. *Proof of Proposition 2*

From Lemma 6 we obtain that, for an impulsive Poisson white noise $X(t)$, $X_0 := \int_0^1 X(t)\,\mathrm{d}t$ is a discrete-continuous random variable; as a result, utilizing Lemma 7, $X_1^{(n)}$ is also discrete-continuous. In order to find the probability of $X_1^{(n)}$ being continuous, observe that, in (86) in Proposition 4, p_{A_n} is an absolutely continuous distribution due to (87), the fact that p_A is absolutely continuous, and Lemma 11. Hence, from Theorem 2, we obtain that $\Pr\left\{X_1^{(n)} \text{ is discrete} = \mathrm{e}^{-\frac{\lambda}{n}}\right\}$, thereby

$$\kappa(n) = n\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right). \tag{32}$$

In order to find $\zeta(n)$ and $m(n)$ we need to write $\mathcal{H}_{m,n}(X)$ as following. Let $P_{X_1^{(n)};m}$ be the pmf of $\left[X_1^{(n)}\right]_m$ as in Definition 6. Therefore,

$$\begin{aligned} n\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right) &= n\sum_{i\in\mathbb{Z}} P_{X_1^{(n)};m}[i]\log\frac{1}{P_{X_1^{(n)};m}[i]} \\ &= nP_{X_1^{(n)};m}[0]\log\frac{1}{P_{X_1^{(n)};m}[0]} \\ &\quad + n\sum_{i\in\mathbb{Z}\setminus\{0\}} P_{X_1^{(n)};m}[i]\log\frac{1}{P_{X_1^{(n)};m}[i]}. \end{aligned}$$

Proving the following limits will imply the first statement of the proposition:

$$\lim_{m,n\to\infty}\frac{n}{\kappa(n)}P_{X_1^{(n)};m}[0]\log\frac{1}{P_{X_1^{(n)};m}[0]} = 1, \tag{33}$$

$$\begin{aligned} \lim_{n\to\infty}\sup_{m:m\geq m(n)}\Bigg|&\frac{n}{\kappa(n)}\sum_{i\in\mathbb{Z}\setminus\{0\}} P_{X_1^{(n)};m}[i]\log\frac{1}{P_{X_1^{(n)};m}[i]} \\ &-\log(mn) - \mathrm{h}\left(A\right) + \log\lambda\Bigg| = 0, \end{aligned} \tag{34}$$

where $\mathrm{h}\left(A\right)$ is the differential entropy of random variable A. Proving the above equalities, will imply the statement of the proposition.

Proof of (33)*:* Observe that from (32) we have that

$$\lim_{n\to\infty}\kappa(n) = \lim_{n\to\infty} n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right) = \lambda. \tag{35}$$

Therefore, it is enough to only prove

$$\lim_{m,n\to\infty} nP_{X_1^{(n)};m}[0]\log\frac{1}{P_{X_1^{(n)};m}[0]} = \lambda. \tag{36}$$

From Proposition 4, we obtain that

$$P_{X_1^{(n)};m}[0] = \mathrm{e}^{-\frac{\lambda}{n}} + (1-\mathrm{e}^{-\frac{\lambda}{n}})P_{A_n;m}[0], \tag{37}$$

for some random variable A_n with features mentioned in (87)-(89). For every m we have

$$\int_{-\frac{1}{2m}}^{\frac{1}{2m}} |p_{A_n}(x) - p_A(x)|\,\mathrm{d}x \leq \int_{\mathbb{R}} |p_{A_n}(x) - p_A(x)|\,\mathrm{d}x.$$

Hence,

$$\begin{aligned} \int_{-\frac{1}{2m}}^{\frac{1}{2m}} p_{A_n}(x)\,\mathrm{d}x &\leq \int_{-\frac{1}{2m}}^{\frac{1}{2m}} p_A(x)\,\mathrm{d}x \\ &\quad + \int_{\mathbb{R}} |p_{A_n}(x) - p_A(x)|\,\mathrm{d}x \\ &\leq \frac{\ell}{m} + 2\frac{\mathrm{e}^{\frac{\lambda}{n}} - \frac{\lambda}{n} - 1}{\mathrm{e}^{\frac{\lambda}{n}} - 1}, \end{aligned} \tag{38}$$

where (38) is true due to (88), and the fact that $p_A(x) \leq \ell$ almost everywhere. Thus, we obtain that

$$\lim_{m,n\to\infty}\int_{-\frac{1}{2m}}^{\frac{1}{2m}} p_{A_n}(x)\,\mathrm{d}x = 0.$$

Therefore, from the definition of $P_{A_n;m}[0]$ defined in Definition 6, we obtain that

$$\lim_{m,n\to\infty} P_{A_n;m}[0] = 0. \tag{39}$$

Therefore, from (37), we can write that

$$\lim_{m,n\to\infty} P_{X_1^{(n)};m}[0] = 1. \tag{40}$$

Hence, if we prove that

$$\lim_{m,n\to\infty} n\log\frac{1}{P_{X_1^{(n)};m}[0]} = \lambda, \tag{41}$$

we can conclude from (40) that

$$\begin{aligned} \lim_{m,n\to\infty} nP_{X_1^{(n)};m}[0]\log\frac{1}{P_{X_1^{(n)};m}[0]} &= \lim_{m,n\to\infty} n\log\frac{1}{P_{X_1^{(n)};m}[0]} \\ &= \lambda. \end{aligned}$$

This will complete the proof of (36). Thus, it only remains to prove (41). Again, by substituting the value of $P_{X_1;m}[0]$ from (37), we have

$$\begin{aligned} &\lim_{m,n\to\infty} n\log\frac{1}{P_{X_1^{(n)};m}[0]} \\ &= -\lim_{m,n\to\infty} n\log\left(\mathrm{e}^{-\frac{\lambda}{n}} + (1-\mathrm{e}^{-\frac{\lambda}{n}})P_{A_n;m}[0]\right) \\ &= -\lim_{m,n\to\infty} n\log\left(1 + (\mathrm{e}^{-\frac{\lambda}{n}} - 1)\left(1 - P_{A_n;m}[0]\right)\right) \\ &= -\lim_{m,n\to\infty} n(\mathrm{e}^{-\frac{\lambda}{n}} - 1)\left(1 - P_{A_n;m}[0]\right) \qquad (42) \\ &= -\lim_{m,n\to\infty} n(\mathrm{e}^{-\frac{\lambda}{n}} - 1) \\ &= \lambda. \qquad (43) \end{aligned}$$

where (42) follows from Taylor series of $\log(1+x)$ near 0 for $x = (\mathrm{e}^{-\lambda/n} - 1)\left(1 - P_{A_n;m}[0]\right)$; observe that $\mathrm{e}^{-\lambda/n} - 1$ is close to zero for large values of n, and (43) is obtained from (39).

Proof of (34)*:* From Proposition 4 we obtain that

$$P_{X_1^{(n)};m}[i] = \left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)P_{A_n;m}[i], \qquad \forall i\in\mathbb{Z}\setminus\{0\}.$$

By substituting the value of $P_{X_1^{(n)};m}[i]$ in terms of $P_{A_n;m}[i]$, we have

$$\begin{aligned} &n\sum_{i\in\mathbb{Z}\setminus\{0\}} P_{X_1^{(n)};m}[i]\log\frac{1}{P_{X_1^{(n)};m}[i]} \\ &= n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)\left(1 - P_{A_n;m}[0]\right)\log\frac{1}{1-\mathrm{e}^{-\frac{\lambda}{n}}} \\ &\quad - n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)P_{A_n;m}[0]\log\frac{1}{P_{A_n;m}[0]} \\ &\quad + n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)\sum_{i\in\mathbb{Z}} P_{A_n;m}[i]\log\frac{1}{P_{A_n;m}[i]}. \end{aligned}$$

Therefore, in order to prove (34), utilizing (32), and it suffices to show that

$$\lim_{m,n\to\infty} P_{A_n;m}[0]\log P_{A_n;m}[0] = 0, \tag{44}$$

$$\begin{aligned} \lim_{n\to\infty}\sup_{m:m\geq m(n)}&\left|\left(1 - P_{A_n;m}[0]\right)\log\frac{1}{1-\mathrm{e}^{-\frac{\lambda}{n}}}\log n + \log\lambda\right| \\ &= 0, \end{aligned} \tag{45}$$

$$\lim_{m,n\to\infty}\sum_{i\in\mathbb{Z}} P_{A_n;m}[i]\log\frac{1}{P_{A_n;m}[i]} - \log m = \mathrm{h}\left(A\right). \tag{46}$$

Proof of (44): It is obtained from (39).

Proof of (45): Let us add and subtract the term $(1 - P_{A_n;m}[0]) \log n$ inside of the absolute value of the left hand side of (45). Hence, we have

$$\begin{aligned}&(1 - P_{A_n;m}[0]) \log \tfrac{1}{1-\mathrm{e}^{-\frac{\lambda}{n}}} - \log n \\ &\quad = (1 - P_{A_n;m}[0]) \log \tfrac{1}{n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)} - P_{A_n;m}[0] \log n,\end{aligned}$$

From (39), and (35), we can write

$$\lim_{m,n\to\infty} (1 - P_{A_n;m}[0]) \log \tfrac{1}{n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)} = -\log \lambda.$$

Hence, we only need to show that

$$\lim_{n\to\infty} \sup_{m: m\geq m(n)} P_{A_n;m}[0] \log n = 0. \tag{47}$$

For (47), from (38), we have that

$$0 \leq P_{A_n;m}[0] \log n \leq \left(\frac{\ell}{m} + 2\frac{\mathrm{e}^{\frac{\lambda}{n}} - \frac{\lambda}{n} - 1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \right) \log n.$$

Thus, (47) is proved due to the choice of $m(n)$, which causes $(\log n)/m$ tends to 0, and the fact that

$$\lim_{n\to\infty} \frac{\mathrm{e}^{\frac{\lambda}{n}} - \frac{\lambda}{n} - 1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \log n = 0. \tag{48}$$

Proof of (46)*:* Remember that in Definition 6, for every arbitrary random variable X, a random variable $\widetilde{X}_m$ with an absolutely continuous distribution $q_{X;m}(x)$ was defined. Thus, corresponding to A_n, we can define random variables $\widetilde{A}_{nm}$ with an absolutely continuous distribution $q_{A_n;m}(x)$. Observe that, from Lemma 2,

$$\sum_{i\in\mathbb{Z}} P_{A_n;m}[i] \log \frac{1}{P_{A_n;m}[i]} = \log m + \mathrm{h}\left(\widetilde{A}_{nm}\right), \tag{49}$$

where $\widetilde{A}_{nm}$ is a continuous random variable with pdf $q_{A_n;m}(x)$. Similarly, for random variable A, using Definition 6, we can define continuous random variable $\widetilde{A}_m$ with pdf $q_{A;m}(x)$. Again, from Lemma 2

$$\begin{aligned}\mathrm{H}\left([A]_m\right) &:= \sum_{i\in\mathbb{Z}} P_{A;m}[i] \log \frac{1}{P_{A;m}[i]} \\ &= \log m + \mathrm{h}\left(\widetilde{A}_m\right).\end{aligned} \tag{50}$$

From (50) and Corollary 1, we obtain

$$\mathrm{h}\left(A\right) = \lim_{m\to\infty} \mathrm{h}\left(\widetilde{A}_m\right) \tag{51}$$

Hence, from (49) and (51), we obtain that in order to prove (46), it suffices to show that

$$\lim_{m,n\to\infty} \mathrm{h}\left(\widetilde{A}_{nm}\right) = \lim_{m\to\infty} \mathrm{h}\left(\widetilde{A}_m\right). \tag{52}$$

To show this, we utilize Lemma 1. This theorem reduces convergence in differential entropy to convergence in total variation distance for a restricted class of distributions. In other words, to show (52), it suffices to show

$$\lim_{m,n\to\infty} \int_{\mathbb{R}} |q_{A_n;m}(x) - q_{A;m}(x)| \,\mathrm{d}x = 0, \tag{53}$$

as long as we can show that $q_{A_n;m}(x)$ and $q_{A;m}(x)$ for all $m, n \in \mathbb{N}$ belong to the class of distributions given in Definition 5. Remember that in the statement of the theorem, in equation (14), we had assumed that

$$p_A \in (\alpha, \ell, v)\text{–}\mathcal{AC}, \tag{54}$$

for some positive values for α, ℓ, v. We show that for all $m, n \in \mathbb{N}$

$$q_{A_n;m}(x), q_{A;m}(x) \in (\alpha, \ell, v')\text{–}\mathcal{AC},$$

where $v' = 2v + 3$. In other words, for all $m, n \in \mathbb{N}$ we have

$$q_{A_n;m}(x), q_{A;m}(x) \in \mathcal{AC}, \tag{55}$$

$$q_{A_n;m}(x), q_{A;m}(x) \leq \ell, \qquad \text{(a.e)}, \tag{56}$$

$$\mathbb{E}\left[\left|\widetilde{A}_{nm}\right|^{\alpha}\right], \mathbb{E}\left[\left|\widetilde{A}_m\right|^{\alpha}\right] \leq 2v + 3. \tag{57}$$

As a result, it remains to show (53), (55), (56) and (57).

Proof of (53): From (95) in Proposition 5, for all $m, n \in \mathbb{N}$ we have

$$\int_{\mathbb{R}} |q_{A_n;m}(x) - q_{A;m}(x)| \,\mathrm{d}x \leq \int_{\mathbb{R}} |p_{A_n}(x) - p_A(x)| \,\mathrm{d}x.$$

Thus,

$$\begin{aligned}&\lim_{m,n\to\infty} \int_{\mathbb{R}} |q_{A_n;m}(x) - q_{A;m}(x)| \,\mathrm{d}x \\ &\quad \leq \lim_{n\to\infty} \int_{\mathbb{R}} |p_{A_n}(x) - p_A(x)| \,\mathrm{d}x \\ &\quad = 0.\end{aligned} \tag{58}$$

where (58) follows from Proposition 4. Hence, (53) is proved.

Proof of (55)*:* Since the pdfs $q_{A_n;m}$ and $q_{A;m}$ are combination of step functions, hence, their cdf are absolutely continuous.

Proof of (56): Since there exists some ℓ such that $p_A(x) \leq \ell$ for almost all $x \in \mathbb{R}$, we obtain that

$$\begin{aligned}q_{A;m}(x) &= m P_{A;m}[i] \\ &= m \int_{\frac{i-\frac{1}{2}}{m}}^{\frac{i+\frac{1}{2}}{m}} p_A(x) \,\mathrm{d}x \\ &\leq m \int_{\frac{i-\frac{1}{2}}{m}}^{\frac{i+\frac{1}{2}}{m}} \ell \,\mathrm{d}x \\ &= \ell.\end{aligned}$$

It remains to show that $q_{A_n;m}(x)$ is bounded too. Similar to the above equation, it suffices to show that $p_{A_n}(x) \leq \ell$. From (87) in Proposition 4, we have that

$$\begin{aligned}p_{A_n}(x) &= \frac{\mathrm{e}^{-\frac{\lambda}{n}}}{1 - \mathrm{e}^{-\frac{\lambda}{n}}} \sum_{k=1}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x) \\ &\leq \frac{\mathrm{e}^{-\frac{\lambda}{n}}}{1 - \mathrm{e}^{-\frac{\lambda}{n}}} \sum_{k=1}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \ell \\ &= \ell,\end{aligned} \tag{59}$$

where (59) follows from the fact that p_A is bounded, and Lemma 12.

Proof of (57): From (96) in Proposition 5, we obtain that there exists $M_1 \in \mathbb{N}$ such that for all $m > M_1$ and $n \in \mathbb{N}$ we have

$$\mathbb{E}\left[\left|\widetilde{A}_m\right|^\alpha\right] \leq 2\mathbb{E}\left[|A|^\alpha\right] + 1, \tag{60}$$

$$\mathbb{E}\left[\left|\widetilde{A}_{nm}\right|^\alpha\right] \leq 2\mathbb{E}\left[|A_n|^\alpha\right] + 1. \tag{61}$$

From (60) we obtain that

$$\mathbb{E}\left[\left|\widetilde{A}_m\right|^\alpha\right] \leq 2v + 1 < 2v + 3. \tag{62}$$

From (89) in Proposition 4 we obtain that there exists $N_1 \in \mathbb{N}$ such that for all $n > N_1$ we have

$$\mathbb{E}\left[|A_n|^\alpha\right] \leq \mathbb{E}\left[|A|^\alpha\right] + 1. \tag{63}$$

Therefore, due to (61) and (63), we obtain that

$$m > M_1, n > N_1 \Longrightarrow \mathbb{E}\left[\left|\widetilde{A}_{nm}\right|^\alpha\right] \leq 2v + 3,$$

Thus, (57) is proved. □

E. Proof of Proposition 3

From (7) and using the fact that $Z(t) = X(t) + Y(t)$, we obtain that

$$\mathcal{H}_{m,n}(Z) = n\mathrm{H}\left(\left[Z_1^{(n)}\right]_m\right) = n\mathrm{H}\left(\left[X_1^{(n)} + Y_1^{(n)}\right]_m\right).$$

Therefore, we aim to prove that

$$\lim_{m,n,\frac{m}{\sqrt[\alpha]{n}}\to\infty}\left(\mathrm{H}\left(\left[X_1^{(n)} + Y_1^{(n)}\right]_m\right) - \log\frac{m}{\sqrt[\alpha]{n}}\right) = \mathrm{h}\left(X_0\right). \tag{64}$$

In the proof of Proposition 1, we compute the quantization entropy of a stable white noise process by showing that

$$\lim_{m,n,\frac{m}{\sqrt[\alpha]{n}}\to\infty}\left(\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right) - \log\frac{m}{\sqrt[\alpha]{n}}\right) = \mathrm{h}\left(X_0\right), \tag{65}$$

Here, we essentially want to prove that we can ignore $Y_1^{(n)}$ in the sum $X_1^{(n)} + Y_1^{(n)}$ when we are computing the quantization entropy. To show the proposition, we only need to prove that

$$\lim_{m,n,\frac{m}{\sqrt[\alpha]{n}}\to\infty}\left(\mathrm{H}\left(\left[X_1^{(n)} + Y_1^{(n)}\right]_m\right) - \mathrm{H}\left(\left[X_1^{(n)}\right]_m\right)\right) = 0. \tag{66}$$

From Lemma 9 we obtain that

$$X_1^{(n)} \stackrel{d}{=} \frac{X_0 - b_n}{\sqrt[\alpha]{n}},$$

where b_n is defined in Lemma 9. Hence,

$$\begin{aligned}\mathrm{H}&\left(\left[X_1^{(n)} + Y_1^{(n)}\right]_m\right)\\ &= \mathrm{H}\left(\left[\frac{X_0 - b_n}{\sqrt[\alpha]{n}} + Y_1^{(n)}\right]_m\right)\\ &= \mathrm{H}\left(\left[X_0 - b_n + \sqrt[\alpha]{n}Y_1^{(n)}\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right),\end{aligned} \tag{67}$$

where the last equality follows from Lemma 10. Similarly,

$$\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right) = \mathrm{H}\left(\left[X_0 - b_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right). \tag{68}$$

For some technical reason that will be needed in the proof, we want to show that $-b_n$ in the (67) and (68) can be replaced by a number $r_n < \frac{\sqrt[\alpha]{n}}{m}$ without changing the value of the entropies. Write $-b_n$ as

$$b_n = r_n + k\frac{\sqrt[\alpha]{n}}{m},$$

for some $k \in \mathbb{Z}$ and

$$0 \leq r_n < \frac{\sqrt[\alpha]{n}}{m}. \tag{69}$$

Then,

$$\begin{aligned}\mathrm{H}&\left(\left[X_1^{(n)} + Y_1^{(n)}\right]_m\right)\\ &= \mathrm{H}\left(k + \left[X_0 + r_n + \sqrt[\alpha]{n}Y_1^{(n)}\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right)\\ &= \mathrm{H}\left(\left[X_0 + r_n + \sqrt[\alpha]{n}Y_1^{(n)}\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right).\end{aligned}$$

Similarly, we have

$$\mathrm{H}\left(\left[X_1^{(n)}\right]_m\right) = \mathrm{H}\left(\left[X_0 + r_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right).$$

Let $V_n := X_0 + r_n$ and $W_n := V_n + \sqrt[\alpha]{n}Y_1^{(n)}$. Then, to prove (66), we only need to show that

$$\lim_{m,n,\frac{m}{\sqrt[\alpha]{n}}\to\infty}\left(\mathrm{H}\left(\left[V_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right) - \mathrm{H}\left(\left[W_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right)\right) = 0. \tag{70}$$

From Lemma 8, and Lemma 12 it can be obtained that V_n, W_n are both continuous random variables with pdfs q_{V_n}, q_{W_n}, respectively.

Define $\widetilde{V}_{n,m/\sqrt[\alpha]{n}} \sim q_{V_n;m/\sqrt[\alpha]{n}}$ and $\widetilde{W}_{n,m/\sqrt[\alpha]{n}} \sim q_{W_n;m/\sqrt[\alpha]{n}}$ according to Definition 6 for random variables V_n and W_n, respectively. From Lemma 2, we have that

$$\begin{cases}\mathrm{H}\left(\left[V_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right) - \log\frac{m}{\sqrt[\alpha]{n}} = \mathrm{h}\left(q_{V_n;m/\sqrt[\alpha]{n}}\right),\\ \mathrm{H}\left(\left[W_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right) - \log\frac{m}{\sqrt[\alpha]{n}} = \mathrm{h}\left(q_{W_n;m/\sqrt[\alpha]{n}}\right).\end{cases} \tag{71}$$

Therefore,

$$\begin{aligned}\mathrm{H}&\left(\left[V_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right) - \mathrm{H}\left(\left[W_n\right]_{\frac{m}{\sqrt[\alpha]{n}}}\right)\\ &= \mathrm{h}\left(q_{V_n;m/\sqrt[\alpha]{n}}\right) - \mathrm{h}\left(q_{W_n;m/\sqrt[\alpha]{n}}\right).\end{aligned} \tag{72}$$

To show (70), we need to show that the difference $\mathrm{h}\left(q_{V_n;m/\sqrt[\alpha]{n}}\right) - \mathrm{h}\left(q_{W_n;m/\sqrt[\alpha]{n}}\right)$ vanishes in the limit. To show this, we utilize Lemma 1: we prove existence of constants $0 < \gamma, \ell, v < \infty$ such that for any m and n satisfying $m/\sqrt[\alpha]{n} \geq 1$ we have

$$\int_{\mathbb{R}}\left|q_{V_n;m/\sqrt[\alpha]{n}}(x) - q_{W_n;m/\sqrt[\alpha]{n}}(x)\right|\mathrm{d}x \leq 2\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right), \tag{73}$$

$$q_{V_n;m/\sqrt[\alpha]{n}}, q_{W_n;m/\sqrt[\alpha]{n}} \in (\gamma, \ell, v)\text{–}\mathcal{AC}. \tag{74}$$

Then, Lemma 1 yields

$$\left|\mathrm{h}\left(q_{W_n;m/\sqrt[\alpha]{n}}\right) - \mathrm{h}\left(q_{V_n;m/\sqrt[\alpha]{n}}\right)\right| \leq c_1\Delta + c_2\Delta\log\frac{1}{\Delta}, \tag{75}$$

for some constants c_1 and c_2 (not depending on m and n) and $\Delta = 2\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right)$. The differential entropy difference in

(75) vanishes as Δ vanishes when n goes to infinity. This completes the proof. It only remains to prove (73) and (74).

Proof of (73): Proposition 5 shows that it is enough to prove that

$$\int_{\mathbb{R}} |p_{V_n}(x) - p_{W_n}(x)| \,\mathrm{d}x \le 2\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right).$$

From (86) in Proposition 4 we obtain that $\sqrt[\alpha]{n} Y_1^{(n)} = W_n - V_n$ has the following pdf:

$$\sqrt[\alpha]{n} Y_1^{(n)} \sim \mathrm{e}^{-\frac{\lambda}{n}} \delta(x) + \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) p_{A_n'}(x),$$

where

$$p_{A_n'}(x) = \frac{1}{\sqrt[\alpha]{n}} p_{A_n}\left(\frac{x}{\sqrt[\alpha]{n}}\right)$$

is absolutely continuous. Therefore, p_{W_n} is

$$\begin{aligned} p_{W_n}(x) &= \left(p_{V_n} * p_{\sqrt[\alpha]{n} Y_1^{(n)}}\right)(x) \\ &= \mathrm{e}^{-\frac{\lambda}{n}} p_{V_n}(x) + \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \left(p_{V_n} * p_{A_n'}\right)(x). \end{aligned}$$

Then, in order to find the total variation, we can write

$$\begin{aligned} |p_{V_n}(x) - p_{W_n}(x)| \le &\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) p_{V_n}(x) \\ &+ \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \left(p_{V_n} * p_{A_n'}\right)(x). \end{aligned}$$

Hence, we can write that

$$\begin{aligned} &\int_{\mathbb{R}} |p_{V_n}(x) - p_{W_n}(x)| \,\mathrm{d}x \\ &\quad\le \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \int_{\mathbb{R}} p_{V_n}(x) \,\mathrm{d}x \\ &\quad\quad + \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \int_{\mathbb{R}} \left(p_{V_n} * p_{A_n'}\right)(x) \,\mathrm{d}x \\ &\quad= 2\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right). \end{aligned}$$

Proof of (74): From Definition 6, it is clear that $\widetilde{V}_{n,m/\sqrt[\alpha]{n}}, \widetilde{W}_{n,m/\sqrt[\alpha]{n}}$ have density and are continuous. It remains to prove the other requirements:

- We show that there exists $\ell < \infty$ that $q_{V_n;m/\sqrt[\alpha]{n}}(x), q_{W_n;m/\sqrt[\alpha]{n}}(x) \le \ell$ for all m, n and $x \in \mathbb{R}$.

  Lemma 8 shows that X_0 has a density which is bounded by some $\ell < \infty$. The same bound applies to the density of $V_n := X_0 + r_n$, which is a shifted version of X_0. Utilizing Lemma 12, W_n also has density bounded by $\ell < \infty$. Thus, we obtain that for some $\ell < \infty$ (not depending on n)

  $$q_{V_n}(x), q_{W_n}(x) \le \ell < \infty, \qquad \forall x \in \mathbb{R}.$$

  From the definition of $q_{V_n;m/\sqrt[\alpha]{n}}(x), q_{W_n;m/\sqrt[\alpha]{n}}(x)$ in Definition 6, the densities of $q_{V_n;m/\sqrt[\alpha]{n}}(x)$ and $q_{W_n;m/\sqrt[\alpha]{n}}(x)$ are the average of the densities of $q_{V_n}(x)$ and $q_{W_n}(x)$ over the quantization interval. Therefore, they are also bounded from above by ℓ:

  $$q_{V_n;m/\sqrt[\alpha]{n}}(x), q_{W_n;m/\sqrt[\alpha]{n}}(x) \le \ell < \infty, \qquad \forall x \in \mathbb{R}.$$

- There exists $v < \infty$ and $\gamma > 0$ such that $\mathbb{E}\left[\left|\widetilde{V}_{n;m/\sqrt[\alpha]{n}}\right|^\gamma\right], \mathbb{E}\left[\left|\widetilde{W}_{n;m/\sqrt[\alpha]{n}}\right|^\gamma\right] \le v$ for all m, n satisfying $m/\sqrt[\alpha]{n} \ge 1$:

  Utilizing Proposition 5 and the fact that $m/\sqrt[\alpha]{n} \ge 1$, we have that

  $$\begin{aligned} \mathbb{E}\left[\left|\widetilde{V}_{n;m/\sqrt[\alpha]{n}}\right|^\gamma\right] &\le \left(\frac{2}{\sqrt{m/\sqrt[\alpha]{n}}}\right)^\gamma + \mathrm{e}^{\frac{\gamma}{\sqrt{m/\sqrt[\alpha]{n}}}} \mathbb{E}\left[|V_n|^\gamma\right] \\ &\le 2^\gamma + \mathrm{e}^\gamma \mathbb{E}\left[|V_n|^\gamma\right] \end{aligned}$$

  A similar equation holds for $\mathbb{E}\left[\left|\widetilde{W}_{n;m/\sqrt[\alpha]{n}}\right|^\gamma\right]$. Therefore, we will be done if we can find $v' < \infty$ and $\gamma > 0$ such that for all m, n satisfying $m/\sqrt[\alpha]{n} \ge 1$, we have

  $$\mathbb{E}\left[|V_n|^\gamma\right], \mathbb{E}\left[|W_n|^\gamma\right] < v'.$$

  To this end, we can write

  $$|V_n| = |X_0 + r_n| \le |X_0| + r_n \le 2\max\{|X_0|, r_n\}.$$

  Since $r_n < \frac{\sqrt[\alpha]{n}}{m} \le 1$, we obtain

  $$\begin{aligned} |V_n|^\gamma &\le 2^\gamma \max\{|X_0|^\gamma, r_n^\gamma\} \\ &\le 2^\gamma \left(|X_0|^\gamma + r_n^\gamma\right) \\ &\le 2^\gamma \left(|X_0|^\gamma + 1\right). \end{aligned}$$

  Note that [27, Property 1.2.16]

  $$\mathbb{E}\left[|X_0|^\gamma\right] < \infty, \qquad \forall \gamma < \alpha.$$

  Therefore, the assertion is proved for V_n. Following a similar argument for W_n, we obtain

  $$\begin{aligned} |W_n|^\gamma &\le 3^\gamma \left(|X_0|^\gamma + r_n^\gamma + n^{\frac{\gamma}{\alpha}} |Y_1^{(n)}|^\gamma\right) \\ &\le 3^\gamma \left(|X_0|^\gamma + 1 + n^{\frac{\gamma}{\alpha}} |Y_1^{(n)}|^\gamma\right). \end{aligned}$$

  Expected value of $|X_0|^\gamma$ is finite for any $\gamma < \alpha$. Thus, we only need to find some $\gamma < \alpha$ such that $\mathbb{E}\left[|Y_1^{(n)}|^\gamma\right]$ is finite. According to (86) in Proposition 4, we have

  $$\begin{aligned} \mathbb{E}\left[|Y_1^{(n)}|^\gamma\right] &= \Pr\left\{Y_1^{(n)} = 0\right\} \mathbb{E}\left[|Y_1^{(n)}|^\gamma \middle| Y_1^{(n)} = 0\right] \\ &\quad + \Pr\left\{Y_1^{(n)} \ne 0\right\} \mathbb{E}\left[|Y_n|^\gamma \middle| Y_1^{(n)} \ne 0\right] \\ &= 0 + \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \mathbb{E}\left[|A_n|^\gamma\right] \end{aligned}$$

  Therefore,

  $$n^{\frac{\gamma}{\alpha}} \mathbb{E}\left[|Y_1^{(n)}|^\gamma\right] = n^{\frac{\gamma}{\alpha}} \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \mathbb{E}\left[|A_n|^\gamma\right]$$

  From (89) in Proposition 4, one can find some $\tilde{v} < \infty$ such that $\mathbb{E}\left[|A_n|^\gamma\right] < \tilde{v}$ for all n provided that $\mathbb{E}\left[|A|^\gamma\right] < \infty$. Furthermore, $\sup_{n \in \mathbb{N}} n^{\frac{\gamma}{\alpha}} \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) < \infty$ for any $\gamma < \alpha$. Therefore, we may choose any $0 < \gamma < \alpha$.

Hence, the proposition is proved. □

F. *Proof of Theorem 3*

Proof of item 1: We can write that

$$\frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)} = \frac{\log n}{\kappa_Y(n)(\zeta_Y(n) + \log m)} \frac{\frac{\mathcal{H}_{m,n}(X)}{\log n}}{\frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)(\zeta_Y(n) + \log m)}}.$$

According to Theorem 2, there exists $c_1, c_2 > 0$ such that

$$\zeta_Y(n) \ge c_1 \log n, \qquad \kappa_Y(n) \ge c_2,$$

where the second inequality is obtained from the fact that $\kappa_Y(n) = n\left(1 - \exp\left[-\frac{\lambda_Y}{n}(1-\alpha_Y)\right]\right)$ tends to the constant $\lambda_Y(1-\alpha_Y)$ as n tends to infinity. Therefore,

$$\frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)} \le \frac{\log n}{c_2(c_1 \log n + \log m)} \frac{\frac{\mathcal{H}_{m,n}(X)}{\log n}}{\frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)(\zeta_Y(n)+\log m)}}.$$

Hence, from the assumption of the theorem, we know that $\log m/\log n$ tends to ∞ as n tends to infinity. As a result, it only suffices to prove that

$$\lim_{n\to\infty} \sup_{m\ge m'(n)} \frac{\mathcal{H}_{m,n}(X)}{\log n} < \infty, \tag{76}$$

$$\lim_{n\to\infty} \inf_{m\ge m'(n)} \frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)(\zeta_Y(n)+\log m)} = 1. \tag{77}$$

Equation (76) is immediate from Theorem 2 since for a discrete X_0 we have

$$\lim_{n\to\infty} \sup_{m:\ m\ge m(n)} |\mathcal{H}_{m,n}(X) - \zeta_X(n)| = 0.$$

Furthermore, $\zeta_X(n)$ has a logarithmic growth from the second part of Theorem 2. To show (77), observe that from Theorem 2, we have

$$\lim_{n\to\infty} \sup_{m>m'(n)} \left|\frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)} - \log m - \zeta_Y(n)\right| = 0,$$

which means that for any $\epsilon > 0$, there exists n_1 such that

$$\left|\frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)} - \log m - \zeta_Y(n)\right| \le \epsilon, \quad \forall n > n_1, \forall m > m'(n).$$

In addition, note that $\zeta_Y(n) \ge c_1 \log n$; therefore $\zeta_Y(n) + \log m$ converges to infinity as n converges to infinity. Thus, for any $\epsilon > 0$ there exists n_2 such that

$$\zeta_Y(n) + \log m > \frac{1}{\epsilon}, \qquad \forall n > n_2, \forall m > m'(n).$$

Hence, by taking $n = \max\{n_1, n_2\}$, we have that

$$\left|\frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)(\zeta_Y(n)+\log m)} - \frac{\log m}{\zeta_Y(n)+\log m} - \frac{\zeta_Y(n)}{\zeta_Y(n)+\log m}\right| \le \epsilon^2, \qquad \forall n > n_1, \forall m > m'(n).$$

Since $c_1 \log n \le \zeta_Y(n) \le c_2 \log n$, and $\log m/\log n$ tends to infinity if $m > m'(n)$, we have

$$\begin{aligned}\lim_{n\to\infty} \sup_{m>m'(n)} \frac{\log m}{\zeta_Y(n)+\log m} &= \lim_{n\to\infty} \inf_{m>m'(n)} \frac{\log m}{\zeta_Y(n)+\log m}\\ &= 1.\end{aligned}$$

Likewise,

$$\begin{aligned}\lim_{n\to\infty} \sup_{m>m'(n)} \frac{\zeta_Y(n)}{\zeta_Y(n)+\log m} &= \lim_{n\to\infty} \inf_{m>m'(n)} \frac{\zeta_Y(n)}{\zeta_Y(n)+\log m}\\ &= 0.\end{aligned}$$

Therefore,

$$\begin{aligned}\lim_{n\to\infty} \sup_{m\ge m'(n)} &\frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)(\zeta_Y(n)+\log m)}\\ &= \lim_{n\to\infty} \inf_{m\ge m'(n)} \frac{\mathcal{H}_{m,n}(Y)}{\kappa_Y(n)(\zeta_Y(n)+\log m)}\\ &= 1\end{aligned}$$

Hence, (77) is proved.

Proof of item 2: We can write that

$$\frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)} = \frac{\frac{\mathcal{H}_{m,n}(X)}{n\log m}}{\frac{\mathcal{H}_{m,n}(Y)}{n\log m}}.$$

It suffices to prove that

$$\lim_{n\to\infty} \sup_{m>m'(n)} \frac{\mathcal{H}_{m,n}(X)}{n\log m} = 0, \tag{78}$$

$$\lim_{n\to\infty} \inf_{m>m'(n)} \frac{\mathcal{H}_{m,n}(Y)}{n\log m} = 1. \tag{79}$$

Proof of (78): Note that from Theorem 2, we obtain that for any $\epsilon > 0$, there exists n_1 such that

$$\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_X(n)} - \log m - \zeta_X(n)\right| \le \epsilon, \quad \forall n > n_1, \forall m > m'(n).$$

In addition, $1/(n\log m) \le \epsilon$ for any $m \ge 2$ and $n \ge \epsilon^{-1}$. Therefore, for any $\epsilon > 0$ we have

$$\left|\frac{\mathcal{H}_{m,n}(X)}{\kappa_X(n) n\log m} - \frac{1}{n} - \frac{\zeta_X(n)}{n\log m}\right| \le \epsilon^2,$$

for all $n > n_2$ and $m > m'(n)$, where $n_2 = \max(n_1, \epsilon^{-1})$. Utilizing the fact that from Theorem 2, $\zeta_X(n) \le c_2 \log n$, we have

$$\begin{aligned}\lim_{n\to\infty} \sup_{m>m'(n)} \frac{\zeta_X(n)}{n\log m} &\le \lim_{n\to\infty} \sup_{m>m'(n)} \frac{c_2 \log n}{n\log m}\\ &\le \lim_{n\to\infty} \frac{c_2\log n}{n}\\ &= 0.\end{aligned}$$

As a result,

$$\lim_{n\to\infty} \sup_{m>m'(n)} \frac{\mathcal{H}_{m,n}(X)}{\kappa_X(n) n\log m} = 0.$$

The value of $\kappa_X(n)$ given in Theorem 2 tends to a constant as n tends to infinity. This completes the proof for (78).

Proof of (79): From the theorem's assumption, the ratio of $\log m'(n)$ and $\zeta_Y(n)$ goes to infinity. We know from Theorem 2 that $\zeta_Y(n)$ is non-decreasing. As a result,

$$\lim_{n\to\infty} \log m'(n) = \infty. \tag{80}$$

Utilizing Theorem 2 and (80), we obtain that for any $\epsilon > 0$, there exists n_1 such that

$$\left|\frac{\mathcal{H}_{m,n}(Y)}{n} - \log m - \zeta_Y(n)\right| \le \epsilon,$$

for all $n > n_1$ and $m > m'(n)$, and

$$\frac{1}{\log(m)} \le \epsilon, \qquad \forall n > n_1, \forall m > m'(n).$$

As a result,

$$\left|\frac{\mathcal{H}_{m,n}(Y)}{n\log m}-1-\frac{\zeta_Y(n)}{\log m}\right|\le\epsilon^2,\quad \forall n>n_1,\forall m>m'(n).$$

We have that $0\le\zeta_Y(n)/\log m\le\zeta_Y(n)/\log m'(n)$ for any $m>m'(n)$. Furthermore, $\zeta_Y(n)/\log m'(n)$ vanishes as n tends to infinity according to the theorem's assumption. Thus,

$$\lim_{n\to\infty}\sup_{m\ge m'(n)}\frac{\mathcal{H}_{m,n}(Y)}{n\log m}=\lim_{n\to\infty}\inf_{m\ge m'(n)}\frac{\mathcal{H}_{m,n}(Y)}{n\log m}=1.$$

Hence, the statement is proved.

Proof of item 3: We can write

$$\begin{aligned}\frac{\mathcal{H}_{m,n}(X)}{\mathcal{H}_{m,n}(Y)}-1=&\frac{\frac{\mathcal{H}_{m,n}(X)}{n}-\log m-\zeta_X(n)}{\frac{\mathcal{H}_{m,n}(Y)}{n}}\\&-\frac{\frac{\mathcal{H}_{m,n}(Y)}{n}-\log\frac{m}{\sqrt{n}}-\mathrm{h}\left(Y_0\right)}{\frac{\mathcal{H}_{m,n}(Y)}{n}}\\&+\frac{\zeta_X(n)-\mathrm{h}\left(X_0\right)-\log\frac{1}{\sqrt{n}}}{\frac{\mathcal{H}_{m,n}(Y)}{n}}\\&+\frac{\mathrm{h}\left(X_0\right)}{\frac{\mathcal{H}_{m,n}(Y)}{n}}-\frac{\mathrm{h}\left(Y_0\right)}{\frac{\mathcal{H}_{m,n}(Y)}{n}}.\end{aligned}$$

In order to prove the theorem, it suffices to show that

$$\lim_{n\to\infty}\sup_{m>m(n)}\left|\frac{\mathcal{H}_{m,n}(X)}{n}-\log m-\zeta_X(n)\right|=0, \tag{81}$$

$$\lim_{n\to\infty}\sup_{m>m(n)}\left|\frac{\mathcal{H}_{m,n}(Y)}{n}-\log\frac{m}{\sqrt{n}}-\mathrm{h}\left(Y_0\right)\right|=0, \tag{82}$$

$$\lim_{n\to\infty}\inf_{m>m(n)}\frac{\mathcal{H}_{m,n}(Y)}{n}=\infty, \tag{83}$$

$$\limsup_{n\to\infty}\zeta_X(n)-\mathrm{h}\left(X_0\right)-\log\frac{1}{\sqrt{n}}\le 0. \tag{84}$$

From Theorem 2 and Proposition 1, the limits of (81) and (82) are proved, respectively.

To prove (83), observe that from Proposition 1

$$\frac{\mathcal{H}_{m,n}(Y)}{n}=\left[\frac{\mathcal{H}_{m,n}(Y)}{n}-\log\frac{m}{\sqrt{n}}\right]+\log\frac{m}{\sqrt{n}},$$

where the first expression converges to $\mathrm{h}\left(Y_0\right)$ according to Proposition 1. The second one tends to infinity because of the assumption of the theorem.

Finally, (84) is the direct consequence of Lemma 13.

Hence, the theorem is proved. □

G. Proof of Theorem 4

Case 1: Stable

Proof of (15): According to Proposition 1, observe that for $i=1,2$

$$\lim_{n\to\infty}\sup_{m\ge m(n)}\log\frac{m}{\sqrt[\alpha_i]{n}}=+\infty.$$

Therefore, we can write

$$\lim_{n\to\infty}\sup_{m\ge m(n)}\left|\frac{\mathcal{H}_{m,n}(X_i)}{n\log\frac{m}{\sqrt[\alpha_i]{n}}}-1\right|=0.$$

Note that

$$\frac{\mathcal{H}_{m,n}(X_1)}{\mathcal{H}_{m,n}(X_2)}=\frac{\frac{\mathcal{H}_{m,n}(X_1)}{n\log(m/\sqrt[\alpha_1]{n})}}{\frac{\mathcal{H}_{m,n}(X_2)}{n\log(m/\sqrt[\alpha_2]{n})}}\frac{\log(m/\sqrt[\alpha_1]{n})}{\log(m/\sqrt[\alpha_2]{n})}.$$

Since the limit of the second fraction depends on how m changes as n tends to ∞, it does not have a limit. However, due to $\alpha_1<\alpha_2$ we can write

$$\frac{\log(m/\sqrt[\alpha_1]{n})}{\log(m/\sqrt[\alpha_2]{n})}<1.$$

As a result, the statement is proved.

Proof of (16): We can write

$$\begin{aligned}&\mathcal{H}_{m,n}(X_1)-\mathcal{H}_{m,n}(X_2)\\&\quad=n\left[\frac{\mathcal{H}_{m,n}(X_1)}{n}-\log\frac{m}{\sqrt[\alpha_1]{n}}-\frac{\mathcal{H}_{m,n}(X_2)}{n}+\log\frac{m}{\sqrt[\alpha_2]{n}}\right]\\&\qquad+n\log\frac{\sqrt[\alpha_2]{n}}{\sqrt[\alpha_1]{n}}.\end{aligned}$$

From Proposition 1 we obtain that there exists n_0 such that for $n>n_0$ we have

$$\frac{\mathcal{H}_{m,n}(X_1)}{n}-\log\frac{m}{\sqrt[\alpha_1]{n}}-\frac{\mathcal{H}_{m,n}(X_2)}{n}+\log\frac{m}{\sqrt[\alpha_2]{n}}\ge C,$$

where $C:=\mathrm{h}\left(X_0^{(\alpha_1)}\right)-\mathrm{h}\left(X_0^{(\alpha_2)}\right)-1$. Therefore, for $n>n_0$, we have that

$$\mathcal{H}_{m,n}(X_1)-\mathcal{H}_{m,n}(X_2)\ge n\left(C+\log\frac{\sqrt[\alpha_2]{n}}{\sqrt[\alpha_1]{n}}\right).$$

Hence, $\mathcal{H}_{m,n}(X_1)-\mathcal{H}_{m,n}(X_2)\to+\infty$.

Case 2: Impulsive Poisson

Proof of (17): Observe that for $i=1,2$ we have

$$\lim_{n\to\infty}\sup_{m\ge m(n)}n\left(1-\mathrm{e}^{-\lambda_i/n}\right)\log(mn)=+\infty.$$

Hence, from Proposition 2, we can write

$$\lim_{n\to\infty}\sup_{m\ge m(n)}\left|\frac{\mathcal{H}_{m,n}(X_i)}{n\left(1-\mathrm{e}^{-\lambda_i/n}\right)\log(mn)}-1\right|=0.$$

Note that

$$\frac{\mathcal{H}_{m,n}(X_1)}{\mathcal{H}_{m,n}(X_2)}=\frac{\frac{\mathcal{H}_{m,n}(X_1)}{n\left(1-\mathrm{e}^{-\lambda_1/n}\right)\log(mn)}}{\frac{\mathcal{H}_{m,n}(X_2)}{n\left(1-\mathrm{e}^{-\lambda_2/n}\right)\log(mn)}}\frac{1-\mathrm{e}^{-\frac{\lambda_1}{n}}}{1-\mathrm{e}^{-\frac{\lambda_2}{n}}}.$$

Hence, we obtain that

$$\lim_{n\to\infty}\sup_{m\ge m(n)}\frac{\mathcal{H}_{m,n}(X_1)}{\mathcal{H}_{m,n}(X_2)}=\frac{\lambda_1}{\lambda_2}<1.$$

Proof of (18): From Proposition 2 and the fact that

$$\lim_{n\to\infty}n\left(1-\mathrm{e}^{-\frac{\lambda}{n}}\right)=\lambda,$$

we can write for $i=1,2$ that

$$\begin{aligned}&\lim_{n\to\infty}\sup_{m\ge m(n)}\mathcal{H}_{m,n}(X_i)-n\left(1-\mathrm{e}^{-\frac{\lambda_i}{n}}\right)\log(mn)\\&\quad=\lim_{n\to\infty}\inf_{m\ge m(n)}\mathcal{H}_{m,n}(X_i)-n\left(1-\mathrm{e}^{-\frac{\lambda_i}{n}}\right)\log(mn)\\&\quad=\lambda_i(\mathrm{h}\left(A_i\right)+\log\lambda_i-1).\end{aligned}$$

Then, we can write that

$$\begin{aligned}&\mathcal{H}_{m,n}(X_1)-\mathcal{H}_{m,n}(X_2)\\&\quad=\left[\mathcal{H}_{m,n}(X_1)-n\left(1-\mathrm{e}^{-\frac{\lambda_1}{n}}\right)\log(mn)\right]\\&\quad\quad-\left[\mathcal{H}_{m,n}(X_2)-n\left(1-\mathrm{e}^{-\frac{\lambda_2}{n}}\right)\log(mn)\right]\\&\quad\quad+n\left(\mathrm{e}^{-\frac{\lambda_2}{n}}-\mathrm{e}^{-\frac{\lambda_1}{n}}\right)\log(mn).\end{aligned}$$

Therefore, the first two parts are converging to a constant. For the last one, note that there exists n_0 such that for $n > n_0$ we have that while the last one converges to $-\infty$. Therefore, the theorem is proved. □

H. Proof of Lemma 6

From Definition 9 and Lemma 4 we obtain that for any function $\varphi(t) \in \mathcal{S}(\mathbb{R})$

$$\begin{aligned}&\mathbb{E}\left[\mathrm{e}^{\mathrm{j}\langle X,\varphi\rangle}\right]\\&\quad=\exp\Bigg(j\mu\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t+\int_{\mathbb{R}\setminus\{0\}}\int\Big(\mathrm{e}^{\mathrm{j}\varphi(t)a}-1\\&\qquad\qquad-\mathrm{j}\varphi(t)a\mathbb{1}_{(-1,1)}(a)\,\mathrm{d}t\Big)v(a)\,\mathrm{d}a\Bigg)\\&\quad=\exp\Bigg(j\mu'\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t\\&\qquad+\lambda\int_{\mathbb{R}\setminus\{0\}}\int\left(\mathrm{e}^{\mathrm{j}\varphi(t)a}-1\right)p_A(a)\,\mathrm{d}t\,\mathrm{d}a\Bigg)\\&\quad=\exp\left(j\mu'\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t\right)\\&\qquad\times\exp\left(\lambda\int_{\mathbb{R}\setminus\{0\}}\int\left(\mathrm{e}^{\mathrm{j}\varphi(t)a}-1\right)p_A(a)\,\mathrm{d}t\,\mathrm{d}a\right).\end{aligned}$$

Hence, from Definition 12, the assertion is proved. □

V. Conclusions and Future Works

In this paper, a definition of quantization entropy for random processes based on quantization in the time and amplitude domains was given. The criterion was applied to a wide class of white noise processes, including stable and impulsive Poisson innovation processes. It was shown that the stable has a higher growth rate of entropy compared to the impulsive Poisson process. In our study, we assumed that the amplitude quantization steps $1/m$ is shrinking sufficiently fast with respect to time quantization steps $1/n$, *i.e.,* m is larger than $m(n)$ for some function $m : \mathbb{N} \to \mathbb{N}$. As a future work, it would be interesting to look at cases where m is restricted grow slowly with n. Characterization of the entropy for other stochastic processes is also left as a future work. Finally, we noted that the asymptotic value of $\kappa(n)/n$ in (8) is a generalization of the concept of Rényi entropy dimension to random processes. Information dimension is known to admit an operational interpretation in the lossy source coding problem [2], [3]. It would be interesting to find a similar operational interpretation of our results from the perspective of lossless source coding or lossy source coding of this type of process.

References

[1] A. Rényi, "On the dimension and entropy of probability distributions," *Acta Mathematica Academiae Scientiarum Hungarica*, vol. 10, no. 1-2, pp. 193–215, Mar. 1959.

[2] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3721–3748, Aug. 2010.

[3] S. Jalali and H. V. Poor, "Universal compressed sensing for almost lossless recovery," *IEEE Transactions on Information Theory*, vol. 63, no. 5, pp. 2933–2953, May 2017.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] E. Candès, "Compressive sampling," in *Proc. International Congress of Mathematicians*, vol. 3, 2006, pp. 1433–1452.

[6] R. Baraniuk, V. Cevher, and M. Wakin, "Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective," *Proc. of the IEEE*, vol. 98, no. 6, pp. 959–971, 2010.

[7] R. A. DeVore, "Nonlinear approximation," *Acta numerica*, vol. 7, pp. 51–150, 1998.

[8] M. Unser, P. Tafti, and Q. Sun, "A unified formulation of Gaussian versus sparse stochastic processes—Part I: Continuous-domain theory," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1945–1962, 2014.

[9] M. Unser, P. Tafti, A. Amini, and H. Kirshner, "A unified formulation of Gaussian versus sparse stochastic processes—Part II:Discrete-domain theory," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 3036–3051, 2014.

[10] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, 2009.

[11] V. Cevher, "Learning with compressible priors," in *Proc. Adv. in Neural Inform. Process. Sys. (NIPS)*, 2009, pp. 261–269.

[12] A. Amini, M. Unser, and F. Marvasti, "Compressibility of deterministic and random infinite sequences," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5193–5201, Nov. 2011.

[13] J. F. Silva and E. Pavez, "Compressibility of infinite sequences and its interplay with compressed sensing recovery," in *Proc. Sig. & Inform. Proc. Assoc. Annual Summit and Conf. (APSIPA ASC), 2012 Asia-Pacific*, Dec. 2012.

[14] R. Gribonval, V. Cevher, and M. E. Davies, "Compressible distributions for high-dimensional statistics," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5016–5034, 2012.

[15] J. F. Silva and M. S. Derpich, "On the characterization of ℓ_p-compressible ergodic sequences," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2915–2928, Jun. 2015.

[16] M. Unser and P. D. Tafti, *An introduction to sparse stochastic processes*. Cambridge University Press, 2014.

[17] A. Amini, U. S. Kamilov, and M. Unser, "The analog formulation of sparsity implies infinite divisibility and rules out Bernoulli-Gaussian priors," in *Inform. Theo. Workshop (ITW), 2012 IEEE*. IEEE, 2012, pp. 682–686.

[18] A. Amini and M. Unser, "Sparsity and infinite divisibility," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2346–2358, Apr. 2014.

[19] J. Fageot, M. Unser, and J. P. Ward, "The n-term approximation of periodic generalized Lévy processes," *arXiv preprint arXiv:1702.03335*, 2017.

[20] G. Lorentz, "Metric entropy and approximation," *Bulletin of the American Mathematical Society*, vol. 72, no. 6, pp. 903–937, 1966.

[21] E. C. Posner and E. R. Rodemich, "Epsilon entropy of stochastic processes with continuous paths," *The Annals of Probability*, pp. 674–689, Aug. 1973.

[22] S. Ihara, *Information theory for continuous systems*. Singapore: World Scientific, 1993.

[23] C. Nair, B. Prabhakar, and D. Shah, "On entropy for mixtures of discrete and continuous variables," *arXiv preprint cs/0607075*, 2006.

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

[25] H. Ghourchian, A. Gohari, and A. Amini, "Existence and continuity of differential entropy for a class of distributions," *IEEE Comm. Letters*, vol. 21, no. 7, pp. 1469–1472, 2017.

[26] I. Gelfand and N. Y. Vilenkin, *Generalized Functions: Applications of Harmonic Analysis*. Academic Press, 1964, vol. 3.

[27] G. Samoradnitsky and M. S. Taqqu, *Stable non-Gaussian random processes: stochastic models with infinite variance*. New York: Chapman & Hall, 1994.

[28] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999. [Online]. Available: https://books.google.com/books?id=tbZPLquJjSoC

[29] H. Tucker, "Absolute continuity of infinitely divisible distributions," *Pacific Journal of Mathematics*, vol. 12, no. 3, pp. 1125–1129, 1962.

[30] J. Blum and M. Rosenblatt, "On the structure of infinitely divisible distributions," *Pacific Journal of Mathematics*, vol. 9, no. 1, pp. 1–7, 1959.

[31] J. Fageot and T. Humeau, "Unified view on lévy white noises: General integrability conditions and applications to linear spde," *arXiv preprint arXiv:1708.02500*, 2017.

Appendix A
Some Useful Propositions and Lemmas

In this section, we provide some lemma and propositions that are utilized in the proof of the main results of the paper. In particular, in Appendix A-A, we provide a useful proposition for Poisson white noises as well as a proposition and a lemma for stable white noises. In Appendix A-B, we consider quantization of a random variable in the amplitude domain and its entropy. In Appendix A-C, we provide a result for the sum of independent discrete-continuous random variables. In Appendix A-D, we state a lemma which is used later to prove the fact that the Gaussian white noise is the least compressible among the white noises we have considered in this paper.

A. Poisson and Stable White Noises

The following proposition states a feature of integrated Poisson white noise in a small interval.

Proposition 4. *Let $X(t)$ be an impulsive Poisson white noise with rate λ and amplitude pdf $p_A \in \mathcal{AC}$. Define Y_n as*

$$Y_n = \langle X, \phi(nt) \rangle = \int_0^{\frac{1}{n}} X(t)\,\mathrm{d}t. \tag{85}$$

Then, we have that

$$p_{Y_n}(x) = \mathrm{e}^{-\frac{\lambda}{n}}\delta(x) + \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) p_{A_n}(x), \tag{86}$$

$$p_{A_n}(x) = \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=1}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x), \tag{87}$$

where p_{Y_n} is the probability density function of Y_n, $*$ is the convolution operator, and A_n is a random variable with probability density function p_{A_n}. Further, we have that

$$\int_{\mathbb{R}} |p_{A_n}(x) - p_A(x)|\,\mathrm{d}x \leq 2\frac{\mathrm{e}^{\frac{\lambda}{n}} - \frac{\lambda}{n} - 1}{\mathrm{e}^{\frac{\lambda}{n}} - 1}, \tag{88}$$

$$\begin{cases} \mathbb{E}\left[|A_n|^{\alpha}\right] \text{ exists} & \forall n \in \mathbb{N}, \\ \lim_{n\to\infty} \mathbb{E}\left[|A_n|^{\alpha}\right] = \mathbb{E}\left[|A|^{\alpha}\right], & \forall \alpha \in (0,\infty), \end{cases} \tag{89}$$

where A is a random variable with the probability density function p_A. Also, the upper bound on the total variation distance in (88) *vanishes as n tends to infinity.*

The main novelties of Proposition 4 are (88) and (89), as (86) and (87) are rather classical results. To make the paper self-contained, we prove all the identities in Appendix B-A.

Lemma 8. *Let X be a stable white noise with parameters $(\alpha, \beta, \sigma, \mu)$. Define random variable X_0 with pdf p_{X_0} as follows:*

$$X_0 = \langle X, \phi \rangle = \int_0^1 X(t)\,\mathrm{d}t,$$

where ϕ is the function defined in (5)*. Then,*

$$p_{X_0} \in \mathcal{AC}, \tag{90}$$

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}} p_{X_0}(x) = L < \infty, \tag{91}$$

$$p_{X_0}(x) \text{ is a piecewise continuous function}, \tag{92}$$

$$\int_{\mathbb{R}} p_{X_0}(x) \left|\log \tfrac{1}{p_{X_0}(x)}\right| \mathrm{d}x < \infty. \tag{93}$$

Moreover, we have that

$$\mathrm{H}\left([X_0]_1\right) < \infty. \tag{94}$$

Except (94), other claims in Lemma 8 are known results in the literature. Again, we include the proof of all parts in Section B-B for the sake of completeness.

Lemma 9. *For a stable white noise with parameters $(\alpha, \beta, \sigma, \mu)$ we have that*

$$X_1^{(n)} \overset{d}{=} \frac{X_0 - b_n}{\sqrt[\alpha]{n}},$$

where $X_1^{(n)}$ was defined in Definition 13,

$$X_0 := \langle X, \phi \rangle = \int_0^1 X(t)\,\mathrm{d}t,$$

and

$$b_n = \begin{cases} \mu\left(1 - \frac{\sqrt[\alpha]{n}}{n}\right) & \alpha \neq 1, \\ \frac{2}{\pi}\sigma\beta\ln n & \alpha = 1. \end{cases}$$

The proof can be found in Appendix B-C

B. Amplitude Quantization

In the following proposition, we show that the total variation distance between two variables decreases by quantizing; furthermore, moments of a quantized random variable tends to the moments of the original random variable, as the quantization step size vanishes.

Proposition 5. *Let random variables $X \sim p_X$ and $Y \sim p_Y$ be continuous, and $[X]_m \sim P_{X;m}$ and $[Y]_m \sim P_{Y;m}$ be their quantized version, defined in Definition 6, respectively. Then for all $m \in (0,\infty)$ we have*

$$\int_{\mathbb{R}} |q_{X;m}(x) - q_{Y;m}(x)|\,\mathrm{d}x \leq \int_{\mathbb{R}} |p_X(x) - p_Y(x)|\,\mathrm{d}x, \tag{95}$$

where $\widetilde{X}_m \sim q_{X;m}$ and $\widetilde{Y}_m \sim q_{Y;m}$ are random variables defined in Definition 6. In addition, for any $\alpha \in (0,\infty)$, and $m \geq 4$ we have

$$\mathbb{E}\left[\left|\widetilde{X}_m\right|^{\alpha}\right] \leq \left(\tfrac{2}{\sqrt{m}}\right)^{\alpha} + \mathrm{e}^{\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^{\alpha}\right], \tag{96}$$

$$\Pr\left\{|X| > \tfrac{1}{\sqrt{m}}\right\} \mathrm{e}^{-2\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^{\alpha}\right] \leq \mathbb{E}\left[\left|\widetilde{X}_m\right|^{\alpha}\right], \tag{97}$$

provided that $\mathbb{E}\left[|X|^{\alpha}\right]$ *exists.*

The proof is given in Appendix B-D.

Corollary 4. *Let $X \sim p_X$ be a continuous random variable, and let $\widetilde{X}_m \sim q_{X;m}$ be the random variable defined in Definition 6. Then, we have that*

$$\lim_{m\to\infty} \mathbb{E}\left[\left|\widetilde{X}_m\right|^{\alpha}\right] = \mathbb{E}\left[|X|^{\alpha}\right],$$

if $\mathbb{E}\left[|X|^{\alpha}\right]$ *exists.*

Proof. Since $p_X \in \mathcal{AC}$, we conclude that $\Pr\{|X| \leq 1/\sqrt{m}\}$ vanishes as m tends to ∞. Hence, the corollary achieved from Proposition 5. □

The following lemma discusses the entropy of quantized version of multiplies of a continuous random variable.

Lemma 10. *Let $X \sim p$ be a continuous random variable, and $m \in (0,\infty)$ be arbitrary. If $\mathrm{H}\left([X]_m\right)$ exists, then for all $a \in (0,\infty)$, we have*

$$\mathrm{H}\left([aX]_{\frac{m}{a}}\right) = \mathrm{H}\left([X]_m\right).$$

The lemma is proved in Appendix B-E.

In the following proposition, we extend Corollary 1 to an arbitrary shifted continuous random variables.

Proposition 6. *Let $X \sim p$ be a continuous random variable with a piecewise continuous pdf $p(x)$. For an arbitrary sequence $\{c_m\}_{m=1}^{\infty} \subset \mathbb{R}$, we have*

$$\lim_{m\to\infty} \mathrm{H}\left([X+c_m]_m\right) - \log m = \mathrm{h}\left(X\right),$$

provided that

$$\mathrm{H}\left([X]_1\right) < \infty,$$
$$\int_{\mathbb{R}} p(x)\left|\log\frac{1}{p(x)}\right| \mathrm{d}x < \infty,$$
$$\operatorname*{ess\,sup}_{x\in\mathbb{R}} p(x) = L < \infty,$$

where $[X]_1$ is the quantized version of X with step size 1.

The proof can be found in Appendix B-F.

C. On Sum of Independent Random Variables

The following lemma is well-known and can be easily proved. Hence, we only mention it without proof.

Lemma 11. *Let X, Y be two independent discrete-continuous random variables where p, q are the probability of X, Y being discrete, respectively. Hence, we can write*

$$X = \begin{cases} X_D & p, \\ X_c & 1-p, \end{cases} \qquad Y = \begin{cases} Y_D & q, \\ Y_c & 1-q, \end{cases}$$

where X_D, Y_D are discrete, and X_c, Y_c are continuous random variables. Therefore, $Z = X+Y$ is also a discrete-continuous random variable such that

$$Z = \begin{cases} X_D + Y_D & pq, \\ X_D + Y_c & p(1-q), \\ X_c + Y_D & (1-p)q, \\ X_c + Y_c & (1-p)(1-q), \end{cases}$$

where the first case makes the discrete part while the other cases make the continuous part of Z.

Proposition 7. *Let X, Y be two independent discrete-continuous random variables where X_D, Y_D are the discrete part, and X_c, Y_c are the continuous part of X and Y, respectively. Then, we have that*

$$\mathrm{h}\left(Z_c\right) \geq \min\{\mathrm{h}\left(X_c\right), \mathrm{h}\left(Y_c\right)\},$$

where $Z = X + Y$.

The proof can be found in Appendix B-G.

Corollary 5. *Using induction, it can be proved that for* i.i.d. *continuous-discrete random variables $X_1, \cdots, X_n$, we have that*

$$\mathrm{h}\left(S_{n,c}\right) \geq \mathrm{h}\left(X_{1,c}\right),$$

where $S_{n,c}$ and $X_{1,c}$ are the corresponding continuous random variables of $X_1 + \cdots + X_n$ and X_1, respectively.

Lemma 12. *Let X be a continuous random variable defined in Definition 1 with probability density function $p_X(x)$ and Y be an arbitrary random variable with probability measure μ_Y, independent of X. Let $Z = X + Y$ with probability measure μ_Z. Then, Z is also continuous random variable with pdf $p_Z(z)$ defined as following:*

$$p_Z(z) = \mathbb{E}\left[p_X\left(z - Y\right)\right].$$

Moreover,

$$p_Z(z) \leq \ell, \qquad \forall z \in \mathbb{R},$$

provided that

$$p_X(x) \leq \ell, \qquad \forall x \in \mathbb{R}.$$

The theorem is proved in Appendix B-H.

D. Maximum of $\zeta(n)$ for a Class of White Noise Processes

In the following lemma we find maximum $\zeta(n)$, defined in Theorem 2 over a class of white noise processes.

Lemma 13. *Let $X(t)$ be a white Lévy noise such that $X_0 := \int_0^1 X(t)\,\mathrm{d}t$ is a continuous random variable (as discussed in Lemma 7). Furthermore, assume that $h(X_0)$ and $\mathrm{H}\left([X_0]_1\right)$ are well-defined. Let $m'(n)$ and $\zeta(n)$ be any function satisfying the statement* (12) *in Theorem 2 for the white noise $X(t)$. Define $m(n)$ an arbitrary function that $m(n) \geq m'(n)$ and $\lim_{n\to\infty} m(n)/\sqrt{n} = \infty$. Then, we have that*

$$\limsup_{n\to\infty} \left(\zeta(n) - \log\frac{1}{\sqrt{n}}\right) \leq \mathrm{h}\left(X_0\right).$$

The lemma is proved in Appendix B-I.

Appendix B
Proof of Some Useful Propositions and Lemmas

A. Proof of Proposition 4

Proof of (86): First, we find the characteristic function of Y_n in terms of the characteristic function of A. From the definition of Y_n in (85), we have that

$$\widehat{p}_{Y_n}(\omega) = \mathbb{E}\left[\mathrm{e}^{\mathrm{j}\omega Y_n}\right] = \mathbb{E}\left[\mathrm{e}^{\mathrm{j}\omega\langle X_n, \phi_{1,n}\rangle}\right],$$

where $\phi_{1,n}$ defined in Definition 13. Due to the definition of white noise in Definition 9, we can write

$$\widehat{p}_{Y_n}(\omega) = \mathbb{E}\left[\mathrm{e}^{\mathrm{j}\langle X_n, \omega\phi_{1,n}\rangle}\right] = \mathrm{e}^{\int_0^{\frac{1}{n}} f(\omega)\,\mathrm{d}t} = \mathrm{e}^{\frac{1}{n}f(\omega)}.$$

Because of the definition of Poisson process in Definition 12, we have that

$$f(\omega) = \lambda \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{j}\omega x} - 1) p_A(x)\,\mathrm{d}x = \lambda \widehat{p}_A(\omega) - \lambda,$$

where $\widehat{p}_A(\omega)$ is the characteristic function of A. Hence, we conclude that

$$\begin{aligned}
\widehat{p}_{Y_n}(\omega) &= \mathrm{e}^{-\frac{\lambda}{n}} \mathrm{e}^{\frac{\lambda}{n}\widehat{p}_A(\omega)} \\
&= \mathrm{e}^{-\frac{\lambda}{n}} \left(1 + \sum_{k=1}^{\infty} \frac{\left(\frac{\lambda}{n}\widehat{p}_A(\omega)\right)^k}{k!}\right) \\
&= \mathrm{e}^{-\frac{\lambda}{n}} + \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) \widehat{p}_{A_n}(\omega), \qquad (98)
\end{aligned}$$

where

$$\widehat{p}_{A_n}(\omega) := \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=1}^{\infty} \frac{\left(\frac{\lambda}{n}\widehat{p}_A(\omega)\right)^k}{k!}. \qquad (99)$$

Therefore, by taking the inverse Fourier transform of (98), (86) is proved.

Proof of (87): It is achieved by taking the inverse Fourier transform from (99).

Proof of (88)*:* We can write

$$\begin{aligned}
p_{A_n}(x) &= \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=1}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x) \\
&= \frac{\frac{\lambda}{n}}{\mathrm{e}^{\frac{\lambda}{n}} - 1} p_A(x) \\
&\quad + \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x).
\end{aligned}$$

Hence, we can write:

$$\begin{aligned}
|p_{A_n}(x) - p_A(x)| \leq{} & \left|\frac{\frac{\lambda}{n}}{\mathrm{e}^{\frac{\lambda}{n}} - 1} - 1\right| p_A(x) \\
& + \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x).
\end{aligned}$$

Since p_A is a probability density, $\overbrace{p_A * \cdots * p_A}^{k}$ is also a probability density; hence,

$$\int_{\mathbb{R}} \big(\overbrace{p_A * \cdots * p_A}^{k}\big)(x)\,\mathrm{d}x = 1.$$

Therefore,

$$\begin{aligned}
\int_{\mathbb{R}} |p_{A_n}(x) - p_A(x)|\,\mathrm{d}x &\leq \left|\frac{\frac{\lambda}{n}}{\mathrm{e}^{\frac{\lambda}{n}} - 1} - 1\right| + \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \\
&= 2\frac{\mathrm{e}^{\frac{\lambda}{n}} - \frac{\lambda}{n} - 1}{\mathrm{e}^{\frac{\lambda}{n}} - 1},
\end{aligned}$$

which vanishes as n tends to ∞.

Proof of (89): Let the sequence of independent and identically distributed random variables $\left\{A^{(i)}\right\}_{i=1}^{\infty}$ have pdf p_A, and let the discrete random variable B be independent to $\left\{A^{(i)}\right\}_{i=1}^{\infty}$ and have the following pmf:

$$\begin{aligned}
P_B[k] &:= \Pr\{B = k\} \\
&= \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \frac{\left(\frac{\lambda}{n}\right)^k}{k!}, \qquad \forall k \in \mathbb{N} := \{1, 2, \cdots\}.
\end{aligned}$$

From (87), we obtain that

$$A_n \overset{d}{=} \sum_{i=1}^{B} A^{(i)}.$$

Now, to find $\mathbb{E}\left[|A_n|^\alpha\right]$, we can write

$$\begin{aligned}
\mathbb{E}\left[|A_n|^\alpha\right] &= \sum_{k=1}^{\infty} P_B[k] \mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^\alpha\right] \\
&= \frac{\frac{\lambda}{n}}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \mathbb{E}\left[|A|^\alpha\right] \\
&\quad + \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^\alpha\right],
\end{aligned}$$

provided that the limit exists. From

$$\lim_{n\to\infty} \frac{\frac{\lambda}{n}}{\mathrm{e}^{\frac{\lambda}{n}} - 1} = 1,$$

we have to prove that

$$\lim_{n\to\infty} \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^\alpha\right] = 0. \qquad (100)$$

In order to prove (100), it suffices to show that

$$\mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^\alpha\right] \text{ exists}, \qquad \mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^\alpha\right] \leq c^k, \qquad (101)$$

where $c = 2^\alpha(1 + \mathbb{E}\left[|A|^\alpha\right])$ is a constant. This is because

$$\begin{aligned}
\frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^\alpha\right] &\leq \frac{1}{\mathrm{e}^{\frac{\lambda}{n}} - 1} \sum_{k=2}^{\infty} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} c^k \\
&= \frac{\mathrm{e}^{\frac{\lambda}{n}c} - \frac{\lambda}{n}c - 1}{\mathrm{e}^{\frac{\lambda}{n}} - 1},
\end{aligned}$$

vanishes as n tends to infinity. Therefore, it only remains to prove (101). We can bound $\left|A^{(1)} + \cdots + A^{(k)}\right|$ as follows:

$$\begin{aligned}
\Big|\sum_{i=1}^{k} A^{(i)}\Big| &\leq \sum_{i=1}^{k} \left|A^{(i)}\right| \\
&\leq 1 + \sum_{i=1}^{k} \left|A^{(i)}\right| \\
&\leq \prod_{i=1}^{k} \left(1 + \left|A^{(i)}\right|\right) \\
&\leq \prod_{i=1}^{k} \left(2\max\left\{1, \left|A^{(i)}\right|\right\}\right).
\end{aligned}$$

By finding the expected value of the α power of both sides, we have

$$\mathbb{E}\left[\Big|\sum_{i=1}^{k} A^{(i)}\Big|^{\alpha}\right] \leq 2^{\alpha k}\,\mathbb{E}\left[\prod_{i=1}^{k} \max\left\{1, \left|A^{(i)}\right|^{\alpha}\right\}\right] \tag{102}$$

$$= 2^{\alpha k} \prod_{i=1}^{k} \mathbb{E}\left[\max\left\{1, \left|A^{(i)}\right|^{\alpha}\right\}\right] \tag{103}$$

$$= \left(2^{\alpha}\,\mathbb{E}\left[\max\left\{1, |A|^{\alpha}\right\}\right]\right)^{k}, \tag{104}$$

where (102) is true because $\max\{1,|x|\}^{\alpha} = \max\{1,|x|^{\alpha}\}$ for all $x \in \mathbb{R}$, (103) is true because $A^{(1)}, \cdots, A^{(n)}$ are independent, and (104) is true because $A^{(1)}, \cdots, A^{(n)}$ are identically distributed. In order to find an upper bound for (104), we can write

$$\mathbb{E}\left[\max\left\{1, |A|^{\alpha}\right\}\right] \leq \mathbb{E}\left[1 + |A|^{\alpha}\right] = 1 + \mathbb{E}\left[|A|^{\alpha}\right].$$

Hence, (101), and as a result, the proposition are proved. □

B. Proof of Lemma 8

Proof of (90)*,* (91)*, and* (92): If we prove that X_0 is a stable random variable, then (90)-(92) are proved. [27] In order to do so, we find the characteristic function of X_0:

$$\widehat{p}_{X_0}(\omega) := \mathbb{E}\left[\mathrm{e}^{\mathrm{j}\omega\langle X,\phi\rangle}\right] = \mathrm{e}^{\int_0^1 f(\omega)\,\mathrm{d}t} = \mathrm{e}^{f(\omega)}, \tag{105}$$

where, $f(\omega)$ is a valid Lévy exponent, defined in (3), and (105) is true because of the definition of stable white noise in Definition 11. Hence, from Definition 10, we obtain that X_0 is a stable random variable.

Proof of (93): Since the functions $x \mapsto p_{X_0}(x)$ and $p \mapsto p|\log(1/p)|$ are continuous, the function $x \mapsto p_{X_0}(x)\,|\log(1/p_{X_0}(x))|$ is also continuous. Thus, for any arbitrary $x_0 > 0$, we have

$$\int_{-x_0}^{+x_0} p_{X_0}(x)\left|\log\tfrac{1}{p_{X_0}(x)}\right| \mathrm{d}x < \infty.$$

Therefore, in order to prove (93), we need to prove the boundedness of the tail of the integral. For sufficiently large x_0, we have

$$p_{X_0}(x) \leq \frac{c}{|x|^{\alpha+1}} \qquad \forall x > x_0, \tag{106}$$

for some positive constant c depending on parameters $(\alpha, \beta, \sigma, \mu)$ [27]. Since $p \mapsto p|\log(1/p)|$ is increasing for $p \in [0, 1/\mathrm{e}]$, it suffices to show that for sufficiently large x_0, the following integral is bounded:

$$\int_{|x|>x_0} \frac{c}{|x|^{\alpha+1}} \log\tfrac{|x|^{\alpha+1}}{c}\,\mathrm{d}x < \infty. \tag{107}$$

By changing variable $y = x/c^{1/(\alpha+1)}$, it suffices to show that

$$\int_{|y|>y_0} \frac{1}{|y|^{\alpha+1}} \log|y|^{\alpha+1}\,\mathrm{d}y < \infty, \tag{108}$$

where $y_0 = x_0/c^{1/(\alpha+1)}$, which holds.

Proof of (94): To show that $\mathrm{H}\left([X_0]_1\right) < \infty$, it suffices to look at the tail of the probability sequence of $[X_0]_1$. From (106), for sufficiently large x_0, we have that for all $|m| > x_0$

$$\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} p_{X_0}(x)\,\mathrm{d}x \leq \int_{m-\frac{1}{2}}^{m+\frac{1}{2}} \frac{a}{|x|^{\alpha+1}}\,\mathrm{d}x < \frac{a}{\left(|m|-\frac{1}{2}\right)^{\alpha+1}}. \tag{109}$$

Hence, it suffices to show that

$$\sum_{|m|-\frac{1}{2}>x_0} \frac{a}{\left(|m|-\frac{1}{2}\right)^{\alpha+1}} \log\frac{\left(|m|-\frac{1}{2}\right)^{\alpha+1}}{a} = 2\sum_{m>x_0+\frac{1}{2}} \frac{a}{\left(m-\frac{1}{2}\right)^{\alpha+1}} \log\frac{\left(m-\frac{1}{2}\right)^{\alpha+1}}{a} < \infty.$$

due that $a/x^{\alpha+1}\log(x^{\alpha+1}/a)$ is a decreasing function for sufficiently large x, we obtain that

$$\sum_{m\geq x_0+2} \frac{a}{\left(m-\frac{1}{2}\right)^{\alpha+1}} \log\frac{\left(m-\frac{1}{2}\right)^{\alpha+1}}{a} < \int_{x>x_0} \frac{a}{x^{\alpha+1}} \log\frac{x^{\alpha+1}}{a}\,\mathrm{d}x < \infty,$$

where the last inequality is true because of (107). □

C. Proof of Lemma 9

Using Definition 11, the characteristic function of $X_1^{(n)}$ is as follows:

$$\widehat{p}_{X_1^{(n)}}(\omega) = \begin{cases} \mathrm{e}^{-\sigma^{\alpha}|\omega|^{\alpha}\left(\frac{1}{n}-\mathrm{j}\frac{1}{n}\beta\,\mathrm{sgn}(\omega)\tan\frac{\pi\alpha}{2}\right)+\mathrm{j}\mu\frac{\omega}{n}} & \alpha \neq 1, \\ \mathrm{e}^{-\sigma|\omega|\left(\frac{1}{n}+\mathrm{j}\frac{1}{n}\beta\frac{2}{\pi}\mathrm{sgn}(\omega)\ln|\omega|\right)+\mathrm{j}\mu\frac{\omega}{n}} & \alpha = 1. \end{cases}$$

From the definition of X_0, we can write

$$\widehat{p}_{X_0}(\omega) = \begin{cases} \mathrm{e}^{-\sigma^{\alpha}|\omega|^{\alpha}\left(1-\mathrm{j}\beta\,\mathrm{sgn}(\omega)\tan\frac{\pi\alpha}{2}\right)+\mathrm{j}\mu\omega} & \alpha \neq 1, \\ \mathrm{e}^{-\sigma|\omega|\left(1+\mathrm{j}\beta\frac{2}{\pi}\mathrm{sgn}(\omega)\ln|\omega|\right)+\mathrm{j}\mu\omega} & \alpha = 1. \end{cases}$$

Thus, we obtain that

$$\widehat{p}_{X_1^{(n)}}(\omega) = \widehat{p}_{X_0}\left(\frac{\omega}{\sqrt[\alpha]{n}}\right)\mathrm{e}^{-\mathrm{j}\omega c_n},$$

where

$$c_n = \begin{cases} \mu\left(\frac{1}{\sqrt[\alpha]{n}} - \frac{1}{n}\right) & \alpha \neq 1, \\ \frac{2}{\pi}\sigma\beta\frac{\ln n}{n} & \alpha = 1. \end{cases}$$

Therefore, $X_1^{(n)}$ can be written with respect to X_0 as follows:

$$X_1^{(n)} \overset{d}{=} \frac{X_0 - b_n}{\sqrt[\alpha]{n}},$$

where $b_n = c_n \sqrt[\alpha]{n}$. □

D. Proof of Proposition 5

Proof of (95): From the definition of $q_{X;m}$ and $q_{Y;m}$, defined in Definition 6, it follows that they are constant in intervals $[(i-1/2)/m,(i+1/2)/m]$ for all $i\in\mathbb{Z}$. Therefore,

$$\begin{aligned}\int_{\frac{i-\frac{1}{2}}{m}}^{\frac{i+\frac{1}{2}}{m}} &|q_{Y;m}(x)-q_{X;m}(x)|\,\mathrm{d}x\\ &=\frac{1}{m}\left|q_{Y;m}\left(\tfrac{i}{m}\right)-q_{X;m}\left(\tfrac{i}{m}\right)\right|\\ &=\left|\int_{\frac{i-\frac{1}{2}}{m}}^{\frac{i+\frac{1}{2}}{m}} p_Y(x)\,\mathrm{d}x-\int_{\frac{i-\frac{1}{2}}{m}}^{\frac{i+\frac{1}{2}}{m}} p_X(x)\,\mathrm{d}x\right| \qquad (110)\\ &\le\int_{\frac{i-\frac{1}{2}}{m}}^{\frac{i+\frac{1}{2}}{m}} |p_Y(x)-p_X(x)|\,\mathrm{d}x,\end{aligned}$$

where (110) is from the definition of q. Hence by summation over $i\in\mathbb{Z}$, we have

$$\int_{\mathbb{R}}|q_{Y;m}(x)-q_{X;m}(x)|\,\mathrm{d}x\le\int_{\mathbb{R}}|p_Y(x)-p_X(x)|\,\mathrm{d}x.$$

Thus, (95) is proved.

Proof of (96) *and* (97): We claim that, it suffices to show that

$$\left|\widetilde{X}_m-X\right|\le\frac{1}{m}. \qquad (111)$$

Because if so, we can write

$$\left||X|-\tfrac{1}{m}\right|^\alpha\le\left|\widetilde{X}_m\right|^\alpha\le\left(|X|+\tfrac{1}{m}\right)^\alpha,$$

and as a result

$$\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\right]\le\mathbb{E}\left[\left|\widetilde{X}_m\right|^\alpha\right]\le\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\right].$$

Therefore, it suffices to show that

$$\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\right]\le\left(\tfrac{2}{\sqrt{m}}\right)^\alpha+\mathrm{e}^{\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^\alpha\right], \qquad (112)$$

$$\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\right]\ge\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathrm{e}^{\frac{-2\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^\alpha\right]. \qquad (113)$$

Proof of (112): By conditioning whether $|X|>1/\sqrt{m}$ or not, we can write

$$\begin{aligned}&\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\right] \qquad (114)\\ &\quad=\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\quad\quad+\Pr\left\{|X|\le\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\middle||X|\le\tfrac{1}{\sqrt{m}}\right]\\ &\quad\le\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\quad\quad+\left(\tfrac{2}{\sqrt{m}}\right)^\alpha, \qquad (115)\end{aligned}$$

where (115) is true since

$$\Pr\left\{|X|\le\tfrac{1}{\sqrt{m}}\right\}\le1,\qquad\tfrac{1}{m}+\tfrac{1}{\sqrt{m}}\le\tfrac{2}{\sqrt{m}}.$$

Note that if $\Pr\left\{|X|>\frac{1}{\sqrt{m}}\right\}=0$, the conditional expected value $\mathbb{E}\big[\left(|X|+\frac{1}{m}\right)^\alpha\big||X|>\frac{1}{\sqrt{m}}\big]$ is not well-defined, but in this case we take the product $\Pr\left\{|X|>\frac{1}{\sqrt{m}}\right\}\mathbb{E}\big[\left(|X|+\frac{1}{m}\right)^\alpha\big||X|>\frac{1}{\sqrt{m}}\big]$ to be zero, without any need for specifying an exact value for $\mathbb{E}\big[\left(|X|+\frac{1}{m}\right)^\alpha\big||X|>\frac{1}{\sqrt{m}}\big]$.

From (115), it only remains to prove that

$$\begin{aligned}\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}&\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\le\mathrm{e}^{\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^\alpha\right].\end{aligned}$$

In order to do that, we can write

$$\left(|X|+\tfrac{1}{m}\right)^\alpha=|X|^\alpha\mathrm{e}^{\alpha\log\left(1+\frac{1}{m|X|}\right)}.$$

Since $|X|>1/\sqrt{m}$, we can write

$$|X|^\alpha\mathrm{e}^{\alpha\log\left(1+\frac{1}{m|X|}\right)}\le|X|^\alpha\mathrm{e}^{\alpha\log\left(1+\frac{1}{\sqrt{m}}\right)}\le|X|^\alpha\mathrm{e}^{\alpha\frac{1}{\sqrt{m}}},$$

where the last inequality is true because $\log(1+x)\le x$. As a result

$$\begin{aligned}\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}&\mathbb{E}\left[\left(|X|+\tfrac{1}{m}\right)^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\le\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[|X|^\alpha\mathrm{e}^{\alpha\frac{1}{\sqrt{m}}}\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\le\mathrm{e}^{\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^\alpha\right].\end{aligned}$$

Proof of (113): Similar to the previous proof, by conditioning whether $|X|>1/\sqrt{m}$ or not, we can write

$$\begin{aligned}&\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\right]\\ &\quad=\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\quad\quad+\Pr\left\{|X|\le\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\middle||X|\le\tfrac{1}{\sqrt{m}}\right]\\ &\quad\ge\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]. \qquad (116)\end{aligned}$$

If $\Pr\{|X|>1/\sqrt{m}\}=0$, (113) is clearly correct. Thus, assume that $\Pr\{|X|>1/\sqrt{m}\}>0$, meaning that $\mathbb{E}\left[||X|-1/m|^\alpha\big||X|>1/\sqrt{m}\right]$ is well-defined. We need to show that

$$\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\ge\mathrm{e}^{-\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^\alpha\right].$$

Observe that when $|x|>1/\sqrt{m}$ and $m\ge4$, we can write

$$\begin{aligned}\left||x|-\tfrac{1}{m}\right|^\alpha&=|x|^\alpha\mathrm{e}^{\alpha\log\left|1-\frac{1}{m|x|}\right|}\\ &\ge|x|^\alpha\mathrm{e}^{\alpha\log\left(1-\frac{1}{\sqrt{m}}\right)}\\ &\ge|x|^\alpha\mathrm{e}^{-\alpha\frac{2}{\sqrt{m}}},\end{aligned}$$

where the last inequality is true because $\log(1-x)\ge-2x$ for $0\le x\le1/2$. As a result,

$$\mathbb{E}\left[\left||X|-\tfrac{1}{m}\right|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\ge\mathrm{e}^{-\frac{\alpha}{\sqrt{m}}}\mathbb{E}\left[|X|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right].$$

Now, we need to show that

$$\mathbb{E}\left[|X|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\ge\mathbb{E}\left[|X|^\alpha\right]. \qquad (117)$$

Without loss of generality, we may assume that $\Pr\{|X|>1/\sqrt{m}\}<1$. Now, we have

$$\begin{aligned}\mathbb{E}\left[|X|^\alpha\right]&=\Pr\left\{|X|>\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[|X|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right]\\ &\quad+\Pr\left\{|X|\le\tfrac{1}{\sqrt{m}}\right\}\mathbb{E}\left[|X|^\alpha\middle||X|\le\tfrac{1}{\sqrt{m}}\right] \qquad (118)\\ &\le\max\left\{\mathbb{E}\left[|X|^\alpha\middle||X|>\tfrac{1}{\sqrt{m}}\right],\right.\\ &\qquad\left.\mathbb{E}\left[|X|^\alpha\middle||X|\le\tfrac{1}{\sqrt{m}}\right]\right\}. \qquad (119)\end{aligned}$$

Since

$$\mathbb{E}\left[|X|^\alpha \middle| |X| \leq \tfrac{1}{\sqrt{m}}\right] \leq \left(\tfrac{1}{\sqrt{m}}\right)^\alpha < \mathbb{E}\left[|X|^\alpha \middle| |X| > \tfrac{1}{\sqrt{m}}\right],$$

we obtain that

$$\begin{aligned}&\max\left\{\mathbb{E}\left[|X|^\alpha \middle| |X| > \tfrac{1}{\sqrt{m}}\right], \mathbb{E}\left[|X|^\alpha \middle| |X| \leq \tfrac{1}{\sqrt{m}}\right]\right\}\\ &\quad= \mathbb{E}\left[|X|^\alpha \middle| |X| > \tfrac{1}{\sqrt{m}}\right].\end{aligned}$$

Therefore, (117) is obtained from (118). Hence, (113) is proved.

Proof of (111): Note that random variable $\widetilde{X}_m$ has the same distribution as the following random variable:

$$\widetilde{X}_m = [X]_m + U_m,$$

where U_m and X are independent, and

$$p_{U_m}(x) = \begin{cases} m & |x| \leq \frac{1}{2m}, \\ 0 & |x| > \frac{1}{2m}. \end{cases}$$

Thus,

$$\begin{aligned}\left|\widetilde{X}_m - X\right| &\leq \left|\widetilde{X}_m - [X]_m\right| + \left|[X]_m - X\right|\\ &\leq |U_m| + \tfrac{1}{2m}\\ &\leq \tfrac{1}{m},\end{aligned}$$

where the last inequality is true because $|[X]_m - X|$ and $|U_m|$ are always less than $1/2m$. Therefore, the proposition is proved. □

E. Proof of Lemma 10

From the definition of entropy, we know that the entropy of a random variable does not depend on the value of the random variable, rather it only depends on the distribution of the random variable. Therefore, if one finds a correspondence between the values of $[aX]_{m/a}$ and the values of $[X]_m$, while they have the same probability, then the entropy of them will be the same. Thus, we define the following correspondence between the values of $[aX]_{m/a}$, which are from the set $\left\{k\frac{a}{m} \,|k \in \mathbb{Z}\right\}$, and the values of $[X]_m$, which are from the set $\left\{k\frac{1}{m} \,|k \in \mathbb{Z}\right\}$.

$$k\tfrac{a}{m} \in \left\{k\tfrac{a}{m} \,|k \in \mathbb{Z}\right\} \leftrightarrow k\tfrac{1}{m} \in \left\{k\tfrac{1}{m} \,|k \in \mathbb{Z}\right\}$$

Now, we show that the correspondent values have the same probability.

$$\begin{aligned}\Pr\left\{[aX]_{\frac{m}{a}} = k\tfrac{a}{m}\right\} &= \Pr\left\{aX \in \left[\left(k - \tfrac{1}{2}\right)\tfrac{a}{m}, \left(k + \tfrac{1}{2}\right)\tfrac{a}{m}\right)\right\}\\ &= \Pr\left\{X \in \left[\left(k - \tfrac{1}{2}\right)\tfrac{1}{m}, \left(k + \tfrac{1}{2}\right)\tfrac{1}{m}\right)\right\}\\ &= \Pr\left\{[X]_m = k\tfrac{1}{m}\right\}.\end{aligned}$$

Therefore, the lemma is proved. □

F. Proof of Proposition 6

Observe that for any given $m, c_m \in \mathbb{R}$, there exist unique $k_m \in \mathbb{Z}$, and $d_m \in \mathbb{R}$ such that

$$c_m = \tfrac{k_m}{m} + d_m, \qquad d_m \in \left[0, \tfrac{1}{m}\right).$$

Therefore, because of the definition of amplitude quantization in Definition 6, we can write

$$[X + c_m]_m = [X + d_m]_m + \tfrac{k_m}{m}.$$

Since entropy is invariant with respect to constant shift, we have

$$\mathrm{H}\left([X + c_m]_m\right) = \mathrm{H}\left([X + d_m]_m\right).$$

As a result, we only need to prove the proposition for $\{d_m\}$, instead of $\{c_m\}$. Take the pdf $q_m(x)$ as follows:

$$\begin{aligned}q_m(x) &= m\Pr\left\{X + d_m \in \left[\tfrac{i-\frac{1}{2}}{m}, \tfrac{i+\frac{1}{2}}{m}\right)\right\}\\ &= m\int_{\frac{i-\frac{1}{2}}{m}-d_m}^{\frac{i+\frac{1}{2}}{m}-d_m} p(y)\,\mathrm{d}y,\end{aligned} \tag{120}$$

where i is the unique number such that $(i - \frac{1}{2})/m \leq x + d_m < (i + \frac{1}{2})/m$. Therefore, from Lemma 2, we obtain that

$$\int_{\mathbb{R}} q_m(x)\log\frac{1}{q_m(x)}\,\mathrm{d}x = \mathrm{H}\left([X + d_m]_m\right) - \log m,$$

where the last equation is true because of (120). So, if we take $\widetilde{X}_m$ a continuous random variable with pdf q_m, we only need to prove

$$\lim_{m\to\infty} \mathrm{h}\left(\widetilde{X}_m\right) = \mathrm{h}\left(X\right).$$

In order to do this, we write $\left|\mathrm{h}\left(\widetilde{X}_m\right) - \mathrm{h}\left(X\right)\right|$ as follows

$$\begin{aligned}&\left|\mathrm{h}\left(\widetilde{X}_m\right) - \mathrm{h}\left(X\right)\right|\\ &\quad\leq \left|\int_{-l}^{l} q_m(x)\log\tfrac{1}{q_m(x)} - p(x)\log\tfrac{1}{p(x)}\,\mathrm{d}x\right|\\ &\qquad+ \left|\int_{|x|>l} q_m(x)\log\tfrac{1}{q_m(x)}\,\mathrm{d}x\right|\\ &\qquad+ \left|\int_{|x|>l} p(x)\log\tfrac{1}{p(x)}\,\mathrm{d}x\right|,\end{aligned}$$

where $l > 0$ is arbitrary. Thus, it suffices to prove that

$$\lim_{m\to\infty}\int_{-l}^{l} q_m(x)\log\tfrac{1}{q_m(x)}\,\mathrm{d}x = \int_{-l}^{l} p(x)\log\tfrac{1}{p(x)}\,\mathrm{d}x, \tag{121}$$

$$\lim_{l\to\infty}\left|\int_{|x|>l} p(x)\log\tfrac{1}{p(x)}\,\mathrm{d}x\right| = 0, \tag{122}$$

$$\lim_{l\to\infty}\left|\int_{|x|>l} q_m(x)\log\tfrac{1}{q_m(x)}\,\mathrm{d}x\right| = 0, \tag{123}$$

hold for all $l > 0$ and uniformly on m.

Proof of (121): From (120), we obtain that since $p(x)$ is a piecewise continuous, the mean value theorem yields that there exists $x_m^* \in \left[(i - \frac{1}{2})/m - d_m, (i + \frac{1}{2})/m - d_m\right)$, such that $q_m(x) = p(x_m^*)$. Since $p(x)$ is piecewise continuous, and $[-l, l]$ is a compact set, $p(x)$ is uniformly continuous over

$[-l, l]$. Therefore, for any ϵ', there exists $M \in \mathbb{R}$ such that for all $x \in [-l, l]$, we have that

$$m > M \ \Rightarrow \ |q_m(x) - p(x)| < \epsilon'.$$

Furthermore, since the function $x \mapsto x \log x$ is continuous and $p(x)$ is uniformly continuous, we have that for all $x \in [-l, l]$

$$\lim_{m\to\infty} q_m(x) \log \tfrac{1}{q_m(x)} = p(x) \log \tfrac{1}{p(x)},$$

uniformly on x. Thus (121) is proved.

Proof of (122): We can write

$$\left| \int_{|x|>l} p(x) \log \tfrac{1}{p(x)} \,\mathrm{d}x \right| \le \int_{|x|>l} p(x) \left| \log \tfrac{1}{p(x)} \right| \mathrm{d}x. \quad (124)$$

By assuming the lemma, we know

$$\begin{aligned} &\int_{\mathbb{R}} p(x) \left| \log \tfrac{1}{p(x)} \right| \mathrm{d}x < \infty \\ &\quad \Longrightarrow \lim_{l\to\infty} \int_{|x|>l} p(x) \left| \log \tfrac{1}{p(x)} \right| \mathrm{d}x = 0. \end{aligned} \quad (125)$$

Hence, (124) implies (122).

Proof of (123): It suffices to show that

$$\begin{aligned} &\int_{|x+d_m|>l} p(x) \log \tfrac{1}{p(x)} \,\mathrm{d}x \\ &\quad \le \int_{|x|>l} q_m(x) \log \tfrac{1}{q_m(x)} \,\mathrm{d}x, \end{aligned} \quad (126)$$

$$\begin{aligned} &\int_{|x|>l} q_m(x) \log \tfrac{1}{q_m(x)} \,\mathrm{d}x \\ &\quad \le \sum_{|i|>l-1} P[i] \log \tfrac{1}{P[i]} + \Pr\{|X| > l-1\} \log 2 \\ &\qquad + o_l(1), \end{aligned} \quad (127)$$

where $o_l(1)$ means that $\lim_{l\to\infty} o_l(1) = 0$, and

$$P[i] = \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} p(x) \,\mathrm{d}x.$$

By changing the variable $y = Lx$, we can write

$$\begin{aligned} &\int_{|x+d_m|>l} p(x) \log \tfrac{1}{p(x)} \,\mathrm{d}x \\ &\quad = \int_{|x+d_m|>l} p(x) \log \tfrac{1}{\frac{1}{L} p(x)} \,\mathrm{d}x \\ &\qquad - \Pr\{|X + d_m| > l\} \log L, \\ &\quad = \int_{|y+Ld_m|>lL} \tfrac{1}{L} p\left(\tfrac{y}{L}\right) \log \tfrac{1}{\frac{1}{L} p\left(\frac{y}{L}\right)} \,\mathrm{d}y \\ &\qquad - \Pr\{|X + d_m| > l\} \log L. \end{aligned}$$

Because $p(x) < L$ almost everywhere, we have that:

$$\frac{1}{L} p\left(\tfrac{y}{L}\right) < 1 \ \Rightarrow \ \frac{1}{L} p\left(\tfrac{y}{L}\right) \log \tfrac{1}{\frac{1}{L} p\left(\frac{y}{L}\right)} > 0.$$

So we obtain that

$$\begin{aligned} &\int_{|y+Ld_m|>lL} \frac{1}{L} p\left(\tfrac{y}{L}\right) \log \tfrac{1}{\frac{1}{L} p\left(\frac{y}{L}\right)} \,\mathrm{d}y \\ &\quad \ge \int_{|y|>lL+L} \frac{1}{L} p\left(\tfrac{y}{L}\right) \log \tfrac{1}{\frac{1}{L} p\left(\frac{y}{L}\right)} \,\mathrm{d}y \\ &\quad = \int_{|x|>l+1} p(x) \log \tfrac{1}{p(x)} \,\mathrm{d}x \\ &\qquad + \Pr\{|X| > l+1\} \log L, \end{aligned}$$

where $x = y/L$. Therefore, we can write that

$$\begin{aligned} \int_{|x+d_m|>l} p(x) \log \tfrac{1}{p(x)} \,\mathrm{d}x \ge &\int_{|x|>l+1} p(x) \log \tfrac{1}{p(x)} \,\mathrm{d}x \\ &- \Pr\{|X| \in [l, l+1]\} \log L. \end{aligned}$$

Thus, from (125) we conclude that the lower bound vanishes as l tends to ∞ uniformly on m. Now, we are going to show that (127) leads that the upper bound vanishes as l tends to ∞ uniformly on m. In order to prove this, note that $\Pr\{|X| > l-1\}$ vanishes as l tends to infinity uniformly on m. Furthermore,

$$\begin{aligned} \mathrm{H}\left([X]_1\right) < \infty &\Longrightarrow \sum_{i\in\mathbb{Z}} P[i] \log \tfrac{1}{P[i]} < \infty \\ &\Longrightarrow \lim_{l\to\infty} \sum_{|i|>l-1} P[i] \log \tfrac{1}{P[i]} = 0. \end{aligned}$$

Thus, in order to prove (123), it only remains to prove (126) and (127). The proof of (126) exists in [1] in the proof of Corollary 1. Thus, we only need to prove (127). Similar to the proof of Corollary 1 in [1], it can be shown that

$$\int_{|x|>l} q_m(x) \log \tfrac{1}{q_m(x)} \,\mathrm{d}x \le \sum_{|i|>l} P_m[i] \log \tfrac{1}{P_m[i]},$$

where $P_m[i] := \Pr\{[X + d_m]_1 = i\}$. This implies

$$\begin{aligned} &\mathrm{H}\left([X + d_m]_1 \big| \, |[X + d_m]_1| > l\right) \\ &\quad = \sum_{|i|>l} \tfrac{P_m[i]}{\Pr\left\{|[X+d_m]_1|>l\right\}} \log \tfrac{\Pr\left\{|[X+d_m]_1|>l\right\}}{P_m[i]}, \end{aligned}$$

where $\mathrm{H}(X|Y = y)$ is defined in [24, p. 29]. Hence, we can write

$$\begin{aligned} &\sum_{|i|>l} P_m[i] \log \tfrac{1}{P_m[i]} \\ &\quad = \Pr\{|[X + d_m]_1| > l\} \\ &\qquad \times \mathrm{H}\left([X + d_m]_1 \big| |[X + d_m]_1| > l\right) \\ &\qquad + \Pr\{|[X + d_m]_1| > l\} \log \tfrac{1}{\Pr\left\{|[X+d_m]_1|>l\right\}}. \end{aligned} \quad (128)$$

Because of the definition of the quantization, we can write

$$[X + d_m]_1 = [X]_1 + E_m,$$

where E_m is a random variable taking values from $\{0, 1\}$. Thus, $[X + d_m]_1$ is a function of $[X]_1$ and E_m; as a result

$$\begin{aligned} &\mathrm{H}\left([X + d_m]_1 \big| |[X + d_m]_1| > l\right) \\ &\quad \le \mathrm{H}\left([X]_1, E_m \big| |[X + d_m]_1| > l\right) \\ &\quad \le \mathrm{H}\left([X]_1 \big| |[X + d_m]_1| > l\right) + \mathrm{H}(E_m) \\ &\quad \le \mathrm{H}\left([X]_1 \big| |[X + d_m]_1| > l\right) + \log 2. \end{aligned}$$

From the definition of $\mathrm{H}\left([X]_1 \middle| |[X+d_m]_1| > l\right)$, we can write that

$$\begin{aligned}\Pr\{|[X+d_m]_1| > l\} \, \mathrm{H}\left([X]_1 \middle| |[X+d_m]_1| > l\right) &= \sum_{|i|>l} P[i]\log\tfrac{1}{P[i]} \\ &\quad -r_m\log r_m - s_m\log s_m + t_m\log t_m \\ &\quad -\Pr\{|[X+d_m]_1| > l\}\log\tfrac{1}{\Pr\{|[X+d_m]_1|>l\}},\end{aligned}$$

where

$$\begin{aligned} r_m &= \Pr\left\{X \in \left(l-\tfrac{1}{2}-d_m, l-\tfrac{1}{2}\right)\right\}, \\ s_m &= \Pr\left\{X \in \left[-l-\tfrac{1}{2}, -l+\tfrac{1}{2}-d_m\right)\right\}, \\ t_m &= \Pr\left\{X \in \left[-l-\tfrac{1}{2}, -l+\tfrac{1}{2}\right)\right\}. \end{aligned}$$

Therefore, (128) can be simplified as follows:

$$\begin{aligned}\sum_{|i|>l} P_m[i]\log\tfrac{1}{P_m[i]} \\ &\leq \sum_{|i|>l} P[i]\log\frac{1}{P[i]} \qquad (129)\\ &\quad -r_m\log r_m - s_m\log s_m + t_m\log t_m \qquad (130)\\ &\quad +\Pr\{|[X+d_m]_1| > l\}\log 2. \qquad (131)\end{aligned}$$

The first term, (129), vanishes as l tends to ∞ because

$$\begin{aligned}\mathrm{H}([X]_1) &= \sum_{i\in\mathbb{Z}} P[i]\log\tfrac{1}{P[i]} < \infty \\ &\implies \lim_{l\to\infty}\sum_{|i|>l} P[i]\log\tfrac{1}{P[i]} = 0.\end{aligned}$$

It can be achieved that (130) and (131) vanish as l tends to ∞, uniformly on m, but we do not write the details here. Thus, (127) is proved and the proof of proposition is complete. □

G. Proof of Proposition 7

According to Lemma 11, we can define random variable U with support $\{\mathrm{dc}, \mathrm{cd}, \mathrm{cc}\}$ and pmf $p_U(u)$ such that

$$p_U(u) = \frac{1}{\Delta}\begin{cases}\Pr\{X \text{ is discrete}\}\Pr\{Y \text{ is continuous}\} & u=\mathrm{dc},\\ \Pr\{X \text{ is continuous}\}\Pr\{Y \text{ is discrete}\} & u=\mathrm{cd},\\ \Pr\{X \text{ is continuous}\}\Pr\{Y \text{ is continuous}\} & u=\mathrm{cc},\end{cases}$$

and

$$p_{Z_c|U}(x|u) = \begin{cases} p_{X_D+Y_c} & u=\mathrm{dc},\\ p_{X_c+Y_D} & u=\mathrm{cd},\\ p_{X_c+Y_c} & u=\mathrm{cc},\end{cases}$$

where

$$\begin{aligned}\Delta := &\Pr\{X \text{ is discrete}\}\Pr\{Y \text{ is continuous}\}\\ &+\Pr\{X \text{ is continuous}\}\Pr\{Y \text{ is discrete}\}\\ &+\Pr\{X \text{ is continuous}\}\Pr\{Y \text{ is continuous}\}.\end{aligned}$$

Thus, we have that

$$\begin{aligned}\mathrm{h}(Z_c) &\geq \mathrm{h}\left(Z_c \middle| U\right)\\ &= p_U(\mathrm{cd})\mathrm{h}(X_c+Y_D) + p_U(\mathrm{dc})\mathrm{h}(X_D+Y_c)\\ &\quad + p_U(\mathrm{cc})\mathrm{h}(X_c+Y_c)\\ &\geq p_U(\mathrm{cd})\mathrm{h}(X_c+Y_D|Y_D)\\ &\quad + p_U(\mathrm{dc})\mathrm{h}(X_D+Y_c|X_D)\\ &\quad + p_U(\mathrm{cc})\mathrm{h}(X_c+Y_c|Y_c)\\ &\geq \min(\mathrm{h}(X_c), \mathrm{h}(Y_c)).\end{aligned}$$

Hence, the proposition is proved. □

H. Proof of Lemma 12

From the Fubini's theorem we obtain that

$$\Pr\{Z \leq z\} = \Pr\{X \leq z - Y\} = \mathbb{E}\left[F_X(z-Y)\right],$$

where $F_X(x) := \Pr\{X \leq x\}$ is the cdf of X. In order to prove the first part of the lemma, we only need to prove that

$$\mathbb{E}\left[F_X(z-Y)\right] = \int_{-\infty}^{z}\mathbb{E}\left[p_X(z-Y)\right]\mathrm{d}z. \qquad (132)$$

The above integral can be written as the following limit:

$$\int_{-\infty}^{z}\mathbb{E}\left[p_X(z-Y)\right]\mathrm{d}z = \lim_{\ell\to\infty}\int_{-\ell}^{z}\mathbb{E}\left[p_X(z-Y)\right]\mathrm{d}z.$$

Note that, from Fubini's theorem, for every finite ℓ, we have that

$$\begin{aligned}\int_{-\ell}^{z}\mathbb{E}\left[p_X(z-Y)\right]\mathrm{d}z &= \mathbb{E}\left[\int_{-\ell}^{z} p_X(z-Y)\,\mathrm{d}z\right]\\ &= \mathbb{E}\left[F_X(z-Y) - F_X(-\ell-Y)\right]\\ &= \mathbb{E}\left[F_X(z-Y)\right] - \mathbb{E}\left[F_X(-\ell-Y)\right].\end{aligned}$$

Hence, to prove (132), it is sufficient to show that

$$\lim_{\ell\to\infty}\mathbb{E}\left[F_X(-\ell-Y)\right] = 0.$$

In order to do so, note that there exists ℓ_X, ℓ_Y such that

$$F_X(\ell_X) \leq \epsilon/2, \qquad F_Y(\ell_Y) \leq \epsilon/2.$$

Thus, we can write

$$\begin{aligned}\mathbb{E}\left[F_X(-\ell-Y)\right] &= \Pr\{Y \leq \ell_Y\}\,\mathbb{E}\left[F_X(-\ell-Y)|Y \leq \ell_Y\right]\\ &\quad +\Pr\{Y > \ell_Y\}\,\mathbb{E}\left[F_X(-\ell-Y)|Y > \ell_Y\right]\\ &\leq \Pr\{Y \leq \ell_Y\}\times 1\\ &\quad +1\times\mathbb{E}\left[F_X(-\ell-\ell_Y)\right]\\ &\leq \epsilon.\end{aligned}$$

Therefore, the the lemma is proved. □

I. *Proof of Lemma 13*

Note that by the entropy power inequality (EPI) [24, Theorem 17.7.3], we have that for n i.i.d. scalar continuous random variables $X_1, \cdots, X_n$

$$\begin{aligned}&\mathrm{e}^{2\mathrm{h}(X_1+\cdots+X_n)} \geq n\mathrm{e}^{2\mathrm{h}(X_1)}\\ &\qquad\Longrightarrow \mathrm{h}\left(X_1\right) \leq \mathrm{h}\left(X_1+\cdots+X_n\right)+\log\frac{1}{\sqrt{n}}.\end{aligned}$$

Therefore, from Lemma 5 we have that

$$\begin{aligned}\mathrm{h}\left(X_1^{(n)}\right) &\leq \mathrm{h}\left(X_1^{(n)}+\cdots+X_n^{(n)}\right)+\log\frac{1}{\sqrt{n}}\\ &= \mathrm{h}\left(X_0\right)+\log\frac{1}{\sqrt{n}},\end{aligned}$$

where $X_1^{(n)}$ was defined in Definition 13. According to Theorem 2, $\zeta(n) = \mathrm{h}\left(X_1^{(n)}\right)$ satisfies (12). This will complete the proof since based on Theorem 1, for any $\zeta'(n)$ satisfying (12), we have that

$$\lim_{n\to\infty} \left|\zeta(n) - \zeta'(n)\right| = 0.$$

Hence,

$$\begin{aligned}\zeta'(n) - \log\frac{1}{\sqrt{n}} &= \zeta'(n) - \zeta(n) + \zeta(n) - \log\frac{1}{\sqrt{n}}\\ &\leq \left|\zeta'(n) - \zeta(n)\right| + \mathrm{h}\left(X_0\right).\end{aligned}$$

Thus, the lemma is proved. $\square$