A PHYSICAL THEORY OF INFORMATION VS. A MATHEMATICAL THEORY OF COMMUNICATION

Manouchehr Amiri

Department of Urology, Tandis Hospital, Tehran, Iran

ABSTRACT

This article introduces a general notion of physical bit information that is compatible with the basics of quantum mechanics and incorporates the Shannon entropy as a special case. This notion of physical information leads to the Binary data matrix model (BDM), which predicts the basic results of quantum mechanics, general relativity, and black hole thermodynamics. The compatibility of the model with holographic, information conservation, and Landauer's principle are investigated. After deriving the "Bit Information principle" as a consequence of BDM, the fundamental equations of Planck, De Broglie, Bekenstein, and mass-energy equivalence are derived.

KEYWORDS

Physical theory of information, Binary data matrix model, Shannon information theory, Bit information principle

1. INTRODUCTION

Information in a broad sense implies a collection of data of unmeasurable concepts or measurable quantities. The usual notion of measurable information in physics invokes the subject of Shannon entropy and information. Claude Shannon in his seminal paper [1] developed a mathematical theory of signal transmission [2]. He denied the semantic aspects of communication and related information theory. According to his theory, the information refers to the opportunity to reduce uncertainty and equals the entropy of the communicated message. He had got the idea of entropy from the second law of thermodynamics [2], [3] and concluded that the information of a *message* could be measured by its predictability, the less predictability the more information it carries [2], [3]. It is clear that Shannon's definition of information was not unique and merely was fitted for his engineering requirements [2], [3]. In this notion of information, the source, channel, and receiver of data are crucial components of communication engineering. Shannon entropy (information) is just concerned with the statistical properties of a given system, independent of the meaning and semantic content of system states [5]. As he emphasized in his seminal article, the meaning of communication and related information content are irrelevant to engineering problem [1]. Subsequently, there have been emerged some critiques around the Shannon notion of physical and biological information [3]. The notion of information independent of its meaning is the subject of main criticism announced by MacKay and others [3], [4]. Subsequently there have been attempts to add a semantic dimension to a formal theory of information, particularly to Shannon's theory [5]-[7]. Shannon's theory is not concerned with the individual message but rather the averages of source messages [8]. Although the physical information basically is related to physical measurable quantities, the current notion of physical information remained as the same definition introduced by Shannon and seems to be insufficient for physical systems. This has been reminded in recent works of Bruckner and Zeilinger [9]. Their main reason for this claim is the measurement problem in quantum mechanics. In other words, there is no definite real

DOI : 10.5121/ijist.2023.13301

value for observables before measurement in quantum physics sense and there is no reality independent of observer or measurement [9]. In quantum information theory, VonNeumann entropy is a candidate for replacing Shannon entropy to measure quantum information [9].However, recent trends toward deterministic quantum mechanics and its local versions [10]-[12] motivates one to investigate the existence of a fundamental version of information which carries the physical meaning of the system and is compatible with quantum mechanics. The main purpose of this article focused on introducing a new version of physical information and its broad consequences. In this version of information theory that I call, “Physical Theory of Information”, a *bit* of information reflects a *real* physical quantity in phase space, and it does not concern with the observer dependent reality of physical parameters. This approach will lead to a new attitude in the theory of information in which each bit of information has a specific meaning and implies a physical value. Just as the mathematical numbers in the governing equations of the laws of physics have a physical meaning, and show the relationship between physical variables, the bits of information in physics must also carry a physical meaning. Some authors showed the short comings of Shannon information content in the field of climatology and its related data. They presented that Shannon information may fail to be applied for certain kinds of useful information yielded by a measurement or observation [18].Introducing the new information model and development of this physical approach to information, and substituting it with physical bit information, could result in more applications of information and entropy in a wide range of physics fields.

1.1 Binary Matrix Model

Does information do anything with physics? The response of J.A. Wheeler was "It from bit," which means that the entire universe is constituted from "bits." Generally, information is defined as “an answer to a specific question” [3]. Therefore, if we restrict answers to “Yes” or “No”, the information could be represented by a set of 0s and 1s or a *binary array*. The formalism of information can be generalized by starting from this more detailed notion of information i.e., the arrays of binary data 0 and 1 for each physical parameter of particles in a system. Any variable (physical parameter)$x_\upsilon$when attributed to a subject or object (particles),carries the information that reflects the “quantity” of that variables. υrefers to the specific parameter and varies between 1 and d (degree of freedom). If we divide the range of this variable into a large number of infinitesimal intervals, then the value of variable is represented by a binary column matrix with 1 entry at the interval where the value of variable x_υ restricted, as depicted in Fig.1

$$x_\upsilon{\uparrow\atop\downarrow}\begin{bmatrix}0\\0\\\vdots\\1\\\vdots\\0\end{bmatrix}\rightarrow\ x_{\upsilon 0}$$

Fig.1. The column matrix’s entries return 0 except for the entry which corresponds to the value$x_{\upsilon 0}$ of parameter x_υ and represented by 1.

The unique entry 1 corresponds to the specific interval that represents the quantity $x_{\upsilon 0}$of variable x_υ.

$$x_{\upsilon 0}-\delta x_\upsilon < x_\upsilon < x_{\upsilon 0}+\delta x_\upsilon$$

Other values can be presented by displacing the 1 entry along the range of variable x_υ of the column matrix. For a set of different variables there are similar column matrices. By consideration of a system of N identical particles (N is a large number), for each particle there is a binary column matrix that specifies thex_{υ} parameter of that particle. Assembling all these column matrices results in a binary data matrix D_{υ} (Fig. 2).As an example, the column matrix specified by the red box, determines the value of x_{υ} for second particle (particle assigned by number2). The rows of D_{υ} are binary arrays $e^{*\upsilon}(x_{\upsilon 0})$ which are defined at any point $x_{\upsilon 0}$ and returns the information of particles at that point, depicted by the blue box in Fig.2:

$$e^{*\upsilon}(x_{\upsilon 0}) = (0110 \dots 10) \tag{1}$$

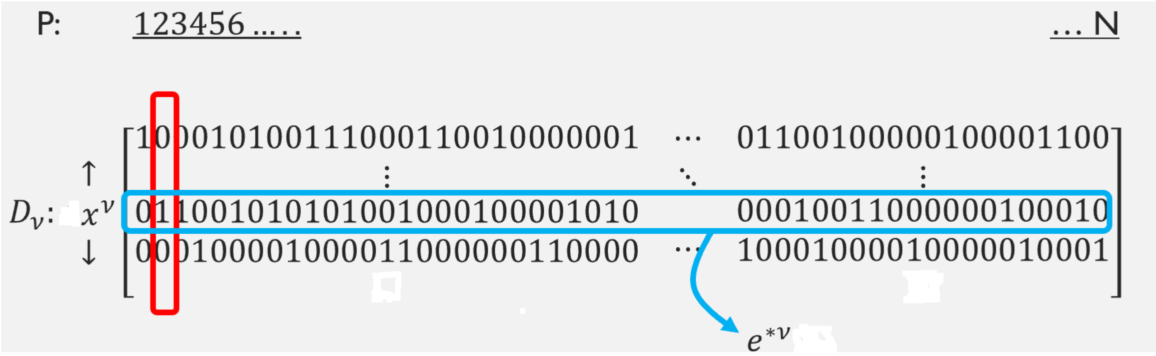

Fig. 2. The structure of matrix D_{υ}. The red box contains a unique “1” at the specific value $x_{\upsilon 0}$

which is the parameter value of particle with label 2. The blue box represents the particles whose parametric values is $x_{\upsilon 0}$. Any “1” entry in this array corresponds the particle with physical value $x_{\upsilon 0}$.

Any particle with the parameter value $x_{\upsilon 0}$is represented by entry 1 in this array. The sum of all 1 entries in $e^{*\upsilon}(x_{\upsilon 0})$ gives the total number of particles whose parameter’s value is $x_{\upsilon 0}$ simultaneously. In phase space, there are d different phase space parameters i.e., degree of freedom (including spatial, linear, and angular momentum), which lead to different D_{υ}. For a comprehensive binary information data system, we embed all D_{υ}s in a columnar matrix D:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_d \end{bmatrix} \tag{2}$$

Where d is the degree of freedom. For any point $(x_{10}, x_{20}, \dots\ x_{d0})$ on phase space, there correspond a set of$e^{*\upsilon}(x_{\upsilon 0})$ that could be gathered in a matrix P:

$$P = \begin{bmatrix} [e^{*1}] \\ [e^{*2}] \\ \vdots \\ [e^{*d}] \end{bmatrix} \tag{3}$$

Where $[e^{*\upsilon}]$ are row binary matrices defined at the point $(x_{10}, x_{20}, \dots\ x_{d0})$.

Theorem 1: The entries of matrix $G = \|g^{\upsilon\mu}\| = PP^T$give the number of particles with the same values of x_{i0} and x_{j0} and $\frac{1}{N} f_{\upsilon\mu}$ represents the joint probabilities.

Proof: With the above definitions, it can be easily proved that the inner products,

$$g^{\upsilon\mu} = [e^{*\upsilon}][e^{*\mu}]^T \tag{4}$$

are the number of particles with the values $x_{\upsilon 0}$and $x_{\mu 0}$, simultaneously. Obviously G is symmetric. Dividing $g_{\upsilon\mu}$ by total particles number obtains the joint probabilities $f_{\upsilon\mu}$:

$$f^{\upsilon\mu} = \frac{1}{N} g^{\upsilon\mu} \tag{5}$$

Matrix Dcontains all physical information of system constituents at any time. For a dynamical system, D evolves over time while the total number of "1" bits is preserved. This number that reflects the information measure of the system is the product of total number of particles and degree of freedom:

$$\mathbb{I} = Nd \tag{6}$$

Which is a constant. This formalism of physical bit information is called BDM (Binary Data Matrix Model) and ensues three immediate consequences:

a) The invariance of$\mathbb{I}$is compatible with *information conservation principle.*

b) If the system's temperature is denoted by T, then the average energy per particle due to equipartition principle reads as:

$$\epsilon = \frac{d}{2} kT \tag{7}$$

Since any particle contains$d$ bits, therefore the energy per bit could be obtained by dividing$\epsilon$by d:

$$\varepsilon \cong \frac{1}{2} kT$$

This is compatible with Landauer's principle i.e.,$\varepsilon = kT\ log\ 2$

c) All information of system represented by a 2-dimensional binary matrix D. This is compatible with Holographic principle.

These are physical principles that correlate the pure physical concepts to bit information and could not be justified in the context of Shannon information theory.

2. CONSEQUENCES OF BDM

If we consider $[e^{*\upsilon}]$ as the dual base vectors in the parameter space (phase space), the inner product (4) gives the dual metric tensor of this space. The author in his previous work [13] proved that the evolution of this metric tensor over time, obeys the quantum Liouville equation and leads to the equation (by assumption natural unit i.e.,$\hbar = 1$) [13]:

$$\langle E \rangle = \frac{1}{2} \frac{\partial}{\partial \tau} \log g \tag{8}$$

g stands for determinant of $\|g^{\upsilon\mu}\|$ and$\langle E \rangle$ for energy per system constituents i.e., particles. On the other hand, for spatial displacement, we obtain the corresponding equation [13]:

$$p_i = -\frac{1}{2} \frac{\partial}{\partial x_i} log\ g \tag{9}$$

The term $-\frac{1}{2}\log g$ shows the Hamilton's principal function F in classical mechanics [15]:

$$E = H = -\frac{\partial F}{\partial \tau} \quad ; \quad p_i = \frac{\partial F}{\partial x_i} \tag{10}$$

2.1. Entropy and Information

In classical statistics, for jointly normal random variables $\boldsymbol{x_1}, \boldsymbol{x_2}, ..\boldsymbol{x_n}$, the entropy is calculated as [14]:

$$\mathcal{H} = \frac{1}{2}\log\Delta + k \tag{11}$$

Where $\Delta$is the determinant of inverse of covariance matrix (σ^{ij}) i.e.,$\Delta = det\,\sigma_{ij}$ and k is constant $k = \log 2\pi e$. We show at the limit $\sigma \to 0$ (σ is the variance of random variables) as can be realized in the case of *black holes*, we can replace the Fisher metric [14] with the metric of the BDM model.

Theorem 2: In the limit as σ→0, which can be realized in the case of black holes [16], the inverse of the Fisher metric determinant and the determinant of the BDM metric are identical.

Proof: In the case of black holes where all the physical variables of its constituent confined to constant infinitesimal intervals δx^i, around the singularity point, if we fix the center of mass of black hole on the origin of spatial coordinates, the expected values of position and momentum of constituents are near zero corresponding BDM metric will be concentrated over these intervals with negligible values out of δx^i and mean values near to zero. Thus, the correlation (covariance) matrix element$\Re_{ij}$while the mean of all random variables vanishes $\bar{x} = 0$ reads as:

$$\Re_{ij} = \sigma_{ij} = \langle g_{ij}\delta x_i \delta x_j\rangle = g_{ij}\delta x_i \delta x_j \tag{12}$$

Determinant of σ_{ij} matrix (denoted by Δ) could be calculated as:

$$\Delta = \sum_{ijk\ldots} \varepsilon_{ijk\ldots}\ \sigma_{i1}\sigma_{j2}\sigma_{k3} \ldots \tag{13}$$

Substitution ofσ_{ij} with $g_{ij}\delta x_i \delta x_j$ gives rise to:

$$\Delta = g\prod_i (\delta x_i)^2 \tag{14}$$

Logarithm of both sides results in:

$$\log\Delta = \log g + 2\sum_i \log\delta x_i = \log g + C \tag{15}$$

On the other hand, definition of Fisher information metric $\mathscr{g}_{ij}$ results in the equivalence of inverse of covariance matrix $\sigma^{ij}$and Fisher metric tensor:

$$\mathscr{g}_{ij} = (\sigma_{ij})^{-1} = \sigma^{ij} \tag{16}$$

If the determinant of $\mathscr{g}_{ij}$is denoted by $\mathscr{g}$, then by equation (15) and (16) we have:

$$\mathscr{g} = \Delta^{-1}; \qquad log\,\mathscr{g}^{-1} = \log g + C' \tag{17}$$

and the theorem is proved.

Therefore, the equation (11) for entropy can be replaced by:

$$\mathcal{H} = \frac{1}{2}\log g^{-1} + C = \frac{1}{2}\log g + C'' \tag{18}$$

Respect to [16], the Euclidean action $\mathcal{H}_B$ of black holes are equivalent to its entropy i.e., Beckenstein-Hawking entropy:

$$S_{\mathrm{BH}} = \mathcal{H}_B$$

Due to the equivalence of action and Hamilton principal function and equations (10) and (18) we have:

$$F = -\frac{1}{2}\log g = \mathcal{H}_B = S_{\mathrm{BH}}$$

This equation reveals the equality of Shannon entropy of black hole and entropy derived by BDM, because it has been proved that the S_{BH} could be derived of Shannon information entropy. This reveals that the **BDM derived information and Shannon information (entropy) are equivalent at the limit**$\sigma \to 0$.

The equivalence of entropy $\mathcal{H}$ and$\frac{1}{2}\log g$ results in "Bit Information Principle".

2.2. Bit Information Principle

The equation (18) implies that the entropy $\mathcal{H}$ is equivalent to $\frac{1}{2}\log g$. By definition, the entropy is equivalent to the measure of information, then we conclude an important result which I call the *"Bit Information principle*"**:**$\boldsymbol{\frac{1}{2}\log\Delta}$ **is a measure of information at the limit** $\boldsymbol{\sigma \to 0}$ **and its spatial and temporal densities give the expected energy and momentum per particle.**

To conclude this principle, first consider Equation (18). It implies the entropic nature of $\frac{1}{2}\log g$ at the limit $\sigma \to 0$. Assuming equivalence of entropy and information, if we denote information by $\mathbb{I}$, we get:

$$\mathbb{I} = \frac{1}{2}\log g \tag{19}$$

Respect to equations (8) and (9) we infer that the time and spatial derivatives of $\frac{1}{2}\log g$ obtains $\langle E\rangle$ and p_i as the expected energy and momentum of system constituents. After returning to MKS units by multiplying the equations by $\hbar$(Planck constant) we get:

$$\langle E\rangle = \hbar\frac{\partial}{\partial\tau}\mathbb{I} \;\; ; \;\; p_i = -\hbar\frac{\partial}{\partial x_i}\mathbb{I} \tag{20}$$

In the BDM model, the information $\mathbb{I}$, is the number of bits. Therefore, Equation (20)reveals the relation between the density of bits over spatial and temporal intervals and the momentum and energy per constituent of the system. In a brief notation:

$$energy\sim\frac{bits}{second}\ , \qquad momentum\sim\frac{bits}{unit\ length}\ , \qquad angular\ momentum\sim\frac{bits}{unit\ angle}$$

Any system with negligible variance of its constituents' parameters exhibits a set of identical bits. For example, a monochromatic electromagnetic or acoustic wave possess a set of identical full wavelengths as bits.

2.3. Outcomes of Bit Information Principle

This principle justifies the Planck and De Broglie equation. When all bits are identical, it means that $\sigma \to 0$ and the required condition for Bit Information Principle is met. The situation of identical bits can be realized in monochromatic light (electromagnetic wave) beam because every full wavelength (pulse) should be considered as a bit of information. For the temporal (time) density of these pulses, the number of bits (full wavelength) per unit time is:

$$f = \frac{1}{T} \tag{21}$$

Where T is the period of monochromatic wave. Then, due to (20) we have:

$$\langle E \rangle = \hbar \frac{1}{T} = \hbar f \tag{22}$$

This is the Planck formula for energy of light wave constituents which is called *photons*.For the spatial density of these pulse, the number of full wavelengths (bit) per unit length is:

$$n = \frac{1}{\lambda} \tag{23}$$

Then, with respect to (20) and ignoring the sign, we obtain:

$$p = \frac{\hbar}{\lambda} \tag{24}$$

This is the De Broglie relation for the wavelength and momentum of a particle. The same relations are also valid for phonons as quanta of mechanical waves. For a black hole, the conditions for the bit information principle are also met because all the mass and its constituents are confined to an infinitesimal interval of space and momentum, and consequently, their variances tend to zero as $\sigma \to 0$. What is the amount of mass that represents a "bit" of a black hole? Theoretically, the mass of the smallest possible black hole, respecting the limitations imposed by the Schwarzschild radius, is the Planck mass:

$$m_p = \sqrt{\frac{\hbar c}{G}} \tag{25}$$

Then, for a macroscopic body like a black hole, the number of contained bits should be derived by dividing its mass by the Planck mass:

$$n = \frac{m}{m_p} \tag{26}$$

According to the bit information principle, energy is equivalent to bit density over time. The Planck time is the smallest time interval and serves as the universal time unit.

$$t_p = \sqrt{\frac{G\hbar}{c^5}} \tag{27}$$

Therefore, the expected energy will be obtained by density of n bits in (25) over the Planck time (27) multiplied by $\hbar$:

$$E = \hbar \frac{n}{t_p} = \hbar \frac{m}{m_p} \frac{1}{t_p} = \hbar m \sqrt{\frac{G}{\hbar c}} \sqrt{\frac{c^5}{G\hbar}} = mc^2 \tag{28}$$

This is the Einstein's mass-energy relation.

The bit information principle is also valid for linear and angular pseudo-momentum in crystals. The ordered arrangement of atoms, ions and molecules in a crystalline material provides the required conditions for the principle. Each atomic plane in a crystal contains atoms that are confined to a small interval of space and momentum. Hence, these atoms can be considered as bits over the interval of the interplanar surfaces of atoms. For bit density along the spatial coordinates in crystal lattices, the total density of a plane of atoms along the axis perpendicular to that plane, is proportional to $\frac{1}{d}$ where d is the distance between atoms planes in crystal. The magnitude of the corresponding reciprocal base lattice is also $\frac{1}{d}$:

$$|G| = \frac{1}{d} \tag{29}$$

The vector with this magnitude perpendicular to the atoms plane is called *crystal momentum* or pseudo-momentum. This momentum appears just in the interactions of atoms lattice with an incident photon or particle waves. With respect to the *bit density principle* from the previous section, this pseudo-momentum is equal to the density of bits (atoms) over the interval d by the equation:

$$\mathcal{P} = \frac{bits}{unit\ length} = \frac{n}{d} \tag{30}$$

The related momentum is the product of density and $\hbar$ i.e.,

$$\mathcal{P} = \hbar\frac{n}{d} \tag{31}$$

Then,the pseudo-momentum per atom reads as:

$$\mathcal{P} = \hbar\frac{1}{d} = \hbar G \tag{32}$$

This is the main relation for crystal momentum which is derived by the bit information principle. Curiously, the similar relation should govern the angular momentum. The suggested relation is as follows:

$$L_\theta = \frac{\hbar}{\theta} \tag{33}$$

Where θ is the angular period of bits over the whole 2π as unit angle. For crystals with N-fold rotational symmetries, it has been verified that the difference of *pseudo angular momentum* of incident and diffracted photons on a crystal with N-fold rotational symmetries obey the relation [17]:

$$\Delta m\hbar = \sigma\hbar + NP\hbar \tag{34}$$

And,for Rayleigh scattering with σ_i and σ_s as incident and scattered helicity of photons, we have [17]:

$$\sigma_i - \sigma_s = NP \tag{35}$$

Where P denotes an integer and N determines the particles or bits density per unit angle 2π (recall the definition of N-fold rotational symmetries). Thus,the equation (33) in the BDM context is compatible with the pseudo-angular momentum relation in (34) and (35).

Bekenstein Bound is another consequence of bit information principle. Due to this bound, the maximum information confined in a region with radius R and energy E due to Bekenstein is:

$$I \leq \frac{2\pi RE}{\hbar c \ln 2} \quad (36)$$

For calculating this bound, we use an example of a black hole, because the information within a black hole bounded to its horizon where the light rays can not pass to exit the black hole and therefore are confined in the radius of horizon of black hole. In this case information bits as the wavelength of the light rays, are spanned over the time interval by which the light travels from center of black hole (singularity point) to horizon at the radius R:

$$\Delta t = \frac{R}{c} \quad (37)$$

Hence, In the language of bit information principle, the maximum bit density I_M over this time interval is proportional to energy E up to a coefficient $\hbar$:

$$E = \hbar \frac{I_M}{\Delta t} = \hbar I_M \frac{c}{R} \Rightarrow I_M = \frac{RE}{\hbar c} \quad (38)$$

Because of the possible information leak from black hole, the true information$I$ confined in this radius is lesser than I_M :

$$I \leq I_M = \frac{RE}{\hbar c} \quad (39)$$

Which is compatible with (36).

3. CONCLUSION

By introducing a physical based theory of information, this article proves the authenticity and benefits of a new approach to the information concept in physics. This notion of information reduces to Shannon information as a special case. The BDM theory approach to information, not only endows a meaning to information, but also is compatible with the well-known principles such asLandauer’s principle, conservation of information and Holographic principles. After deriving the crucial “bit information principle”, the theory predicts the diverse principles such as, De Broglie, Planck and Bekenstein equations and mass-energy equivalence.

REFERENCES

[1] Shannon, Claude Elwood. "A mathematical theory of communication." *ACM SIGMOBILE mobile computing and communications review* 5.1 (2001): 3-55.
[2] Griffin, Emory A. *A first look at communication theory*. McGraw-Hill, 2003.
[3] Logan, Robert K. "What is information?: Why is it relativistic and what is its relationship to materiality, meaning and organization." *Information* 3.1 (2012): 68-91.
[4] MacKay, D. M. "Proceedings of the information theory symposium." (1950).
[5] Lombardi, Olimpia, Federico Holik, and Leonardo Vanni. "What is Shannon?" *Synthese* 193.7 (2016): 1983-2012.
[6] Barwise, Jon, and Jerry Seligman. *Information flow: the logic of distributed systems*. Cambridge University Press, 1997.
[7] Dretske, Fred. "Knowledge and the Flow of Information." (1981).
[8] Timpson, Christopher Gordon. "Quantum information theory and the foundations of quantum mechanics." *arXiv preprint quant-ph/0412063* (2004).
[9] Brukner, Časlav, and Anton Zeilinger. "Conceptual inadequacy of the Shannon information in quantum measurements." *Physical Review A* 63.2 (2001): 022113.

[10] Lindgren, Jussi, and Jukka Liukkonen. "The Heisenberg Uncertainty Principle as an Endogenous Equilibrium Property of Stochastic Optimal Control Systems in Quantum Mechanics." *Symmetry* 12.9 (2020): 1533.

[11] Haag, Rudolf. *Local quantum physics: Fields, particles, algebras*. Springer Science & Business Media, 2012.

[12] Hooft, Gerard T. "Deterministic quantum mechanics: the mathematical equations." *arXiv preprint arXiv:2005.06374* (2020).

[13] Amiri, M. Binary Data Matrix Theory. Preprints.org 2016, 2016120049. https://doi.org/10.20944/preprints201612.0049.v2.

[14] Papoulis, Athanasios, and S. Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[15] H. Goldstein, *Classical mechanics*: Pearson Education India, 1965

[16] C.Castro, (2008). The Euclidean gravitational action as black hole entropy, singularities, and spacetime voids. *Journal of Mathematical Physics*, *49*(4), 042501.

[17] Tatsumi, Yuki, Tomoaki Kaneko, and Riichiro Saito. "Conservation law of angular momentum in helicity-dependent Raman and Rayleigh scattering." *Physical Review B* 97.19 (2018): 195444.

[18] Petty, Grant W. "On some shortcomings of Shannon entropy as a measure of information content in indirect measurements of continuous variables." *Journal of Atmospheric and Oceanic Technology* 35.5 (2018): 1011-1021.