Mathematics Interdisciplinary Research **x** (202x) $xx - yy$

On the Entropy Rate of a Random Walk on t-Designs

Reza Kahkeshani *

Abstract

In this paper, a random walk on t-designs are considered. We assign a weight to each block and walk randomly on the vertices with a probability proportional to the weight of blocks. This stochastic process is a Markov chain. We obtain a stationary distribution for this process and compute its entropy rate. It is seen that, when the blocks have the same weight, the uniform distribution on the vertices is a stationary distribution and the entropy rate depends only on the number of vertices.

Keywords: random walk, Markov chain, design, entropy rate, stationary distribution.

2010 Mathematics Subject Classification: Primary 60J10, Secondary 60G50, 05B99, 94A17.

How to cite this article
R. Kahkeshani, On the Entropy Rate of a Random Walk on t-Designs, *Math. Interdisc. Res.* **x** (202x) xx-yy.

1. Introduction

Let X be a discrete random variable with alphabet $\mathcal{X}$ and probability mass function $p(x) = \Pr\{X = x\}$, $x \in \mathcal{X}$. The entropy $H(X)$ of X is defined as

$$H(X) = -\sum_{x\in\mathcal{X}} p(x) \log p(x),$$

where logarithm is to the base 2 and entropy is expressed in bits. Here, the convention $0 \log 0 = 0$ will be used. The entropy $H(X)$ is a measure of the uncertainty of X and moreover, it is a measure of the amount of information required on the

*Corresponding author (E-mail: kahkeshanireza@kashanu.ac.ir)
Academic Editor: Mohammad Reza Darafsheh
Received 28 May 2020, Accepted 2 August 2020
DOI: 10.22052/mir.2020.233084.1214

average to describe X. Let (X,Y) be a pair of discrete random variables with a joint distribution $p(x,y)$, $(x,y)\in\mathcal{X}\times\mathcal{Y}$. The joint entropy $H(X,Y)$ is defined by

$$H(X,Y)=-\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}p(x,y)\log p(x,y).$$

Similarly, the entropy of a collection of random variables, such as $H(X_1,X_2,\ldots,X_n)$, is defined.

A stochastic process $\{X_i\}_{i\in\mathbb{N}}$ can be defined as an indexed sequence of random variables. This process is characterizied by the probability mass functions

$$\Pr\{(X_1,X_2,\ldots,X_n)=(x_1,x_2,\ldots,x_n)\}=p(x_1,x_2,\ldots,x_n),$$

where $(x_1,x_2,\ldots,x_n)\in\mathcal{X}^n$ and $n\in\mathbb{N}$. This process is called to be stationary if $\Pr\{(X_1,X_2,\ldots,X_n)=(x_1,x_2,\ldots,x_n)\}$ is equal to $\Pr\{(X_{l+1},X_{l+2},\ldots,X_{l+n})=(x_1,x_2,\ldots,x_n)\}$, for all $x_1,x_2,\ldots,x_n\in\mathcal{X}$ and every shift l. A Markov chain is a stochastic process $\{X_i\}_{i\in\mathbb{N}}$ such that $\Pr\{X_{n+1}=x_{n+1}|X_n=x_n,X_{n-1}=x_{n-1},\ldots,X_1=x_1\}$ is equal to $\Pr\{X_{n+1}=x_{n+1}|X_n=x_n\}$, for all $x_1,x_2,\ldots,x_{n+1}$ in $\mathcal{X}$. A Markov chain $\{X_i\}_{i\in\mathbb{N}}$ is called to be time invariant if $\Pr\{X_{n+1}=b|X_n=a\}=\Pr\{X_2=b|X_1=a\}$, for all $n\in\mathbb{N}$ and $a,b\in\mathcal{X}$. It is easy to see that a time invariant Markov chain with alphabet $\mathcal{X}=\{1,2,\ldots,m\}$ can be characterized by an initial state and a probability transition matrix $P=(p_{ij})$, where $p_{ij}=\Pr\{X_{n+1}=j|X_n=i\}$. A distribution μ on $\mathcal{X}$ is said to be stationary if $\mu P=P$. In other words, μ is a distribution on the states such that the distributions at the successive times are the same. The entropy rate of a stochastic process $\{X_i\}_{i\in\mathbb{N}}$ is

$$H(\mathcal{X})=\lim_{n\to\infty}\frac{1}{n}H(X_1,X_2,\ldots,X_n),$$

when the limit exists. Also, a related quantity for entropy rate is defined by

$$H'(\mathcal{X})=\lim_{n\to\infty}H(X_n|X_{n-1},X_{n-2},\ldots,X_1),$$

when the limit exists. These two quantities correspond to different notions. It can be shown that if $\{X_i\}_{i\in\mathbb{N}}$ is a stationary Markov chain then $H(\mathcal{X})=H'(\mathcal{X})=H(X_2|X_1)$. See [2, 6, 7] for more details and examples.

In this paper, motivated by a random walk on a weighted graph [2], a random walk on t-designs are considered. We assign a weight to each block and walk randomly on the vertices with a probability proportional to the weight of blocks. This stochastic process is a Markov chain. We obtain a stationary distribution for this process and compute its entropy rate. It is seen that, when the blocks have the same weight, the uniform distribution on the vertices is a stationary distribution and the entropy rate depends only on the number of vertices. For more information and some new results on random walks, entropy rates and their applications, please see [3, 5, 8].

2. t-Designs

Let $\mathcal{S} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ be an incidence structure which consists of point set $\mathcal{P}$, block set $\mathcal{B}$ and an incidence relation $\mathcal{I} \subseteq \mathcal{P} \times \mathcal{B}$. The elements of $\mathcal{I}$ are called flags and the notation $p\mathcal{I}B$ means that $(p, B) \in \mathcal{I}$. A block $B \in \mathcal{B}$ is sometimes identified with the set of points p incident with it. Here, $\mathcal{I}$ is in fact the membership relation $\in$. If we replace each block of $\mathcal{S}$ by its complement then we obtain the complement of the structure, denoted by $\overline{\mathcal{S}}$. The dual of $\mathcal{S} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ is the incidence structure $\mathcal{S}^\top = (\mathcal{B}, \mathcal{P}, \mathcal{I}^\top)$, where $B\mathcal{I}^\top p$ if and only if $p\mathcal{I}B$. The incidence matrix of $\mathcal{S}$ is a matrix M of size $|\mathcal{P}| \times |\mathcal{B}|$ whose rows and columns are labled by points and blocks, respectively, such that the entry (p, B) is 1 if and only if p is incident with B, and 0 otherwise. The incidence structure $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ is called a t-(v, k, λ) design if $|\mathcal{P}| = v$, $|B| = k$ for any $B \in \mathcal{B}$, and every t distinct points are incident with precisely λ blocks. It is known that the number of blocks, denoted by b, is equal to $\lambda\binom{v}{t}/\binom{k}{t}$. The design $\mathcal{D}$ is called trivial if $\mathcal{B}$ consists of all the k-subsets of $\mathcal{P}$. If $v = b$ then $\mathcal{D}$ is called symmetric. It is well-known that the number of blocks incident with s points ($s \leq t$), denoted by λ_s, is independent of the set and $\lambda_s = \lambda\binom{v-s}{t-s}/\binom{k-s}{t-s}$. Therefore, every t-(v, k, λ) design is also an s-(v, k, λ_s) design, where $s \leq t$. The complement of a t-(v, k, λ) design $\mathcal{D}$ is also a design $\overline{\mathcal{D}}$ with parameters t-$(v, v-k, \overline{\lambda})$, where $\overline{\lambda} = \sum_{s=0}^{t}(-1)^s\binom{t}{s}\lambda_s$. The number of blocks incident with any point, λ_1, is also denoted by r and called the replication number. If $\mathcal{D}$ is a t-(v, k, λ) design then $\mathcal{D}^\top$ is a design with b points such that its block size is r. If M is the incidence matrix of $\mathcal{D}$ then the incidence matrix of $\mathcal{D}^\top$ is $M^\top$. It can be shown that if $\mathcal{D}$ is a 2-(v, k, λ) design then $bk = vr$ and $\lambda(v-1) = r(k-1)$. For more details, see [1, 4].

3. Results

Let the incidence structure $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ be a t-(v, k, λ) design with the vertex set $\{1, 2, \ldots, v\}$. To each block $B \in \mathcal{B}$, we assign a weight $\omega(B) \geq 0$ in $\mathbb{R}$ and set

$$\omega = \sum_{B \in \mathcal{B}} \omega(B),$$
$$\omega_i = \sum_{i \in B \in \mathcal{B}} \omega(B),$$
$$\omega_{ij} = \sum_{i,j \in B \in \mathcal{B}} \omega(B),$$

where $i, j \in \mathcal{P}$ and $i \neq j$. In other words, ω_i is the sum of the weights of all blocks containing the vertex i and ω_{ij} is also the sum of the weights of all blocks containing the points i and j. Note that for any vertex i, we have

$$\begin{aligned}\sum_{\substack{j\in\mathcal{P}\\ j\neq i}}\omega_{ij} &= \sum_{\substack{j\in\mathcal{P}\\ j\neq i}}\sum_{\substack{B\in\mathcal{B}\\ i,j\in B}}\omega(B)\\ &= \sum_{\substack{B\in\mathcal{B}\\ i\in B}}\sum_{\substack{j\in\mathcal{P}\\ i\neq j\in B}}\omega(B)\\ &= \sum_{\substack{B\in\mathcal{B}\\ i\in B}}(k-1)\omega(B)\\ &= (k-1)\omega_i.\end{aligned}$$

A random walk $\{X_n\}_{n=1}^{\infty}$ in $\mathcal{D}$ is a sequence of points of $\mathcal{D}$ in such a way that $X_n = i$ and $X_{n+1} = j$ if there exists a block B containing the points i and j. Moreover, we walk from i to j with the probability $p_{ij} = \omega_{ij}/\left((k-1)\omega_i\right)$. As it is seen, we walk randomly from the vertex i to the vertex j with a probability proportional to the weight of the blocks containing i and j, and the values $\{p_{ij}\}_{1\leq j\leq v}$ form a mass probability function. By definition, this stochastic process is a Markov chain with the probability transition matrix $P = \left(p_{ij}\right)_{v\times v}$. Set $\mu = (\mu_1, \mu_2, \ldots, \mu_v)$, where $\mu_i = \omega_i/(k\omega)$ for any $1 \leq i \leq v$. Since

$$\begin{aligned}\sum_{i=1}^{v}\mu_i &= \sum_{i=1}^{v}\frac{\omega_i}{k\omega}\\ &= \frac{1}{k\omega}\sum_{i=1}^{v}\sum_{i\in B\in\mathcal{B}}\omega(B)\\ &= \frac{1}{k\omega}\sum_{B\in\mathcal{B}}\sum_{\substack{i\in\mathcal{P}\\ i\in B\in\mathcal{B}}}\omega(B)\\ &= \frac{1}{k\omega}\sum_{B\in\mathcal{B}}k\omega(B)\\ &= 1,\end{aligned}$$

μ is a probability distribution on the points $\mathcal{P}$. Moreover, for any $1 \leq j \leq v$,

$$\begin{aligned}\sum_{l=1}^{v}\mu_l p_{lj} &= \sum_{l=1}^{v}\frac{\omega_l}{k\omega}\frac{\omega_{lj}}{(k-1)\omega_l}\\ &= \frac{1}{k(k-1)\omega}\sum_{l=1}^{v}\omega_{lj}\\ &= \frac{\omega_j}{k\omega}\\ &= \mu_j.\end{aligned}$$

Therefore, μ is also a stationary distribution. Now, the entropy rate of this process is

$$\begin{aligned}
H(\mathcal{X}) &= H(X_2|X_1) \\
&= -\sum_{i=1}^{v} \mu_i \sum_{j=1}^{v} p_{ij} \log p_{ij} \\
&= -\sum_{i=1}^{v} \frac{\omega_i}{k\omega} \sum_{j=1}^{v} \frac{\omega_{ij}}{(k-1)\omega_i} \log \frac{\omega_{ij}}{(k-1)\omega_i} \\
&= -\sum_{i=1}^{v}\sum_{j=1}^{v} \frac{\omega_{ij}}{k(k-1)\omega} \log \frac{\omega_{ij}}{(k-1)\omega_i} \\
&= -\sum_{i=1}^{v}\sum_{j=1}^{v} \frac{\omega_{ij}}{k(k-1)\omega} \log \frac{\omega_{ij}}{k(k-1)\omega} + \sum_{i=1}^{v}\sum_{j=1}^{v} \frac{\omega_{ij}}{k(k-1)\omega} \log \frac{\omega_i}{k\omega} \\
&= -\sum_{i=1}^{v}\sum_{j=1}^{v} \frac{\omega_{ij}}{k(k-1)\omega} \log \frac{\omega_{ij}}{k(k-1)\omega} + \sum_{i=1}^{v} \frac{\omega_i}{k\omega} \log \frac{\omega_i}{k\omega} \\
&= H\left(\cdots, \frac{\omega_{ij}}{k(k-1)\omega}, \cdots\right) - H\left(\cdots, \frac{\omega_i}{k\omega}, \cdots\right).
\end{aligned}$$

So, the following theorem is implied:

Theorem 3.1. Let $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ be a t-(v, k, λ) design. Assign a non-negative real number $\omega(B)$ to each block $B \in \mathcal{B}$ and set

$$\begin{aligned}
\omega &= \sum_{B \in \mathcal{B}} \omega(B), \\
\omega_i &= \sum_{i \in B \in \mathcal{B}} \omega(B), \\
\omega_{ij} &= \sum_{i,j \in B \in \mathcal{B}} \omega(B),
\end{aligned}$$

for any $i \neq j \in \mathcal{P}$. Let $\{X_n\}_{n=1}^{\infty}$ be a random walk on the points of $\mathcal{D}$ with the probabiity transition matrix $P = \left(p_{ij}\right)_{v \times v}$, where $p_{ij} = \omega_{ij} / \left((k-1)\omega_i\right)$. Set $\mu_i = \omega_i/(k\omega)$, where $1 \leq i \leq v$. Then, $\{X_n\}_{n=1}^{\infty}$ is a Markov chain with the stationary distribution $\mu = (\mu_1, \mu_2, \ldots, \mu_v)$ and the entropy rate

$$H(\mathcal{X}) = H\left(\cdots, \frac{\omega_{ij}}{k(k-1)\omega}, \cdots\right) - H\left(\cdots, \frac{\omega_i}{k\omega}, \cdots\right).$$

Note that if all the blocks have equal weight then $p_{ij} = \lambda_2/(r(k-1))$ and $\mu_i = r/(kb) = 1/v$. Also,

$$\frac{\omega_i}{k\omega} = \frac{r}{kb} = \frac{1}{v}$$

and

$$\frac{\omega_{ij}}{k(k-1)\omega} = \frac{\lambda_2}{k(k-1)b} = \frac{1}{v(v-1)}.$$

Hence, in this case, the uniform distribution on $\mathcal{P}$ is a stationary distribution and the entropy rate is

$$\begin{aligned} H(\mathcal{X}) &= H(\cdots, \frac{1}{v(v-1)}, \cdots) - H(\cdots, \frac{1}{v}, \cdots) \\ &= \log(v-1). \end{aligned}$$

Acknowledgement. The author would like to thank the referee for his/her suggestions and comments. This work is partially supported by the University of Kashan under grant number 985983/1.

Conflicts of Interest. The author declares that there are no conflicts of interest regarding the publication of this article.

References

[1] T. Beth, D. Jungnickel and H. Lenz, *Design Theory*, Volume I, 2nd edition, Cambridge University Press, Cambridge, 1999.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edition, John Wiley & Sons, Inc., New Jersey, 2006.

[3] R. Li and W. Heidrich, Hierarchical and view-invariant light field segmentation by maximizing entropy rate on 4D ray graphs, *ACM Trans. Graph.* **38** (6) (2019) Article 167.

[4] J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, 2nd edition, Cambridge University Press, Cambridge, 2001.

[5] H. Ni and X. Niu, Agglomerative oversegmentation using dual similarity and enropy rate, *Pattern Recogn.* **93** (2019) 324 − 336.

[6] N. Privault, *Understanding Markov Chains: Examples and Applications*, 2nd edition, Springer, Singapore, 2018.

[7] S. M. Ross, *Stochastic Processes*, 2nd edition, John Wiley & Sons, Inc., New York, 1996.

[8] R. J. H. Ross, C. Strandkvist and W. Fontana, Compressibility of random walker trajectories on growing networks, *Phys. Lett. A* **383** (17) (2019) 2028−2032.

Reza Kahkeshani
Department of Pure Mathematics,
Faculty of Mathematical Sciences,
University of Kashan,
Kashan, I. R. Iran
e-mail: kahkeshanireza@kashanu.ac.ir