MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **93** (2025) 713–724
ISSN: 0340–6253
doi: 10.46793/match.93-3.713V

Hamming Matrix and Hamming Energy of a Graph

Nemanja Vučićević[a,*], Izudin Redžepović[b], Nenad Stojanović[a]

[a] *Faculty of Science, University of Kragujevac, Kragujevac, Serbia*

[b] *Department of Natural Sciences and Mathematics, State University of Novi Pazar, Novi Pazar, Serbia*

nemanja.vucicevic@pmf.kg.ac.rs, iredzepovic@np.ac.rs, nenad.stojanovic@pmf.kg.ac.rs

(Received July 23, 2024)

Abstract

Hamming distance is a highly valuable quantity in computer science. In this work, we establish the Hamming matrix H of a graph G, $H(G)$. This is a square matrix, where the elements of the $H(G)$ are Hamming distances. Also, we define the Hamming energy of a graph, $HE(G)$, which is a sum of the absolute eigenvalues of $H(G)$. Finally, we present some bounds on the $HE(G)$ and its predictive potential.

1 Introduction

Let $G = (V, E)$ be a simple, undirected graph with n vertices and m edges. Let $V(G) = \{v_1, v_2, \ldots, v_n\}$ be the vertex set of G and $E(G) = \{e_1, e_2, \ldots, e_m\}$ be the edge set of G. If the vertices v_i and v_j are adjacent then we write $v_i \sim v_j$ and if they are not adjacent then we write $v_i \nsim v_j$. The edge and its end vertex are said to be incident to each other. The

*Corresponding author.

degree of a vertex v_i, denoted by $deg_G(v_i)$ is the number of edges incident to it.

The concept of graph energy was introduced in 1978 [3]. The energy of a graph G, $E(G)$, is defined as the sum of the absolute values of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_n$ of adjacency matrix A, i.e.,

$$E(G) = \sum_{i=1}^{n} |\lambda_i|.$$

Also, Nikiforov [5] extended the concept of energy to all (not necessarily square) matrices, defining the energy of a matrix A as the sum of the singular values of M. Recall that the singular values of a matrix M are equal to the square roots of the eigenvalues of the (square) matrix MM^T. There are a large number of results related to the energy of a graph. Graph energy has significant applications in chemistry, for example, in encoding information on molecular structure using topological molecular descriptors. Topological molecular descriptors are mathematical values calculated from a graph representation of a molecule, also known as topological indices. These descriptors are proven to be very valuable in chemistry. Today, they find extensive application, as can be seen in the works [1, 2, 4, 9–11, 13].

Here, we will present results related to eigenvalue estimates, which can be found in [12].

The spectrum $\sigma(A)$ of a square matrix $A = [a_{ij}]_{n\times n}$ is the collection of all eigenvalues of A, i.e.,

$$\sigma(A) := \{\lambda \mid \det(A - \lambda I) = 0\},$$

and with

$$N := \{1, 2, \ldots, n\},$$

we call

$$r_i(A) := \sum_{j\in N\setminus\{i\}} \mid a_{ij} \mid, \ (i \in N)$$

the i-th deleted absolute row sum of A. Further, we set

$$\begin{cases} \Gamma_i(A) := \{z \mid |z - a_{ii}| \leq r_i(A)\}, & i \in N; \\ \Gamma(A) := \bigcup\limits_{i \in N} \Gamma_i(A), \end{cases} \tag{1}$$

$\Gamma_i(A)$ is called the ith-Geršgorin disk of A, and the union of the n Geršgorin disks is called the Geršgorin set.

Theorem 1. *(Geršgorin theorem) For any $A = [a_{ij}]_{n \times n}$ and any $\lambda \in \sigma(A)$, there is a positive integer $k \in N$ such that*

$$| \lambda - a_{kk} | \leq r_k(A).$$

Consequently (from (1)), $\lambda \in \Gamma_k(A)$, and hence, $\lambda \in \Gamma(A)$. As this is true for each $\lambda \in \sigma(A)$, then

$$\sigma(A) \subseteq \Gamma(A).$$

On the other hand, let $\mathbb{Z}_2 = \{0, 1\}$ and $(\mathbb{Z}_2, +)$ be the additive group, where $+$ denotes addition modulo 2. For any positive integer n,

$$\mathbb{Z}_2^n = \{(x_1, x_2, \ldots, x_n) | x_1, x_2, \ldots, x_n \in \mathbb{Z}_2\}.$$

Element of $\mathbb{Z}_2^n$ is an n-tuple $(x_1, x_2, \ldots, x_n)$ written as $x = x_1 x_2 \ldots x_n$ where every x_i is either 0 or 1 and it is called a string.

Let $x = x_1 x_2 \ldots x_n$ and $y = y_1 y_2 \ldots y_n$ be the elements of $\mathbb{Z}_2^n$. Then the sum $x \oplus y$ is computed by adding the corresponding components of x and y under addition modulo 2. That is, $x_i + y_i = 0$ if $x_i = y_i$ and $x_i + y_i = 1$ if $x_i \neq y_i$, $i = 1, 2, \ldots, n$.

The Hamming distance $H_d(x, y)$ between the strings $x = x_1 x_2 \ldots x_n$ and $y = y_1 y_2 \ldots y_n$ is the number of is such that $x_i \neq y_i$, $i = 1, 2, \ldots, n$. Thus, $H_d(x, y)$ is a number of positions in which x and y differ. For example, if $x = 01001$ and $y = 11010$, $H_d(x, y) = 3$, see [7, 8]. Additionally, Hamming distance is employed in various applications such as cryptography, bioinformatics, and information retrieval, where it helps in matching patterns or sequences by quantifying their similarity.

2 Preliminaries

The incidence matrix of graph G is a matrix $B(G) = [b_{ij}]$ of order $n \times m$, in which $b_{ij} = 1$ if the vertex v_i is incident to the edge e_j and $b_{ij} = 0$, otherwise. Denote by $s(v)$, the row of the incidence matrix corresponding to the vertex v. It is a string in the set $\mathbb{Z}_2^n$ of all n-tuples over the field of order two. Sum of Hamming distances between all pairs of strings generated by the incidence matrix of a graph G is denoted by $H_B(G)$ and is called the Hamming index of G, i.e.,

$$H_B(G) = \sum_{\{u,v\} \subseteq V(G)} H_d(s(u), s(v)). \tag{2}$$

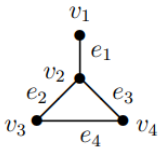

Figure 1. An example graph.

Example 1. For a graph G given in Figure 1, the incidence matrix is

$$B(G) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Therefore, the Hamming index is $H_B(G) = 2 + 3 + 3 + 3 + 3 + 2 = 16$.

Now, we present some important consequences for the Hamming distance, which will be utilized in this work, based on the findings from [7].

Theorem 2. *(Theorem 2.1. [7]) Let G be a graph with n vertices and m edges. Let u and v be the vertices of G and l be the number of edges which are neither incident to u nor incident to v. Then*

$$H_d(s(u), s(v)) = \begin{cases} m - 1 - l, & \text{if } u \sim v; \\ m - l, & \text{if } u \nsim v. \end{cases}$$

Theorem 3. *(Theorem 2.2. [7]) Let u and v be the vertices of G. Then*

$$H_d(s(u), s(v)) = \begin{cases} deg_G(u) + deg_G(v) - 2, & \text{if } u \sim v; \\ deg_G(u) + deg_G(v), & \text{if } u \nsim v. \end{cases}$$

Theorem 4. *(Corollary 3.1. [7]) If G is a complete graph K_n, then $H_B(K_n) = n(n-1)(n-2)$.*

Theorem 5. *(Corollary 3.2. [7]) If G is a cycle graph C_n, then $H_B(C_n) = 2n(n-2)$.*

3 Hamming matrix and its energy

In this section, we introduce the Hamming matrix and the energy derived from a such matrix.

Specifically, the Hamming matrix of graph G, $H(G) = [h_{ij}]_{n\times n}$, is formed as a square matrix, where the elements of the square matrix are Hamming distances, i.e.,

$$h_{ij} = H_d(s(v_i), s(v_j)).$$

Thus, we obtain a symmetric matrix with zeros on its main diagonal, because $h_{ii} = H_d(s(v_i), s(v_i)) = 0$. We also obtain that it holds

$$\sum_{ij} h_{ij} = 2H_B(G).$$

Following the concept of the energy of a graph, we define the Hamming energy of a graph G as the sum of the absolute eigenvalues of the Hamming matrix, i.e., if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of $H(G)$, the Hamming energy denoted by $HE(G)$ is

$$HE = HE(G) = \sum_{i=1}^{n} |\lambda_i|. \tag{3}$$

For instance, based on Example 1, we derive the Hamming matrix

$$H(G) = \begin{bmatrix} 0 & 2 & 3 & 3 \\ 2 & 0 & 3 & 3 \\ 3 & 3 & 0 & 2 \\ 3 & 3 & 2 & 0 \end{bmatrix}.$$

The eigenvalues of the matrix are $\lambda_1 = -4$, $\lambda_2 = -2$, $\lambda_3 = -2$, $\lambda_4 = 8$, therefore, the Hamming energy is

$$HE(G) = \sum_{i=1}^{4} |\lambda_i| = 16$$

and, in this case, it coincides with $H_B(G)$.

4 Some bounds on the hamming energy of a graph

In this section, we will present some results on the Hamming energy of the graph G and the Hamming matrix.

Theorem 6. *Let G be a graph of order n with no isolated vertices. Then*

$$HE(G) \leq 2\sqrt{n} \cdot H_B(G). \tag{4}$$

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of the Hamming matrix of graph G. Using the inequality between the arithmetic mean and the quadratic mean

$$\sum_{i=1}^{n} |\lambda_i| \leq \sqrt{n \sum_{i=1}^{n} \lambda_i^2} = \sqrt{n \cdot \mathrm{tr}(H^T(G)H(G))}.$$

Therefore, from $\mathrm{tr}(H^T(G)H(G)) = \sum_{ij} h_{ij}^2$ we obtain

$$\sqrt{n \cdot \mathrm{tr}(H^T(G)H(G))} = \sqrt{n \cdot \sum_{ij} h_{ij}^2} \leq \sqrt{n \cdot \left(\sum_{ij} h_{ij}\right)^2}$$
$$= \sqrt{n \cdot 4H_B^2(G)} = 2\sqrt{n}H_B(G).$$

So, we get the upper bound for Hamming energy by the Hamming index

$$HE(G) = \sum_{i=1}^{n} |\lambda_i| \leq 2\sqrt{n}H_B(G). \qquad \blacksquare$$

In the special case of the complete graph K_n, there is a distinct relationship between the Hamming energy and the Hamming index. The complete connectivity of K_n allows for explicit formulas that directly relate these two invariants, showcasing how the symmetrical structure of K_n influences their values. This helps us understand K_n's properties and sets a standard for more complex graphs.

Theorem 7. *If K_n is complete graph then*

$$HE(G) = 4(n-1)(n-2).$$

Proof. If K_n is complete graph, then $deg_{K_n}(u) = n-1$, $u \in V(G)$. Also from Theorem 4 we have Hamming distances

$$H_d(s(u), s(v)) = \begin{cases} deg_G(u) + deg_G(v) - 2, & \text{if } u \sim v \\ deg_G(u) + deg_G(v), & \text{if } u \nsim v \end{cases}$$
$$= \begin{cases} 2n-4, & \text{if } u \sim v; \\ 0, & \text{if } u \nsim v. \end{cases}$$

Therefore, Hamming matrix is

$$H(G) = \begin{bmatrix} 0 & 2n-4 & \dots & 2n-4 & 2n-4 \\ 2n-4 & 0 & \dots & 2n-4 & 2n-4 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2n-4 & 2n-4 & \dots & 0 & 2n-4 \\ 2n-4 & 2n-4 & \dots & 2n-4 & 0 \end{bmatrix}.$$

From $det(H(G) - \lambda I) = 0$ we obtain $\lambda_1 = (n-1)(2n-4)$ and $\lambda_i = 2n-4$, for $i = 2, \ldots, n$. Given this,

$$HE(G) = \sum_{i=1}^{n} |\lambda_i| = 2 \cdot (n-1)(2n-4) = 4(n-1)(n-2). \qquad \blacksquare$$

Corollary. *For the complete graph K_n with $n \geq 4$, we have $HE(G) = \frac{4}{n} H_B(G)$.*

Considering the relationship between the Hamming index $H_B(G)$ of graph G and its largest eigenvalue λ_1, the next inequality provides a significant insight into the spectral characteristics of G. This inequality shows how important the Hamming index is in limiting spectral properties.

Theorem 8. *Let G be a graph on n vertices and let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of the Hamming matrix $H(G)$. If $H_B(G)$ is the Hamming index of the graph G, then*

$$\frac{2H_B(G)}{n} \leq \lambda_1.$$

Proof. Given that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ and applying the Cauchy-Schwarz inequality, we obtain

$$n\lambda_1^2 \geq \sum_{i=1}^{n} \lambda_i^2 = \operatorname{tr}\left(H(G)^T H(G)\right) = \sum_{ij} h_{ij}^2 \geq \frac{\left(\sum\limits_{ij} h_{ij}\right)^2}{n} = \frac{4H_B^2(G)}{n}.$$

Therefore,

$$\frac{2H_B(G)}{n} \leq \lambda_1. \qquad \blacksquare$$

By applying Geršgorin's circular theorem (1), we provide meaningful estimates for the Hamming energy of certain special classes of graphs. This represents a significant result and application in graph energy. Therefore, we illustrate the use of the theorem through our examples.

Theorem 9. *If C_n is cycle graph, then*

$$HE(G) \leq 4n(n-2).$$

Proof. If C_n is cycle graph, then $deg(u) = 2$, $u \in V(G)$. Also from Theorem 4 we have Hamming distances

$$H_d(s(u), s(v)) = \begin{cases} 2, & \text{if } u \sim v; \\ 4, & \text{if } u \nsim v. \end{cases}$$

Therefore, Hamming matrix is

$$H(G) = \begin{bmatrix} 0 & 2 & 4 & \dots & 4 & 2 \\ 2 & 0 & 2 & \dots & 4 & 4 \\ 4 & 2 & 0 & \dots & 4 & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 4 & 4 & 4 & \dots & 0 & 2 \\ 2 & 4 & 4 & \dots & 2 & 0 \end{bmatrix}.$$

Hence, in each row, we have one 0, two 2s, and the rest are 4s. One of the most significant results we use is Geršgorin's theorem. In our case, the Hamming matrix has a special form, with all zeros on the main diagonal. Therefore, the application of this theorem is significant because it is directly relevant to our evaluation.

$$|\lambda - h_{ii}| = |\lambda| \leq r_i(H(G)).$$

Therefore, for $\lambda \in \sigma(H(G))$ we have

$$|\lambda| \leq r_i(H(G)) = 2 \cdot 2 + 4 \cdot (n-3) = 4(n-2).$$

Given this, the energy is

$$HE(G) = \sum_{i=1}^{n} |\lambda_i| \leq \sum_{i=1}^{n} r_i(H(G)) = \sum_{i=1}^{n} 4(n-2) = 4n(n-2). \qquad \blacksquare$$

Theorem 10. *If S_n $(n > 3)$ is a star graph and let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of the Hamming matrix $H(G)$, then*

$$\lambda_1 \leq (n-1)(n-2).$$

Proof. For star graph S_n we have Hamming distances

$$H_d(s(u), s(v)) = \begin{cases} n-2, & \text{if } u \sim v; \\ 2, & \text{if } u \nsim v. \end{cases}$$

Therefore, Hamming matrix is

$$H(G) = \begin{bmatrix} 0 & n-2 & n-2 & \dots & n-2 & n-2 \\ n-2 & 0 & 2 & \dots & 2 & 2 \\ n-2 & 2 & 0 & \dots & 2 & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ n-2 & 2 & 2 & \dots & 0 & 2 \\ n-2 & 2 & 2 & \dots & 2 & 0 \end{bmatrix}.$$

Hence we have two cases. For the first row, we have one 0 and the rest are $n-2$'s. So, form Geršgorin's theorem 1 we obtain

$$|\lambda| \leq r_1(H(G)) = (n-1) \cdot (n-2).$$

The following applies to the rest of the species

$$|\lambda| \leq r_i(H(G)) = n-2+(n-2) \cdot 2 = 3(n-2), \;\; i \neq 1.$$

So, for $n > 3$ we have $r_1(H(G)) \geq r_i(H(G))$. Therefore,

$$\lambda_1 \leq (n-1)(n-2).$$ ■

5 Predictive potential of the hamming energy of a graph

To give a glimpse into the possible chemical applicability of the Hamming graph energy, we present here its predictive potential.

The correlation between $HE(G)$ and the entropy (S), the heat of vaporization (H_{vap}), and the heat of formation (H_f) of octane molecules is depicted in Figure 2. To obtain a better perspective into the predic-

tive potential of $HE(G)$, the correlation between graph energy $E(G)$ and physicochemical properties is also depicted.

Figure 2 shows that the Hamming energy of a graph is better correlated with the S and H_{vap} of molecules, compared to the graph energy. Namely, in both cases, the coefficient of correlation (R) for the Hamming energy of a graph is significantly higher indicating that $HE(G)$ may be applied to model these physicochemical properties of molecules. However, the graph energy shows a somewhat better correlation with the heat of formation.

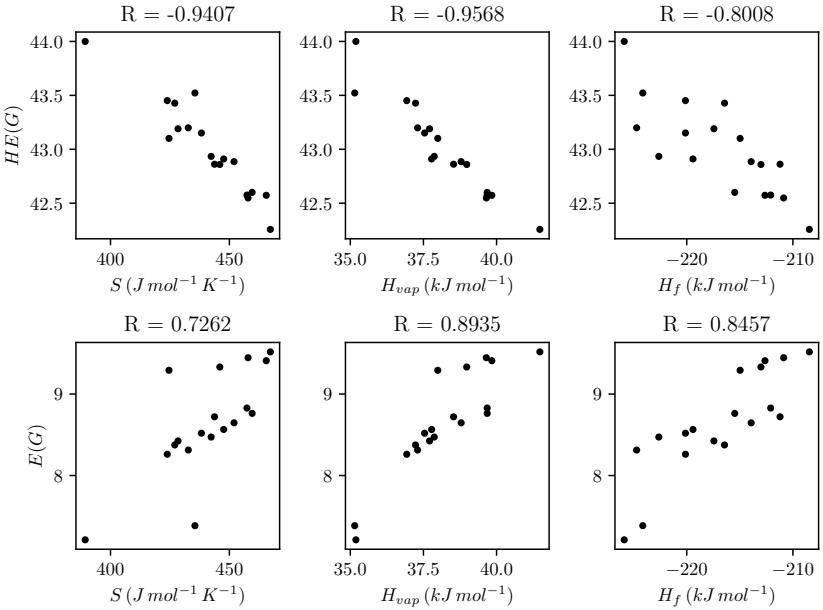

Figure 2. Correlation between the Hamming energy $HE(G)$ and energy $E(G)$ and the physicochemical properties of octanes.

***Acknowledgment*:** This work was supported by the Serbian Ministry of Science, Technological Development and Innovation (Agreement No. 451-03-65/2024-03/200122 and 451-03-65/2024-03/200252). I.R. gratefully acknowledges the financial support of the State University of Novi Pazar.

References

[1] M. V. Diudea, I. Gutman, L. Jantschi, *Molecular Topology*, Nova Sci. Pub., Huntington, 2001.

[2] O. Ivanciuc, Chemical graphs, molecular matrices and topological indices in chemoinformatics and quantitative structure–activity relationships, *Curr. Comput. Aided Drug Des.* **9** (2013) 153–163.

[3] I. Gutman, The energy of a graph, *Ber. Math. Stat. Sekt. Forschungszent. Graz* **103** (1978) 1–22.

[4] J. M. Amigó, J. Gálvez, V. M. Villar, A review on molecular topology: applying graph theory to drug discovery and design, *Naturwissenschaften* **96** (2009) 749–761.

[5] V. Nikiforov, The energy of graphs and matrices, *J. Math. Anal. Appl.* **326** (2007) 1472–1475.

[6] X. Li, Y. Shi, I. Gutman, *Graph Energy*, Springer, New York, 2012.

[7] H. S. Ramane, I. B. Baidari, R. B. Jummannaver, V. V. Manjalapur, G. A. Gudodagi, A. S.Yalnaik, A. S. Hanagawadimath, Hamming index of graphs with respect to its incidence matrix, *Indonesian J. Combin.* **6** (2022) 120–129.

[8] H. S. Ramane, A. B. Ganagi, Hamming index of class of graphs. *Int. J. Curr. Engg. Tech* (2013) 205–208.

[9] J. C. Dearden, The use of topological indices in QSAR and QSPR modeling, in: K. Roy (Ed.), *Advances in QSAR Modeling. Challenges and Advances in Computational Chemistry and Physics*, Springer, Cham, 2017, pp. 57–88.

[10] I. Redžepović, Chemical applicability of Sombor indices, *J. Serb. Chem. Soc.* **86** (2021) 445–457.

[11] I. Redžepović, B. Furtula, Predictive potential of eigenvalue-based topological molecular descriptors *J. Comput. Aided Mol. Des.* **34** (2020) 975–982.

[12] R. S. Varga, *Geršgorin and His Circles*, Springer, Kent 2011.

[13] S. Wagner, H. Wang, *Introduction to Chemical Graph Theory*, Chapman & Hall/CRC, New York, 2018.