EURASIP Journal on
Advances in Signal Processing
a SpringerOpen Journal

Rate distortion performance analysis of nested sampling and coprime sampling

Junjie Chen[1*] and Qilian Liang[1,2]

Abstract

In this paper, rate distortion performance of nested sampling and coprime sampling is studied. It is shown that with the increasing of distortion, the data rate decreases. With these two sparse sampling algorithms, the data rate is proved to be much less than that without sparse sampling. With the increasing of sampling spacings, the data rate decreases at certain distortion, which is because with more sparse sampling, less number of bits is required to represent the information. We also prove that with the same sampling pairs, the rate of nested sampling is less than that of coprime sampling at the same distortion. The reason is that nested sampling collects a little less number of samples than coprime sampling with the same length of data, which is a little sparser than coprime sampling.

Keywords: Sparse sampling; Rate distortion; Information theory; Nested sampling; Coprime sampling

1 Introduction

The twenty-first century is awash with data. Data are flooding in at rates never seen before, doubling almost every 18 months [1], as result of new information gathered from Radar, Web communities, newly deployed smart assets, and customer data from public, proprietary, purchased sources, and so forth. For example, oil companies, telecommunication companies, and other data-centric industries have had huge data for long time. Data is being collected and transmitted at unprecedented scale [2,3] in a wide range of application areas nowadays. The phrase 'Big Data' as defined by US National Science Foundation in its recent solicitation, refers to large, diverse, complex, distributed data sets generated from instruments, sensors, Internet transactions, email, video, click streams, and all other digital sources available today and in the future.

Unstructured data is data that does not follow a specified format, which is really most of the Big Data. Radar or sonar data is one typical example, which includes meteorological, vehicular, and oceanographic seismic information, such as in [4], Big Data from O'Reilly radar was described.

And many efforts have been made to develop suitable compression techniques for Big Data. However, traditional compression methods [5,6] are all based on Nyquist rate, which will have poor efficiency in terms of both sampling rate and computational complexity. Unlike traditional compression techniques, some sparse sampling algorithms have been proposed to overcome Nyquist sampling requirement, like compressive sensing, nested sampling, and coprime sampling.

Nested sampling [7] is an non-uniform sampling, using two different samplers in each period. Although the signal is sampled sparsely and non-uniformly, the autocorrelation of signal could be estimated at all lags. Therefore, although the samples can be arbitrarily sparse, it keeps the signal's statistical information [8]. While coprime sampling uses two uniform samplers, with sample spacings P and Q coprime integers. The authors in [8] have already proved that these two sets of samples of the signal could fully estimate all lags of autocorrelation of the original signal. As both nested sampling and coprime sampling could keep the statistical property of the original signal, these two sampling algorithms could be applied to Big Data to highly reduce the transmission or storage cost of Big Data.

Information rate distortion function is a measure of distortion between the original source and its representation. In this paper, we will provide theoretical rate distortion performance, because of these two sparse sampling algorithms, either nested sampling (NS) or coprime sampling (CS). We will show that with these two sparse sampling algorithms, the data rate is much less than that without

*Correspondence: junjie.chen@mavs.uta.edu
[1] Department of Electrical Engineering, University of Texas at Arlington, 416 Yates St, Arlington, TX 76019, USA
Full list of author information is available at the end of the article

sparse sampling for a given distortion. With the increasing of sampling spacings, the data rate decreases at certain distortion, which is because with more sparse sampling, less number of bits is required to represent the information. We will also prove that with the same sampling pairs, the rate of nested sampling is less than that of coprime sampling at the same distortion.

The rest of this paper is organized as follows. In Section 2, we give a brief introduction of nested sampling and coprime sampling separately. Theoretical derivation of rate distortion performance of nested sampling and coprime sampling is detailed in Sections 3.1 and 3.2. Also, the theoretical analysis and comparison of these two sparse sampling is given in Section 3.3. In Section 4, numerical results are provided to verify the theoretical rate distortion results derived in Section 3. Conclusions are given in Section 5.

2 Preliminaries
2.1 Nested sampling
The nested array was first introduced in [7] as an effective approach to array processing with enhanced degrees of freedom [9,10]. In time domain, the signal's autocorrelation could also be obtained from nested sampling structure [11]. And although the samples from this nested sparse sampling are sparsely and non-uniformly located, the samples of the autocorrelation can be computed at any specified rate. Some applications which depend on the difference co-array, or autocorrelation, like Direction-of-arrival (DOA) estimation and beamforming could be done based on nested sampling.

In the simplest form, there are two levels of sampling density in nested sampling [11-14], with the level 1 samples at the N_1 locations and the level 2 samples at the N_2 locations.

$1 \leq l \leq N_1$, for level 1

$(N_1 + 1)m, 1 \leq m \leq N_2$, for level 2

An example of periodic sparse sampling using nested sampling structure is shown in Figure 1, with $N_1 = 3$ and $N_2 = 5$. The cross differences are given by

$$k = (N_1 + 1)m - l, 1 \leq m \leq N_2, 1 \leq l \leq N_1. \qquad (1)$$

The cross differences (2) are in the following range with the maximum value $(N_1 + 1)N_2 - 1$ [9,11], except the integers and the corresponding negated versions shown in (3).

$$-[(N_1 + 1)N_2 - 1] \leq k \leq [(N_1 + 1)N_2 - 1] \qquad (2)$$

$$(N_1 + 1), 2(N_1 + 1), \cdots, (N2 - 1)(N_1 + 1) \qquad (3)$$

For example, consider the example in Figure 1, where $N_1 = 3$, $N_2 = 5$, $1 \leq m \leq 5$ and $1 \leq l \leq 3$, the cross differences $k = (N_1 + 1)m - l$ will achieve these values

$$1, 2, 3, (), 5, 6, 7, (), 9, 10, 11, (), 13, 14, 15, (), 17, 18, 19.$$

The difference 0 is also missing, for the reason that m and l are nonzero. Meanwhile, we notice that all of the missing differences could be covered by the self differences among the second array,

$$(N_1 + 1)(m_1 - m_2), 1 \leq m_1, m_2 \leq N_2. \qquad (4)$$

Combining the cross differences and the self-differences is the difference-co-array, which is a filled difference co-array as shown in (2). This indicates that using nested array structure, although sparse samples are obtained, the degrees of freedom is enhanced,

$$2[(N_1 + 1)N_2 - 1] + 1 = 2(N_1 + 1)N_2 - 1. \qquad (5)$$

Based on the principle above, a sparse non-uniform sampling using nested sampling structure could be performed as in Figure 1. There are two levels of nesting, with N_1 level-1 samples and N_2 level-2 samples in each period, with period $(N_1 + 1)N_2$. It is obvious that nested sampling is non-uniform and the samples obtained are very sparse.

We could notice that, in $(N_1 + 1)N_2 T$ seconds, there are totally $N_1 + N_2$ samples. Therefore, the average sampling rate is

$$f_{s,\text{nested}} = \frac{N_1 + N_2}{(N_1 + 1)N_2 T} \approx \frac{1}{N_1 T} + \frac{1}{N_2 T} < \frac{1}{T} \qquad (6)$$

Here, $T = 1/f_n$, $f_n \geq 2f_{\max}$ is the Nyquist sampling frequency. As the Nyquist sampling rate is $1/T$, the
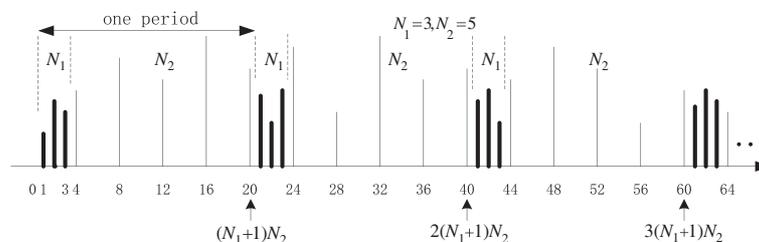

Figure 1 Nested sampling with $N_1 = 3, N_2 = 5$.

average sampling rate of nested sampling is smaller than the conventional Nyquist sampling rate.

2.2 Co-prime sampling

Different from nested sampling, coprime sampling involves two sets of uniformly spaced samplers [8,12-14] as shown in Figure 2, each of them are sparsely sampled.

The coprime sampling uniformly sample the source signal using two sub-Nyquist samplers, with sample spacing PT and QT respectively, where P and Q are coprime integers with $P < Q$. $1/T$ Hz is the Nyquist rate for a bandlimited process, i.e., $1/T = 2f_{\max}, f_{\max}$ being the highest frequency.

$$x(n) = x_c(nT). \tag{7}$$

Consider the product

$$x(Pn_1)x^*(Qn_2), \tag{8}$$

where $x(Pn_1)$ and $x(Qn_2)$ comes from the first and the second sampler. Let the difference as

$$k = Pn_1 - Qn_2. \tag{9}$$

It has been shown that k can achieve any integer value in the range $0 \leq k \leq PQ - 1$ in [11], if n_1 and n_2 are in the ranges $0 \leq n_1 \leq 2Q - 1$ and $0 \leq n_2 \leq P - 1$.

For coprime sampling, the two samplers totally collect $P + Q$ samples in PQT seconds, so the average sampling rate is

$$f_{s,\text{coprime}} = \frac{P + Q}{PQT} = \frac{1}{PT} + \frac{1}{QT} < \frac{1}{T}. \tag{10}$$

Same as in nested sampling, $T = 1/f_n, f_n \geq 2f_{\max}$ is the Nyquist sampling frequency. It is obvious the average sampling rate of coprime sampling is also much smaller than the conventional Nyquist sampling rate. However, the signal's second-order statistics, like the autocorrelation, is kept, which allows us to sample a signal sparsely and estimate some aspects of the signal (spectra, DOA, and so on) at a significantly higher resolution.

3 Rate distortion performance

Information rate distortion function is a measure of distortion between the original source and its representation. Our purpose is to construct a distortion function

which can measure the distortion because of these two sparse sampling algorithms, either nested sampling (NS) or coprime sampling (CS). Sparse sampling can cause possible distortion because less number of samples are used. A wide variety of distortion functions, such as Euclidean distance, Hamming distance, Mahalanobis distance, and Itakura-Saito distance have been used. In this paper, squared error distortion is used. The original samples are denoted as $x_i, i = 1, \cdots, L$, where L is the total number of samples. Assume that all original information from L samples is $X^L = [x_1, x_2, \cdots, x_L]$, the selected information after sparse sampling can be represented as [15]

$$\hat{X}^{L'} = S(X^L), \tag{11}$$

where $S(\cdot)$ denotes sparse sampling, either nested sampling or coprime sampling. $\hat{X}^{L'} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{L'}]$ and $L' < L$. The distortion associated with the sparse sampling between all original samples and the selected samples is

$$D = Ed(X^L, \hat{X}^{L'}), \tag{12}$$

where $d(\cdot)$ is the distortion function.

The expectation in (12) is with respect to the probability distribution on X^L. The rate distortion function $R(D)$ is the minimum of data rates R such that (R, D) is in the rate distortion region for a given distortion. From [16,17], we know that information rate distortion function is defined as

$$R(D) = \min_{Ed(X^L, \hat{X}^{L'}) \leq D} I(X^L; \hat{X}^{L'}), \tag{13}$$

where $I(X^L; \hat{X}^{L'})$ is the mutual information between X^L and $\hat{X}^{L'}$.

$$I(X^L; \hat{X}^{L'}) = H(X^L) - H(X^L | \hat{X}^{L'})$$

$$= H(X^L) - H(X^L - \hat{X}^{L'} | \hat{X}^{L'}) \tag{14}$$

$$\overset{(a)}{\geq} H(X^L) - H(X^L - \hat{X}^{L'}),$$

where inequality (a) follows from the fact that condition reduces the entropy.

From formula (13), we know that

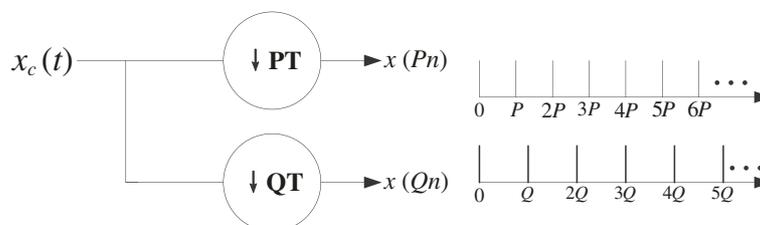$$Ed(X^L, \hat{X}^{L'}) \leq D \tag{15}$$

Figure 2 Coprime sampling.

For squared error distortion,

$$
\begin{aligned}
Ed(X^L, \hat{X}^{L'}) &= E(X^L - \hat{X}^{L'})^2 \\
&= \sum E(x_i - \hat{x}_j)^2 \\
&\overset{(b)}{=} \sum_k D_k \le D,
\end{aligned}
\tag{16}
$$

where $i = 1, \cdots, L$ and $j = 1, \cdots, L'$, and (b) follows from the definition that $E(x_i - \hat{x}_j)^2 = D_k$.

Since Gaussian assumption is a classical modeling assumption heavily used in areas such as signal processing and communication system [18], from [16], the rate distortion function for a single Gaussian source $N(0, \sigma^2)$ with squared error distortion is

$$
R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} & 0 \le D \le \sigma^2, \\ 0 & D > \sigma^2 \end{cases}
\tag{17}
$$

For L-independent zero mean Gaussian sources $x_1, \cdots, x_L$ with variance $\sigma_1^2, \sigma_2^2, \cdots, \sigma_L^2$, the rate distortion performance with squared error distortion is given by [16,17,19,20]

$$
R(D) = \sum_{i=1}^{L} \frac{1}{2} \log \frac{\sigma_i^2}{D_i},
\tag{18}
$$

where

$$
D_i = \begin{cases} \lambda & \text{if } \lambda < \sigma_i^2, \\ \sigma_i^2 & \text{if } \lambda \ge \sigma_i^2 \end{cases}
\tag{19}
$$

where λ is chosen so that $\sum_i^L D_i = D$, and $D_i = E(x_i - \hat{x}_i)^2$. This gives rise to a kind of reverse waterfilling. We choose a constant λ and only describe those random variables with variance greater than λ, and no bits are used to describe random variables with variance less than λ.

3.1 For nested sampling

Theorem 1. *(Rate distortion for nested sampling of Gaussian source) Let $x_i \sim N(0, \sigma_i^2)$, $i = 1, 2, \cdots, L$, be independent Gaussian random variables, and under squared error distortion. The rate distortion between the original Gaussian source and after nested sampling of these Gaussian random variables is given by*

$$
R_{NS}(D) = \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 - \sum_{k=1}^{K_{NS}} \frac{1}{2} \log 2\pi e D_k,
\tag{20}
$$

where K_{NS} is given in (24) and

$$
D_k = \begin{cases} \lambda & \text{if } \lambda < \sigma_k^2, \\ \frac{K_{NS}}{L} \sigma_k^2 & \text{if } \lambda \ge \sigma_k^2. \end{cases}
\tag{21}
$$

where λ is chosen so that $\sum_{k=1}^{K_{NS}} D_k = D$.

Proof 1. *For nested sampling (NS), all L original information is $X^L = [x_1, x_2, \cdots, x_L]$.*

And less number of samples L' will be selected based on nested sampling as described,

$$
\begin{aligned}
\hat{X}_{NS}^{L'} &= [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{L'}] \\
&= [x_1, \cdots, x_{N_1}, x_{(N_1+1)}, \cdots, x_{N_2(N_1+1)}, \cdots]
\end{aligned}
\tag{22}
$$

Therefore, (16) becomes

$$
Ed(X^L, \hat{X}_{NS}^{L'}) = \sum_{k=1}^{K_{NS}} D_k,
\tag{23}
$$

where the length of K_{NS} could be determined based on the following formula, here we assume $Y = L(\mod (N_1+1)N_2)$

$$
K_{NS} = \begin{cases} \lfloor \frac{L}{(N_1+1)N_2} \rfloor N_1(N_2 - 1), & \text{if } Y \le (N_1+1) \\ \lfloor \frac{L}{(N_1+1)N_2} \rfloor N_1(N_2 - 1) + Z, & \text{if } Y > (N_1+1), \end{cases}
\tag{24}
$$

where

$$
Z = \lfloor (\frac{Y}{N_1+1} - 1) \rfloor * N_1 + U
\tag{25}
$$

in which $U = (Y - (N_1+1))(\mod (N_1+1))$.

If all samples are assumed to be independent Gaussian $N(0, \sigma_i^2)$, hence, the corresponding rate distortion function for nested sampling will be

$$
\begin{aligned}
R_{NS}(D) &= \min_{Ed(X^L, \hat{X}_{NS}^{L'}) \le D} I(X^L; \hat{X}_{NS}^{L'}) \\
&\ge \min_{Ed(X^L, \hat{X}_{NS}^{L'}) \le D} H(X^L) - H(X^L - \hat{X}_{NS}^{L'}) \\
&\overset{(c)}{\ge} \min_{\sum_{k=1}^{K_{NS}} D_k = D} H(X^L) - H(N(0, E(X^L - \hat{X}_{NS}^{L'})^2)) \\
&= \min_{\sum_{k=1}^{K_{NS}} D_k = D} \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 - \sum_{k=1}^{K_{NS}} \frac{1}{2} \log 2\pi e D_k,
\end{aligned}
\tag{26}
$$

where inequality (c) follows from the fact that the normal distribution maximizes the entropy for a given second moment, and $\sum_{k=1}^{K_{NS}} D_k = D$.

To find the minimum value, we could use Lagrange multipliers

$$
J(D) = \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 - \sum_{k=1}^{K_{NS}} \frac{1}{2} \log 2\pi e D_k + \sum_{k=1}^{K_{NS}} D_k
\tag{27}
$$

and differentiating with respect to D_k and setting equal to 0, we have

$$\frac{\partial J}{\partial D_k} = -\frac{1}{D_k 2 \ln 2} + \lambda = 0 \tag{28}$$

or

$$D_k = \frac{1}{\lambda 2 \ln 2} = \lambda' \tag{29}$$

which results in an equal distortion for each random variable, if the constant λ' is less than σ_i^2 for all i. As the increase of the total allowable distortion D, the constant λ' increases until it exceeds σ_i^2 for some i. Kuhn-Tucker conditions could be used to find the minimum in (26) if we increase the total distortion D. In this case, the Kuhn-Tucker conditions yield

$$\frac{\partial J}{\partial D_k} = -\frac{1}{D_k 2 \ln 2} + \lambda. \tag{30}$$

Therefore,

$$D_k = \begin{cases} \lambda & \text{if } \lambda < \sigma_k^2 \\ \frac{K_{\text{NS}}}{L} \sigma_k^2 & \text{if } \lambda \geq \sigma_k^2, \end{cases} \tag{31}$$

where λ is chosen so that $\sum_{k=1}^{K_{\text{NS}}} D_k = D$.

3.2 For coprime sampling

Theorem 2. *(Rate distortion for coprime sampling of Gaussian source) Let $x_i \sim N(0, \sigma_i^2)$, $i = 1, 2, \cdots, L$, be independent Gaussian random variables, and under squared error distortion. The rate distortion between the original Gaussian source and after coprime sampling of these Gaussian random variables is given by*

$$R_{\text{CS}}(D) = \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 - \sum_{k=1}^{K_{\text{CS}}} \frac{1}{2} \log 2\pi e D_k, \tag{32}$$

where K_{CS} is given in (36) and

$$D_k = \begin{cases} \lambda & \text{if } \lambda < \sigma_k^2 \\ \frac{K_{\text{CS}}}{L} \sigma_k^2 & \text{if } \lambda \geq \sigma_k^2, \end{cases} \tag{33}$$

where λ is chosen so that $\sum_{k=1}^{K_{\text{CS}}} D_k = D$.

Proof 2. *For coprime sampling (CS), we still assume the original information with length L, i.e., $X^L = [x_1, x_2, \cdots, x_L]$.*

And based on coprime sampling, less number of samples L'' will be selected,

$$\begin{aligned} \hat{X}_{\text{CS}}^{L''} &= [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{L''}] \\ &= [x_P, x_Q, x_{2P}, x_{2Q}, \cdots]. \end{aligned} \tag{34}$$

Similarly, (16) becomes

$$Ed(X^L, \hat{X}_{\text{CS}}^{L''}) = \sum_{k=1}^{K_{\text{CS}}} D_k, \tag{35}$$

where the length of K_{CS} could be determined based on the following formula

$$K_{\text{CS}} = L - \lfloor \frac{L}{P} \rfloor - \lfloor \frac{L}{Q} \rfloor + \lfloor \frac{L}{PQ} \rfloor. \tag{36}$$

Therefore, the corresponding rate distortion function for coprime sampling of independent Gaussian source $N(0, \sigma_i^2)$ is

$$\begin{aligned} R_{\text{CS}}(D) &= \min_{Ed(X^L, \hat{X}_{\text{CS}}^{L''}) \leq D} I(X^L; \hat{X}_{\text{CS}}^{L''}) \\ &\geq \min_{Ed(X^L, \hat{X}_{\text{CS}}^{L''}) \leq D} H(X^L) - H(N(0, E(X^L - \hat{X}_{\text{CS}}^{L''})^2)) \\ &= \min_{\sum D_k = D} \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 - \sum_{k=1}^{K_{\text{CS}}} \frac{1}{2} \log 2\pi e D_k. \end{aligned} \tag{37}$$

The minimum value could be obtained using the similar procedure as described in nested sampling.

3.3 Theoretical analysis

Without sparse sampling, the rate distortion function would be

$$\begin{aligned} R_{\text{WSS}}(D) &= \min_{Ed(X^L, X^L) \leq D} I(X^L; X^L) \\ &\geq H(X^L) - H(N(0, E(X^L - X^L)^2)) \\ &= \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 - 0 \\ &= \sum_{i=1}^{L} \frac{1}{2} \log 2\pi e \sigma_i^2 \end{aligned} \tag{38}$$

which is much greater than that with sparse sampling.

From the above derivation of rate distortion function of nested sampling and coprime sampling, we could notice that if the sampling spacings are assumed to be the same, i.e., $N_1 = P$ and $N_2 = Q$ for these two sparse sampling methods, then the minimum value of $K_{\text{NS}_{\min}}$ could be achieved when $Y = L(\text{mod}\,(N_1 + 1)N_2) = 0$, therefore

$$K_{\text{NS}_{\min}} = \frac{N_1(N_2 - 1)L}{(N_1 + 1)N_2} = \frac{P(Q-1)L}{(P+1)Q} = \frac{(P^2 Q - P^2)L}{PQ(P+1)}. \tag{39}$$

Table 1 K_{NS} and K_{CS} with respect to sampling intervals when $N_1 = P$, $N_2 = Q$, and $L = 1{,}000$

$N_1 = P$	$N_2 = Q$	K_{NS}	K_{CS}
3	4	561	500
3	5	600	533
3	7	642	572
3	11	681	607
3	13	690	616
3	17	705	628
3	23	717	638
5	23	794	765
7	23	833	821
11	23	873	870

While for coprime sampling, the minimum value of $K_{CS_{min}}$ could be achieved when $L \pmod P = 0$, $L \pmod Q = 0$, and $L \pmod{PQ} = 0$, therefore

$$K_{CS_{min}} = \frac{(P-1)(Q-1)L}{PQ} = \frac{(P^2Q - P^2 - Q + 1)L}{PQ(P+1)}.$$
(40)

As we know that for these two sparse sampling algorithms, the sampling interval is for sure greater than Nyquist sampling spacing, which indicates that $Q > 1$, therefore,

$$K_{NS_{min}} > K_{CS_{min}}$$
(41)

which indicates that in most cases, $K_{NS} > K_{CS}$. Table 1 shows some example of K_{NS} and K_{CS} with respect to sampling intervals when $N_1 = P$, $N_2 = Q$, and $L = 1{,}000$. It is clear that with the increase of sampling spacings, samples are selected more sparsely by both nested sampling
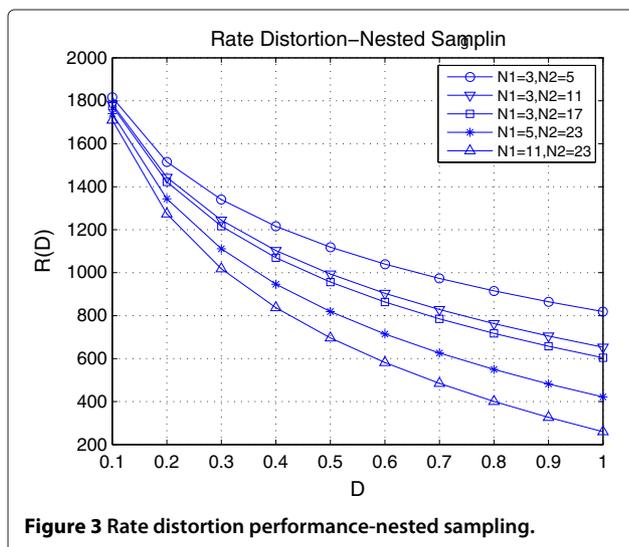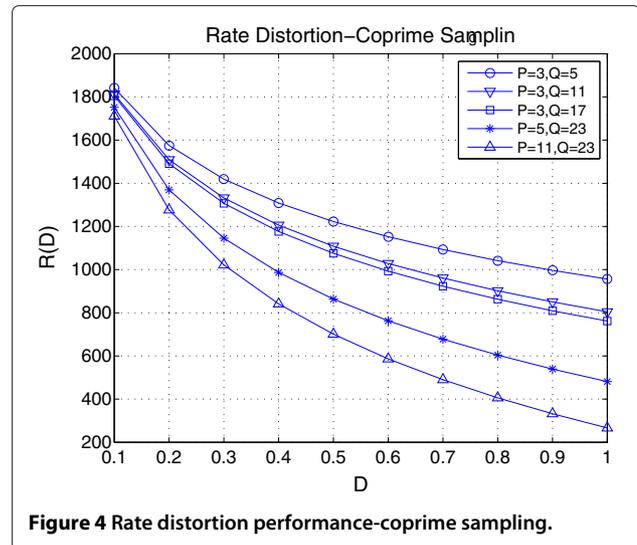

Figure 4 Rate distortion performance-coprime sampling.

and coprime sampling, which results in a increase of K_{NS} and K_{CS}. In addition, we could notice that $K_{NS} > K_{CS}$ as proved.

With our assumption that all samples are independent Gaussian $N(0, \sigma_i^2)$, we could conclude that

$$R_{NS}(D) < R_{CS}(D) < R_{WS}(D)$$
(42)

which indicates that both nested sampling and coprime sampling use less number of bits to describe the information compared that without sparse sampling (WS).

As we know from the introduction part, in $(N_1 + 1)N_2T$ seconds, there are totally $N_1 + N_2$ samples for nested sampling, while coprime sampling totally collect $P + Q$ samples in PQT seconds. If the sampling intervals are the same, i.e., $N_1 = P$ and $N_2 = Q$, it is obvious that nested sampling is a little sparser than coprime sampling method.
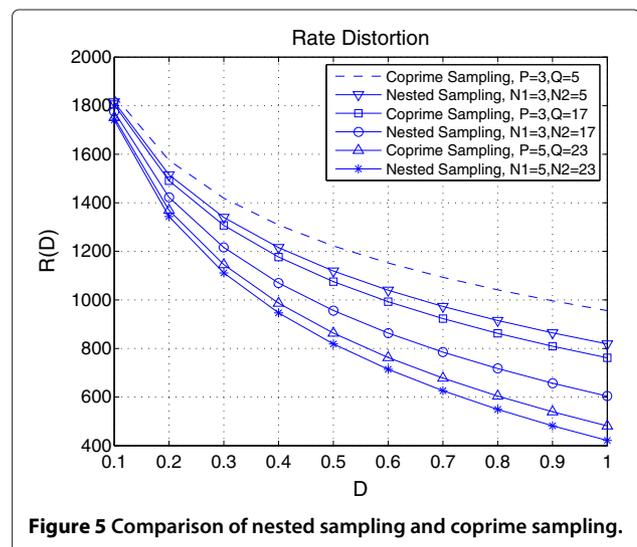

Figure 3 Rate distortion performance-nested sampling.

Figure 5 Comparison of nested sampling and coprime sampling.

$R_{NS}(D) < R_{CS}(D)$ is because nested sampling collects a little less number of samples than coprime sampling with the same length L of data. The rate $R(D)$ at a given distortion for both sparse sampling algorithms is less than that without sparse sampling. The reason is that with sparse sampling, less number of bits is used to describe the original information.

4 Numerical results

The total length of the information is set to be $L = 1,000$. Each sample is assumed to follow a Gaussian distribution $N(0, 1)$ with zero mean and unit variance. We also assume $D_k = \lambda < \sigma^2 = 1$, which is equal distortion for each random variable.

Figure 3 shows the rate distortion performance of nested sampling with different sampling spacings. It is clear that with the increasing of distortion, the rate decreases. When the sampling intervals N_1 and N_2 becomes larger, i.e., less samples are acquired, the rate becomes smaller. For example, when $D = 0.3$, $N_1 = 3, N_2 = 5$, the data rate $R(D) \approx 1,350$, while with the increase of sampling pairs to $N_1 = 3, N_2 = 11$, then $R(D) \approx 1220$, which is much smaller. This is because with more sparse sampling, less number of bits is required to represent the information.

The rate distortion performance of coprime sampling with different sampling spacings is shown in Figure 4. Similarly as nested sampling, with the increasing of distortion, the rate $R(D)$ decreases. When the sampling intervals P and Q becomes larger, the rate becomes smaller.

Figure 5 compares the rate distortion performance between nested sampling and coprime sampling, where D is the distortion between the original source and its sparse-sampled representation, and $R(D)$ is the corresponding rate at a particular distortion D. With the same sampling spacings chosen, $N_1 = P$, and $N_2 = Q$, at the same distortion, the rate of nested sampling is less than that of coprime sampling. For example, when $N_1 = P = 3$, and $N_2 = Q = 17$, when $D = 0.3$, the rate for nested sampling is $R_{NS}(D) \approx 1,200$, while the rate for coprime sampling is $R_{CS}(D) \approx 1,300$. This verifies the result that $R_{NS}(D) < R_{CS}(D)$, because nested sampling collects a little less number of samples than coprime sampling with the same length L of data, which is a little sparser than coprime sampling.

5 Conclusions

Information rate distortion function is a measure of distortion between the original source and its representation. Our purpose in this paper is to construct a distortion function which can measure the distortion because of these two sparse sampling algorithms. Information theoretical rate distortion performance for these two sparse sampling methods, nested sampling and coprime sampling, is studied in this paper. It is showed that with these two sparse

sampling algorithms, the data rate is proved to be much less than that without sparse sampling at a given distortion. With the increasing of sampling spacings, i.e., data are more sparsely acquired, the data rate decreases at certain distortion. This is because with sparser sampling, less number of bits is required to represent the information. We also show that with the same sampling pairs, the rate of nested sampling is less than that of coprime sampling at the same distortion.

Author details
[1]Department of Electrical Engineering, University of Texas at Arlington, 416 Yates St, Arlington, TX 76019, USA. [2]College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China.

References
1. J Bughin, M Chui, J Manyika, *Clouds, Big Data, and Smart Assets: Ten Tech-enabled Business Trends to Watch*, vol. 56. (McKinsey Quarterly, New York, 2010), pp. 75–86
2. A Labrinidis, HV Jagadish, Challenges and opportunities with big data. Proc. VLDB Endowment. **5**(12), 2032–2033 (2012)
3. D Bollier, CM Firestone, *The Promise and Peril of Big Data* (Aspen Institute, Communications and Society Program, Washington, DC, 2010)
4. O'Reilly Radar Team, *Big Data Now: Current Perspectives from O'Reilly Radar* (O'Reilly Media, Sebastopol, 2011)
5. J Chen, Q Liang, J Paden, P Gogineni, in *Proceedings of the IEEE International Conference on Communications (ICC'12)*. Compressive sensing analysis of synthetic aperture radar raw data (Ottawa, ON, 10–15 June 2012), pp. 6362–6366
6. J Chen, Q Liang, Efficient sampling for radar sensor networks. Int. J. Sensor Netw., in press
7. P Pal, PP Vaidyanathan, Nested Arrays: a novel approach to array processing with enhanced degrees of freedom. IEEE Trans. Signal Process. **58**(8), 4167–4181 (2010)
8. P Pal, PP Vaidyanathan, in *Proceedings of the Digital Signal Process. Workshop, IEEE Signal Process. Educ. Workshop*. Coprime sampling and the MUSIC algorithm (Sedona, AZ, 4–7 January 2011), pp. 289–294
9. P Pal, PP Vaidyanathan, in *Proceedings of the Acoustics Speech Signal Process*. A novel array structure for directions-of-arrival estimation with increased degrees of freedom (Dallas, TX, 14–19 March 2010), pp. 2606–2609
10. P Pal, Piya, PP Vaidyanathan, in *Proceedings of the IEEE Int. Conf. Acoustics, Speech Signal Process. (ICASSP)*. Two dimensional nested arrays on lattices (Prague, 22–27 May 2011, 2011), pp. 548–2551
11. PP Vaidyanathan, P Pal, Sparse sensing with co-prime samplers and arrays. IEEE Trans. Signal Process. **59**(2), 573–586 (2011)
12. J Chen, Q Liang, B Zhang, X Wu, Spectrum efficiency of nested sparse sampling and co-prime sampling. EURASIP J. Wireless Commun. Netw. **2013**, 47 (2013)
13. J Chen, Q Liang, J Wang, H-A Choi, in *Wireless Algorithms, Systems, and Applications*. Spectrum efficiency of nested sparse sampling (Springer, Berlin Heidelberg, 2012), pp. 574–583
14. J Chen, Q Liang, J Wang, Secure transmission for big data based on nested sampling and coprime sampling with spectrum efficiency. Secur. Commun. Netw. Wiley Security Comm. Networks (2013). doi:10.1002/sec.785

15. Q Liang, X Cheng, SC Huang, D Chen, Opportunistic sensing in wireless sensor networks: theory and application. IEEE Trans. Comput. (2013). doi:10.1109/TC.2013.85
16. TM Cover, JA Thomas, *Elements of Information Theory, Second Edition*. (Wiley, Hoboken, 2006)
17. J Chen, Q Liang, in *Proceedings of the IEEE Global Telecommun. Conf. (GLOBECOM 2011)*. Rate distortion performance analysis of compressive sensing (Houston, TX, 5–9 December 2011), pp. 1–5
18. M Capdevila, OWM Florez, A communication perspective on automatic text categorization. IEEE Trans. Knowl. Data Eng. **21**, 1027–1041 (2009)
19. J Chen, Q Liang, B Zhang, X Wu, Information theoretic performance bounds for noisy compressive sensing. Paper presented at the ICC'2013, Budapest, 09–13 June 2013
20. J Chen, Q Liang, Theoretical performance limits for compressive sensing with random noise. Paper presented at the IEEE Global Communications Conference 2013 (GLOBECOM'13), Atlanta, 09–13 December 2013