# An Improvement of the Upper Bound on the Entropy of Information Sources 

Y. Sayyari*<br>Sirjan University Of Technology


#### Abstract

Theory of zeta functions and fractional calculus plays an important role in the statistical problems and Shannon's entropy. There is a close relationship between the maximum entropy values and fractional equations. Estimation of Shannon's entropies of information sources from numerical simulation of long orbits is difficult. Our aim within this paper is to present a strong upper bound for the Shannon's entropy of information sources and estimate the numerical entropy value by figuring out entropy-fitted bounds.


AMS Subject Classification: 37B40; 94A17; 26A51.
Keywords and Phrases: Entropy, Shannon's entropy, information source, stochastic process, zeta function.

## 1 Introduction

In the last years we witnessed an increasing interest in the generalization of the concepts of fractional calculus and of entropy [1, 2, 3, 6, 16]. We provide a brief introduction to entropy and fractional calculus in the following: If $s>1$, then Riemann function is defined as

$$
\zeta(s)=\sum_{n=1}^{\infty} \frac{1}{n^{s}} .
$$

[^0]The subject of fractional calculus has emerged as a powerful mathematical instrument during the past years, and is used in every branch of the statistics, engineering, and in other fields. S. Golomb showed that Riemann's zeta function $\zeta$ induces a probability distribution $\pi(n)=\frac{n^{-s}}{\zeta(s)}$ on the positive integers, for every $s>1$ [12]. In Guiasu [14], the author proved that the probability distribution mentioned above is the unique solution of an entropy-maximization problem. Fractional calculus of zeta functions can also be used to maximize

$$
H=-\sum_{n} \pi(n) \log \pi(n)
$$

where $\{\pi(n): n \in \mathbb{N}\}$ is a probability distribution on $\mathbb{N}$ [13].
Theorem 1.1. [13] Let $\alpha \in \mathbb{R} \backslash \mathbb{Z}, \pi(n)>0$ and $\sum_{n} \pi(n)=1$. The maximization of Shannon entropy $H=-\sum_{n} \pi(n) \log \pi(n)$ and

$$
\sum_{n \in \mathbb{N}} \pi(n) \log D_{f}^{\alpha} n^{-x}=\chi_{\alpha}, \quad x>1+\alpha
$$

has a solution given by

$$
\pi(n)=\frac{D_{f}^{\alpha} n^{-x}}{\zeta^{(\alpha)(x)}}, \quad n \in \mathbb{N}
$$

where the forward Grunwald-Letnikov fractional derivative of $f$ is defined as follows:

$$
D_{f}^{\alpha} f(x)=\lim _{h \rightarrow 0^{+}} \frac{\sum_{m=0}^{\infty}\binom{\alpha}{m}(-1)^{m} f(x-m h)}{h^{\alpha}}
$$

In differential equations, fractional equations are used to model the behavior of diseases $[7,8,18]$ and fraction equations are used in optimization too [15]. In [19], the authors introduced a new mathematical model for the transmission of Zika virus between humans as well as between humans and mosquitoes by the use of fractional-order Caputo derivative. In [9], the authors give the numerical simulations of the fractional model, a new model which is based on Caputo fractional derivative. There is a close relationship between the maximum entropy values and fractional
equations $[10,12,13,14]$. Entropy and mutual information for random variables play important roles in dynamical systems and information theory. The entropy actually measures the degree of irregularities of a dynamic system, and researchers have done so much to calculate this concept, which is often successful [4, 11, 21, 22], but numerical calculations of entropy are still difficult. Tapus and Popescu presented a strong upper bound for the classical Shannon entropy [24]. In [24, 20, 21, 23], the authors presented a strong upper bound for the classical Shannon entropy. In [17], the authors presented the algebraic and Shannon entropies for hypergroupoids and commutative hypergroups, respectively, and studies their fundamental properties. In [20], the author applying Jensen's inequality in information theory and obtained some results for the Shannon's entropy of random variables and Shannon's entropy of information sources. Our purpose within this work is to present a strong upper bound for the Shannon entropy of information sources, refining recent results from the literature.

Let $X$ be a non-empty set, $\mathcal{F}$ is an $\sigma$-algebra of subsets of $X, \mu$ is a measure on $X$ and $\mu(X)=1$, then $(X, \mathcal{F}, \mu)$ is called measure probability space. A finite set of measurable sets $\alpha=\left\{A_{1}, \ldots, A_{n}\right\}$ is called a finite partition if the following properties are fulfilled [25]:

$$
\bigcup_{i=1}^{n} A_{i}=X, \quad \text { and } \quad A_{i} \cap A_{j}=\emptyset \text { for every } i, j(1 \leq i \neq j \leq n) .
$$

For a partition $\alpha=\left\{A_{1}, \ldots, A_{n}\right\}$, the entropy of $\alpha$ is defined by

$$
\begin{equation*}
H_{\mu}(\alpha):=-\sum_{i=1}^{n} \mu\left(A_{i}\right) \log \left(\mu\left(A_{i}\right)\right) . \tag{1}
\end{equation*}
$$

Definition 1.2. [5] Let $S$ be a random variable on $X$ with discrete finite state space $A=\left\{a_{1}, \ldots, a_{N}\right\}$. We define $p: A \rightarrow[0,1]$ by $p(s)=\mu\{\omega \in$ $X: S(\omega)=s\}$. The Shannon's entropy of $S$ is defined by

$$
\begin{equation*}
H_{\mu}(S):=-\sum_{s \in A, p(s) \neq 0} p(s) \log p(s) . \tag{2}
\end{equation*}
$$

An information sources $\mathbf{S}$ is a sequence $\left(S_{n}\right)_{n=1}^{\infty}$ of the random variables $S_{n}: X \longrightarrow A$, where $n \in \mathbb{N}$. For given $L \geq 1$ we define a mapping
$p: A^{L} \rightarrow[0,1]$ by $p\left(s_{1}^{L}\right)=\mu\left\{\omega \in X: S_{1}(\omega)=s_{1}, \ldots, S_{L}(\omega)=s_{L}\right\}$. The Shannon entropy of order $L$ and the Shannon entropy of source $\mathbf{S}$ are respectively defined by

$$
H_{\mu}\left(S_{1}^{L}\right)=-\frac{1}{L} \sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}, \ldots, s_{L}\right) \log p\left(s_{1}, \ldots, s_{L}\right)
$$

and $h_{\mu}(\mathbf{S})=\lim _{L \rightarrow \infty} H_{\mu}\left(S_{1}^{L}\right)$, where the the summation is taken over the collection $\left\{s_{1}^{L} \in A^{L}: p\left(s_{1}^{L}\right) \neq 0\right\}$. In this paper we use the symbol $s_{1}^{L}$ instead of notation $\left(s_{1}, \ldots, s_{L}\right)$ and Let $p\left(s_{1}^{L}\right) \neq 0$ for every $L \in \mathbb{N}$.

Theorem 1.3. [20] Let $I=[a, b]$ be an interval, $H: A^{L} \longrightarrow I$ be $a$ function, and $f: I \longrightarrow \mathbb{R}$ be a convex function, then

$$
\begin{align*}
& \sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) f\left(H\left(s_{1}^{L}\right)\right)-f\left(\sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right) \\
& \geq \max \left\{p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)\right. \\
& \left.-\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}\right)\right\}, \tag{3}
\end{align*}
$$

where the maximum is taken over all $r_{1}^{L} \neq t_{1}^{L} \in A^{L}$.

## 2 Main results

In this section, we continue with a refinement of Theorem 1.3, as follows:

Theorem 2.1. Let $I=[a, b]$ be an interval, $H: A^{L} \longrightarrow I$ be a function,
and $f: I \longrightarrow \mathbb{R}$ be a convex function. Then

$$
\begin{aligned}
& \sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) f\left(H\left(s_{1}^{L}\right)\right)-f\left(\sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right) \\
& \geq \max \left\{p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)+p\left(u_{1}^{L}\right) f\left(H\left(u_{1}^{L}\right)\right)\right. \\
& -\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) \times \\
& \left.f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right)\right\}, \\
& \geq \max \left\{p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)+p\left(u_{1}^{L}\right) f\left(H\left(u_{1}^{L}\right)\right)\right\} \\
& -\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) \times \\
& \left.f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right)\right\},
\end{aligned}
$$

where the maximum is taken over all distinct $r_{1}^{L}, t_{1}^{L}, u_{1}^{L} \in A^{L}$.
Proof. Choose arbitrary $t_{1}^{L}, r_{1}^{L}, u_{1}^{L} \in A^{L}$. So,

$$
\begin{aligned}
& f\left(\sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right)=f\left(\sum_{s_{1}^{L} \neq r_{1}^{L}, t_{1}^{L}, u_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right) \\
& +\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right)\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right) \\
& \leq \sum p\left(s_{1}^{L}\right) f\left(H\left(s_{1}^{L}\right)\right) \\
& +\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right),
\end{aligned}
$$

where $s_{1}^{L} \neq r_{1}^{L}, t_{1}^{L}, u_{1}^{L} \in A^{L}$. Therefore,

$$
\begin{aligned}
& \sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) f\left(H\left(s_{1}^{L}\right)\right)-f\left(\sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right) \\
& \geq p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)+p\left(u_{1}^{L}\right) f\left(H\left(u_{1}^{L}\right)\right) \\
& -\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right) .
\end{aligned}
$$

Since $s_{1}^{L}, t_{1}^{L} \in A^{L}, u_{L}^{1}$ are arbitrary,

$$
\begin{aligned}
& \sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) f\left(H\left(s_{1}^{L}\right)\right)-f\left(\sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right) \\
& \geq \max \left\{p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)+p\left(u_{1}^{L}\right) f\left(H\left(u_{1}^{L}\right)\right)\right\} \\
& \left.-\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right)\right\},
\end{aligned}
$$

where the maximum is taken over all distinct $r_{1}^{L}, t_{1}^{L}, u_{1}^{L} \in A^{L}$. On the other hand,

$$
\begin{aligned}
& f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right) \\
& =f\left(\frac{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)} \frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}\right. \\
& \left.+\frac{p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right) \\
& \leq \frac{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)} f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}\right) \\
& +\frac{p\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)} f\left(H\left(u_{1}^{L}\right)\right) .
\end{aligned}
$$

So,

$$
\begin{aligned}
& \left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right) \\
& \leq\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}\right)+\left(p\left(u_{1}^{L}\right)\right) f\left(H\left(u_{1}^{L}\right)\right)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)+p\left(u_{1}^{L}\right) f\left(H\left(u_{1}^{L}\right)\right) \\
& -\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right) H\left(u_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)+p\left(u_{1}^{L}\right)}\right) \\
& \left.\geq p\left(r_{1}^{L}\right) f\left(H\left(r_{1}^{L}\right)\right)+p\left(t_{1}^{L}\right) f\left(H\left(t_{1}^{L}\right)\right)\right\} \\
& -\left(p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)\right) f\left(\frac{p\left(r_{1}^{L}\right) H\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right) H\left(t_{1}^{L}\right)}{p\left(r_{1}^{L}\right)+p\left(t_{1}^{L}\right)}\right)
\end{aligned}
$$

which completes the proof.
In order to present the generalization, we define some notation, as follows:

$$
T_{k}:=\max \left\{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right) f\left(H\left(r_{i 1}^{L}\right)\right)-\left(\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)\right) f\left(\frac{\sum_{i=1}^{k} p\left(r_{i} L\right) H\left(r_{i 1}^{L}\right)}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right)\right\}
$$

where $2 \leq k \leq N^{L}-1$, the maximum is taken over all distinct $r_{i}^{L} \in A^{L}$.

Theorem 2.2. Let $I=[a, b]$ be an interval, $H: A^{L} \longrightarrow I$ be a function, $|A|=N$ and $f: I \longrightarrow \mathbb{R}$ be a convex function, then

$$
\begin{aligned}
0 \leq T_{2} \leq T_{3} \leq \ldots \leq T_{N^{L}-1} \leq & \sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) f\left(H\left(s_{1}^{L}\right)\right) \\
& -f\left(\sum_{s_{1}^{L} \in A^{L}} p\left(s_{1}^{L}\right) H\left(s_{1}^{L}\right)\right) .
\end{aligned}
$$

Proof. The proof is similar to the proof of Theorem 2.1.

## 3 The sources entropy upper bound

In this section we present a strong upper bound for the Shannon's entropy of information sources.

Theorem 3.1. Let $\mathbf{S}$ be an information source. Then

$$
\begin{aligned}
h_{\mu}(\mathbf{S}) & \leq \log N-\max _{k}\left\{\lim _{L \rightarrow \infty} \frac{1}{L} \log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right]\right. \\
& \left.\times\left[\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right]\right\} .
\end{aligned}
$$

Proof. Since

$$
\begin{aligned}
& -L H_{\mu}\left(S_{1}^{L}\right)+\log \left(N^{L}\right) \geq \max _{k}\left\{-\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right) \log \left(\frac{1}{p\left(r_{i 1}^{L}\right)}\right)\right. \\
& \left.+\left(\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)\right) \times \log \left(\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right)\right\} \\
& =\max _{k}\left\{\log \left(\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right)+\log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right]\right\}, \\
& \log N-H_{\mu}\left(S_{1}^{L}\right) \geq \max \left\{\frac{1}{L} \log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right]\right. \\
& \left.\times\left[\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right]\right\},
\end{aligned}
$$

and

$$
\begin{aligned}
H_{\mu}\left(S_{1}^{L}\right) \leq \log N-\max _{k} & \left\{\frac{1}{L} \log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right]\right. \\
\times & {\left.\left[\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right]\right\} . }
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& h_{\mu}(\mathbf{S}) \\
& \leq \log N-\lim _{L \rightarrow \infty} \max _{k}\left\{\frac{1}{L} \log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right]\right. \\
& \left.\times\left[\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right]\right\} \\
& \leq \log N-\max _{2 \leq k \leq N^{L}-1}\left\{\lim _{L \rightarrow \infty} \frac{1}{L} \log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right]\right. \\
& \left.\times\left[\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right]\right\},
\end{aligned}
$$

which completes the proof.
Entropy of information sources is very important in synamical systems and information theory. Let $(X, \mathcal{F}, \mu)$ me a probability measure space. For a partition

$$
\alpha=\left\{A_{0}, \ldots, A_{N}\right\}
$$

and measure-preserving dynamical system $f: X \longrightarrow X$, the maps

$$
S_{n}: X \longrightarrow T_{N}:=\{0, \ldots, N\}
$$

defined as

$$
S_{n}(x)=i \text { if and only if } f^{n}(x) \in A_{i}
$$

are random variables on the probability measure space $X$. In this case we have

$$
p(i)=\mu\left(A_{i}\right),
$$

for every $i(0 \leq i \leq N)$, and $h_{\mu}\left(\mathbf{S}_{\alpha}\right)=h_{\mu}(f, \alpha)$ where $\mathbf{S}_{\alpha}=\left\{S_{n}\right\}$ [5]. Since The metric entropy of $f$ is then the supremum of $h_{\mu}(f, \alpha)$ over all finite partitions of $(X, \mathcal{F}, \mu)$ (i.e.

$$
\begin{equation*}
\left.h_{\mu}(f)=\sup _{\alpha} h_{\mu}(f, \alpha)=\sup _{\alpha} h_{\mu}\left(\mathbf{S}_{\alpha}\right) .\right) \tag{4}
\end{equation*}
$$

Thus, an approximation of entropy $f$ is obtained by using 4 .

## 4 Conclusion

In this paper, we have obtained some mathematical inequalities for entropy of information sources. For the entropy of an of information sources, this paper discovered suitable bounds with the help of which the Shannon's entropy value could be approximated. Theorem 3.1, shows that in general,

$$
\log N-\frac{1}{L} \log \left[\left\{\frac{k}{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right\}^{\sum_{i=1}^{k} p\left(r_{i 1}^{L}\right)}\right] \times\left[\prod_{i=1}^{k}\left\{p\left(r_{i 1}^{L}\right)\right\}^{p\left(r_{i 1}^{L}\right)}\right]
$$

can only be expected to be an upper bound of $h_{\mu}(\mathbf{S})$, we will try to extend it in the future.

## References

[1] M. Adeel, K.A. Khan, D. Peĉarić and J. Peĉarić, Generalization of the Levinson inequality with applications to information theorey, Journal of Inequalities and Applications, 212 (2019), (DOI: 10.1186/s13660-019-2166-8).
[2] M. Adeel, K.A. Khan, D. Peĉarić and J. Peĉarić, Levinson type inequalities for higher order convex functions via Able-Gontscharo interpolation, Advances in Difference Equations, 430 (2019), (DOI: 10.1186/s13662-019-2360-5).
[3] M. Adeel, K.A. Khan, D. Peĉarić and J. Peĉarić, Estimation of f-divergence and Shannon entropy by Levinson type inequalities via new Green's functions and Lidstone polynomial, Advances in Difference Equations, 27 (2020), (DOI: 10.1186/s13662-020-2501$\mathrm{x})$.
[4] J.M. Amigo and M.B. Kennel, Topological permutation entropy, Physica D, 231 (2007), 137-142.
[5] J.M. Amigo, Permutation Complexity in Dynamical Systems "Ordinal Patterns, Permutation Entropy, and All That", SpringerVerlag, Berlin, (2010).
[6] D. Baleanu, B. Ghanbari, J.H. Asad, A. Jajarmi and H.M. Pirouz, Planar System-Masses in an Equilateral Triangle: Numerical Study within Fractional Calculus. CMES-Computer Modeling in Engineering and Sciences, 124 (3) (2020), 953-968.
[7] D. Baleanuab, A. Jajarmic, H. Mohammadi and Sh. Rezapoure, A new study on the mathematical modelling of human liver with Caputo-Fabrizio fractional derivative, Chaos, Solitons and Fractals, 134 (2020), (DOI: 10.1016/j.chaos.2020.109705).
[8] D. Baleanu, S.S. Sajjadi, J.H. Asad, et al. Hyperchaotic behaviors, optimal control, and synchronization of a nonautonomous cardiac conduction system. Advances in Difference Equations, 157 (2021), (DOI: 10.1186/s13662-021-03320-0).
[9] D. Baleanu, S.S. Sajjadi, A. Jajarmi, et al. The fractional dynamics of a linear triatomic molecule, Romanian Reports in Physics, 73 (1) (2021), 1-13.
[10] C. Cattani and E. Guariglia, Fractional derivative of the Hurwitz $\zeta$-function and chaotic decay to zero, Journal of King Saud University-Science 28 (1) (2016), 75-81.
[11] Ch. Corda, M. FatehiNia, M.R. Molaei and Y. Sayyari, Entropy of Iterated Function systems and their relations with black holes and Bohr-like black holes entropies, Entropy, 2056 (2018).
[12] S.W. Golomb, A class of probability distributions on the integers, Journal of Number Theory, 2 (1970), 189-192.
[13] E. Guariglia, Fractional calculus, zeta functions and Shannon entropy, Open Mathematics, 19 (1) (2021),(DOI: 10.1515/math-20210010).
[14] S. Guiasu, An optimization problem related to the zeta function, Canadian Mathematical Bulletin 29 (1) (1986), 70-73.
[15] A. Jajarmi and D. Baleanu, On the fractional optimal control problems with a general derivative operator, Asian Journal of Control, 23 (2) (2021), 1062-1071.
[16] R.L. Magin and C. Ingo, Entropy and information in a fractional order model of anomalous diffusion, 16th IFAC Symposium on System Identification The International Federation of Automatic Control Brussels, Belgium, (2012), 11-13.
[17] A. Mehrpooya, Y. Sayyari and M.R. Molaei, Algebraic and Shannon entropies of commutative hypergroups and their connection with information and permutation entropies and with calculation of entropy for chemical algebras, Soft Computing, 23 (24) (2019), 13035-13053.
[18] H. Mohammadi, S. Rezapour and A. Jajarmi. On the fractional SIRD mathematical model and control for the transmission of

COVID-19: The first and the second waves of the disease in Iran and Japan. ISA Trans, (2021), (DOI: 10.1016/j.isatra.2021.04.012).
[19] S. Rezapour, H. Mohammadi and A. Jajarmi, A new mathematical model for Zika virus transmission. Advances in Difference Equations, 589 (2020). (DOI:10.1186/s13662-020-03044-7).
[20] Y. Sayyari, New bounds for entropy of information sources, Wavelet and Linear Algebra, 7 (2) (2020), 1-9.
[21] Y. Sayyari, New entropy bounds via uniformly convex functions, Chaos, Solitons and Fractals, 141 (1) (2020), (DOI: 10.1016/j.chaos.2020.110360).
[22] Y. Sayyari, M.R. Molaei and S.M. Moghayer, Entropy of continuous maps on quasi-metric spaces, Journal of Advanced Research in Dynamical and Control Systems, 7 (4) (2015), 1-10.
[23] S. Simic, Jensen's inequality and new entropy bounds, Applied Mathematics Letters, 22 (8) (2009), 1262-1265.
[24] N. Tapus and P.G. Popescu, A new entropy upper bound, Applied Mathematics Letters, 25 (11) (2012), 1887-1890.
[25] P. Walters, An Introduction to Ergodic Theory, Springer Verlag, New York, (2000).

## Yamin Sayyari

Assistant Professor of Mathematics
Department of Mathematics
Sirjan University Of Technology
Sirjan, Iran
E-mail: ysayyari@gmail.com, y.sayyari@sirjantech.ac.ir


[^0]:    Received: March 2021; Published: July 2021
    *Corresponding Author

