

A black hole-aided deep-helix channel model for DNA.

SARKER, M.A.L., KADER, M.F., SARKER, M.M.K., LEE, M.H. and HAN, D.S.

2022

This preprint is under consideration at Scientific Reports. A preprint is a preliminary version of a manuscript that has not completed peer review at a journal. Research Square does not conduct peer review prior to posting preprints. The posting of a preprint on this server should not be interpreted as an endorsement of its validity or suitability for dissemination as established information or for guiding clinical practice.

A Black Hole-Aided Deep-Helix Channel Model for DNA

Md Abdul Latif Sarker

Kyungpook National University

Md Fazlul Kader

University of Chittagong

Md Mostafa Kamal Sarker

Robert Gordon University

Moon Lee

Jeonbuk National University

Dong Han (✉ dshan@knu.ac.kr)

Kyungpook National University

Research Article

Keywords:

Posted Date: January 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1026992/v3>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Black Hole-Aided Deep-Helix Channel Model for DNA

Md. Abdul Latif Sarker^{1, 5}, Md. Fazlul Kader², Md. Mostafa Kamal Sarker³, Moon Ho Lee⁴, and Dong Seog Han^{5,*}

¹Centre for ICT & Automotive Convergence, Kyungpook National University, Daegu 41566, Korea

²Department of Electrical and Electronic Engineering, University of Chittagong, Chittagong, Bangladesh

³National Subsea Centre, Robert Gordon University, Aberdeen AB10 7GJ, United Kingdom

⁴Division of Electronics Engineering, Chonbuk National University, Jeonju-si 54896, Korea

⁵School of Electronics Engineering, Kyungpook National University, Daegu 41566, Korea

*corresponding.dshan@knu.ac.kr

ABSTRACT

In this article, we present a *black-hole-aided deep-helix* (bh-dh) channel model to enhance information bound and mitigate a multiple-helix directional issue in Deoxyribonucleic acid (DNA) communications. The recent observations^{1,2,3,4} of DNA do not match with Shannon^{5,6,7} bound due to their multiple-helix^{8,9} directional issue. Hence, we propose a bh-dh channel model in this paper. The proposed bh-dh channel model follows a similar fashion of DNA and enriches the earlier DNA observations^{1,2,3,4} as well as achieving a composite like information bound. To do successfully the proposed bh-dh channel model, we first define a *black-hole-aided Bernoulli*-process and then consider a symmetric bh-dh channel model. After that, the geometric and graphical insight shows the resemblance of the proposed bh-dh channel model in DNA and Galaxy layout. In our exploration, the proposed bh-dh symmetric channel geometrically sketches a *deep-pair-ellipse* when a *deep-pair* information bit or digit is distributed in the proposed channel. Furthermore, the proposed channel graphically shapes as a *beautiful circulant ring*. The ring contains a *central-hole*, which looks like a *central-black-hole*^{10,11} of a Galaxy. The coordinates of the *inner-ellipses* denote a *deep-double* helix and the coordinates of the *outer-ellipses* sketch a *deep-parallel* strand. Finally, the proposed bh-dh symmetric channel significantly outperforms the traditional binary^{5,6,7} symmetric channel and is verified by computer simulations in terms of Shannon entropy and capacity bound.

Introduction

Black¹² hole plays an important role in the part of nature, for example, a large galaxy contain a supermassive^{10,11} black-hole at its center. It is not completely black and the area of each individual black hole cannot decrease^{13,14}. A white¹⁵ hole also play a similar manner in the part of nature. There is a connection investigated^{15,16} between holes (black or white) and entropy. *It is to be noted that a black or a white hole is a constant blind source dependence natural strategy, which is widely used to expand the space of a region and help execute a multiple-helix directional issue in the part of nature. Since DNA^{1,2,3,4} is a part of nature, it carries out a multiple-helix^{8,9} directional issue. Hence, a black-hole strategy is not only present in the Galaxy, its existence is also noticeable in the DNA or other parts of nature such as a Cyclone.*

DNA is a fundamental appearance of life, which assists to trace out the hereditary information within chromosomes. It is mainly an information stream that relies on the nitrogenous bases^{1,2,17,18}: adenine (A), cytosine (C), guanine (G), and thymine (T). The quantity of each nucleobase is the key factor to trace out the hereditary information within chromosomes. The main idea of DNA was first introduced in 1869 by Friedrich Miescher. Almost 7 to 8 decades later, many scientists were investigated the physical structure of DNA molecules. For example, in 1950, Erwin Chargaff's team investigated the composition of human DNA and proposed Chargaff's rules of DNA base pairing¹⁷. Zamenhof et al showed that different living organisms contain different quantities of each DNA nucleobase¹⁸. In 1953, Linus Pauling and Robert Corey studied an unsatisfactory structure of nucleic acid in which their model consisted of three intertwined chains with the phosphates near the fiber axis and the bases on the outside^{2,3}. Around the same time, the X-ray fiber diffraction pattern was obtained by Rosalind Franklin¹⁹, and finally, on April 25, 1953, James Watson and Francis Crick suggested a satisfactory right-handed double¹-helix structure of DNA. After Watson and Crick model, several researchers have significantly studied the physical structure of DNA, for example, a twisted⁹-circular structure and a four⁴-stranded structure. It is observed that, most of the relevant above literature's investigated the DNA model utilizing the chemical and biological method. To the best of our knowledge, the concept of the proposed *black-hole-aided deep-helix* (bh-dh) channel model for DNA communications has not yet been studied in the literature.

In this paper, we propose a bh-dh channel model to expand information bound and conquer a multiple-helix directional issue of DNA communications. The geometrical form of the proposed bh-dh channel is a deep-pair ellipse, which graphically illustrates as a beautiful circulant ring. It is observed that the proposed bh-dh channel model follows a similar fashion of DNA and enriches the double-helix¹ DNA model as well as achieving a composite like information bound. In general, a stochastic process is a Bernoulli process. Hence, we first define a *black hole*-aided Bernoulli (bh-B) process. Based on the bh-B process, we then consider a symmetric bh-dh channel model and measure the composite-type of entropy and capacity bound of the channel. Through computer simulations, we finally verify the effectiveness of the proposed bh-dh symmetric channel with the conventional binary^{5,6,7} symmetric channel, in terms of Shannon entropy and capacity bound.

Proposed bh-dh channel model

In this Section, we will design the proposed bh-dh channel considering a black hole-aided technique. To do successfully design the proposed bh-dh channel, we call **Proposition 1** as below:

Proposition 1: A black or a white hole is a constant blind source, which produces a central-hole to help expand a circular symmetric channel space.

Proof of Proposition 1: To proof **Proposition 1**, we first define a black-hole-aided Bernoulli (bh-B) process and then consider a bh-dh symmetric channel model to analyze the geometric and graphical illustration as bellow:

Definition of a bh-B process

Consider η be a constant blind or *black-hole* source and $x \in \mathcal{X}$ be a binary or a discrete random source, which construct a composite input source x_d . The composited input source x_d satisfy $\eta \leq x_d \leq 1 + \eta$ with $0 \leq x \leq 1$. Now, using^{20,21}, we define a bh-B multiplicative form as follows.

$$\Pr(X_d = x_d) = \begin{cases} \frac{(1-p_d)^\eta}{p_d^\eta} p_d^{x_d} (1-p_d)^{1-x_d} & 0 < \eta \leq 1 \text{ for the right-handed bh-B case,} \\ \frac{(1-p_d)^{-\eta}}{p_d^{-\eta}} p_d^{x_d} (1-p_d)^{1-x_d} & -1 \leq \eta < 0 \text{ for the left-handed bh-B case,} \\ p^x (1-p)^{1-x} & \eta = 0 \text{ for otherwise,} \end{cases} \quad (1)$$

where, p_d denotes the composited probability mass function satisfy $\eta \leq p_d \leq 1 + \eta$ with $0 \leq p \leq 1$ and p denotes the probability mass function.

A symmetric bh-dh channel model

If we set $x \in \{0, 1\}$ in (1), then using^{6,7} and (1), we can model a symmetric bh-dh channel with transition matrix \mathbf{P}_d , which is given by

$$\mathbf{P}_d = \begin{bmatrix} 1-p_d & p_d \\ p_d & 1-p_d \end{bmatrix}. \quad (2)$$

It is noted that the proposed channel \mathbf{P}_d leads two different circulant matrices when two different bits or digits are applied in (2). We also noted that if p_d satisfies $0 \leq p_d \leq 1$ in (2), then the \mathbf{P}_d matrix be a circulant and if p_d satisfies $p_d < 0$ or $p_d > 1$ in (2), then the \mathbf{P}_d matrix become an inverse-circulant. Now, we analyze the geometrical and graphical illustration of the proposed \mathbf{P}_d matrix utilizing the DNA examples as follows.

(a) Geometrical construction of the proposed bh-dh symmetric channel: Let a *black-hole* constant $\eta = TG - AC$, where η is measured by the DNA nitrogenous bases^{1,12,13} (C, G, A and T) following the DNA physical structure. According to C, G, A and T values, we provide the following examples:

Example 1: For a Human^{18,22} DNA case, if any Human contains the percentage of the DNA nucleobases are $A = 29.3$, $T = 30.0$, $C = 20.0$ and $G = 20.7$, we then obtain a *black-hole* constant $\eta_h = 0.35$.

Example 2: For an Octopus^{18,22} DNA case, if any Octopus contains the percentage of the DNA nucleobases are $A = 33.2$, $T = 31.6$, $C = 17.6$ and $G = 17.6$, we then get a *black-hole* constant $\eta_o = -0.28$.

Example 3: For a hole-free case, if any organisms or viruses satisfy the percentage of the DNA nucleobases are $A = T$ and $C = G$, then the *black-hole* constant η be zero, i.e, $\eta_f = 0$.

Now, we set $p \in \{0, 1\}$, $\eta_h = 0.35$, $\eta_o = -0.28$ and $\eta_f = 0$ in (2) and make TABLE I. From (1), (2) and TABLE I, we can summarize the following points as:

(i) The proposed \mathbf{P}_d matrix pursues a right-handed DNA channel when η satisfies $0 < \eta \leq 1$ with $TG > AC$.

- (ii) The proposed \mathbf{P}_d matrix becomes a left-handed DNA channel when η satisfies $-1 \leq \eta < 0$ with $TG < AC$ and
- (iii) Otherwise \mathbf{P}_d matrix be a \mathbf{P}_s when $\eta = 0$ with $TG = AC$ and \mathbf{P}_s denotes the central hole-free^{5,6,7} symmetric circulant channel.

(b) Graphical illustration of the proposed bh-dh symmetric channel: According to the geometrical construction of the bh-dh symmetric channel, we analyze the graphical illustration in this subsection. The graphical illustration of the proposed bh-dh symmetric channel model is depicted in Fig. 1. In Fig. 1, we observe that the area of both inner and outer ellipses depend on p_d values, where constant η generates a *center-hole* in Fig. 2. In Fig. 2, we also observe that the proposed bh-dh symmetric channel model conducts a deep-pair ellipse when a deep-pair information bit or digit is applied in (2). For example, when we set $\eta = 0$ and a 1×198 binary-chain in (2), the proposed bh-dh symmetric channel model turns back into a binary symmetric channel and draws a central hole-free circle in Fig. 2a. In contrast, when we set $\eta = 0.35$ and a 1×198 binary-chain in (2), the proposed bh-dh symmetric channel model illustrates a *beautiful-circulant-ring* with central-hole as shown in Fig. 2b.

The most important graphical view describes in Fig. 3 and Fig. 4. Fig. 3 shows the graphical inspection of the proposed bh-dh channel model, which follows the physical structure of DNA. An X-ray pattern of the double¹-helix DNA model is investigated in Fig. 3a. A polar diagram depicts the proposed bh-dh channel model using by a 1×81 binary-chain in Fig. 3b. A 1×81 and 1×91 discrete-chain is applied with the proposed bh-dh channel model in Fig. 3c and Fig. 3d, respectively.

Fig. 4 presents the graphical exposition of the Milky way²³, the galaxy^{10,11} and the proposed bh-dh channel model. A black hole around the Milky way as shown in Fig. 4a. A polar diagram depicts the proposed bh-dh channel model utilizing a 1×198 binary-chain with $\eta = 0.35$ in Fig. 4b. a supermassive^{10,11} black hole residing at the galaxy's core in Fig. 4c. A polar diagram shows the proposed bh-dh channel model considering a 1×270 binary-chain with $\eta = 0.6180$ in Fig. 4d. According to Fig. 4b and Fig. 4d, it can be noticed that the area of the *center-hole* of a circle relies on the η values.

From the above definition, geometrical and graphical illustration, we can see that the (2) is a bh-dh symmetric channel, which has expanded a circular symmetric space by producing a central hole. Thus, **Proposition 1** has been proven.

Information entropy and capacity bound of the proposed bh-dh channel model

Particularly, the information capacity is the maximum amount of information that can pass through a channel without error. In contrast, information entropy is a function of a transitional error probability that is usually measured uncertainty in Shannons⁵ (bits) or natural units (nats) or decimal digits (dits). Since the Shannon bound does not match with the DNA information bound, we investigate the entropy and capacity bound of the proposed bh-dh channel model in this Section.

Consider random input and output variables X and Y for a communication channel where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Hence, the well-known capacity formula^{5,6,7} is given by

$$C_s = \max_{p(x)} I(X;Y). \quad (3)$$

where $p(x)$ denotes the input distribution function, C_s satisfy $0 \leq C_s \leq 1$ with $0 \leq x \leq 1$ and $0 \leq p \leq 1$, and $I(X;Y)$ denotes the mutual information⁶, that is given by

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(Y) - \sum p(x)H(Y|X=x) \\ &\leq 1 - H_s(p), \end{aligned} \quad (4)$$

where $H(X)$ and $H(Y)$ are the marginal entropies, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies, the p is the error probability satisfy $\sum_{x \in \{0,1\}} p(x) = 1$, and the Shannon entropy $H_s(p)$ is given by^{6,7} the following equation

$$H_s(p) = -p \log_2 p - (1-p) \log_2 (1-p) \quad \text{bits}, \quad (5)$$

where $H_s(p)$ satisfy $0 \leq H_s(p) \leq 1$. Similarly, based on (1), (3) and (4), we can write the following capacity formula for the proposed bh-dh channel model as

$$C_d = H_d(Y_d) - H_d(p_d) = \begin{cases} 1 - H_d(p_d, rh) & \text{for right-handed DNA case when } 0 < \eta \leq 1 \text{ with } TG > AC, \\ 1 - H_d(p_d, lh) & \text{for left-handed DNA case when } -1 \leq \eta < 0 \text{ with } TG < AC, \\ 1 - H_s(p) & \text{for Shannon case when } \eta = 0 \text{ with } TG = AC, \end{cases} \quad (6)$$

where $H_d(Y_d) = 1$ when $p_d = 1/2$ is applied in (6), $H_d(p_d)$ denotes the entropy of the proposed bh-dh channel model. Similarly in (5), the $H_d(p_d)$ is given by

$$H_d(p_d) = -p_d \log_2 p_d - (1 - p_d) \log_2 (1 - p_d) \quad \text{bits.} \quad (7)$$

Results and Discussion

In this section, we compare the proposed bh-dh symmetric channel model over the binary-symmetric channel via computer simulations. Throughout the simulations, we consider a variable p for both right and left-handed η cases in TABLE II and TABLE III. In TABLE II, when we set $p = 0$ for both $\eta = 1$ and $\eta = -1$ in (6) and (7), the measured entropy $H_d(p_d)$ is given by 0 bit and -2 bit and the measured capacity C_d is given by 1 bit and 3 bits, respectively. In contrast, when we set $p = 1$ for both $\eta = 1$ and $\eta = -1$, the measured entropy $H_d(p_d)$ is given by -2 bit and 0 bit and the measured capacity C_d is 3 bits and 1 bit, respectively.

In TABLE III, when we set $p = 0$ for both $\eta_h = 0.35$ and $\eta_o = -0.28$ in (6) and (7), the measured entropy $H_d(p_{d,h})$ is 0.93 bit and $H_d(p_{d,o})$ is -0.97 bit, and the measured capacity $C_{d,h}$ is 0.0659 bit and $C_{d,o}$ is 1.97 bits, respectively. In contrast, when we set $p = 1$ for both $\eta = 0.35$ and $\eta = -0.28$ in (6) and (7), the measured entropy $H_d(p_{d,h})$ is -1.11 bit and $H_d(p_{d,o})$ is 0.85 bit, and the measured capacity $C_{d,h}$ is 2.11 bits and $C_{d,o}$ is 0.15 bit, respectively.

Fig. 5 shows the maximum expandable entropy and capacity bound of the proposed bh-dh symmetric channel model when we set $p = 0 : 0.0001 : 1$, $\eta = 1$ and $\eta = -1$ in (6) and (7). In Fig. 5a and Fig. 5b, the composited entropy and capacity bound is given by $-2 \leq H_d(p_d) \leq 1$ and $0 \leq C_d \leq 3$ for the proposed bh-dh symmetric channel, respectively where the maximum expandable entropy and capacity is bounded by $-2 \leq H_d(p_d) \leq 0$ and $1 \leq C_d \leq 3$, respectively.

Fig. 6 illustrates the entropy and capacity of the proposed bh-dh channel model using a Human and an Octopus DNA case. Fig. 6a and Fig. 6c shows the entropy and capacity for Octopus DNA case when $p = 0 : 0.0001 : 1$, $\eta_o = -0.28$ is applied in (6) and (7). Fig. 6b and Fig. 6d presents the entropy and capacity for Human DNA case when $p = 0 : 0.0001 : 1$, $\eta_h = 0.35$ is applied in (6) and (7).

Conclusions

In this paper, we proposed the bh-dh channel model to expand information bound and help mitigate the multiple-helix directional issue of DNA communications. We observe that the geometrical shape of the proposed bh-dh channel comprised a deep-double helix DNA model, which enriched the double-helix DNA model. The graphical insight of the proposed bh-dh channel is also comprised a beautiful circulant ring, which showed a similar exposition of the galaxy. The composited entropy and capacity bound of the proposed bh-dh channel is $-2 \leq H(p_d) \leq 1$ and $0 \leq C_d \leq 3$, which can outperform the Shannon entropy and capacity bound. The maximum expandable entropy and capacity bound of the proposed bh-dh channel model are $-2 \leq H(p_d) \leq 0$ and $1 \leq C_d \leq 3$, respectively. Hence, the capacity improvement of the proposed bh-dh channel model is about 2 bits per channel use. In future work, the proposed bh-dh channel model can be extended further to the DNA healthcare multiple-input and multiple-output communication systems.

References

1. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids : A structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).
2. Pauling, L. & Corey, R. B. A proposed structure for the nucleic acid. *Proc Natl Acad. Sci., U. S. A* **39**, 84-97 (1953).
3. Pauling, L. & Corey, R. B. Structure of the nucleic acids. *Nature* **171**, 346 (1953).
4. Giulia, B. David, T. John, M. & Shankar, B. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chemistry* **5**, 182-186 (2013).
5. Shannon, C. E. A mathematical theory of communication. *The Bell System Technical Journal*, **27**, 379-423 (1948).
6. Cover, T. M. & Thomas, J. A. *Elements of Information Theory Ch. 8*. (John Wiley & Sons Inc, New York, 1991).
7. Hamming, R. W. *Coding and Information Theory Ch. 6 & Appen. B*. (Prentice-Hall, New Jersey, 1986).
8. Vinograd, J. Bruner, R. Kent, R. & Weigle, J. Band-centrifugation of macromolecules and viruses in self-generating density gradients. *Proc Natl Acad. Sci., U. S. A*, **49**, 902-910 (1963).
9. Vinograd, J. Lebowitz, J. Radloff, R. Watson, R. & Laipis P. The twisted circular form of polyoma virus DNA. *Proc Natl Acad. Sci., U. S. A*, **53**, 1104-1111 (1965).

10. John, K. & Luis, C. H. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *Annual Review of Astronomy and Astrophysics*, **51**, 511–653 (2013).
11. Online at <https://www.nasa.gov/feature/goddard/2019/hubble-uncovers-black-hole-that-shouldnt-exist> (2019).
12. Misner, C.W. *Gravitation*. (Freeman, San Francisco, 1973).
13. Hawking, S. W. Gravitational Radiation from Colliding Black Holes. *Phys. Rev. Lett.* **26**, 1344-1346 (1971).
14. Hawking, S. W. Black hole explosions?. *Nature* **248**, 30–31 (1974).
15. Hawking, S. W. Black holes and thermodynamics. *Phys. Rev. D* **13**, 191 –197 (1976).
16. Bekenstein, J. D. Black holes and Entropy. *Phys. Rev. D* **15**, 2333 –2346 (1976).
17. Chargaff, E. Zamenhof, S. & Green, C. Composition of human desoxypentose nucleic acid. *Nature* **165**, 756-757 (1950).
18. Zamenhof, S. Chargaff, E. & Brawerman, G. On the Desoxypentose acids from several microorganisms. *Biochimica et Biophysica Acta*, **9**, 402-405 (1952).
19. Franklin, R. E. & Gosling, R. Molecular configuration in sodium thymonucleate. *Nature*, **171**, 740-741 (1953).
20. Uspensky, J. V. Introduction to Mathematical Probability. *Nature* **141**, 769 (1938)
21. Klenke, A. *Probability Theory*. (Springer-Verlag. 2006).
22. Bansal, M. DNA structure: Revisiting the Watson-Crick double helix. *Current Science*, **85**, 1556-1563 (2003).
23. Online at <https://www.space.com/43024-rare-black-hole-in-milky-way.html> (2019).

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (2021R1A6A1A03043144).

Author contributions statement

Conceptualization, methodology, visualization, writing, M. A. Latif Sarker; validation, formal analysis, review and editing, Md. F. Kader, M. M. Kamal Sarker, M. H. Lee and D. S. Han. All authors reviewed the manuscript.

Conflicts of interest

The authors declare no conflict of interest

Figures

Figure 1

See image above for figure legend.

Figure 2

See image above for figure legend.

Figure 3

See image above for figure legend.

Figure 4

See image above for figure legend.

Figure 5

See image above for figure legend.

Figure 6

See image above for figure legend.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableFile.pdf](#)