Average Distances in Square-Cell Configurations

IVAN GUTMAN,1 SANDI KLAZAR,2 AMAL RAJAPAKSE3,*
1Faculty of Science, University of Kragujevac, P.O. Box 60, 34000 Kragujevac, Yugoslavia
2Department of Mathematics, PEF, University of Maribor, Koroska 160, 2000 Maribor, Slovenia
3Department of Mathematics and Applied Geometry, Institute für Angewandte Mathematik, Montanuniversität Leoben, A-8700 Leoben, Austria

Received 10 November 1998; revised 19 June 1999; accepted 31 August 1999

ABSTRACT: A square-cell configuration ("square animal") is a subgraph of the square lattice in which all inner faces are 4-cycles. We determine explicit expressions for the sum (W) of the (topological) distances between all pairs of vertices of a square-cell configuration, as well as for the related average distance W. Such expressions are deduced for several families of symmetric square-cell configurations. For instance, if O(n) stands for the octagonal square-cell configuration with n circular levels, then
\[ W(O(n)) = \left(\frac{211}{5}\right)n^3 - \left(\frac{181}{3}\right)n^2 + \left(\frac{109}{3}\right)n^2 - \left(\frac{35}{3}\right)n^2 + \left(\frac{22}{15}\right)n \]
and
\[ W(O(n)) = 2\left(7n^2 - 10n + 4\right)\left(7n^2 - 10n + 3\right)\left(7n^2 - 10n + 3\right)^{-1}W(O(n)). \]

Key words: square-cell configuration; square lattice; distance (in square-cell configuration); lattice animals; square animals

Introduction

The square lattice and its finite subgraphs (defined below and called square-cell configurations) are frequently encountered in statistical physics, where they are used for construction of mathematical theories—so-called lattice models—of a variety of phenomena, such as magnetization, phase transition, random walks, percolation, fractal growth, surface phenomena, and characterization (see, for instance, [1–11] and the references cited therein). Therefore, the properties of the square lattice and the square-cell configurations have attracted considerable attention. In this work, we examine one such property, namely, the average topological distance between the vertices of a square-cell configuration. This quantity is a convenient and intuitively plausible measure of the compactness of the respective square-cell configuration and may be useful whenever random walks on it are considered (see, in particular, section 5.4 in [2]).

Let S be a square-cell configuration, and V(S), its vertex set. Let |V(S)| = N. The distance d(x,y|S) between two vertices x and y of S is equal to the smallest number of steps in which one can get...
from \(x\) to \(y\) (or vice versa). The sum of all such distances will be denoted by \(W = W(S)\):

\[
W(S) = \sum_{\{x,y\} \subseteq V(S)} d(x,y(S))
\]

and the average (or mean) distance in \(S\) by \(\bar{W} = \frac{W(S)}{\binom{N}{2}} = \frac{2W(S)}{N(N - 1)}\). (1)

At this point, it is worth mentioning that \(W\) is just the same as the Wiener index (for a recent review, see [12, 13]), a structure-descriptor defined on molecular graphs, the first time used 50 years ago by Wiener for predicting physicochemical properties of alkanes [14]. Mathematicians have independently examined the same quantity, calling it graph distance [15] or graph admittance [16], but were not concerned with square-cell configurations. The closely related average distance of a graph has also been much studied in the mathematical literature [17–19]. To the authors’ knowledge, Peter John’s article [20] is the only published work dealing with the calculation of \(W\) of square-cell configurations.

**Square-cell Configurations: Definition and Basic Properties**

A complete grid (also called mesh [11, 21]) is the Cartesian product of two paths. It can also be described as a rectangular (sub)lattice. A grid graph is a subgraph of a complete grid. A square-cell configuration is a grid graph having the property that, when embedded in the plane, all its inner faces are 4-cycles. In this article, we concentrate on some families of symmetric and compact square-cell configurations.

Recall that the names square animal or square lattice animal [3, 11, 21, 22] and polyomino [23] are synonymous to what here is called a square-cell configuration.

It has been shown in [24] that any grid graph is a median graph and thus a partial binary Hamming graph as well. For the purpose of this work, it is not necessary to specify median and Hamming graphs, and we refer to [25] for definitions and additional information.

So, square-cell configurations are partial binary Hamming graphs. For these graphs, an important relation \(\Theta\) was introduced by Djoković [26] (see also [25, 27–29]). Let \(G\) be a connected graph with vertex set \(V(G)\) and edge set \(E(G)\). If \(e = xy \in E(G)\) and \(f = uv \in E(G)\), then \(e \Theta f\) holds if \(d(x,u) + d(y,v) \neq d(x,v) + d(y,u)\). The relation \(\Theta\) is reflexive and symmetric, yet it needs not be transitive. We denote its transitive closure by \(\Theta^+\) and call the equivalence classes of \(\Theta^+\) the cuts of \(G\). Hence, the cuts of \(G\) are pairwise disjoint subsets of \(E(G)\). Winkler [27] proved that a connected graph is a partial binary Hamming graph if and only if it is bipartite and \(\Theta^+ = \emptyset\).

As we already mentioned, our square-cell configurations are embedded into the plane (in the natural way). From this point of view, there are two kinds of cuts: those embracing horizontal edges and those with vertical edges. We call the cuts with vertical edges \(C_1\)-cuts, whereas those with horizontal edges will be referred to as \(C_2\)-cuts. In Figure 1, a square-cell configuration is shown possessing five \(C_1\)-cuts and seven \(C_2\)-cuts. The five \(C_1\)-cuts are indicated by parallel lines.

If \(G\) is a partial binary Hamming graph, then it is well known that the graph obtained from \(G\) by removing all edges of an arbitrary cut has exactly two connected components. With this in mind, we formulate the following proposition from [30]:

**Proposition 1.** Let \(G\) be a partial binary Hamming graph on \(N\) vertices and with \(k\) cuts. For \(i = 1, 2, \ldots, k\), let \(N_i\) be the number of vertices of \(G\) in one of the components of the graph obtained from \(G\) by removing the \(i\)th cut. Then,

\[
W(G) = \sum_{i=1}^{k} N_i(N - N_i).
\]

Proposition 1 provides a simple method for the calculation of the sum of distances and average distances of partial binary Hamming graphs, which is particularly suitable for the chemically very important class of hexagonal systems. This was elaborated in more detail in [31] and several explicit expressions are given in [32].

**FIGURE 1.** A square-cell configuration and its \(C_1\)-cuts.
The main purpose of this article was to demonstrate that this approach is also quite simple to apply to other partial binary Hamming graphs, for instance, to square-cell configurations. We show this by computing general expressions for $W$ for several families of these graphs. We think that for additional square-cell configurations or other partial binary Hamming graphs of interest to the readers it will not be difficult to obtain the corresponding formulas for $W$ along the same lines.

In what follows, $N$ will denote the number of vertices of the square-cell configuration considered. In addition, expressions of the form

$$\sum_i \left[(i(2n+2) - 1)(N - (i(2n+2) - 1))\right]$$

will be written as

$$\sum_i \left[(i(2n+2) - 1)(N - i)\right],$$

that is, $F$ will stand for the value of the “first bracket.” Moreover, for a square-cell configuration $S$, we will use $W_1$ and $W_2$ to denote the partial sums from Proposition 1, corresponding to the $C_1$-cuts and $C_2$-cuts, respectively. Hence, $W(S) = W_1 + W_2$.

**Rectangulars, Octagons, and Hexagons**

We begin our computations with three relatively simple representatives of square-cell configurations: rectangular square-cell configurations, octagonal square-cell configurations, and hexagonal square-cell configurations.

**Rectangulars**

For $n \geq 1$ and $1 \leq k \leq n$, let $R(n,k)$ be the rectangular square-cell configuration. The definition of $R(n,k)$ should be clear from the example $R(8,5)$ shown in Figure 2.

For $R(n,k)$, we have

$$N = (n+1)(k+1)$$

$$W_1 = \sum_{i=1}^{k} \left[(i(n+1))(N - i)\right]$$

$$W_2 = \sum_{i=1}^{k} \left[(i(k+1))(N - i)\right].$$

The sum of distances of $R(n,k)$ is equal to $W_1 + W_2$, and simplifying the expression, we obtain

**Proposition 2.** For any $n \geq 1$ and $1 \leq k \leq n$, we have

$$W(R(n,k)) = \frac{1}{6}(k+1)(n+2)(n+1)[(n+1) + n]$$

$$\overline{W}(R(n,k)) = \frac{1}{3}(k+n+2).$$

By setting $n = r-1$ and $k = s-1$, where $r \geq 2$, $s \geq 2$, the above formulas reduce to

$$W(R(r-1,s-1)) = \frac{1}{6}(rs)(r+s)(rs-1)$$

$$\overline{W}(R(r-1,s-1)) = \frac{1}{3}(r+s).$$

These are the well-known formulas for the Wiener number of the Cartesian product of two paths on $r$ and $s$ vertices, respectively; cf. [33-35].

**Octagons**

For $n \geq 2$, let $O(n)$ be the octagonal square-cell configuration. For instance, the octagonal square-cell configuration $O(3)$ is depicted in Figure 3.
GUTMAN, KLAČAR, AND RAJAPAKSE

For $O(n)$, we have

$$N = 7n^2 - 4n + 1.$$  

For the first $n$ and the last $n$ cuts of $C_1$, we have

$$W_1 = 2 \sum_{i=1}^{n} \left[ (i(n + i))(N - \mathcal{F}) \right],$$

whereas for the middle $n - 2$ cuts from $C_1$,

$$W_2 = \sum_{i=1}^{n-2} \left[ (2n^2 + (3n - 1))(N - \mathcal{F}) \right].$$

Clearly, $W_1 = W_1' + W_1''$, and, by symmetry, $W_2 = W_1$. Simplifying the expression $2(W_1' + W_1'')$, we arrive at

**Proposition 3.** For any $n \geq 2$,

$$W(O(n)) = \frac{211n^5}{5} + \frac{181n^4}{3} + \frac{109n^3}{3} - \frac{35n^2}{3} + \frac{22n}{15},$$

$$\overline{W}(O(n)) = \frac{2W(O(n))}{(7n^2 - 4n + 1)(7n^2 - 4n)}.$$

**HEXAGONS**

For $n \geq 2$, let $H(n)$ be the hexagonal square-cell configuration. The definition of $H(n)$ should be clear from the example $H(4)$ which is shown in Figure 4.

For $H(n)$, we have $N = 4n^2$. For the first $(n - 1)$ and the last $(n - 1)$ cuts from $C_1$, we have

$$W_1 = 2 \sum_{i=1}^{n-1} \left[ (i(i + 1))(N - \mathcal{F}) \right],$$

whereas for the middle $n$ cuts from $C_1$,

$$W_1' = 2 \sum_{i=1}^{n} \left[ (n(n-1) + i(2n))(N - \mathcal{F}) \right].$$

Clearly, $W_1 = W_1' + W_1''$. Now, for the first $(n - 1)$ and the last $(n - 1)$ cuts from $C_2$, we have

$$W_2 = \sum_{i=1}^{n-1} \left[ (i(n + i))(N - \mathcal{F}) \right].$$

Since $W_2 = W_2' + (N/2)^2$, we finally get

**Proposition 4.** For any $n \geq 2$,

$$W(H(n)) = \frac{1}{15} \frac{158n^5}{N} - \frac{7n^3}{3} - \frac{n}{5},$$

$$\overline{W}(H(n)) = \frac{1}{30} \frac{79n^3}{N} - \frac{7n}{12} - \frac{1}{20n}.$$

**Rectangular Trapeziums and Trapeziums**

In this section, we obtain the formulas for the average distance number of certain slightly more involved square-cell configurations: rectangular trapeziums and trapeziums. The same results can also be obtained as special cases of the bitrapeziums, which are outlined in the subsequent section. We, nevertheless, deem that it is worth considering these special cases separately, in order to demonstrate how the computations are quite simple.

In what follows, we give only expressions for $W(S)$ and $N = N(S)$ and skip the respective formula for $\overline{W}(S)$ The latter is readily obtained using Eq. (1).

**RECTANGULAR TRAPEZIUMS**

Let $n \geq 1$. Then, for $1 \leq k \leq n$, let $RT(n,k)$ denote the rectangular trapezium square-cell configuration. For instance, $RT(6,6)$ is shown in Figure 5.

For $RT(n,k)$, we have

$$N = (k + 1)(n - k + 1) + \frac{k}{2}(k + 3).$$

For $C_1$-cuts (from the bottom), we have

$$W_1 = \sum_{i=1}^{k} \left[ \left( i(n + 1) - \frac{1}{2}(i - 1)(i - 2) \right)(N - \mathcal{F}) \right].$$

Now, for the first $n-k$ cuts (from the left) from $C_2$, we have

$$W_2 = \sum_{i=1}^{n-k} \left[ (i(k + 1))(N - \mathcal{F}) \right].$$
whereas for the last $k$ cuts from $C_2$, 
\[ W_1' = \sum_{i=1}^{k} \left[ \left( \frac{1}{2}(i+3) \right)(N - \mathcal{F}) \right] \]

As $W_2 = W_3 + W_4$, simplifying the expression $W_1 + 2(W_3 + W_4)$, we obtain

**Proposition 5.** For any $n \geq 1$ and $1 \leq k \leq n$,

\[
W(\mathbf{RT}(n, k)) = \frac{n^2(k^2 + 2k + 1)}{6} - \frac{n^2(k + 1)(k^2 - 13k - 6)}{12} - \frac{n(k + 1)(5k^2 - 21k - 4)}{12} + \frac{k(k^4 + 5k^3 + 5k^2 - 35k - 36)}{60}. 
\]

For instance, for $k = 1$, the above formula reduces to

\[
W(\mathbf{RT}(n, 1)) = \frac{1}{3}(2n^3 + 9n^2 + 10n + 3). \tag{3} 
\]

Clearly, the latter formula can also be obtained from Eq. (2) by substituting $r = n + 1$ and $s = k + 1 = 2$.

Moreover, for $k = n$, the above expression for $W(\mathbf{RT}(n, k))$ reduces to the formula for the sum of distances of rectangular-triangle square-cell configurations. Thus, if for $n \geq 1$, $T(n)$ denotes the rectangular-triangle square-cell configuration, then we have

**Corollary 5.1.** For any $n \geq 1$,

\[
W(T(n)) = \frac{n(2n^4 + 25n^3 + 90n^2 + 95n + 28)}{30}. 
\]

**DISTANCES IN SQUARE-CELL CONFIGURATIONS**

**FIGURE 6.** The trapezium square-cell configuration $T(11, 5)$.

**TRAPEZIUMS**

For $n \geq 1$ and $1 \leq k \leq \lfloor n/2 \rfloor$, let $T(n, k)$ stand for the trapezium square-cell configuration. The definition of $T(n, k)$ should be clear from the example $T(11, 5)$, shown in Figure 6.

For $T(n, k)$, we have

\[
N = (n + 1)(k + 1) - k(k - 1) 
\]

\[
W_1 = \sum_{i=1}^{k} \left[ \frac{i(n + 1) - (i - 1)(i - 2)}{2}(N - \mathcal{F}) \right] 
\]

\[
W_2 = 2 \sum_{i=1}^{k} \left( \frac{i}{2}(i+3) \right)(N - \mathcal{F}) 
\]

\[
+ \sum_{i=1}^{n-k} \left( \frac{k}{2}(k+3) + i(k+1) \right)(N - \mathcal{F}). 
\]

As $W(T(n,k)) = W_1 + W_2$, by simplifying the above expressions, we arrive at

**Proposition 6.** For any $n \geq 1$ and $1 \leq k \leq \lfloor n/2 \rfloor$,

\[
W(T(n, k)) = \frac{n^3(k^2 + 2k + 1)}{6} - \frac{n^2(2k^3 - 6k^2 - 11k - 3)}{6} - \frac{n(3k^4 - 20k^3 + 19k^2 + 34k + 4)}{12} + \frac{k(4k^4 - 10k^3 + 35k^2 - 35k - 24)}{30}. 
\]

For instance, for $k = 1$, the above formula, as previously, reduces to Eq. (3).

**BITRAPEZIUMS**

In this section, we generalize Proposition 6 to bitrapeziums. A similar generalization is also possible for rectangular trapeziums from Proposition 5, but we leave it to the reader.
Denote a bitrapezium system as $BT(n, k_1, k_2)$, where $0 \leq k_1 \leq \lfloor n/2 \rfloor - 1$ and $0 \leq k_1 \leq \lfloor n/2 \rfloor - 1$. For instance, $BT(13, 3, 4)$ is shown in Figure 7.

For $BT(n, k_1, k_2)$, we have
$$N = k_1(n - k_1) + k_2(n - k_2) + 2(n + 1).$$

Let $N_1$ and $N_2$ denote the number of vertices of the upper trapezium, lying above the middle cut, and the number of vertices of the lower trapezium, lying below the middle cut, respectively. Then,
$$N_1 = k_1(n - k_1) + (n + 1)$$
and
$$N_2 = k_2(n - k_2) + (n + 1).$$

Then, for the first $k_1$ cuts from the set $C_1$, lying in the upper trapezium, we have
$$W'_1 = \sum_{i=0}^{k_1-1} \left[ (N_1 - (i + 1)(n + 1) + \sum_{j=0}^{i} 2j) (n - F) \right].$$

For the next $k_2$ cuts from the set $C_1$, lying in the lower trapezium,
$$W'_2 = \sum_{i=0}^{k_1-1} \left[ (N_2 - (i + 1)(n + 1) + \sum_{j=0}^{i} 2j) (n - F) \right].$$

Clearly, $W_1 = W'_1 + W'_2 + N_1 N_2$.

Now, for the first $k_1 + 1$ cuts from the set $C_2$ that are on the leftmost part of the trapezium (by the symmetry on the right most part), we have
$$W'_2 = \sum_{i=1}^{k_1+1} \left[ (i(i + 1))(n - F) \right].$$

Without loss of generality, it may be assumed that $k_1 \leq k_2$. Then, for the next $(k_2 - k_1)$ cuts from $C_2$ (starting from $k_1 + 2$ to $k_2 + 1$),
$$W_2 = \sum_{i=1}^{k_2-k_1} \left[ \frac{(k_1 + 1)(k_1 + 2)}{(k_1 + 1)(i + 1)} (n - F) \right].$$

It remains to consider the next $n - 2(k_1 + 1)$ cuts from the set of elementary cuts $C_2$. For these cuts,
$$W''_2 = \sum_{i=1}^{n - 2(k_1 + 1)} \left[ \frac{1}{2} (k_2 - k_1)(3k_1 + k_2 + 5) + \frac{1}{2} (k_2 - k_1)(3k_1 + k_2 + 5) \right].$$

By symmetry, $W_2 = 2W'_2 + 2W''_2 + W_2''$. Clearly, $W(BT(n, k_1, k_2)) = W_1 + W_2$ and by simplifying we get

\textbf{Proposition 7.} For any $n \geq 1$ and $0 \leq k_1 \leq \lfloor n/2 \rfloor - 1$, and $0 \leq k_1 \leq \lfloor n/2 \rfloor - 1$, where $k_1 \leq k_2$, the sum of distances of $BT(n, k_1, k_2)$ is equal to
$$\frac{n^3(k_1 + k_2 + 2)^2}{6} - \frac{n^2(k_1 + k_2 + 2)(k_1 + 1)(k_1 + 2) + 2k_1^2 - 17k_2 - 18}{6} - \frac{n^2k_1^3 - k_1(12k_2 + 17) + 2k_1^2 - 17k_2 - 18}{6} + \frac{3k_1^4 - 4k_1^2(k_2 + 2) + k_1^2(6k_2 + 24k_2 + 23)}{12} - \frac{2k_1(2k_2 + 12k_2^2 + k_2 - 12)}{12} + \frac{3k_2^4 - 8k_2^2 - 23k_2^2 + 24k_2 + 40}{12} - \frac{5k_1^2 + 10k_1^2 + 10k_2^2(3 - k_2)}{30} + \frac{5k_1^2(4k_2^2 + 12k_2^2 + 3 - k_2 - 10) - 5k_1^2(2k_2 - 1 - 1)}{30} - \frac{4k_2^2 + 10k_2^2 + 35k_2^2 + 50k_2^2 - 9k_2 - 30}{30}.$$

Trapeziums can be described as $T(n, k) = BT(n, 0, k - 1)$. Hence, we can use the formula for $W$ of bitrapezium systems obtained above, in order to obtain the expression for the trapeziums from Proposition 6:
$$W(BT(n, 0, k - 1)) = \frac{n^2(k_1^2 + 2k_1 + 1)}{6} - \frac{n^2(2k_1^2 - 6k_2^2 - 11k_2 - 3)}{6}.$$
\[
\frac{n(3k^4 - 20k^3 + 19k^2 + 34k + 4)}{12k(4k^4 - 10k^3 + 35k^2 - 35k - 24)}
\]

Notice that due to our assumption \(k_1 \leq k_2\) we may only set \(k_1 = 0, k_2 = k - 1 (k \geq 1)\) in the above expression to get the desired equation for \(W\) of the trapeziums. If we also set \(k_1 = 0\), then the above formula reduces to the previous Eq. (3).

In conclusion, we mention two additional special cases: If \(n\) is even, then for \(k_1 = k_2 = n/2 - 1\), we obtain the formula for even symmetric bitrapeziums, that is,

**Corollary 7.1.** For any even \(n \geq 2\),

\[
W(\beta T(n; n/2 - 1, n/2 - 1)) = \frac{n(7n^2 + 70n^3 + 190n^2 + 80n - 92)}{120}.
\]

Finally, if \(n\) is odd, we set \(k_1 = k_2 = (n - 1)/2\) to obtain \(W\) of odd symmetric bitrapeziums:

**Corollary 7.2.** For any odd \(n \geq 1\),

\[
V(\beta T(n; (n - 1)/2, (n - 1)/2)) = \frac{(7n^2 + 70n^3 + 250n^2 + 380n^2 + 223n + 30)}{120}.
\]

**References**

Hill, T. L. Statistical Thermodynamics; Addison-Wesley, Reading, M.A. 1960.


16. Plesnik, J. J Graph Theory 1984, 8, 1.


