Run-length compression of quantized Gaussian stationary signals

Oleg Seleznjev and Mykola Shykula

Communicated by Nikolai Leonenko

Abstract. We consider quantization of random continuous-valued signals. In practice, analogue signals are quantized at sampling points with further compression. We study probabilistic models for run-length encoding (RLE) algorithm applied to quantized sampled random signals (Gaussian processes). This compression technique is widely used in digital signal and image processing. The mean (inverse) RLE compression ratio (or data rate savings) and its statistical inference are considered. In particular, the asymptotic normality for some estimators of this characteristic is shown. Numerical experiments for synthetic and real data are presented.

Keywords. Asymptotic normality, Gaussian process, compression ratio, run-length encoding, uniform quantization.

2010 Mathematics Subject Classification. Primary 60G15; secondary 94A29, 94A34.

1 Introduction

Quantization in signal processing, or discretization in amplitude of continuous-valued signals, is intensively studied in the last decades. For a comprehensive overview of quantization techniques, we refer to [9] (see also [8] and the references therein). Nowadays the interest to statistical problems for quantization is renewed due to the wide spread of various types of sensors and increased performance requirements. Two necessary steps are needed to digitalize a signal: sampling (or time discretization) and quantization. Quantization maps the continuous-valued measurements of the signal to a set of discrete values (say, quantization levels). Quantization is a part of digital signal representation for various conversion techniques such as source coding, data compression, archiving, restoration etc., where data compression refers to reducing the amount of data required to represent a given quantity of information. The underlying basis of this reduction is the removal of redundant data. Later, the compressed signal is decompressed to

The first author is partly supported by the Swedish Research Council grant 2009-4489.
reconstruct an original sampled signal which is often can be used for exact recovering of the original signal (e.g., by Nyquist sampling technique). In lossless data compression, an original signal can be recovered exactly. Run-length encoding (RLE) algorithm (see, e.g., [12]) is one of the simplest lossless coding techniques used in data compression. In RLE, a signal that consists of repeated symbols is compressed by encoding each repetition (or run) as a pair (“symbol”, “length of the run”). This technique is widely used in signal and image processing independently (e.g., BMP and PCX compression formats) or as a part of more advanced methods (see, e.g., [4] and the references therein).

The prediction of the memory capacity (or quantization rate) needed for encoded (quantized) signals is an important problem especially nowadays for huge digital signal and image databases in various applications. Probabilistic modelling for random signals (processes) allows us to analyze this problem in average case setting (cf. [14]). In order to demonstrate the general approach, we consider the RLE compression in more detail by using correlation structure of an original Gaussian stationary process when sampling and quantization steps vary accordingly.

Some statistical problems for quantized data are considered in [20]. General quantization problems for Gaussian processes are investigated in [11, 15].

We consider probabilistic modelling and statistical inference for the mean memory capacity of quantized in value/time realizations of a Gaussian process when RLE compression is used. Throughout the paper it is assumed, without comment, that $X(t), t \in [0, T]$, is a Gaussian stationary process with mean $\mu$, variance $\sigma^2 > 0$, and covariance and correlation functions $K(t)$ and $k(t) = K(t)/\sigma^2$, respectively, $t \in [0, T]$. Let $k(t)$ be continuous and for any $h > 0$, $|k(h)| < 1$, i.e., the joint distribution for $X(0)$ and $X(h)$ is nonsingular. In particular, $k(t) \to 1$ if and only if $t \to 0$.

In applications, signals are observed at sampling points. For a sampling step $h > 0$, let

$$\mathcal{S}_{n,h}(X) := \{X(kh), k = 0, 1, \ldots, n\}$$

denote a sample from $X(t)$, $n = T/h$ is the number of sampling points. For a positive $\varepsilon$ and a real $x$, define the uniform quantizer $q_{\varepsilon}(x) := \varepsilon[x/\varepsilon]$, where $[x]$ denotes the integer part of $x$. Then for $\mathcal{S}_{n,h}(X)$ and $q_{\varepsilon}(x)$, the random memory capacity needed for RLE compressed quantized signal (RLE quantization rate) up to the fixed multiplicative constant equals

$$l_{\varepsilon,h}(X) = l_{\varepsilon,h}(X, T) := \sum_{j=0}^{n-1} I(q_{\varepsilon}(X(jh)) \neq q_{\varepsilon}(X((j + 1)h))),$$

where $I(\cdot)$ is the indicator function. Henceforth we use the term RLE quantization.
rate for \( l_{\varepsilon,h}(X) \) if no confusion can arise. Denote by
\[
L_{\varepsilon,h}(X) = L_{\varepsilon,h}(X, T) := \mathbb{E}(l_{\varepsilon,h}(X)),
\]
the mean RLE quantization rate for \( \mathcal{S}_{n,h}(X) \). It follows by the stationarity of \( X(t) \) that
\[
L_{\varepsilon,h}(X) = \frac{T}{h} \left( 1 - \mathbb{P}(q_{\varepsilon}(X(0)) = q_{\varepsilon}(X(h))) \right) = \frac{T}{h} r_{\varepsilon,h}(X),
\]
where the mean RLE (inverse) compression ratio \( r_{\varepsilon,h}(X), 0 < r_{\varepsilon,h}(X) \leq 1 \), characterizes the compression efficiency of RLE (in average). A related equivalent characteristics is data rate savings \( s_{\varepsilon,h} = 1 - r_{\varepsilon,h} \). Henceforth, we consider the ratio \( r_{\varepsilon,h}(X) \) as a main RLE compression parameter for a given sampling step \( h \).

Let \( \overset{d}{\rightarrow} \) denote convergence in distribution, \( \phi(x) \) and \( \Phi(x) \) be the density and distribution functions of the standard normal distribution \( N(0,1) \), respectively. The asymptotic normality for an estimator \( \hat{\theta}_m \) of a parameter \( \theta \) means that
\[
\sqrt{m}(\hat{\theta}_m - \theta) \overset{d}{\rightarrow} N(0, \sigma^2) \quad \text{as } m \to \infty.
\]

The paper is organized as follows. In Section 2, we study the behavior of the mean RLE inverse compression ratio for a given probabilistic model. In Section 3, we consider statistical inference for \( r_{\varepsilon,h}(X) \), when a standard time series model with increasing number of observations is assumed. Some properties (e.g., asymptotic normality) of sample estimators for \( r_{\varepsilon,h}(X) \) are investigated. Section 4 provides numerical experiments, illustrating the rate of convergence in the obtained results. Section 5 proves the statements from Sections 2 and 3.

## 2 Run-length encoding. Probabilistic modelling

In this section, we consider some properties of the mean RLE inverse compression ratio (or the mean RLE quantization rate) assuming a known probabilistic model. First, we study the behavior of \( r_{\varepsilon,h}(X) \) for any positive cell width \( \varepsilon \) and sampling step \( h \), then we consider the asymptotic behavior of \( r_{\varepsilon,h}(X) \), when \( \varepsilon \) and \( h \) vary in an according way. We begin with some notation. Let for a sampling step \( h > 0 \),
\[
v(h) = v(K, h) := \frac{1}{\sigma (1 - k(h))^{-1/2}} = \frac{K(0)^{1/2}}{(K(0)^2 - K(h)^2)^{1/2}}.
\]
Let \( Y \) and \( U \) be independent standard normal and \([0,1]\)-uniform random variables, respectively, and \( Z := |Y|/U \),
\[
g(x) := 2 \int_0^x (1 - y/x) \phi(y) dy = \mathbb{P}(Z \leq x), \quad x > 0,
\]
i.e., \( g(x) \) is the probability distribution function of \( Z \), \( g(0) = 0 \). We can interpret \( g(x) \) as following. First, we note that

\[
g(x) = P(Y \leq x(1 - U)) - P(Y \leq -xU) = P(0 \leq Y + xU < x).
\]  

(2.2)

From the additive noise model for the uniform quantizer \( q_\varepsilon \) (see, e.g., \([17, 20]\)) we have for sufficiently small positive \( \varepsilon \)

\[
X(0) \xrightarrow{d} q_\varepsilon(X(0)) + \varepsilon U,
\]

(2.3)

where \( U \) is a \([0, 1]\)-uniform random variable independent of \( X(0), X(h) \). Let

\[
\sigma(h)^2 := E(X(h) - X(0))^2 = 2(K(0) - K(h))
\]

and \( Y = (X(h) - X(0))/\sigma(h) \), i.e., a standard normal variable independent of \( U \). By definition, we get

\[
P(q_\varepsilon(X(0)) = q_\varepsilon(X(h))) = P(0 \leq X(h) - \varepsilon[X(0)/\varepsilon] < \varepsilon).
\]

(2.4)

It follows from (2.3) and (2.4) that for sufficiently small \( \varepsilon \),

\[
P(q_\varepsilon(X(0)) = q_\varepsilon(X(h))) \simeq P(0 \leq X(h) - X(0) + \varepsilon U < \varepsilon)
= P(0 \leq \sigma(h)Y + \varepsilon U < \varepsilon)
= g(\varepsilon/\sigma(h)) \sim g(\varepsilon v(h)) \quad \text{as } h \to 0,
\]

and one may expect that

\[
r_{\varepsilon,h}(X) \sim 1 - g(\varepsilon v(h)) \quad \text{as } \varepsilon, h \to 0.
\]

(2.5)

Now we proceed with the main result of this section. In the following theorem we evaluate \( r_{\varepsilon,h}(X) \). Write \( w(h) := ((1 + k(h))/2)^{1/2} \sim 1 \) as \( h \to 0 \).

**Theorem 2.1.** For any positive \( \varepsilon \) and \( h \),

\[
r_{\varepsilon,h}(X) = 1 - g(\varepsilon v(h))\left(w(h) + \frac{\varepsilon}{\sigma}\delta_{\varepsilon,h}(X)\right),
\]

where

\[
|\delta_{\varepsilon,h}(X)| \leq \left(\frac{2}{\pi}\right)^{1/2} \left(1 + \frac{\varepsilon^2}{4 w(h)^2 \sigma^2} \exp\left\{\frac{\varepsilon^2}{2 w(h)^2 \sigma^2}\right\}\right).
\]

If \( \mu \in \{k \varepsilon, k \in \mathbb{Z}\} \), then

\[
|\delta_{\varepsilon,h}(X)| \leq (2/\pi)^{1/2}.
\]
From Theorem 2.1, we immediately obtain the lower and upper bounds for $r_{\varepsilon, h}(X)$, for any positive $\varepsilon$ and $h$, since $g(x), x > 0$, is monotone and positive,

$$r_{\varepsilon, h}^l \leq r_{\varepsilon, h}(X) \leq r_{\varepsilon, h}^u,$$

where

$$r_{\varepsilon, h}^l := \rho_{\varepsilon, h} - \frac{\varepsilon}{\sigma} \Delta_{\varepsilon, h}, \quad r_{\varepsilon, h}^u := \rho_{\varepsilon, h} + \frac{\varepsilon}{\sigma} \Delta_{\varepsilon, h},$$

$$\rho_{\varepsilon, h} := 1 - g(\varepsilon v(h))w(h),$$

$$\Delta_{\varepsilon, h} := g(\varepsilon v(h))\left(\frac{2}{\pi}\right)^{1/2} \left(1 + \frac{\varepsilon^2}{4w(h)^2\sigma^2} \exp\left\{\frac{\varepsilon^2}{2w(h)^2\sigma^2}\right\}\right).$$

Notice that these bounds do not depend on the mean $\mu$ of the original process $X(t)$. The parameters $\varepsilon$ and $\sigma$ correspond to quantization accuracy and deviation around the mean of the process, respectively. In many applications, in order to obtain an informative quantized process realization, it is reasonable to suppose that $\varepsilon/\sigma$ is sufficiently small and therefore, the obtained bounds, (2.6), are precise enough. Moreover, we have

$$r_{\varepsilon, h}(X) \sim \rho_{\varepsilon, h} \quad \text{as } \varepsilon \to 0,$$

and for sufficiently small $h$, we obtain (cf. (2.5))

$$r_{\varepsilon, h}(X) \sim 1 - g(\varepsilon v(h)) \quad \text{as } \varepsilon, h \to 0.\quad (2.7)$$

By definition, heuristically, the relationship between the quantization cell width and quadratic mean (q.m.) smoothness of $X(t)$ determines the asymptotic behavior of $r_{\varepsilon, h}(X)$. Denote by

$$\gamma(t) := \frac{1}{2} E\{(X(t) - X(0))^2\} = \sigma^2(1 - k(t)) = K(0) - K(t),$$

the semivariogram of $X(t), t \in [0, T]$. Semivariograms are commonly used for modelling random functions in geosciences (see, e.g., [5]). In the following theorem we investigate the asymptotic behavior of $r_{\varepsilon, h}(X)$, when $\varepsilon$ and $\gamma(h)$ vary in an according way, i.e., $h = h(\varepsilon)$. Write $v(\varepsilon, h) := \varepsilon \gamma(h)^{-1/2}$.

**Theorem 2.2.** When $\varepsilon \to 0$,

$$r_{\varepsilon, h}(X) \sim \begin{cases} 
1 - \frac{v(\varepsilon, h)}{2\sqrt{\pi}}, & \text{if } v(\varepsilon, h) \to 0, \\
1 - g(\tau/\sqrt{2}), & \text{if } v(\varepsilon, h) \to \tau, \ 0 < \tau < \infty, \\
\frac{2}{v(\varepsilon, h)\sqrt{\pi}}, & \text{if } v(\varepsilon, h) \to \infty \text{ and } \varepsilon v(\varepsilon, h) \to 0.
\end{cases} \quad (2.9)$$
Remark 2.3. (i) The additional condition \( \varepsilon \nu(\varepsilon, h) \to 0 \) for \( \nu(\varepsilon, h) \to \infty \) is technical and we claim that the corresponding result is valid in a more general case.

(ii) The corresponding asymptotics for the mean RLE quantization rate \( L_{\varepsilon, h}(X) \) now follows from the definition

\[
L_{\varepsilon, h}(X) = \frac{T}{\varepsilon} r_{\varepsilon, h}(X).
\]

For a continuously (q.m.) differentiable \( X(t) \), we have \( \gamma(h) \sim \lambda_2 h^2/2 \), where \( \lambda_2 = -K^{(2)}(0) \) is the second spectral moment of \( X(t) \), and

\[
\nu(\varepsilon, h) \sim \frac{\varepsilon}{h} \left( \frac{\lambda_2}{2} \right)^{-1/2} \quad \text{as } h \to 0.
\]

We rewrite more explicitly (2.9) when \( \varepsilon \) and \( h \) vary in an according way. Let \( \varepsilon \to 0 \). Then

\[
 r_{\varepsilon, h}(X) \sim \begin{cases} 
 1 - \frac{(2\pi \lambda_2)^{-1/2}}{h} \varepsilon, & \text{if } \frac{\varepsilon}{h} \to 0, \\
 1 - g(\varepsilon \lambda_2^{-1/2}/h), & \text{if } \frac{\varepsilon}{h} \to c, \ 0 < c < \infty, \\
 \left( \frac{2\lambda_2}{\pi} \right)^{1/2} \frac{h}{\varepsilon}, & \text{if } \frac{\varepsilon}{h} \to \infty \text{ and } \frac{\varepsilon^2}{h} \to 0.
\end{cases}
\] (2.10)

In particular,

\[
L_{\varepsilon, h}(X) \sim \frac{T}{\varepsilon} \left( \frac{2\lambda_2}{\pi} \right)^{1/2} = o \left( \frac{T}{h} \right), \quad \text{if } \frac{\varepsilon}{h} \to \infty \text{ and } \frac{\varepsilon^2}{h} \to 0 \text{ as } \varepsilon \to 0, \ (2.11)
\]

(cf. [15]). We claim that the asymptotic relationship (2.10) between q.m. smoothness of the process, the quantization cellwidth, and the sampling step is general up to a constant for various lossless coding algorithms.

3 Statistical inference for RLE

In the previous section, we study the behavior of the mean RLE inverse compression ratio \( r_{\varepsilon, h}(X) \) provided that a stochastic structure of the original Gaussian process \( X(t) \) is known. Now we assume that the covariance function is unknown and there are observed data \( X(jh), \ j = 0, \ldots, m \), for statistical evaluation of \( r_{\varepsilon, h}(X) \) with a fixed sampling step \( h > 0 \) and let \( X_j := X(jh) \). Note that this sample can be different from \( \mathcal{S}_{n,h}(X) \) and is used as in time-series modelling for estimation of unknown parameters \( K(0), K(h) \) when the number of observation \( m \) increases and \( h \) is fixed. In particular, for these realizations, we admit \( m \geq n, \ i.e., \ m h \geq T \). Let

\[
\hat{\mu} := \frac{1}{m} \sum_{j=0}^{m} (X_j - \hat{\mu})^2, \quad \hat{\mu} := \frac{1}{m} \sum_{j=0}^{m-1} (X_j - \hat{\mu})(X_{j+1} - \hat{\mu}),
\] (3.1)
be standard sample estimators of \( K(0) \) and \( K(h) \), respectively, and
\[
\hat{\mu} := \frac{1}{m + 1} \sum_{j=0}^{m} X_j.
\]
In a similar way, we introduce
\[
\hat{\gamma}(h) := \frac{1}{2m} \sum_{j=0}^{m-1} (X_{j+1} - X_j)^2, \quad \hat{\gamma}_K(h) := \hat{K}(0) - \hat{K}(h),
\]
(3.2)
sample estimators for \( \gamma(h) \) (see, e.g., [5, 18]). Various statistical properties of the above estimators are well known. Specifically, the asymptotic joint normality of \( \hat{K}(0) \) and \( \hat{K}(h) \) has been established by several authors under appropriate conditions on the covariance function \( K(t) \) that ensure the dependence in the process decreases sufficiently fast as the distance (time) between observations increases (see, e.g., [1], [7, Section 6.3] and [6]). We use the following conditions:

(A) Assume that \( X(jh) \) can be represented as follows:
\[
X_j = \mu + \sum_{i \in \mathbb{Z}} a_i(h) V_{j-i}, \quad j \geq 0, \quad \sum_{i \in \mathbb{Z}} |a_i(h)| < \infty,
\]
where \( V_i, i \in \mathbb{Z} \), are independently and identically distributed standard normal variables.

(B) \( \sum_{j \in \mathbb{Z}} |K(jh)| < \infty. \)

Note that (B) follows from (A) and (A) says that \( \{X_j\} \) is a linear time series. For completeness and further references, we give the following proposition on asymptotic properties of \( \hat{K}(0), \hat{K}(h), \hat{\gamma}(h), \) and \( \hat{\gamma}_K(h) \), when the number of sample points \( m \to \infty \).

**Proposition 3.1.**  
(i) If (A) holds, then \( \hat{K}(0) \) and \( \hat{K}(h) \) are asymptotically joint normal estimators of \( K(0) \) and \( K(h) \), respectively, as \( m \to \infty \).

(ii) If (B) holds, then \( \hat{\gamma}(h) \) and \( \hat{\gamma}_K(h) \) are asymptotically normal estimators of \( \gamma(h) \) as \( m \to \infty \).

Applying the results from the previous section, we evaluate the mean RLE inverse compression ratio \( r_{\varepsilon,h}(X) \). Denote by \( \hat{\rho}_{\varepsilon,h} := \rho_{\varepsilon,h}(\hat{K}(0), \hat{K}(h)) \) a sample (plug-in) estimator of \( \rho_{\varepsilon,h} \) based on \( \hat{K}(0) \) and \( \hat{K}(h) \). Then a natural estimator of \( r_{\varepsilon,h}(X) \) is provided by \( \hat{r}_{\varepsilon,h} \), (2.8), for small enough \( \varepsilon \). In a similar way,
\[
\hat{r}_{\varepsilon,h}^l := r_{\varepsilon,h}^l(\hat{K}(0), \hat{K}(h)) \quad \text{and} \quad \hat{r}_{\varepsilon,h}^u := r_{\varepsilon,h}^u(\hat{K}(0), \hat{K}(h))
\]
define the corresponding sample estimators of the lower $r^l_{\varepsilon,h}$ and upper $r^u_{\varepsilon,h}$ bounds for $r_{\varepsilon,h}(X)$, respectively, (2.6). In the following theorem we study the asymptotic behavior of $\hat{\rho}_{\varepsilon,h}, \hat{r}^l_{\varepsilon,h},$ and $\hat{r}^u_{\varepsilon,h}$.

**Theorem 3.2.** If (A) holds, then for any positive $\varepsilon$ and $h$, $\hat{\rho}_{\varepsilon,h}, \hat{r}^l_{\varepsilon,h},$ and $\hat{r}^u_{\varepsilon,h}$ are asymptotically normal estimators of $\rho_{\varepsilon,h}, r^l_{\varepsilon,h},$ and $r^u_{\varepsilon,h}$, respectively, as $m \to \infty$.

**Remark 3.3.** The asymptotic variances of $\hat{\rho}_{\varepsilon,h}, \hat{r}^l_{\varepsilon,h},$ and $\hat{r}^u_{\varepsilon,h}$ can be found by using those of $b_{K.0/}$ and $b_{K.h}$ given in [1, Chapter 8.4.2] (see also [13, Chapter 6a.2]). Alternatively, data-resampling methods for stationary time series can be used (see, e.g., [16, Chapter 9]). Then, Theorem 3.2 together with (2.6) allows us to construct asymptotic confidence intervals for the mean RLE inverse compression ratio $r_{\varepsilon,h}(X)$ for any positive $\varepsilon$ and $h$.

By Theorem 2.2, the asymptotic behavior of $r_{\varepsilon,h}(X)$ as $\varepsilon \to 0$ is determined through the semivariogram $\gamma(h)$. Thus, applying, for instance, the sample estimator $\hat{\gamma}_K(h) = \hat{K}(0) - \hat{K}(h)$ (or $\hat{\gamma}(h)$) of $\gamma(h)$, we get plug-in (asymptotic) estimators of $r_{\varepsilon,h}(X)$ for various cases, e.g., (2.9),

$$\rho^1_{\varepsilon,h}(\gamma(h)) := 1 - \frac{\varepsilon}{2(\pi \gamma(h))^{1/2}}.$$  

Then the asymptotic normality of $\hat{\rho}^1_{\varepsilon,h} := \rho^1_{\varepsilon,h}(\hat{\gamma}_K(h))$ (or $\rho^1_{\varepsilon,h}(\hat{\gamma}(h))$) follows from the condition (B) in a standard way.

In fact, the observed data are already quantized. However, for a certain class of Gaussian processes $X(t)$, $t \in [0, T]$, and the uniform quantizer $q_{\varepsilon}(x)$, the normalized error process $Z_{\varepsilon}(t)$ in the additive noise model $X(t) = q_{\varepsilon}(X(t)) + \varepsilon Z_{\varepsilon}(t)$, $t \in [0, T]$, behaves asymptotically like a white noise (see [17]). In particular, the process $Z_{\varepsilon}(t), t \in [0, T]$, has asymptotically independent values for $t \neq s$; $Z_{\varepsilon}(t)$ is asymptotically independent of $X(t)$ as $\varepsilon \to 0$; the q.m. quantization error is of order $\varepsilon$. In this section until now we neglect the quantization error effect. For a detailed analysis of the influence of quantization effects in some statistical problems we refer to [20].

Let assume that the observed data $Q^\varepsilon_{m,h}(X) := \{q_{\varepsilon}(X(jh)) : j = 1, \ldots, m\}$ are quantized now. Based on $Q^\varepsilon_{m,h}(X)$, a plain (nonparametric) estimator of the mean RLE inverse compression ratio $r_{\varepsilon,h}(X)$ is

$$\hat{r}_{\varepsilon,h}(X) := \frac{1}{m} \sum_{j=0}^{m-1} I(q_{\varepsilon}(X_j) \neq q_{\varepsilon}(X_{j+1})).$$  

**Theorem 3.4.** Let $\mu \in \{k\varepsilon, k \in \mathbb{Z}\}$. If (B) holds, then for any positive $\varepsilon$ and $h$, $\hat{r}_{\varepsilon,h}(X)$ is an asymptotically normal estimator of $r_{\varepsilon,h}(X)$ as $m \to \infty$.  

Brought to you by | Lulea University of Technology
Authenticated
Download Date | 12/17/14 6:20 PM
Remark 3.5. The variance for this estimator can be evaluated by data-resampling methods for stationary time series (cf. Remark 3.3).

4 Numerical experiments

The first two numerical examples illustrate the rate of convergence for the results obtained in the previous sections. While in Example 4.1 we assume that the stochastic structure of an original process is known, in Example 4.2 this structure is supposed to be unknown and there are (simulated) observed data for evaluation. In Example 4.3 we compare estimators $\hat{\rho}_{\varepsilon,h}$, $m = 50$ (Theorem 3.2), and $\hat{r}_{\varepsilon,h}$, $m = 700$ (Theorem 3.4) of the mean inverse compression ratio for the real data related to paper production process.

We approximate the (theoretical) mean RLE quantization rates

$$L_{\varepsilon,h}(X) = \frac{T}{h} r_{\varepsilon,h}$$

in Examples 4.1 and 4.2 by the standard Monte-Carlo (MC) method and let $N$ denote the number of MC simulations.

Example 4.1. Let $X_1(t)$, $t \in [0, T]$, $T = 4.8$, be a stationary Gaussian process with zero mean and covariance function $K_1(t) = e^{-t^2}$. For the sampling step $h = 0.3$, we have, by Theorem 2.1 and (2.6),

$$L_{\varepsilon,h}(X_1) \sim \frac{T}{h} \rho_{\varepsilon,h}(K_1(0), K_1(h)) \quad \text{as } \varepsilon \to 0.$$ (4.1)

Figure 1 illustrates (4.1), i.e., performance of the approximation of $L_{\varepsilon,h}(X_1)$ by $(T/h) \rho_{\varepsilon,h}(K_1(0), K_1(h))$. Note the accuracy of the approximation even for moderate values of $\varepsilon \in [0, 2]$.

Example 4.2. Let for the sampling step $h = 0.1$,

$$\mathcal{D}_{m,h}(X_i) = \{X_i(jh), j = 1, \ldots, m\}, \quad i = 1, 2,$$

be two samples (or observed data) of zero mean Gaussian stationary processes $X_i(t)$, $t \geq 0$, $i = 1, 2$, with covariance functions $K_1(t) = e^{-t^2}$ and $K_2(t) = e^{-|t|}$, respectively. Let $T = 5$. Applying Theorems 3.2 and 3.4, we obtain the estimates

$$\frac{T}{h} \hat{\rho}_{\varepsilon,h}(X_i) = \frac{T}{h} \rho_{\varepsilon,h}(\hat{K}_i(0), \hat{K}_i(h))$$

and $(T/h) \hat{r}_{\varepsilon,h}(X_i)$, respectively, for the mean RLE quantization rate $L_{\varepsilon,h}(X_i, T)$ by using the data $\mathcal{D}_{m,h}(X_i), i = 1, 2$. 
Figure 1. The mean RLE quantization rate $L_{ε,h}(X_1)$ (solid line) and estimated $(T/h)ρ_{ε,h}$ (dashed line), $K_1(t) = e^{-t^2}$, $T = 4.8$, $h = 0.3$, $N = 10000$.

Figures 2 and 3 compare the approximations of $L_{ε,h}(X_i, T)$ by the estimates $(T/h)\hat{ρ}_{ε,h}(X_i)$ and $(T/h)\tilde{ρ}_{ε,h}(X_i)$, $i = 1, 2$, respectively, and show increasing estimation accuracy when the number of observations increases, $m = 5, 50, 500$. Notice also that the values of the mean RLE quantization rates for the more smooth process $X_1(t)$ are less than for those of the Ornstein–Uhlenbeck process $X_2(t)$, $t \geq 0$, cf. (2.10).

Figure 2. The mean RLE quantization rate $L_{ε,h}(X_i)$ (solid line) and the estimates $(T/h)ρ_{ε,h}(\hat{K}_i(0), \hat{K}_i(h))$, $i = 1$ (left), 2 (right), $m = 5, 50, 500$, $T = 5$, $h = 0.1$, $N = 5000$. 
Figure 3. The mean RLE quantization rate $L_{\varepsilon,h}(X_i)$ (solid line) and the estimates $(T/h)\hat{R}_{\varepsilon,h}(X_i)$, $i = 1$ (left), 2 (right), $m = 5, 50, 500$, $T = 5$, $h = 0.1$, $N = 5000$.

**Example 4.3** (Paper machine data). The paper production process includes many steps starting from wood chipping and ending with rolling final outcome paper. Wood pulping is one of the intermediate steps. During *thermomechanical pulping* (TMP) wood chips are softened by steam. This facilitates separation of cellulose and reduces damage of individual fibres (see, e.g., [3]). Various characteristics (e.g., TMP conductivity) are observed during the paper production process (e.g., for further analysis and prediction paper breaks). In this example, we consider TMP conductivity data during 12 hours with no breaks in additional 12 hours before and after, the time series $\{Y_k, k = 1, \ldots, 721\}$, one observation per minute, Figure 4.

Figure 4. TMP conductivity time series.
First, we fit the corresponding autoregressive AR(2) model for the data \( \{Y_k\} \) (see, e.g., [19]). Then we compare inverse compression ratio estimates \( \hat{\rho}_{\varepsilon,h}(Y) \) and \( \hat{r}_{\varepsilon,h}(Y) \) for various \( \varepsilon \).

Figure 5 illustrates the closeness of the estimates \( n, \hat{\rho}_{\varepsilon,h}(Y), m = 50, \) and \( n\hat{r}_{\varepsilon,h}(Y), m = 700, \) for the mean RLE quantization rate \( L_{\varepsilon,h}(Y) = nr_{\varepsilon,h}(Y), n = 721 \) and the TMP conductivity data.

![Figure 5. The estimators of the mean RLE quantization rate \( n\rho_{\varepsilon,h}(\hat{K}(0), \hat{K}(h)) \) (solid line) and \( n\hat{r}_{\varepsilon,h}(Y) \) (dashed line), TMP conductivity time series, \( n = 721 \).](image)

Note that the estimates are close even for a small enough sample size, \( m = 50 \). The predicted mean RLE quantization rates are not significant due to irregularity of the original data, Figure 4. The results of numerical experiments confirm that the plug-in estimator \( \hat{\rho}_{\varepsilon,h} \) is smoother as a function of \( \varepsilon \) than the nonparametric estimator \( \hat{r}_{\varepsilon,h} \) and with smaller variance, but biased. Note that the unbiased nonparametric \( \hat{r}_{\varepsilon,h} \) can be applicable for a more wide class of stochastic models.

### 5 Proofs

**Proof of Theorem 2.1.** We prove first the assertion for \( \sigma = 1 \). By definition it is enough to show that

\[
P(q_\varepsilon(X(0)) = q_\varepsilon(X(h))) = g(\varepsilon v(h))w(h) P(\varepsilon, h, \mu),
\]

for some \( P(\varepsilon, h, \mu) \) such that

\[
|P(\varepsilon, h, \mu) - 1| \leq \left( \frac{2}{\pi} \right)^{1/2} \left( 1 + \frac{\varepsilon^2}{4w(h)^2\sigma^2} \right). \tag{5.2}
\]
Let \( p_h(x, y) \) be the joint Gaussian density function of \( X(0) \) and \( X(h) \). We represent \( p_h(x, y) \) as follows:

\[
p_h(x, y) = \frac{1}{2\pi(1 - k(h)^2)^{1/2}} e^{-\frac{(x-y)^2}{2(1-k(h)^2)}} e^{-\frac{(x-\mu)(y-\mu)}{1+k(h)}}. \tag{5.3}
\]

Let \( I_k(\varepsilon) = [\varepsilon k, (\varepsilon k + 1)], k \in \mathbb{Z} \). By definition of the uniform quantizer \( q_\varepsilon(x) \),

\[
P(q_\varepsilon(X(0)) = q_\varepsilon(X(h))) = \sum_{k \in \mathbb{Z}} P(X(0) \in I_k(\varepsilon), X(h) \in I_k(\varepsilon))
\]

\[
= \sum_{k \in \mathbb{Z}} \iint_{I_k(\varepsilon)^2} p_h(x, y) dx dy = \frac{\theta(\varepsilon, h)}{\varepsilon} \int_{\mathbb{R}} f_{\varepsilon, h}(z) dz,
\]

where a piecewise constant function

\[
f_{\varepsilon, h}(z) := \theta(\varepsilon, h)^{-1} \iint_{I_k(\varepsilon)^2} p_h(x, y) dx dy \quad \text{for } z \in I_k(\varepsilon), k \in \mathbb{Z},
\]

\[
\theta(\varepsilon, h) := \frac{1}{2\pi(1 - k(h)^2)^{1/2}} \iint_{I_0(\varepsilon)^2} e^{-\frac{(x-y)^2}{2(1-k(h)^2)}} dx dy.
\tag{5.4}
\]

It follows by calculus that, for \( v(h) = (1 - k(h)^2)^{-1/2} \),

\[
\frac{\theta(\varepsilon, h)}{\varepsilon} = \frac{\varepsilon v(h)}{2\pi} \iint_{[0,1]^2} e^{-\frac{(x-y)^2}{2}v(h)^2} dx dy
\]

\[
= \frac{\varepsilon v(h)}{\pi} \int_0^1 (1 - x) e^{-\frac{x^2}{2}v(h)^2} dx
\]

\[
= (2\pi)^{-1/2} g(\varepsilon v(h)). \tag{5.5}
\]

The integral mean-value theorem applied to \( f_{\varepsilon, h}(z) \) together with (5.3) and (5.4) give that there exist \( x_k, y_k \in I_k(\varepsilon) \), \( x_k \leq y_k \), such that

\[
f_{\varepsilon, h}(z) = e^{-\frac{(x_k-\mu)(y_k-\mu)}{1+k(h)}} \quad \text{for } z \in I_k(\varepsilon), k \in \mathbb{Z}. \tag{5.6}
\]

Recall that \( w(h) = ((1 + k(h))/2)^{1/2} \). Let \( k_0 \in \mathbb{Z} \) be such that \( \mu \in I_{k_0}(\varepsilon) \). If \( x_{k_0} < y_{k_0} \), then it is clear that either \( \mu \in (x_{k_0}, y_{k_0}) \) or \( \mu \in I_{k_0}(\varepsilon) \setminus (x_{k_0}, y_{k_0}) \) holds. If \( \mu \in I_{k_0}(\varepsilon) \setminus (x_{k_0}, y_{k_0}) \), then \( (x_{k_0} - \mu)(y_{k_0} - \mu) \geq 0 \) and, by (5.6), we get

\[
\int_{\mathbb{R}} f_{\varepsilon, h}(z) dz = w(h) \sum_{k \in \mathbb{Z}} \frac{e^{-z_k^2/2\varepsilon}}{w(h)}, \tag{5.7}
\]

here \( w(h)z_k := ((x_k - \mu)(y_k - \mu))^{1/2} \in [k\varepsilon - \mu, (k + 1)e - \mu] \), \( k \in \mathbb{Z} \). Let

\[
P(\varepsilon, h, \mu) := (2\pi)^{-1/2} \sum_{k \in \mathbb{Z}} \frac{e^{-z_k^2/2\varepsilon}}{w(h)}.
\]
By the monotonicity of $e^{-y^2/2}$, $y \in \mathbb{R}$, we have

$$P(\varepsilon, h, \mu) = (2\pi)^{-1/2} \sum_{k \neq -1,0} \frac{e^{-z_k^2/2} \varepsilon}{w(h)}$$

$$+ (2\pi)^{-1/2} \left( \frac{e^{-z^2_{-1}/2} \varepsilon}{w(h)} + \frac{e^{-z^2_0/2} \varepsilon}{w(h)} \right)$$

$$\leq (2\pi)^{-1/2} \int_{\mathbb{R}} e^{-y^2/2} dy + 2(2\pi)^{-1/2} \frac{\varepsilon}{w(h)}$$

$$= 1 + \left( \frac{2}{\pi} \right)^{1/2} \frac{\varepsilon}{w(h)}$$ \hspace{1cm} (5.8)

and, similarly,

$$P(\varepsilon, h, \mu) \geq 1 - \left( \frac{2}{\pi} \right)^{1/2} \frac{\varepsilon}{w(h)},$$ \hspace{1cm} (5.9)

that is, inequality (5.2) holds (cf. the Euler–Maclaurin summation formula, [10]). If $\mu \in (x_{k_0}, y_{k_0})$, then $(x_{k_0} - \mu)(y_{k_0} - \mu) < 0$ and by (5.6), we obtain

$$\int_{\mathbb{R}} f_{\varepsilon, h}(z) dz = \left( \pi (1 + k(h)) \right)^{1/2} \left( P_1(\varepsilon, h, \mu) + P_2(\varepsilon, h, \mu) \right),$$ \hspace{1cm} (5.10)

where as in (5.8) and (5.9), we get

$$|P_1(\varepsilon, h, \mu) - 1| \leq \left( \frac{2}{\pi} \right)^{1/2} \frac{\varepsilon}{w(h)},$$

and

$$|P_2(\varepsilon, h, \mu)| = (2\pi)^{-1/2} \left| \exp \left\{ - \frac{(x_{k_0} - \mu)(y_{k_0} - \mu)}{1 + k(h)} \right\} \right.$$ 

$$- \exp \left\{ - \frac{(x_{k_0} - \mu)^2}{1 + k(h)} \right\} \frac{\varepsilon}{w(h)}$$

$$\leq (2\pi)^{-1/2} \left| \frac{(x_{k_0} - \mu)(y_{k_0} - x_{k_0})}{1 + k(h)} \right| \exp \left\{ \frac{\varepsilon^2}{1 + k(h)} \right\} \frac{\varepsilon}{w(h)}$$

$$\leq (2\pi)^{-1/2} \exp \left\{ \frac{\varepsilon^2}{2w(h)^2} \right\} \frac{\varepsilon^3}{2w(h)^3},$$ \hspace{1cm} (5.11)

by the Lagrange Theorem, so (5.2) holds also for

$$P(\varepsilon, h, \mu) = P_1(\varepsilon, h, \mu) + P_2(\varepsilon, h, \mu).$$
Now, combining (5.11) with (5.5) and (5.7) (or (5.10)), and taking into account (5.8) and (5.9), we obtain (5.1). Note that if \( \mu \in \{k\varepsilon, k \in \mathbb{Z}\} \), then \( q_\varepsilon(X(t)) = q_\varepsilon(X(t) - \mu) \) and the assertion follows by applying (5.8) and (5.9).

For \( \sigma > 0, \sigma \neq 1 \), the assertion follows immediately from (5.1) together with (5.2), if we notice that

\[
P(q_\varepsilon(X(0)) = q_\varepsilon(X(h))) = P(q_\varepsilon(\tilde{X}(0)) = q_\varepsilon(\tilde{X}(h)))
\]

for \( \tilde{X}(t) = X(t)/\sigma \) and \( \tilde{\varepsilon} = \varepsilon/\sigma \). This completes the proof of the theorem.

Proof of Theorem 2.2. First, we give the following elementary calculus result for further references.

Lemma 5.1. The function \( g(x) \) is continuous on \((0, +\infty)\) and

\[
g(x) = \begin{cases} 
\frac{x}{\sqrt{2\pi}} (1 + O(x^2)) & \text{as } x \to 0+, \\
1 - \frac{2}{\sqrt{\pi}} \frac{1}{x} + O\left(\frac{\phi(x)}{x^3}\right) & \text{as } x \to +\infty.
\end{cases}
\]

Proof of Lemma 5.1. By definition and Taylor expansion, we obtain

\[
g(x) = \sqrt{\frac{2}{\pi}} x \int_0^1 (1-u)e^{-x^2 u^2} \frac{du}{x} = \frac{x}{\sqrt{2\pi}} (1 + O(x^2)) \quad \text{as } x \to 0 .
\]

Now integration by part gives

\[
g(x) = 1 - 2 \left( \overline{\Phi}(x) - \frac{\phi(x)}{x} \right) - \left( \frac{2}{\pi} \right)^{1/2} \frac{1}{x}, \quad x > 0,
\]

\[
\overline{\Phi}(x) - \frac{\phi(x)}{x} = O\left(\frac{\Phi(x)}{x^2}\right) = O\left(\frac{\phi(x)}{x^3}\right) \quad \text{as } x \to \infty,
\]

where \( \overline{\Phi}(x) := 1 - \Phi(x) \). This completes the proof of the lemma.

We prove the statement when \( \sigma = 1 \). For \( \sigma > 0, \sigma \neq 1 \), it is enough to recall (5.12). Note that \( v(h) \sim (2\gamma(h))^{-1/2} \) as \( h \to 0 \), \( \gamma(h) = 1 - k(h) \). Therefore, all three cases follow straightforward from Theorem 2.1 and Lemma 5.1. Indeed, if \( \varepsilon v(h) \to 0 \) as \( \varepsilon \to 0 \), then

\[
r_{\varepsilon, h}(X) = 1 - \frac{\varepsilon v(h)}{\sqrt{2\pi}} (1 + O(\varepsilon^2 v(h)^2)) (w(h) + \varepsilon \delta(\varepsilon, h, \mu))
\]

\[
= 1 - \frac{w(h)}{\sqrt{2\pi}} \varepsilon v(h) + O(\varepsilon^2 v(h))
\]

\[
\sim 1 - \frac{\varepsilon}{2(\pi(1-k(h)))^{1/2}} \quad \text{as } \varepsilon \to 0.
\]
Note that $v(h) \to \infty$ is equivalent to $k(h) \to 1$, that is, $h \to 0$. Further, if $\varepsilon v(h) \to \tau$ as $\varepsilon \to 0$, $0 < \tau < \infty$, then $v(h) \to \infty$ and

$$r_{\varepsilon,h}(X) = 1 - (g(\tau) + o(1))(w(h) + \varepsilon \delta(\varepsilon, h)) = 1 - w(h)g(\tau) + o(1)$$

$$\sim 1 - g(\tau) \quad \text{as } \varepsilon \to 0.$$

Finally, let $\varepsilon v(h) \to \infty$ as $\varepsilon \to 0$. Then $v(h) \to \infty$ and

$$1 - w(h) = \frac{1 - k(h)}{2}(1 + w(h))^{-1} = O\left(\frac{1}{v(h)^2}\right)$$

$$= O\left(\frac{\varepsilon^2}{\varepsilon^2 v(h)^2}\right) = o(\varepsilon^2) \quad \text{as } \varepsilon \to 0.$$

Thus, by Lemma 5.1 we get

$$r_{\varepsilon,h}(X) = 1 - \left(1 - \sqrt{\frac{2}{\pi} \frac{1}{\varepsilon v(h)}} + O\left(\frac{\phi(\varepsilon v(h))}{\varepsilon^3 v(h)^{3/2}}\right)\right)(1 + O(\varepsilon))$$

$$= \sqrt{\frac{2}{\pi} \frac{1}{\varepsilon v(h)}} + O(\varepsilon) + o\left(\frac{1}{\varepsilon v(h)}\right)$$

$$= \sqrt{\frac{2}{\pi} \frac{1}{\varepsilon v(h)}} \left(1 + O(\varepsilon^2 v(h)) + o(1)\right)$$

$$= \frac{T}{h} \left(\frac{1 - k(h)}{\pi}\right)^{1/2} \frac{2}{\varepsilon} (1 + o(1)) \quad \text{as } \varepsilon \to 0,$$

if, additionally, $\varepsilon^2 v(h) \to 0$ as $\varepsilon \to 0$. This completes the proof of the theorem. □

Proof of Proposition 3.1. Assertion (i) of the proposition follows immediately by [1, Corollary 8.4.1]. Further, it follows by [2, Theorem 3] applied to the vector Gaussian time series $Y_j := (X(jh), X((j + 1)h))$, $j \geq 1$, that $\hat{\gamma}(h)$ is an asymptotically normal estimator of $\gamma(h)$. In fact, three conditions should be verified in this case,

(a) $E(X(h) - X(0))^4 < \infty$.

There exist limits

(b) $\lim_{m \to \infty} 1/m \sum_{i, j = 0}^m K((i - j)h)$,

(c) $\lim_{m \to \infty} 1/m \sum_{i, j = 0}^m K((i - j)h)^2$.

Condition (b) follows from simple algebra and (B),

$$\frac{1}{m} \sum_{i, j = 0}^m K((i - j)h) = \sum_{k = 0}^m \left(1 - \frac{k}{m}\right) K(kh) \to \sum_{k = 0}^\infty K(kh) \quad \text{as } m \to \infty,$$
and similarly we get (c). We can assume that $Y_j$ is zero-mean since the estimators are shift invariant.

By the Cramér–Slutsky Theorem, we get the asymptotic normality of $\hat{\gamma}_K(h)$, since

$$m^{1/2}(\hat{\gamma}_K(h) - \gamma(h)) = m^{1/2}(\hat{\gamma}(h) - \gamma(h)) + \frac{m^{1/2}}{2(m-1)}d_m,$$

where

$$E|d_m| = E|(X(h) - \hat{\mu})^2 + (X(mh) - \hat{\mu})^2 - 2\hat{K}(0)| \leq 16\sigma^2.$$

so, the assertion (i i) holds. This completes the proof of the proposition. \hfill \Box

**Proof of Theorem 3.2.** By Proposition 3.1 (i), $\hat{K}(0)$ and $\hat{K}(h)$ are asymptotically joint normal estimators. Thus, $\hat{\rho}_{\varepsilon,h} = \rho_{\varepsilon,h}(\hat{K}(0), \hat{K}(h))$, as a continuously differentiable two-dimensional function of $\hat{K}(0)$ and $\hat{K}(h)$, is also an asymptotically normal estimator of $\rho_{\varepsilon,h}$ (see, e.g., [13, Chapter 6a.2]). In a similar way, we obtain the asymptotic normality of $\hat{B}^l_{\varepsilon,h}$ and $\hat{B}^u_{\varepsilon,h}$. This completes the proof. \hfill \Box

**Proof of Theorem 3.4.** The assertion is an immediate consequence of [2, Theorem 3] applied to the stationary Gaussian vector time series

$$Y_j = (X(jh), X((j + 1)h)), \quad j \geq 1,$$

with a similar argument as in Proposition 3.1. This completes the proof. \hfill \Box

**Acknowledgments.** The authors would like to thank Professor Patrik Eklund, Computer Science Institute, Umeå University, for the provided datasets of paper production process measurements and helpful discussions.

**Bibliography**


Received June 27, 2011; accepted August 7, 2012.

**Author information**

Oleg Seleznjev, Department of Mathematics and Mathematical Statistics, Umeå University, 90187 Umeå, Sweden.
E-mail: oleg.seleznjev@math.umu.se

Mykola Shykula, Department of Statistics, Umeå University, 90187 Umeå, Sweden.
E-mail: mykola.shykula@stat.umu.se