

PAPER • OPEN ACCESS

Genetic alphabets, unitary matrixes and quantum-algorithmic genetics

To cite this article: S Petoukhov *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **489** 012048

View the [article online](#) for updates and enhancements.

You may also like

- [Efficient methods for determining folding free energies in single-molecule pulling experiments](#)

A Severino, A M Monge, P Rissone et al.

- [Designing synthetic RNA for delivery by nanoparticles](#)

Dominika Jedrzejczyk, Edyta
Gendaszewska-Darmach, Roza
Pawlowska et al.

- [Identification of unique interactions between the flexible linker and the RecA-like domains of DEAD-box helicase Mss116](#)

Yuan Zhang, Mirkó Palla, Andrew Sun et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Genetic alphabets, unitary matrixes and quantum-algorithmic genetics

S Petoukhov^{1,2}, E Petukhova¹, V Svirin^{1,2}

¹Mechanical Engineering Research Institute of RAS, 4, M. Kharitonyevskiy Lane, Moscow, 101990, Russia

²Moscow State Conservatory, 13/6 Bolshaya Nikitskaya Street, Moscow, 125009, Russia

spetoukhov@gmail.com

Abstract. Unitary operators are the basis of all calculations in quantum computers and play an important role in quantum mechanics. The presented results show that matrix forms of representation of structured DNA and RNA alphabets can be considered as a set of sparse unitary matrices. These results are useful for the development of models of quantum-algorithmic genetics, from the point of view of which living organisms are quantum-information entities.

1. Introduction

Information molecules of DNA and RNA follow the principles of quantum mechanics. It can be assumed that genetic informatics uses quantum computing algorithms based on unitary (or orthogonal, in the case of real numbers) matrices. Any unitary matrix can be used as a logic gate in quantum computers [1]. The article presents some authors' results for the development of quantum-algorithmic genetics models [2-8].

2. The genetic code and binary oppositions

Why DNA alphabet is built by nature on the basis of only four letters: adenine A, guanine G, cytosine C and thymine T? Why is this alphabet based on these very simple molecules? Modern science does not know the answer to these questions. But it knows that this set of 4 "letters" has symmetric properties because it is endowed with three pairs of binary-opposition signs or indicators:

- 1) two letters are purines (A and G) and the other two are pyrimidines (C and T). In terms of these features the following can be taken: $C = T = 0$, $A = G = 1$;
- 2) two complementary letters (A and T) are connected by 2 hydrogen bonds, and other complementary letters (C and G) are connected by 3 hydrogen bonds. From the point of view of this feature, the following can be taken: $A = T = 0$, $C = G = 1$;
- 3) two letters are keto-molecules (G and T) and the other two are amino-molecules (A and C).

From the point of view of this feature, the following can be taken $A = C = 0$, $G = T = 1$. With regard to these three types of binary-oppositional traits in DNA letters any DNA text can be presented as a combination of three parallel messages in three binary languages. At the same time, these messages are mutually connected on the basis of the logical operation of modulo-2 addition used in quantum and conventional computers: the sum of modulo-2 addition of any two of these three binary representations of the DNA text always gives its third binary representation [4].



It is easy to check that each of the letters A, C, G, T is uniquely determined by any two of these three indicators. This allows presenting DNA alphabets of 4 nucleotides, 16 doublets, 64 triplets, etc. in a convenient form of square tables (Fig. 1-3), in which lines and columns are numbered with binary numbers according to the following principle. The components of each column are binary numbered in accordance with the binary opposition signs “purine or pyrimidine” (for example, the CAG triplet and all other triplets in the same column are a bunch of “pyrimidine-purine-purine”, and therefore this column is numbered 011). The components of each line are numbered with binary numbers in accordance with the signs of “keto- or amino-” (for example, the same CAG triplet and all other triplets in its line are a bunch of “amino-amino-keto”, and therefore this line is numbered 001). In such tables (Fig. 1-3) each of the members of the alphabets of 4 nucleotides, 16 doublets, 64 triplets, ... takes its individual place in a strict order

	000	001	010	011	100	101	110	111
000	CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
001	CCT	CCG	CAT	CAG	ACT	ACG	AAT	AAG
010	CTC	CTA	CGC	CGA	ATC	ATA	AGC	AGA
011	CTT	CTG	CGT	CGG	ATT	ATG	AGT	AGG
100	TCC	TCA	TAC	TAA	GCC	GCA	GAC	GAA
101	TCT	TCG	TAT	TAG	GCT	GCG	GAT	GAG
110	TTC	TTA	TGC	TGA	GTC	GTA	GGC	GGA
111	TTT	TTG	TGT	TGG	GTT	GTG	GGT	GGG

Figure 1. Table of DNA alphabet of triplets with the arrangement of all members in accordance with the binary-oppositional signs of the bases A, C, G, T..

	0	1
0	C	A
1	T	G

	00	01	10	11
00	CC	CA	AC	AA
01	CT	CG	AT	AG
10	TC	TA	GC	GA
11	TT	TG	GT	GG

Figure 2. Table of DNA alphabet of nucleotides with the arrangement of all members in accordance with the binary-oppositional signs of the bases A, C, G, T

Figure 3. . Table of DNA alphabet of doublets with the arrangement of all members in accordance with the binary-oppositional signs of the bases A, C, G, T

It is essential that these 3 separate genetic tables form a single tensor family of matrices based on the tensor (or Kronecker) product [2, 4]. Indeed, the second tensor degree of (2x2)-nucleotide matrix [C, A; T, G] automatically gives the (4x4)-matrix of 16 doublets, and its third tensor degree gives the (8x8)-matrix of 64 triplets with the same arrangement of components as in the tables in Fig. 1-3, which appear now as matrices. Thus, the structural organization of the system of DNA alphabets is connected with matrices and algebraic operation of the tensor product of matrices, which plays an important role in quantum mechanics and quantum computing in connection with the following: the state space of a multicomponent quantum system is the tensor product of the state spaces of its components [1, p. 71].

It is well known that the degeneracy of the genetic code is associated with the division of the whole set of 64 triplets by the nature into two equal subsets:

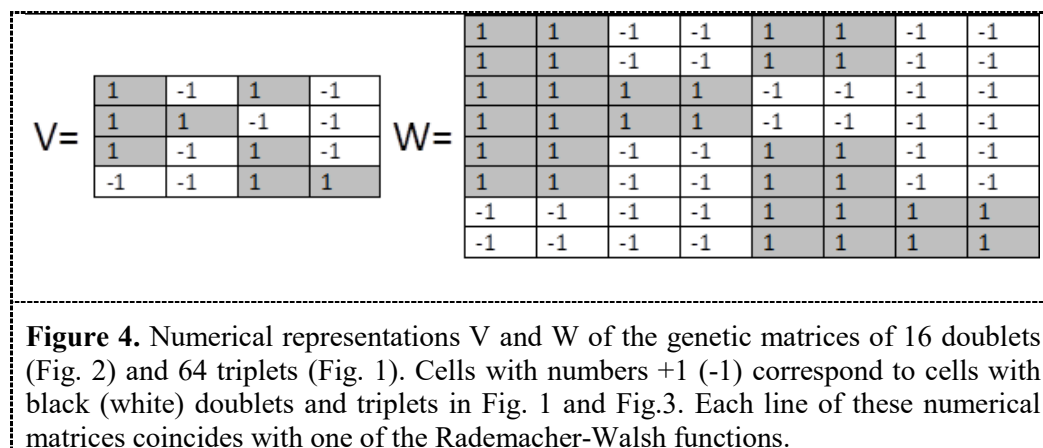
- 32 triplets have "strong roots" (indicated in black in the matrixes in Fig. 1), i.e. start with 8 "strong" doublets AC, CC, CG, CT, GC, GG, GT, TC;
- 32 triplets have "weak roots" (white in Fig. 1), i.e. start with 8 "weak" doublets CA, AA, AG, AT, GA, TA, TG, TT.

Code values of triplets with strong roots do not depend on the letters in the third position, but triplets with weak roots depend on the letters in the third position.

What is the location of these "black" and "white" members of the DNA alphabets in the genetic matrixes of 16 doublets and 64 triplets in Fig. 1? An unexpected phenomenological fact is their symmetrical arrangement in these matrices, built quite formally without mentioning the degeneracy of the genetic code. The symmetric properties of the resulting mosaics in these matrices are as follows:

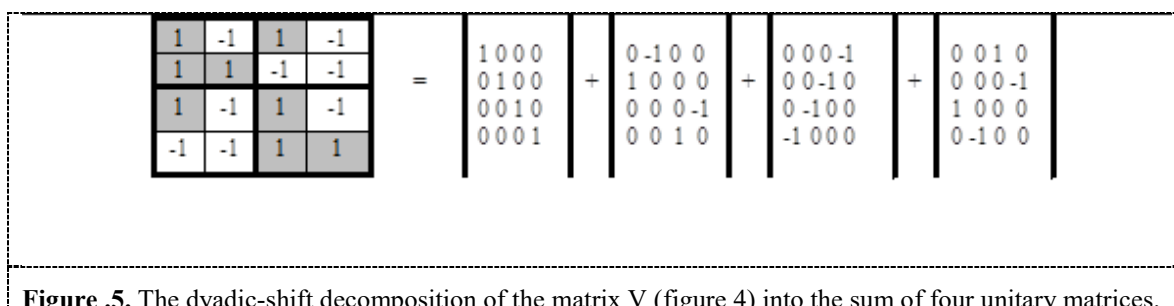
- the left and right halves of the matrices are mirror-antisymmetric to each other;
- both quadrants along each diagonal are the same on the mosaic;
- the mosaic of each line has a meander character of the known Rademacher functions, which are a special case of the Walsh functions and contain only the numbers +1 and -1.

Through this analogy with the Rademacher-Walsh functions from the theory of noise-immunity coding of information, the symbolic genomatrix from Fig. 1-3 are converted to numerical matrixes V and W with components +1 and -1, where components +1 (-1) represent black (white) doublets and triplets respectively (Fig. 4).



3. On unitary matrices

These matrices V and W (Fig. 4) represent the phenomenological features of the degeneracy of the genetic code. Below the properties of genetic matrices V and W will be analyzed more deeply. The analysis begins with (4x4)-matrix V, which turns out to be the sum of four sparse unitary (orthogonal) matrixes $V = V_0 + V_1 + V_2 + V_3$ (Fig. 5), for which the unitarity condition is fulfilled: $V_0 \cdot V_0^T = I$, $V_1 \cdot V_1^T = I$, $V_2 \cdot V_2^T = I$, $V_3 \cdot V_3^T = I$, where I is the unit matrix (this decomposition is not arbitrary, but is based on the principle of dyadic shifts from the theory of digital signal processing [2]).



Unitary transformations preserve the norms of vectors. The tensor product of two unitary matrixes always generates a new unitary matrix. According to the described result, the degeneracy of the genetic code is associated with sparse unitary matrices, which can serve as quantum gates in biological quantum computers.

A set of these unitary matrices V_0, V_1, V_2 and V_3 (Fig. 5) is closed with respect to multiplication. Their multiplication table (Fig. 4) coincides with the multiplication table of Cockle's split-quaternions (<http://en.wikipedia.org/wiki/Split-quaternion>), known in particular in connection with the Poincare disk model of Lobachevski geometry.

*	V_0	V_1	V_2	V_3
V_0	V_0	V_1	V_2	V_3
V_1	V_1	$-V_0$	V_3	$-V_2$
V_2	V_2	$-V_3$	V_0	$-V_1$
V_3	V_3	V_2	V_1	V_0

Figure 6. Table of multiplication of unitary genetic matrixes V_0, V_1, V_2, V_3 from Fig. 5

It is to be added that the dyadic-shift decomposition of (8x8)-matrix W (Fig. 4) gives a set of 8 sparse unitary matrices, which is also closed with respect to multiplication and has a Cockle's bi-split quaternion multiplication table [4].

The reasons for the connection of structured DNA and RNA alphabets in their different matrix representations with Cockle's split-quaternions should be explained. It seems to be related to the connection of split-quaternions and Poincare disc model of Lobachevski geometry with known experimental data on Lobachevski geometry of the space of visual perception, which has resulted in some authors' model approaches to the brain on the basis of Clifford algebra of quantum computers (see details in [4]). Our results on the relationship between the genetic system and split-quaternions support the importance of developing algebra-geometric approaches to understanding the inherited properties of sensory systems and biological informatics in general.

The lines and columns of each of the unitary matrices V_0, V_1, V_2 and V_3 form an orthogonal system of functions that can be used for spectral analysis of physiological and other processes in their vector representations. The construction of each of these matrices to tensor powers gives new unitary matrixes of increased order with orthogonal systems of functions in their lines and columns.

4. Conclusions

The results presented in the article together with our other works [2-9] allow us to develop the elements of quantum-algorithmic genetics as one of the promising new directions of science. This scientific direction concerns also works of some other authors about quantum biology, which suggests that biological phenomena such as photosynthesis, enzyme catalysis, avian navigation or olfaction may use of a number of the non-trivial features of quantum mechanics, such as coherence, tunnelling and entanglement.

References

- [1] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information*. (New York: Cambridge University Press)
- [2] Petoukhov S V 2008 *Matrix Genetics, Algebras of the Genetic Code, Noise Immunity* (Moscow: RHD)
- [3] Petoukhov S V 2017 The rules of long DNA-sequences and tetra-groups of oligonucleotides *Preprint* 1709.04943
- [4] Petoukhov S V 2018 The genetic coding system and unitary matrices *Preprint* 201804.0131
- [5] Petoukhov S V and Svirin V I *Biologia Serbica*, **40**, 51

- [6] Petoukhov S, Petukhova E and Svirin V New Symmetries and Fractal-Like Structures in the Genetic Coding System. In: Z. Hu, S. Petoukhov, Dychka I., He, M. (eds) *Advances in Computer Science for Engineering and Education. ICCSEEA 2018. Advances in Intelligent Systems and Computing*, vol. 754. Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-91008-6_59
- [7] Petoukhov S V January 2016 The system-resonance approach in modeling genetic structures *Biosystems* **139** 1-11
- [8] Petoukhov S. V. and Petukhova E. S. 2017 Symmetries in genetic systems and the concept of geno-logical coding. - *Information*, **8**, <https://doi.org/10.3390/info8010002>
- [9] . Darvas G. 2018 Petoukhov's rules on symmetries in long DNA-texts *Symmetry: Culture and Science* **29** 318-20 <https://doi.org/10.26830/>