Some Improvements on Low-Discrepancy Matrices

Akitoshi KAWAMURA
Department of Computer Science, University of Tokyo

August 12, 2005*

We consider Aronov et al.'s *dither matrix problem*: put n^2 consecutive numbers on an $n \times n$ board in as "balanced" a way as possible in the sense that the numbers in each $k \times k$ region on the board add up to approximately the same sum. We construct a fully balanced configuration for some cases, completing the investigation to identify all (n, k) for which this is possible. For some other cases, we give a heuristic method attaining lower discrepancy than existing techniques. This method obtains the entries of the desired matrix by ranking the values of a continuous function at grid points.

1 The Problem

Fill an $n \times n$ board with the first n^2 nonnegative integers, and balance the sum of numbers in each $k \times k$ region on the board, where the edges of the board are linked to the opposite side (Figure 1). This optimization problem was posed by Aronov et al. [1] with the aim of giving good dither matrices for digital halftoning, a technique to convert a continuous-tone image to a binary image for printing or display on binary devices.

Let us state the problem more precisely. Fix a positive integer n. By a *configuration* we mean an $n \times n$ matrix in which each of the first n^2 nonnegative integers appears exactly once. Any subset $R \subseteq [n]^2$ is called a *region*,

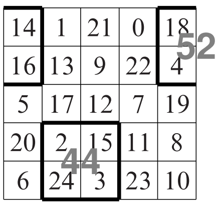

Figure 1: An example with discrepancy 8 ($n = 5$, $k = 2$)

*Presented at the Eighth Korea-Japan Joint Workshop on Algorithms and Computation (WAAC 2005).

where $[n]$ denotes the set $\{0, 1, \dots, n-1\}$. Its translation by $(a, b) \in \mathbf{Z}^2$ is written

$$(1) \quad R + (a, b) = \big\{ \big((i + a) \bmod n, (j + b) \bmod n\big) \bigm| (i, j) \in R \big\},$$

where $x \bmod n$ denotes the number $x' \in [n]$ for which $x - x'$ is a multiple of n. The set of all translations of R is denoted by $\overline{R} = \big\{ R + (a, b) \bigm| (a, b) \in [n]^2 \big\}$.

To avoid too many subscripts, we denote the $(i, j)^{\text{th}}$ entry of an integer matrix M by $M(i, j)$ (rather than $m_{i,j}$), regarding a matrix as a function on pairs of indices. We denote the sum of the numbers in a region R by

$$(2) \qquad M(R) = \sum_{(i,j) \in R} M(i, j).$$

The *discrepancy* of M with respect to a class $\mathscr{R}$ of regions is defined to be the difference between the maximum and minimum of $M(R)$ as R varies in $\mathscr{R}$. Our goal is to find a configuration with small discrepancy. In particular, we are interested in discrepancy wrt $\overline{[k]^2}$, the class of square regions of fixed size $k \in [n] \setminus \{0\}$. We reserve the symbol k for this purpose, and simply say discrepancy to mean discrepancy wrt $\overline{[k]^2}$.

This paper reports two contributions to this problem. In Section 2 we identify all (n, k) for which zero-discrepancy configurations exist, completing Aronov et al.'s investigation [1]. Section 3 gives a heuristic method to find a low-discrepancy configuration for $k = 2$, refuting Asano et al.'s conjecture to the contrary [2].

2 Zero Discrepancy

Along the lines of the previous works [1], we determine all (n, k) admitting a zero-discrepancy configuration.

0	1	2	3	4
7	8	9	5	6
14	12	10	13	11

$M_{3,5}$

0	1	2	3	4
9	8	7	6	5

$M_{2,5}$

$\Longrightarrow$

0	1	2	3	4
7	8	9	5	6
14	12	10	13	11
15	16	17	18	19
24	23	22	21	20

$M_{5,5}$

Figure 2: Building $M_{5,5}$ from $M_{3,5}$ and $M_{2,5}$, incrementing the latter by 15

2.1 Cases Defying Zero Discrepancy

Although the following Lemmata 1 and 2 have been known [1, Theorem 1], we restate them with proofs for completeness. A pair (a, b) of integers is said to be *coprime* if no integer greater than 1 divides both a and b.

Lemma 1. *If k is odd and n is even, every configuration has positive discrepancy.*

Proof. For $M\big([k]^2 + (a, b)\big)$ to be constant, it should be k^2 times the average of all entries $0, \dots, n^2 - 1$. But this is not an integer if k is odd and n is even. □

Lemma 2. *If (n, k) is coprime, every configuration has positive discrepancy.*

Proof. Assume that M is a zero-discrepancy configuration. Then for each $a \in [n]$ and $b \in [n]$ we have

$$M\big([1] \times [k] + (a, b)\big) = M\big([1] \times [k] + (a + k, b)\big), \tag{3}$$

since these k-ominos are the difference between two square regions $[k]^2 + (a, b)$ and $[k]^2 + (a + 1, b)$. From this it follows, by coprimality of (n, k), that $M\big([1]\times[k]+(a, b)\big)$ does not depend on a. In fact, it does not depend on b either, since

$$\sum_{a\in[n]} M\big([1]\times[k]+(a, b)\big) = \frac{1}{k} \sum_{a\in[n]} M\big([k]^2+(a, b)\big) \tag{4}$$

is constant. Hence $M(a, b) = M(a, b+k)$, because these locations are the difference between two k-ominos. □

2.2 Cases Admitting Zero Discrepancy

We prove that there is a zero-discrepancy configuration in all cases not shown otherwise above. We begin with a lemma providing a building block for the construction:

Lemma 3. *Let $k > 1$ and $u > 0$ be integers. If either k is even or u is odd, then there exists a $k \times u$ matrix $M_{k,u}$ in which each number in $[ku]$ appears exactly once and each column has the same sum.*

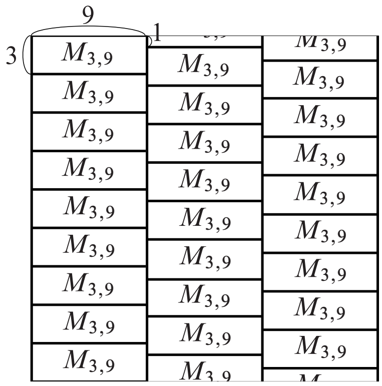

Figure 3: Construction of M^- for $k = 3, u = 9, s = 1$

Proof. If such matrices $M_{k,u}$ and $M_{k',u}$ are given, we can get a required matrix $M_{k+k',u}$ of larger size by

$$M_{k+k',u}(i, j) = \begin{cases} M_{k,u}(i, j) & \text{if } i < k \\ M_{k',u}(i - k, j) + ku & \text{if } i \geq k \end{cases} \tag{5}$$

as illustrated in Figure 2. Hence, we only need to prove the statement for $k = 2, 3$. This is done by defining

$$M_{2,u}(i, j) = \begin{cases} j & \text{if } i = 0 \\ 2u - 1 - j & \text{if } i = 1 \end{cases} \tag{6}$$

for any u and

$$M_{3,u}(i, j) = \begin{cases} j & \text{if } i = 0 \\ 3u' + 1 + j & \text{if } i = 1 \text{ and } j \leq u' \\ u' + j & \text{if } i = 1 \text{ and } j > u' \\ 6u' + 2 - 2j & \text{if } i = 2 \text{ and } j \leq u' \\ 8u' + 3 - 2j & \text{if } i = 2 \text{ and } j > u' \end{cases} \tag{7}$$

for odd $u = 2u' + 1$. □

Now we state the main result:

Theorem 4. *There exists a configuration with discrepancy zero if and only if (n, k) is not coprime and either k is even or n is odd.*

Proof. One direction is shown by Lemmata 1 and 2. We prove the other direction by constructing a zero-discrepancy configuration M. Note that a matrix has discrepancy zero if it does so with k replaced by any positive divisor of k. Hence, we may assume that k is a prime and $n = ku$ for some positive integer u.

We construct M using $M_{k,u}$ from Lemma 3. Let s be an integer for which k does not divide $u^2 - s^2$. For example, put $s = 1$ if k divides u, and $s = 0$ otherwise. Let M^- be the $n \times n$ matrix made up of ku copies of $M_{k,u}$, aligned vertically and descending by s towards the right (Figure 3). Thus,

$$(8)\quad M^-(a, ub_0 + b_1) = M_{k,u}\big((a - sb_0) \bmod k, b_1\big)$$

for $a \in [n]$, $b_0 \in [k]$ and $b_1 \in [u]$. We claim that the desired configuration M is obtained by

$$(9)\qquad M(a, b) = nM^-(a, b) + M^-(b, a).$$

Since M^- has discrepancy zero by the property of $M_{k,u}$, so does M. It remains to show that M is injective. Suppose

$$(10)\quad M(ua_0 + a_1, ub_0 + b_1) = M(ua'_0 + a'_1, ub'_0 + b'_1)$$

for some $a_0, b_0, a'_0, b'_0 \in [k]$ and $a_1, b_1, a'_1, b'_1 \in [u]$. By (8), (9) and bijectivity of $M_{k,u}$, this means that

$$(11)\qquad \begin{cases} \big((ua_0 + a_1 - sb_0) \bmod k, b_1\big) \\ \qquad = \big((ua'_0 + a'_1 - sb'_0) \bmod k, b'_1\big), \\ \big((ub_0 + b_1 - sa_0) \bmod k, a_1\big) \\ \qquad = \big((ub'_0 + b'_1 - sa'_0) \bmod k, a'_1\big). \end{cases}$$

We immediately have $(a_1, b_1) = (a'_1, b'_1)$. This and (11) yield $ua_0 - sb_0 \equiv ua'_0 - sb'_0 \pmod k$ and $ub_0 - sa_0 \equiv ub'_0 - sa'_0 \pmod k$. Eliminating b_0 and b'_0 we get $(u^2 - s^2)a_0 \equiv (u^2 - s^2)a'_0 \pmod k$, whence $a_0 = a'_0$ because $u^2 - s^2$ does not divide k. Similarly, $b_0 = b'_0$. We have shown $(a_0, a_1, b_0, b_1) = (a'_0, a'_1, b'_0, b'_1)$, as desired. □

The general idea of the above construction was as follows. Let M_0 and M_1 be $n \times n$ matrices containing each of the first n nonnegative integers exactly n times. If they have discrepancy zero, so does $M = nM_0 + M_1$. For M to be a configuration, M_0 and M_1 must be *mutually orthogonal* in the sense that the pairs $\big(M_0(a, b), M_1(a, b)\big)$ are all distinct for $a, b \in [n]$. To meet this requirement, we used as M_0 and M_1 the matrix M^- and its transposition. A similar technique has been well-known for discrepancy wrt $\overline{[1] \times [n]} \cup \overline{[n] \times [1]}$, in which case a zero-discrepancy configuration is called a *semimagic square* and M_0 and M_1 are called *Latin squares*.

3 Real Functions Help Find Low-Discrepancy Matrices

In this section we confine ourselves to the case $k = 2$. Theorem 4 tells us that in this case we have zero discrepancy if and only if n is even. For odd n's, Asano et al. constructed a configuration with discrepancy $2n$ using the idea of mutually orthogonal matrices mentioned above, and conjectured that this was optimal [2]. In this section we give a heuristic method attaining smaller discrepancy.

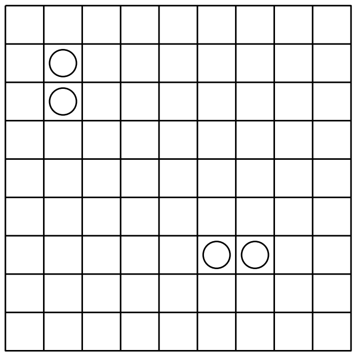

Figure 4: A region in $\overline{t([2]^2)}$ when $n = 9$.

3.1 Our Technique

Assume that n is odd. Define $f : [0, 1]^2 \to \mathbf{R}$ by

$$(12)\quad f(x, y) = g(x) + g(y)$$
$$\text{where}\quad g(x) = \begin{cases} 1 - |4x - 1|^2 & \text{if } x \le \frac{1}{2} \\ -1 + |4x - 3|^2 & \text{if } x \ge \frac{1}{2} \end{cases}$$

and a function $z : [n]^2 \to [0, 1]^2$ by

$$(13)\qquad z(i, j) = \left(\frac{i}{n} + \frac{1}{6n}, \frac{j}{n}\right).$$

Let $\hat{M}$ be the configuration such that the $\hat{M}(i, j)^{\text{th}}$ smallest entry in the real matrix $f \circ z$ (under some tie-breaking rule) is at (i, j):

$$(14)\quad \hat{M}(i, j) = \sharp\big\{ (i', j') \in [n]^2 \;\big|\; f(z(i', j')) < f(z(i, j)) \text{ or } \big(f(z(i', j')) = f(z(i, j)) \text{ and } ni' + j' < ni + j \big) \big\}.$$

We obtain the desired configuration by $M = \hat{M} \circ t$, where $t : [n]^2 \to [n]^2$ is the bijection

$$(15)\qquad t\big((i + j) \bmod n, (i - j) \bmod n\big) = (i, j).$$

Experiments show that this configuration has smaller discrepancy than the existing lower bound $2n$ (Section 3.3).

3.2 The Idea

The intuitive ideas behind the above formula are as follows. The transformation t in (15) means that we seek, instead of M, a configuration $\hat{M}$ whose discrepancy wrt

$$(16)\qquad t\big([2]^2\big) = \big\{ t(a, b) \;\big|\; (a, b) \in [2]^2 \big\}$$

is small (Figure 4). Thus, we want to average the sum

$$(17)\quad \hat{M}\big(t[2]^2 + (a, b)\big) = \hat{M}(a, b) + \hat{M}(a + 1, b) + \hat{M}\left(a + \frac{n+1}{2}, b + \frac{n-1}{2}\right) + \hat{M}\left(a + \frac{n+1}{2}, b + \frac{n+1}{2}\right)$$

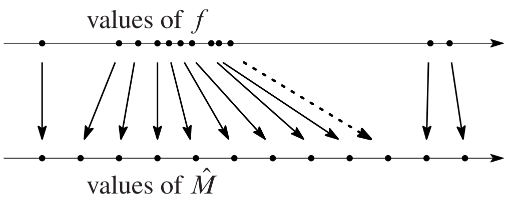

Figure 5: Congestion of values results in undeserved ranking (dotted line)

(indices modulo n). With $f \circ z$ replacing $\hat{M}$, this equals

$$(18) \quad f\left(x - \frac{1}{2n}, y\right) + f\left(x + \frac{1}{2n}, y\right) + f\left(x + \frac{1}{2}, y + \frac{1}{2} - \frac{1}{2n}\right) + f\left(x + \frac{1}{2}, y + \frac{1}{2} + \frac{1}{2n}\right)$$

with $x = a/n + 1/(2n) + 1/(6n)$ and $y = b/n$, where arguments of f are read modulo 1. Instead of averaging (17), we look for a function f that makes (18) constant as (x, y) varies, and obtain $\hat{M}$ as the rank, as defined in (14), of values of f at the grid points given by (13). This simplification of course relies on the following (informally stated) assumption:

Assumption 1 Values of f at grid points are "uniformly distributed" over an interval.

To see what this means, note that ranking may assign $\hat{M}(i, j)$ an "undeservedly big (or small)" value when the corresponding value of f is barely above (or below) a crowd of rivals (Figure 5). Assumption 1 says that this occurs so rarely that replacing $\hat{M}$ by $f \circ z$ does little harm.

The function f in (12) is designed so that

$$(19) \qquad f(x, y) = -f\left(x + \frac{1}{2}, y + \frac{1}{2}\right)$$

for each $x, y \in [0, 1]$. In replacing (18) by (19) we are making another assumption:

Assumption 2 The function f is so flat that the terms $1/(2n)$ in (18) are insignificant.

3.3 Experiment

As long as f satisfies (19), the discrepancy is attributed to the extent to which Assumptions 1 and 2 are violated. We experimented on several choices of such f, including

$$(20) \quad f(x, y) = \sin(2\pi x) + \sin(2\pi y)$$

$$(21) \quad f(x, y) = \begin{cases} |2x - 1| + |2y - 1| - 1 \\ \qquad \text{if } |2x - 1| + |2y - 1| \leq 1 \\ \text{defind by (19)} \quad \text{otherwise} \end{cases}$$

Table 1: Discrepancy attained by different choices of f and z

z from (13) (with shift)			
n \ f	(12)	(20)	(21)
5	14	15	20
7	16	17	27
11	17	21	37
15	21	24	70
31	28	37	122
51	32	42	136
101	38	60	259
151	50	76	784
201	52	83	1052

f from (12)		
n \ z	with shift	without shift
5	14	14
7	16	19
11	17	26
15	21	34
31	28	66
51	32	106
101	38	206
151	50	306
201	52	406

in place of (12). We find that (12) performs best, beating Asano et al.'s construction when $n > 7$ (Table 1, left). The poor performance of (21) may be ascribed to the violation of Assumption 2: the graph of (21) has some creases, on which the $1/(2n)$ shifts in (18) tend to take much effect. We also tried replacing (13) by $z(i, j) = (i/n, j/n)$, deleting the $1/(6n)$ shift (Table 1, right). The reason this works poorly may be explained by Assumption 1: without the shift, f would vanish at many grid points, giving some of them an undeserved rank through tie-breaking.

However, we have not succeeded in making these intuitive arguments rigorous. Although it is tempting to conjecture from Table 1 that there always exist $o(n)$-discrepancy configurations, we do not as yet have a proof.

References

[1] B. Aronov, T. Asano, Y. Kikuchi, S. Nandy, S. Sasahara, and T. Uno. A generalization of magic squares with applications to digital halftoning. In *Proceedings of ISAAC 2004*, LNCS **3341**, 89–100. Springer.

[2] T. Asano, S. Choe, S. Hashima, Y. Kikuchi, and S.-C. Sung. Distributing distinct integers uniformly over a square matrix with application to digital halftoning. IPSJ Technical Reports 26 (AL-100), March 2005.