Index assignment for two-channel quantization

József Balogh, János A. Csirik

Abstract

This paper concerns the design of a multiple description scalar quantization system for two symmetric channels for an unbounded discrete information source. This translates to the combinatorial problem of finding an arrangement of the integers into the infinite plane square grid so that each row and each column contains exactly N numbers, such that the difference between any two numbers in the same row (or column) is at most d, with d to be minimized for a given N. The best previous bounds for the lowest d were $N^2/3+O(N)$ and $N^2/2+O(N)$. We give new lower and upper bounds, both of the form $3N^2/8+O(N)$. We also consider minimizing the maximal variance in any row or column and show that it must be at least $0.0167N^4$, and that it does not have to be more than $0.0188N^4$.

I. Introduction

A *diversity system* provides several different channels for transmitting information from the source to the user. This way, if a channel breaks down, an alternate path is available between the source and the user. Consider a diversity system with two channels. If identical information is sent over each channel and if both channels work, half of the received information is worthless. We consider sending different information over each channel in such a way that if only one channel works the information received over it is sufficient to achieve a minimum fidelity. On the other hand, should both channels work, the information from one channel can be used to augment the information from the other channel to achieve a higher fidelity than would each channel alone. The problem of designing codes of this kind is known as the multiple descriptions problem ([1]) and is a generalization of the problem of source coding subject to a fidelity criterion ([2]).

A multiple description scalar quantizer (MDSQ) is a scalar quantizer designed for operation in a diversity system. The encoder of an MDSQ sends information over each channel of the diversity system subject to a rate constraint. The decoder reconstructs the source sample based on the information received from the channels that are currently working. The objective is to design a decoder-encoder pair that minimizes the distortion when both channels work, subject to constraints on the distortion when only one channel works (either channel may break down). Thus, in the event that exactly one of the channels is broken, a minimum fidelity is guaranteed.

Applications of multiple description source codes arise in speech and video coding over packet-switched networks, where packet losses can result in a degradation in signal quality and there are significant delay constraints. For details, and more examples, see [3], [4].

The design of a MDSQ system involves two steps: quantization of the source, and the index assignment problem of distributing the quantized signal over multiple channels. In this paper, we will only consider the index assignment problem. That is, we will consider MDSQ design for an unbounded discrete information source. We consider the case of two symmetric channels, where the rate of both channels is equal and the *maximum distortion* of either channel is to be minimized (this corresponds to wanting the smallest possible distortion if either channel fails). This translates to the combinatorial problem of finding an arrangement of the integers into the infinite plane square grid so that each row and each column contains exactly N numbers. In this setup, the integers written into the grid represent symbols emitted by the source, and the coordinates of a number are what is sent on the two channels for that source symbol. The parameter N serves to control the amount ($\log_2 N$) by which the rate of each channel is less than that of the source.

If only one of the channels works then we only know that the source symbol is one of the N possibilities in the row or column selected by the working channel. Let the maximal occuring difference between any two numbers in the same row (or column) be d. Then minimizing the absolute distortion given the rate amounts to finding, for a given N, an arrangement with the smallest possible *spread* d. This problem was considered in [4] and in [5], where it was proved that the best d is at least $N^2/3+O(N)$ and at most $N^2/2+O(N)$ (these are the best known results not contained in this paper, as far as we know). In this paper we show that the best d is in fact $3N^2/8+O(N)$. Specifically, in Sections III-B and III-C we show that

Theorem I.1: Assume that the integers have been arranged within a plane square grid, with each row and column containing exactly N numbers, and the difference between any two numbers in the same row (or column) is at most d. Then

$$3N^2/8-1/2 \leq d.$$

For even N, this bound is optimal, that is, an arrangement with $d=\lceil 3N^2/8-1/2\rceil$ exists. For odd N, an arrangement exists with $d=3N^2/8+N/4+O(1)$.

JB is at the Institute for Advanced Study in Princeton, New Jersey, USA. JAC is at AT&T Labs–Research in Florham Park, New Jersey, USA. This work was carried out while both authors were at AT&T Labs–Research.

Both the upper bound (construction using large square blocks) and the lower bound (proof using a particular combination of local and global methods) of this paper contain new ideas that will extend to the case of asymmetric channels (with different rates), and to more than two channels.

The index assignment problem for MDSQ also has connections to the theory of graph bandwidths. For more details, see [6].

Later in the paper, we consider the MDSQ problem where we aim to reduce the *mean squared error* of the reconstituted signal rather than the biggest possible absolute distortion. In the combinatorial setup, this corresponds to considering the supremum of the variances of the rows and columns in our arrangements. We are not aware of any previously proved lower bounds in this context.

Since our construction for the upper bound involves putting together densely filled squares (just like in the case of minimizing spread), it will be a useful and interesting intermediate step to investigate the same question for arrangements of the integers 1 through N^2 in an N-by-N square.[1] This will be carried out in Sections IV-B through IV-D, where we prove

Theorem I.2: Assume that the integers 1 through N^2 have been arranged in an N-by-N matrix P, and the variance of numbers in any row or column is at most V. Then

$$(1/24 + 10^{-7})N^4 - 1/24 \leq V.$$

For any N, there exists an arrangement with $V = N^4/16 + O(N^3)$.

The arrangements in the whole plane are then investigated in Sections IV-E and IV-F. We obtain

Theorem I.3: Assume that the integers have been arranged within a plane square grid, with each row and column containing exactly N numbers, and the variance of numbers in any row or column is at most V. Then

$$N^4/60 - 1/60 \leq V.$$

For any N, an arrangement exists with $V = N^4/53\frac{1}{3}$.
The best previous upper bound known to the authors is $N^4/48 + O(N^3)$ in [7].

It is also useful and interesting to determine the maximal *spread* within rows and columns when the integers 1 through N^2 are arranged in an N-by-N square. This question has already been resolved fully by Tanya Y. Berger-Wolf and Edward M. Reingold in [5]. The smallest possible spread that can be attained is $N(N+1)/2 - 1$.

II. Notation

We shall always deal with arrangements of numbers into square grids, where each row or column contains exactly N numbers. For any set of integers T, let $H(T)$ denote the set of *horizontal neighbors* of T, i.e., the set of integers who are in the same row as some element of T. Similarly, define $V(T)$ to be the set of *vertical neighbors* of T. For example, if R is any row of an arrangement, the cardinality of $V(R)$ is N^2, since R has N elements and each one has N numbers in its column.

For any integer x in the arrangement, let $C(x)$ denote the (numbers in the) column of x and $R(x)$ the (numbers in the) row of x.

For any set Z we use $\#Z$ to denote its cardinality. As usual, we let the *variance* of a list $(X_1, X_2, \ldots, X_n)$ of real numbers be denoted by

$$\mathrm{Var}(X_1, X_2, \ldots, X_n) = \frac{1}{n}\sum_i (X_i - \overline{X})^2,$$

where $\overline{X} = \frac{1}{n}\sum_i X_i$ is the *mean.*

For any random variable X, we use $\mathbb{E}(X)$ and $\mathrm{Var}(X)$ to denote its mean and variance.

III. Minimizing the absolute distortion

A. Outline of argument

For the upper bound, we exhibit a construction. The crucial idea is that once we decided on the *footprint* of an arrangement (the fields in the infinite grid where we are to put integers), it is relatively straightforward to permute the integers within the selected positions so as to achieve the smallest possible maximal spread. The optimal footprint for even N turns out to be composed of solid $N/2$-by-$N/2$ squares arranged along a diagonal, plus another solid $N/2$-by-$N/2$ placed immediately below each of the diagonal ones. Details, and a description of the case of odd N, are given in Section III-B.

To give an idea of the argument for the lower bound, consider first the lower bound $d \geq (N^2-1)/3$ due to Diggavi, Orlitsky and Vaishampayan ([7]). Pick an arbitrary row R of the arrangement, and let m and M be the smallest and

[1] In the MDSQ setup, this corresponds to splitting a bounded discrete information source into two channels which both have exactly half the rate of the source.

largest elements of R. Then clearly $M-m \leq d$. Furthermore, every element of $V(R)$ is at most $m+2d$, since any element of R is at most $m+d$, and any element of $V(R)$ is at most d bigger than the element of R in its column. Similarly, all elements of $V(R)$ are at least $M-2d$. Thus we have at least N^2 distinct integers in the interval $[M-2d, m+2d]$ of length at most $3d$. It follows that $d \geq (N^2-1)/3$.

This argument can be improved in the following direction. Assume that d is approximately $N^2/3$, and let g and G be the minimal and maximal elements of $V(R)$. It can then be deduced that $H(C(g)) \cap H(C(G))$ must contain approximately $N/3$ rows, and that it must contain essentially all the numbers in $[m, M]$. Running through a similar argument starting with $C(g)$ instead of R, we get $N/3$ columns in $V(R)$ that contain essentially all of the integers in $[m-d, m]$. However, the $N^2/9$ squares where those rows and columns intersect must then be essentially empty, implying that a lot of the elements of $H(C(g)) \cap H(C(G))$ do not lie in $V(R)$. But since all elements of $H(C(g)) \cap H(C(G))$ do lie in $[M-2d, m+2d]$, we can get a stronger lower estimate for d than before, since $V(R)$ *plus* all of $H(C(g)) \cap H(C(G))$ and must lie in $[M-2d, m+2d]$. Working through this argument, we can obtain a lower bound of the form $d \geq cN^2$, where $1/3 < c < 3/8$.

In Section III-C, we show how to extend this argument by carefully considering the contents of a large set of special sets of rows and columns and the ways in which their contents can intersect.

B. The upper bound for the spread

Proof: (Proof of Theorem I.1, upper bound.) We will first explain our construction for even N. Examples will be given for $N = 8$. Start by filling in the numbers 1 through $N^2/4$ in an $N/2$-by-$N/2$ square as follows. Start by placing the numbers 1, 2, …, $N^2/8 - N/4$ below the northwest–southeast diagonal by filling the available spaces in the last row, then the second-last row, and so on, progressing from left to right in each row. Then place the numbers from $N^2/4$ down to $N^2/8 + N/4 + 1$ above the diagonal, by filling the available spaces in the last column, then the second-last column, and so on, progressing from top to bottom in each column. At this point, our square looks like this:

$$\begin{array}{cccc} . & 11 & 13 & 16 \\ 6 & . & 12 & 15 \\ 4 & 5 & . & 14 \\ 1 & 2 & 3 & . \end{array}$$

Of the remaining numbers ($N^2/8 - N/4 + 1$ through $N^2/8 + N/4$), put the central one[2] into the upper left corner, distribute the rest arbitrarily. We get

$$\begin{array}{cccc} 8 & 11 & 13 & 16 \\ 6 & 7 & 12 & 15 \\ 4 & 5 & 10 & 14 \\ 1 & 2 & 3 & 9 \end{array}$$

Let $P_{i,j}$ denote the (i, j) entry of the $N/2$-by-$N/2$ matrix just constructed. Let $Q_{i,j}$ denote the number (if any) that is placed into the (i, j)th position of the infinite plane square grid in the arrangement we are currently constructing.[3] We define $Q_{i,j}$ as follows. Let

$$\begin{aligned} i &= \frac{N}{2}a + i', && \text{with } i' \in \{1, 2, \ldots, \tfrac{N}{2}\}, \\ j &= \frac{N}{2}b + j', && \text{with } j' \in \{1, 2, \ldots, \tfrac{N}{2}\}, \end{aligned}$$

where a and b are integers. Then

$$Q_{i,j} = \begin{cases} P_{i,j} + aN^2/2 & \text{if } a = b, \\ P_{j,i} + (a + \frac{1}{2})N^2/2 & \text{if } a + 1 = b, \\ \text{empty} & \text{otherwise.} \end{cases}$$

For $N = 8$, we obtain

[2] If N is divisible by 4, there are two central ones: just pick one.
[3] For consistency, the first coordinate increases toward the north in both P and Q.

						40	43	45	48	.	.
						38	39	44	47	.	.
						36	37	42	46	.	.
						33	34	35	41	.	.
		8	11	13	16	25	30	31	32		
		6	7	12	15	19	26	28	29		
		4	5	10	14	18	21	23	27		
		1	2	3	9	17	20	22	24		
.	.	-7	-2	-1	0						
.	.	-13	-6	-4	-3						
.	.	-14	-11	-9	-5						
.	.	-15	-12	-10	-8						

It is easy to check that the maximal spread is $\lceil 3N^2/8 - 1/2 \rceil$. Note that no smaller maximum spread can be expected with this layout, since three adjacent squares must contain $3N^2/4$ numbers, all within distance $2d$ of each other.

For odd N, we can achieve $d = 3N^2/8 + N/4 + O(1)$. The construction starts by carrying out the above construction for $N+1$ (which is even). Letting $i = \frac{N+1}{2}a + i'$ and $j = \frac{N+1}{2}b + j'$, with a, b integers and $i', j' \in \{1, 2, \ldots, \frac{N}{2}\}$, replace $Q_{i,j}$ by an empty square if

- *either* $a = b$, $i' + j' = \frac{N+3}{2}$, and $i' \leq \frac{N+1}{4}$;
- *or* $a + 1 = b$, $i' + j' = \frac{N+3}{2}$, and $i' > \frac{N+1}{4}$.

For example, for $N = 7$, we get

							43	45	48	.	.
						38		44	47	.	.
						36	37	42	46	.	.
						33	34	35	41	.	.
			11	13	16	25	30	31	32		
		6		12	15	19	26	28	29		
		4	5	10	14	18	21		27		
		1	2	3	9	17	20	22			
.	.	-7	-2	-1	0						
.	.	-13	-6	-4	-3						
.	.	-14	-11		-5						
.	.	-15	-12	-10							

To complete the construction, apply an order-preserving bijection between the remaining numbers and the set of all integers:

							39	41	44	.	.
						36		40	43	.	.
						34	35	38	42	.	.
						31	32	33	37	.	.
			11	13	16	23	28	29	30		
		8		12	15	19	24	26	27		
		6	7	10	14	18	21		25		
		3	4	5	9	17	20	22			
.	.	-5	0	1	2						
.	.	-9	-4	-2	-1						
.	.	-10	-7		-3						
.	.	-11	-8	-6							

The maximal difference in this case is easily seen to be $3N^2/8 + N/4 + O(1)$. ■

C. The lower bound for the spread

We will prove the lower bound in Theorem I.1 indirectly. The following convention embodies the opposite of the conclusion of the theorem:

Convention III.1: From now on, we shall assume that

$$d < 3N^2/8 - 1/2.$$

Definition III.2: Let R_0 be an arbitrary row in the arrangement. Let m_1 be the smallest element of $V(R_0)$. If i is a positive odd integer and m_i has already been defined, let R_i be the column containing m_i, let $N_i = H(R_i)$. If i is a positive even integer and m_i has already been defined, let R_i be the row containing m_i, let $N_i = V(R_i)$. For any positive integer i, if N_i has already been defined, let m_{i+1} be the smallest element of N_i.

Unwinding Definition III.2, we get columns R_1, R_3, R_5, …, and rows R_2, R_4, R_6, …. For all positive integers i, we have $m_i \in R_i \subset N_i$. Also note that N_i always has exactly N^2 elements.

Lemma III.3: For all $x \in N_i$, we have

$$m_{i+1} \leq x \leq m_i + 2d.$$

Proof: The lower bound is valid by the definition of m_{i+1}.

Note that $m_i \in R_i$ and let x_0 be the element of R_i that is in the same row (respectively column) as x, if i is odd (respectively even). By the definition of d, we have $x - x_0 \leq d$ and $x_0 - m_i \leq d$. The sum of these two inequalities yields the upper bound of the lemma. ■

Lemma III.4: For any positive integer i,

$$m_i - m_{i+1} \leq d.$$

Proof: Let x_0 be the unique element of $R_i \cap R_{i+1}$. Then $x_0 - m_{i+1} \leq d$ by the definition of d, and $m_i \leq x_0$ by the definition of m_i. Thus, the lemma follows. ■

Lemma III.5: For any positive integer i,

$$N^2 - 2d - 1 \leq m_i - m_{i+1}.$$

Proof: By Lemma III.3,

$$\begin{aligned} \max N_i &\leq m_i + 2d, \\ m_{i+1} &\leq \min N_i. \end{aligned}$$

Since N_i has N^2 elements, we also have

$$\min N_i + N^2 - 1 \leq \max N_i.$$

Summing the three displayed inequalities and rearranging yields the claim of the lemma. ■

Corollary III.6: If i and j are any pair of distinct non-negative integers, then $m_i \neq m_j$. The same statement applies to R_i and N_i.

Proof: By Convention III.1, $N^2 - 2d + 1$ is positive. Thus, $m_i \neq m_j$ follows by Lemma III.5. The similar statements for R_i and N_i of course follow directly from the one for m_i. This demonstrates that the construction in Definition III.2 never bogs down: the successive N_i always contain new numbers that have not occured in any N_i before. ■

Lemma III.7: For any positive integer i, and any integer $k \geq 2$,

$$N_i \cap N_{i+2k} = \emptyset.$$

Proof: For an indirect proof, assume that there is a row (respectively, column) Z that contains elements of both R_i and R_{i+2k}, if i is odd (respectively, even). Let x_0 be the unique element of $R_{i+2k} \cap Z$, and let x_1 be the unique element of $R_i \cap Z$. Then $x_0 - m_{i+2k} \leq d$ and $x_1 - x_0 \leq d$. However, since x_1 is in N_i, we also have $x_1 \geq m_i$. Combining these inequalities, we obtain that

$$m_i - m_{i+2k} \leq 2d.$$

Summing the conclusion of Lemma III.5 applied $2k$ times (for i, …, $i + 2k - 1$), we obtain

$$2k(N^2 - 2d + 1) \leq m_i - m_{i+2k}.$$

Combining the two displayed inequalities above, we obtain

$$2kN^2 + 2k \leq (4k + 2)d,$$

which contradicts Convention III.1, thus proving our lemma. ■

Definition III.8: For any set X, we shall use $\#X$ to denote the number of elements in X.

Lemma III.9: For any positive integer i,

$$\#(N_i \cap N_{i+3}) \leq 2d + 1 - (m_{i+1} - m_{i+3}).$$

Proof: In this proof, for brevity's sake, denote $N_i \cap N_{i+3}$ by F. If F is empty then there is nothing to prove, since the right hand side of our inequality is always at least 1 by Lemma III.4. Let us now assume that F is not empty.

Since any element of F is also in N_i, by Lemma III.3 we have

$$m_{i+1} \leq \min F.$$

Since any element of F is also in N_{i+3}, by Lemma III.3 we also have

$$\max F \leq m_{i+3} + 2d.$$

Finally, it is clearly also true that

$$\min F + \#F - 1 \leq \max F.$$

Combining the three displayed inequalities proves this lemma. ■

Lemma III.10: For any positive integer i, and any integer $k \geq 2$,

$$N_i \cap N_{i+2k+1} = \emptyset.$$

Proof: For an indirect proof, assume that $N_i \cap N_{i+2k+1}$ is not empty. Then the method of the proof of Lemma III.9 goes through to show that

$$1 \leq \#(N_i \cap N_{i+2k+1}) \leq 2d + 1 - (m_{i+1} - m_{i+2k+1}).$$

However, by repeated application of Lemma III.5, this implies

$$2k(N^2 - 2d - 1) \leq 2d.$$

This contradicts Convention III.1, thus proving our lemma. ■

Definition III.11: For any integer $i \geq 2$, let

$$H_i = N_{i-1} \cap N_{i+1}.$$

For any integer $i \geq 3$, let

$$G_i = N_i - N_{i-2} - N_{i+2}.$$

Lemma III.12: For any integer $i \geq 2$,

$$\#H_i = \#(N_{i+1} \cap N_{i-1}) \leq N^2/2.$$

Proof: Note that H_i is not empty as $R_i \subset H_i$. Since any element of H_i is also in N_{i-1}, by Lemma III.3 we have

$$m_i \leq \min H_i.$$

Since any element of H_i is also in N_{i+1}, by Lemma III.3 we also have

$$\max H_i \leq m_{i+1} + 2d.$$

Finally, it is clearly also true that

$$\min H_i + \#H_i - 1 \leq \max H_i.$$

Combining the three displayed inequalities and using Lemma III.5, we obtain

$$\#H_i \leq 4d + 2 - N^2,$$

which implies our lemma, since by Convention III.1 the right hand side is at most $N^2/2$. ■

Corollary III.13: For any integer $i \geq 3$,

$$\#(G_i \cap H_i) \leq \#G_i/2.$$

Proof: Assume that i is odd (respectively even). Every column (resp. row) of H_i contains exactly N elements. Therefore, Lemma III.12 implies that H_i has at most $N/2$ columns (resp. rows). Therefore, each row (resp. column) of G_i can contain no more than $N/2$ elements of H_i. Since each row (resp. column) of G_i contains exactly N elements, this implies that the size of $G_i \cap H_i$ is at most half the size of G_i, as asserted. ■

Lemma III.14: For any integer $i \geq 4$,

$$N^2 \leq 8d + 4 + m_{i+3} + m_{i+2} - m_{i-1} - m_{i-2} + \#G_{i+1}/2 + \#G_i + \#G_{i-1}/2.$$

Proof: We will prove the lemma by giving an upper bound on $\#N_i = N^2$.

By Lemmas III.7 and III.10, the set N_i is entirely contained in the union of the sets N_{i-3}, N_{i-1}, N_{i+1} and N_{i+3}, which union is in turn equal to the (disjoint) union of the sets N_{i-3}, G_{i-1}, H_i, G_{i+1} and N_{i+3}. From Lemma III.9, we immediately obtain

$$\begin{aligned}
\#(N_i \cap N_{i-3}) &\leq 2d + 1 - (m_{i-2} - m_i), \\
\#(N_i \cap N_{i+3}) &\leq 2d + 1 - (m_{i+1} - m_{i+3}).
\end{aligned}$$

In order to estimate the remaining part of N_i, namely its intersection with $G_{i-1} \cup H_i \cup G_{i+1}$, we need to decompose N_i itself as the (disjoint) union $H_{i-1} \cup G_i \cup H_{i+1}$. This way, we broke the remaining part of N_i into nine fragments. Three of them can be disposed of in the following way:

$$\#(G_i \cap (G_{i-1} \cup H_i \cup G_{i+1})) \leq \#G_i.$$

Four other fragments can be estimated from above using Lemma III.9 as follows:

$$\begin{aligned}
\#(H_{i-1} \cap (H_i \cup G_{i+1})) \le \#(N_{i-2} \cap N_{i+1}) &\le 2d + 1 - (m_{i-1} - m_{i+1}), \\
\#(H_{i+1} \cap (H_i \cup G_{i-1})) \le \#(N_{i+2} \cap N_{i-1}) &\le 2d + 1 - (m_i - m_{i+2}).
\end{aligned}$$

Finally, we can invoke Corollary III.13 twice to obtain

$$\begin{aligned}
\#(G_{i-1} \cap H_{i-1}) &\le \#G_{i-1}/2, \\
\#(G_{i+1} \cap H_{i+1}) &\le \#G_{i+1}/2.
\end{aligned}$$

Summing all the displayed equations in this proof yields the assertion of the lemma. ■

Lemma III.15: For any integers $t \ge 3$ and $k \ge 2$, we have

$$2(\#G_t + \#G_{t+4} + \cdots + \#G_{t+4k-4}) \le 4d + 2 + 2(m_t - m_{t+4k-3}) - 2(k-1)N^2.$$

Proof: Let X denote the (disjoint) union of the sets G_t, N_{t+2}, G_{t+4}, N_{t+6}, …, N_{t+4k-6} and G_{t+4k-4}. By Lemma III.3 (and since the m_i form a decreasing sequence),

$$\begin{aligned}
\max X &\le m_t + 2d, \\
m_{t+4k-3} &\le \min X.
\end{aligned}$$

We also have

$$\min X + \#X - 1 \le \max X.$$

Since each N_i has size N^2, we also know that

$$\#X = \#G_t + \#G_{t+4} + \cdots + \#G_{t+4k-4} + (k-1)N^2.$$

Combining the displayed (in)equalities and multiplying by two, we obtain our claim. ■

Proof: (Proof of Theorem I.1, lower bound.) Take some integer $k \ge 2$.

First of all, sum the conclusion of Lemma III.14 with i taking each of the values 4, 5, 6, …, $4k+3$ to obtain

$$4kN^2 \le 4k(8d+4) + 2(\#G_4 + \#G_5 + \cdots + \#G_{4k+3}) + \mu_1^+ - \mu_1^- + \gamma. \tag{1}$$

Here μ_1^+ and μ_1^- involve values of various m_i, specifically

$$\begin{aligned}
\mu_1^+ &= m_{4k+6} + 2m_{4k+5} + 2m_{4k+4} + 2m_{4k+3} + m_{4k+2}, \\
\mu_1^- &= m_6 + 2m_5 + 2m_4 + 2m_3 + m_2,
\end{aligned}$$

whereas γ involves various values of $\#G_i$,

$$\gamma = \#G_3/2 - \#G_4/2 - \#G_{4k+3}/2 + \#G_{4k+4}/2.$$

The sum of $\#G_i$ in (1) can be estimated from above by summing the conclusion of Lemma III.15 for each of $t = 7$, $t = 6$, $t = 5$ and $t = 4$, to obtain

$$2(\#G_4 + \#G_5 + \cdots + \#G_{4k+3}) \le 16d + 8 - 8(k-1)N^2 + \mu_2^+ - \mu_2^-, \tag{2}$$

where

$$\begin{aligned}
\mu_2^+ &= 2m_7 + 2m_6 + 2m_5 + 2m_4, \\
\mu_2^- &= 2m_{4k+4} + 2m_{4k+3} + 2m_{4k+2} + 2m_{4k+1}.
\end{aligned}$$

Now combining (1) and (2), we obtain

$$(12k-8)N^2 \le d(32k+16) + (16k+8) + \gamma + \mu_1^+ + \mu_2^+ - \mu_1^- - \mu_2^-. \tag{3}$$

Note that by Lemma III.4, the absolute value of $\mu_1^+ - \mu_2^-$ is bounded above by $12d$ (as is the absolute value of $\mu_2^+ - \mu_1^-$). It is also clear that the absolute value of γ is no bigger than N^2. Since none of these bounds depends on k, we may let k tend to (positive) infinity in (3) and divide by $4k$ to obtain

$$3N^2 \le 8d + 4,$$

which implies

$$3N^2/8 - 1/2 \leq d.$$

This contradicts Convention III.1, thereby completing our indirect proof of Theorem I.1.

In summary, we have shown that

$$3N^2/8 - 1/2 \leq d.$$

■

Remark III.16: For odd N, the methods of Theorem I.1 extend to show that

$$3N^2/8 + N/8 - 1/2 \leq d.$$

The proof of Theorem I.1 can be modified to prove Remark III.16 as follows. Changing the assertion in Convention III.1 according to Remark III.16, instead of the assertion of Lemma III.12 we can prove

$$\#H_i = \#(N_{i+1} \cap N_{i-1}) < N(N+1)/2.$$

However, since $\#H_i$ must be divisible by N, this in fact implies

$$\#H_i \leq N(N-1)/2. \tag{4}$$

Using this, we can strenghten the claim of Lemma III.13 to

$$\#(G_i \cap H_i) \leq \frac{N-1}{2N} \#G_i.$$

Recall that by Lemmas III.7 and III.10, the set N_i of N^2 elements is the disjoint union of H_{i-1}, G_i and H_{i+1}. Therefore (4) also implies that

$$\#G_i \geq N,$$

which, with Lemma III.15, implies

$$N^2/4 + N/4 \leq \lim_{k\to\infty} \frac{m_{t+4k} - m_t}{k}. \tag{5}$$

Then the proof can proceed as before, though using (2) with a factor of $2 - 1/N$ instead of 2, and finally utilizing (5) to conclude that

$$3N^2/8 + N/8 - 1/4 \leq d,$$

which contradicts the modified Convention III.1, thus proving Remark III.16.

IV. Minimizing the mean squared error

A. Outline of argument

We shall begin in Section IV-B by giving a lower bound on the row and column variances in an N-by-N square arrangement. This proof is based on an algebraic inequality relating the average row or column variance to the variance of the set of all numbers in the square.

In Section IV-C, we give an upper bound for the row an column variances by constructing appropriate square arrangements. First we exhibit an explicit arrangement with relatively low row and column variances. Then we prove a lemma that allows us to combine two arrangements with lower N into a single one with higher N, such that the row and column variances of the bigger arrangement are well controlled by the row and column variances of the smaller arrangements. We then use this combination lemma, along with arrangements for $N \leq 10$ constructed by hand, to construct square arrangements for any N with low row and column variances.

In Section IV-D, we revisit the issue of a lower bound for the row and column variances in square arrangements by showing that the algebraic inequality used in Section IV-B cannot be sharp in this application. Upon refining our inequality, this non-sharpness turns out to be related to the fact that the sum of two independent random variables of (nearly) equal variance cannot be a (nearly) uniformly distributed random variable. We prove a theorem in this direction and then use it to improve the lower bound given in Section IV-B slightly.

Section IV-E gives an upper bound for the row and column variances in the infinite case by construction. The footprint of this construction is the same as the one in Section III-B, and we use the combination lemma and the square arrangements of Section IV-C to fill the footprint.

Finally, in Section IV-F we prove a lower bound for the row and column variances in the infinite case by a more involved version of the argument of Section IV-B combined with some ideas from Section III-C.

B. A lower bound for the variance in a square

Theorem IV.1: Assume that the integers 1 through N^2 have been arranged in an N-by-N matrix P, and the variance of numbers in any row or column is at most V. Then

$$(N^4-1)/24 \le V.$$

The following theorem yields the crucial ingredient for the proof of Theorem IV.1.

Theorem IV.2: Let $X_{i,j}$ $(1 \le i, j \le N)$ denote the elements of an N-by-N matrix. Then

$$\mathrm{Var}(X_{1,1}, \ldots, X_{N,N}) \le \frac{1}{N} \sum_i \mathrm{Var}(X_{i,1}, X_{i,2}, \ldots, X_{i,N}) + \frac{1}{N} \sum_j \mathrm{Var}(X_{1,j}, X_{2,j}, \ldots, X_{N,j}). \tag{6}$$

Proof: Substitute the definitions of all the variances into (6), multiply by N^4, and move all terms to the right side. It is easy to check that the coefficients of the monomials on the right hand side are as follows. The coefficients of the terms of form $X_{i,j}^2$ (where i and j need not be distinct) are $(N-1)^2$, the coefficients of the terms of form $X_{i,j}X_{l,k}$ (with $i \ne l$ and $j \ne k$) are 2, and the coefficients of the remaining terms (either $X_{i,j}X_{i,k}$ or $X_{i,j}X_{l,j}$) are all $2(1-N)$. There are no other terms.

This transformation shows that our assertion is equivalent to stating that a certain quadratic form in the variables $X_{i,j}$ is positive semi-definite. Let G denote the N^2-by-N^2 matrix whose rows (resp. columns) are labeled with the variables $X_{i,j}$ that represents the quadratic form in question. The entries of G can be read off from the calculation above. The diagonal entries are all $(N-1)^2$. The off-diagonal entry corresponding to row $X_{i,j}$ and column $X_{l,k}$ is 1 or $1-N$ (respectively) depending on whether $(i \ne l) \wedge (j \ne k)$ is true or not (respectively).

The matrix G is the Kronecker (tensor) square of an N-by-N matrix H whose diagonal elements are equal to $N-1$, and whose other elements are -1. Since this matrix H has eigenvalues 0 (singly) and N (repeated $N-1$ times), the matrix G has eigenvalues 0 ($2N-1$ of them) and N^2 ($(N-1)^2$ of them). This proves our result. ■

Theorem IV.2 can be generalized to higher dimensions. This will be described in a forthcoming paper.

Proof: (Proof of Theorem IV.1.) An easy calculation shows that the variance of the integers from 1 through N^2 is

$$\mathrm{Var}(1, 2, \ldots, N^2) = (N^4-1)/12. \tag{7}$$

Now we can apply Theorem IV.2 to our matrix P:

$$(N^4-1)/12 \le \frac{1}{N} \sum_i \mathrm{Var}(P_{i,1}, P_{i,2}, \ldots, P_{i,N}) + \frac{1}{N} \sum_j \mathrm{Var}(P_{1,j}, P_{2,j}, \ldots, P_{N,j}). \tag{8}$$

Since every row or column variance on the right hand side is at most V, the claim of the theorem follows immediately. ■

C. The upper bound for the variance in a square

Proof: (Proof of Theorem I.2, upper bound.) First, let N be any even integer and define the N-by-N matrix P (containing the integers 1 through N^2) by

$$P_{i,j} = \begin{cases} (i-1)N/2 + j & \text{if } j \le N/2, \\ (i-1)N/2 + j + (N^2-N)/2 & \text{otherwise.} \end{cases} \tag{9}$$

It is easy to check that the variance of each row of P is $N^4/16 + N^2/48 - 1/12$, and the variance of each column is $N^4/48 - N^2/48$.

This construction can also be used to give a construction with maximal variance $N^4/16 + O(N^3)$ for odd N. For any odd N,

- *Step A.* take the construction just given for $N+1$;
- *Step B.* omit a row and a column to get to an N-by-N square; and
- *Step C.* apply the unique order-preserving bijection between the remaining numbers and the set $\{1, 2, \ldots, N^2\}$.

In Step A, each row and column variance is bounded above by $(N+1)^4/16 + O((N+1)^3) = N^4/16 + O(N^3)$, as described for the case of even N above. In Step B, the variance of each row or column increases by no more[4] than a factor of $1 + 1/N$, and is therefore still bounded above by $N^4/16 + O(N^3)$ (of course the constant hidden in the O factor increased). Finally, Step C cannot increase variances either since no pairwise distances are increased between the numbers (as stated in Lemma IV.10(iii) in Section IV-F), and therefore we can conclude that in our new arrangement, the maximal variance obtained in any row or column will not exceed $N^4/16 + O(N^3)$. ■

[4]This follows from the general fact that if a set of n numbers has variance v then any subset of $n-1$ numbers in it has variance no more than $nv/(n-1)$.

By checking all the possible arrangements, we can verify that the construction given above is optimal for $N \leq 2$. However, for any specific N over 2, it is easy to find an arrangement that has a lower maximal variance than the general construction given above. For example, for N between 3 and 10, we have found arrangements[5] with the following maximal variances:

N	3	4	5	6	7	8	9	10
V	4.7	12.2	29.6	59.8	109.7	184.5	299.9	455.2
$10^2V/N^4$	5.80	4.77	4.74	4.61	4.57	4.50	4.57	4.55

Concrete arrangements can be combined to form larger arrangements with relatively low maximal row/column variances. The following lemma (which is stated with enough generality to be of use in Section IV-E for infinite arrangements) explains how.

Lemma IV.3: Assume that an arrangement A (resp. B) has exactly N (resp. N') integers in every row and column, and that the maximal variance within a row or column is at most V (resp. V'). The arrangement A must be an N-by-N square, whereas the arrangement B can be either an N'-by-N' square or infinite. Then there is an arrangement C with exactly NN' integers in every row and column, and the maximal variance within a row or column is at most $V + N^4V'$. The arrangement C is square (of size NN'-by-NN') or infinite depending as B is a square or is infinite.

Proof: Define $C_{i,j}$ as follows. For

$$\begin{aligned} i &= Na + i', \qquad \text{with } i' \in \{1, 2, \ldots, N\}, \\ j &= Nb + j', \qquad \text{with } j' \in \{1, 2, \ldots, N\}, \end{aligned}$$

where a and b are integers, set

$$C_{i,j} = A_{i',j'} + N^2(B_{a,b} - 1).$$

It is then clear that

$$\begin{aligned} \mathrm{Var}(C_{i,\cdot}) &= \mathrm{Var}(A_{i',\cdot}) + N^4\mathrm{Var}(B_{a,\cdot}) \leq V + N^4V', \\ \mathrm{Var}(C_{\cdot,j}) &= \mathrm{Var}(A_{\cdot,j'}) + N^4\mathrm{Var}(B_{\cdot,b}) \leq V + N^4V', \end{aligned}$$

which proves our lemma. ∎

For example, if we apply Lemma IV.3 with A being the optimal arrangement for $N = 2$ (with $V = 1$), and B being the optimal arrangement for $N = 3$ (with $V' = 14/3$), namely

$$A: \begin{array}{cc} 3 & 4 \\ 1 & 2, \end{array} \qquad B: \begin{array}{ccc} 6 & 7 & 9 \\ 3 & 5 & 8 \\ 1 & 2 & 4, \end{array}$$

we obtain

$$\begin{array}{cccccc} 23 & 24 & 27 & 28 & 35 & 36 \\ 21 & 22 & 25 & 26 & 33 & 34 \\ 11 & 12 & 19 & 20 & 31 & 32 \\ 9 & 10 & 17 & 18 & 29 & 30 \\ 3 & 4 & 7 & 8 & 15 & 16 \\ 1 & 2 & 5 & 6 & 13 & 14, \end{array}$$

with a maximal row/column variance of $227/3 \approx 75.7$.

Theorem IV.4: For any positive integer N, let $V(N)$ be the smallest possible maximal variance in a row or column when the numbers 1 through N^2 are arranged in an N-by-N square. Then let $c(N)$ be defined as

$$c(N) = V(N)/N^4.$$

Then

(a) $c(2N) \leq c(N) + 1/(16N^4)$,

(b) $c(N) \leq c(N+1)(1 + 1/N)^5$.

Proof: Part (a) follows from applying Lemma IV.3 to the optimal construction for 2 playing the role of A and the optimal construction for N playing the role of B.

Part (b) follows from the argument given in the proof of the upper bound of Theorem I.2: take the optimal construction for an $(N+1)$-by-$(N+1)$ square, delete a row and a column, and adjust the remaining numbers (bijectively) to lie in the interval $[1, N^2]$. ∎

[5]These arrangements are available at `www.csirik.net/square-variances.html`.

It is clear that the results of Theorem IV.4 can be used to improve the upper bound on all $c(N)$ if sufficiently good constructions are available for small N. For example, given arrangements for all integers from 1 through N_0, we can use Theorem IV.4(a) to give good arrangements for all even integers in the range $[N_0+1, 2N_0]$, and then use Theorem IV.4(b) to give good arrangements for all odd integers in the same range, while controlling the corresponding $c(N)$. This procedure can be iterated and will clearly give a universal upper bound on $c(N)$, which would depend on how good the initial set of arrangements were. Using this method and the constructions for $N \leq 10$ given above, we can get

$$c(N) < 1/20 \qquad (\text{for } N \geq 4).$$

For example,

$$\begin{aligned} c(16) &\leq c(8) + 1/(16 \cdot 8^4), \\ c(32) &\leq c(16) + 1/(16 \cdot 16^4) \leq c(8) + 1/(16 \cdot 8^4) + 1/(16 \cdot 16^4), \end{aligned}$$

and so on, from which it follows that

$$c(N) \leq 1/22 \qquad (\text{for } N = 2^k \geq 8,)$$

where k is of course meant to be an integer.

D. A better lower bound for the variance in a square

Since our construction does not agree with the lower bound given in Theorem IV.1, it is natural to ask whether the theorem could be improved. The matrix P given in Section IV-C illustrates that it is possible for the result of Theorem IV.1 to not be sharp, even though the result of Theorem IV.2 is sharp. The problem is that Theorem IV.1 allows us to give a lower bound for the *average* of the row and column variances of any square matrix P containing the numbers 1 through N^2, but for any particular matrix, either the inequality of Theorem IV.2 in not sharp, or the average and the maximal row/column variances are not equal.

In this section, we will develop these ideas to improve the lower bound of Theorem IV.1 by a little bit. Larger improvements should be possible by more tightly controlling the various estimates given below. First, we need the following more precise version of Theorem IV.2.

Theorem IV.5: Let $X_{i,j}$ $(1 \leq i,j \leq N)$ denote the elements of an N-by-N matrix. Assume that $\sum_{i,j} X_{i,j} = 0$. Then

$$\mathrm{Var}(X_{1,1},\ldots,X_{N,N}) + \frac{1}{N^2}\sum_{i,j} R_{i,j}^2 = \frac{1}{N}\sum_i \mathrm{Var}(X_{i,1},\ldots,X_{i,N}) + \frac{1}{N}\sum_j \mathrm{Var}(X_{1,j},\ldots,X_{N,j}), \tag{10}$$

where

$$R_{i,j} = X_{i,j} - \frac{1}{N}\sum_k X_{i,k} - \frac{1}{N}\sum_k X_{k,j},$$

for all $1 \leq i,j \leq N$.

Proof: Adding $(\sum_{i,j} X_{i,j})^2/N^4$ to the right-hand side of (10), we get a polynomial identity that is easy to verify. However, by assumption we have $\sum_{i,j} X_{i,j} = 0$, which yields the statement of the theorem.[6] ■

Let us now return to our matrix P. We will scale P to simplify our calculations. Define an N-by-N matrix U by $U_{i,j} = (X_{i,j} - (N^2+1)/2)/N^2$. Thus $\sum_{i,j} U_{i,j} = 0$ and the elements of U all lie within $[-1/2, 1/2]$. Clearly, the biggest row or column variance α in U will be equal to V/N^4 and all other interesting properties of P will be represented in U too. We can write

$$U = F + G + R,$$

where each element of F is the average of the elements in the corresponding row of U, each element of G is the average of the elements in the corresponding column of U, and R is defined as $U - F - G$.

Let us now consider the matrices U, F, G and R as random variables on the probability space $S_R \times S_C$ (where S_R and S_C are N-element sets representing rows and columns, respectively) endowed with the uniform distribution. Because P contains the integers from 1 through N^2, U is nearly uniformly distributed[7] on the interval $[-1/2, 1/2]$. In this case, Theorem IV.5 implies (recall that α denotes the maximal row/column variance that occurs in the matrix U)

$$\mathrm{Var}(U) + \mathbb{E}(R^2) \leq 2\alpha. \tag{11}$$

[6]We discovered the identity (10) by calculating the contribution of the component of X that lies outside of the 0-eigenspace of the matrix G (in Theorem IV.1) to the value of the quadratic form defined by G.

[7]This approximation becomes more and more accurate as N tends to infinity.

Let $\overline{U_{\bullet,j}}$ denote the average of all elements of row j in the matrix U. Then, for any $1 \le k \le N$, the inequality between the arithmetic and the quadratic means implies that

$$\left(\frac{1}{N}\sum_j U_{k,j} - \frac{1}{N}\sum_j U_{\bullet,j}\right)^2 = \left(\frac{1}{N}\sum_j (U_{k,j} - U_{\bullet,j})\right)^2 \le \frac{1}{N}\left(\sum_j (U_{k,j} - U_{\bullet,j})^2\right).$$

Summing both sides over k and dividing by N, we obtain

$$\mathrm{Var}(F) \le \frac{1}{N}\sum_j \mathrm{Var}(U_{1,j}, U_{2,j}, \ldots, U_{N,j}) \le \alpha, \tag{12}$$

and similarly

$$\mathrm{Var}(G) \le \alpha. \tag{13}$$

Note that F and G are independent as random variables, since values of F are constant within rows and values of G are constant within columns.

We can now see clearly why Theorem IV.1 cannot be sharp. If that bound was sharp, then (11) would imply that $R = 0$ and thus

$$U = F + G.$$

Then (12) and (13) together with $\mathrm{Var}(U) = \mathrm{Var}(F) + \mathrm{Var}(G)$ imply that

$$\mathrm{Var}(F) = \mathrm{Var}(G) = \frac{1}{2}\mathrm{Var}(U)$$

The non-sharpness of Theorem IV.1 in this point of view is related to (but does not yet immediately follow from) the following general fact.

Fact IV.6: A random variable that is uniformly distributed on some interval of the real numbers cannot be a sum of two independent random variables of equal variances.

This fact can be proved using the following lemma.

Lemma IV.7: Let F and G be independent real-valued random variables with $\mathbb{E}F = \mathbb{E}G = 0$, and let $n_2 = \mathbb{E}(F+G)^2$ and $n_4 = \mathbb{E}(F+G)^4$. Then $a = \mathbb{E}F^2$ satisfies

$$4a^2 + n_4 \ge n_2(n_2 + 4a).$$

Proof: We have (using $\mathbb{E}F = 0$ to drop the cross term)

$$n_2 = \mathbb{E}(F+G)^2 = \mathbb{E}F^2 + \mathbb{E}G^2,$$

and hence

$$\mathbb{E}G^2 = n_2 - a.$$

Clearly

$$\mathbb{E}F^4 \ge (\mathbb{E}F^2)^2 = a^2$$

and

$$\mathbb{E}G^4 \ge (\mathbb{E}G^2)^2 = (n_2 - a)^2.$$

This implies that (again using $\mathbb{E}F = 0$ and $\mathbb{E}G = 0$ to drop some terms)

$$n_4 = \mathbb{E}(F+G)^4 = \mathbb{E}F^4 + 6(\mathbb{E}F^2)(\mathbb{E}G^2) + \mathbb{E}G^4 \ge a^2 + 6a(n_2 - a) + (n_2 - a)^2,$$

which rearranges to yield the statement of the lemma. ∎

Proof: (Proof of Fact IV.6.) Without loss of generality, we can assume that our uniform distribution has mean 0, and that it is written as a sum of independent random variables F and G of mean 0. Applying Lemma IV.7 with $a = \mathbb{E}F^2 = \mathbb{E}G^2 = n_2/2$, we obtain

$$n_4/n_2^2 \ge 2.$$

However, for a uniform distribution we have $n_4/n_2^2 = 9/5$, which shows that a decomposition of the proposed type is not possible. ∎

The general question of how a uniformly distributed random variable can be broken up as a sum of two independent random variables has been studied extensively. For details, and a general characterization of those pairs of independent random variables that sum to a uniform distribution, the reader is referred to [8, Section 1.4], [9], [10] and the references contained therein. From the general characterization we can infer that if F and G are independent random variables with $\mathrm{Var}(F) \geq \mathrm{Var}(G)$ and $F+G$ is uniformly distributed, then $\mathrm{Var}(F)/\mathrm{Var}(G) \geq 3$, and this is achieved only when F is uniformly distributed over two discrete values, and G is uniformly distributed over a shorter interval.[8]

The following theorem is a more general version of Lemma IV.7 that allows us to cope with the fact that in our application U is not exactly uniformly distributed, and that the variances of F and G are not exactly equal. It will allow us to improve the lower bound given in Theorem IV.1.

Theorem IV.8: Let μ, λ be arbitrary positive real numbers. Let U, R, F and G be random variables of mean 0 with

$$U - R = F + G,$$

such that F and G are independent, $\mathbb{E}U^2 = m_2$, $\mathbb{E}U^4 = m_4$, and assume that almost surely $|U| < \mu$ and $|R| < \lambda$.

If $0 < 2m_2^2 - m_4$ then there is an $\varepsilon = \varepsilon(m_2, m_4, \mu, \lambda) > 0$ such that the following inequalities cannot be satisfied simultaneously:

$$\mathbb{E}F^2 = \frac{1}{2}m_2 + a, \qquad \text{with } |a| < \varepsilon, \tag{14}$$

$$\mathbb{E}R^2 \leq 2\varepsilon. \tag{15}$$

(The proof of the theorem provides information about a way to determine such an $\varepsilon(m_2, m_4, \mu, \lambda)$.)

Proof: We will apply Lemma IV.7 to the random variables F and G. Accordingly, let us define

$$\begin{aligned} n_2 &= \mathbb{E}(F+G)^2 = \mathbb{E}(U-R)^2, \\ n_4 &= \mathbb{E}(F+G)^4 = \mathbb{E}(U-R)^4. \end{aligned}$$

Note that $\mathbb{E}|R| \leq \sqrt{\mathbb{E}R^2} \leq \sqrt{2\varepsilon}$, so we can bound the difference between n_2 and m_2 as follows:

$$|n_2 - m_2| = |\mathbb{E}(U-R)^2 - \mathbb{E}U^2| \leq 2|\mathbb{E}UR| + \mathbb{E}R^2 \leq 2\mu\sqrt{2\varepsilon} + 2\varepsilon.$$

Similarly, we can obtain

$$|n_4 - m_4| \leq 4|\mathbb{E}U^3R| + 6\mathbb{E}U^2R^2 + 4|\mathbb{E}UR^3| + \mathbb{E}R^4 \leq 4\mu^3\sqrt{2\varepsilon} + 6\mu^2 2\varepsilon + 4\mu\lambda 2\varepsilon + \mu^2 2\varepsilon.$$

The crucial point in these inequalities is that

$$n_2 = m_2 + b, \qquad \text{with } |b| < f_1(\varepsilon), \tag{16}$$

$$n_4 = m_4 + c, \qquad \text{with } |c| < f_2(\varepsilon), \tag{17}$$

where f_1 and f_2 are continuous functions of ε, with value 0 when $\varepsilon = 0$.[9]

Now applying Lemma IV.7, we get

$$4(\mathbb{E}F^2)^2 + n_4 \geq n_2(n_2 + 4\mathbb{E}F^2).$$

Using (14), (16), and (17), this transforms to

$$4a^2 + c - 4m_2b - b^2 - 4ab \geq 2m_2^2 - m_4. \tag{18}$$

Estimating the left hand side from above (we can drop $-b^2$ which is never positive, and estimate the absolute value of the other four terms from above one-by-one, for example by $|4a^2| < 4(2\varepsilon)^2 = 16\varepsilon^2$), we get

$$16\varepsilon^2 + f_2(\varepsilon) + 4\mu|f_1(\varepsilon)| + 8\varepsilon f_2(\varepsilon) \geq 2m_2^2 - m_4.$$

By the hypothesis of the theorem, the right hand side here is positive. On the other hand, the left hand side is a continuous function of ε, and is 0 when $\varepsilon = 0$, so we get a contradiction if ε is too small. This proves our theorem. ■

Theorem IV.9: Let $N > 1$. Assume that the integers 1 through N^2 have been arranged in an N-by-N matrix P, and the variance of numbers in any row or column is at most V. Then

$$N^4(1/24 + 10^{-7}) - 1/24 \leq V.$$

Proof: In terms of our scaled matrix U, we need to show that

$$1/24 + 10^{-7} - 1/(24N^4) \leq \alpha.$$

[8] The best result in this direction implied by Lemma IV.7 is $\mathrm{Var}(F)/\mathrm{Var}(G) \geq (3+\sqrt{5})/2 = 2.62$.

[9] The functions f_1 and f_2 also depend on μ and λ, but our notation supresses that for simplicity.

Theorem IV.8 applies exactly, we just need to determine all the parameters. Since all elements of U lie in $[-1/2, 1/2]$, we have

$$|U| \le 1/2 = \mu.$$

By construction, all elements of F and G must lie in $[-1/2, 1/2]$ too, so we have

$$|R| \le 3/2 = \lambda.$$

A quick calculation shows that

$$2m_2^2 - m_4 = \frac{1}{720} + \frac{1}{72N^2} - \frac{11}{720N^4},$$

which is positive for all $N > 1$ as required.

Now use (18) the proof of Theorem IV.8. We can substitute the values $\lambda = 3/2$ and $\mu = 1/2$ and use the upper bounds on a, b and c obtained in the proof of Theorem IV.8 to obtain

$$\frac{5\sqrt{2}}{2}\sqrt{\varepsilon} + \frac{27}{2}\varepsilon + 8\sqrt{2}\varepsilon\sqrt{\varepsilon} + 32\varepsilon^2 > \frac{1}{720} + \frac{1}{72N^2} - \frac{11}{720N^4}.$$

Since the right hand side is a decreasing function of N for $N > 3/2$, and it tends to $1/720$ at infinity, it is also true that

$$\frac{5\sqrt{2}}{2}\sqrt{\varepsilon} + \frac{27}{2}\varepsilon + 8\sqrt{2}\varepsilon\sqrt{\varepsilon} + 32\varepsilon^2 > \frac{1}{720}.$$

This implies that $\varepsilon > 1.53859 \times 10^{-7}$, thereby proving our theorem. ∎

Since the last theorem proved was just the lower bound of Theorem I.2, we conclude that the proof of Theorem I.2 is now also complete.

E. The upper bound for the variance

Proof: (Proof of Theorem I.3, upper bound.) We shall use Lemma IV.3 to construct assignments with small maximal variances V. Suppose that N is even (odd N can be handled as in Section IV-C). Let A be an $(N/2)$-by-$(N/2)$ matrix with small maximal variance $c_A \cdot (N/2)^4$ (containing the numbers 1 through $N^2/4$). Let B be an (optimal) infinite assignment of the integers to the infinite square grid where each line and column contains exactly two numbers.[10] The variance in any row or column of B is of course $1/4$. Let us apply Lemma IV.3 to A and B. The result is an assignment containing N numbers in each row and column, with a maximal row or column variance of

$$\left(\frac{1}{64} + \frac{c_A}{16}\right) N^4.$$

Let us now consider what this gave us. In Theorem I.2, we gave an explicit construction for each N for an N-by-N square of maximal variance $N^4/16 + O(N^4)$. Choosing this for A we get an upper bound of $(5/256)N^4 + O(N^3) = (1/51.2)N^4 + O(N^3)$ for the infinite case. In Section IV-C we also showed that by blowing up examples for small N, we can get squares with maximal variance no more than $N^4/20$. This gives us for the infinite case a construction with an upper bound of $(3/160)N^4 = N^4/53\frac{1}{3}$. ∎

Remark. For N which are powers of two, we can construct squares with maximal variances $N^4/22$ (see Section IV-C). For the infinite case, this yields constructions with maximal variances of $N^4/54.15$. However, with this method we cannot get close to the lower bound, because by Theorem I.2, the maximal variance for an N-by-N square is always at least $N^4/24$, but even a square of maximal variance $N^4/24$ would give an infinite construction with $(7/384)N^4 = N^4/54.857$. It is far from clear if the method described in this section gives the optimal construction. On the other hand, we believe that the lower bound in Theorem I.3 is not sharp, either.

F. The lower bound for the variance

Let us fix N. Let us consider an assignment to the infinite grid, where each column and row contains N numbers, and the maximum of the variances is (finite and) as small as possible. Let d denote the maximal spread of this assignment and V the maximum of the variances of rows and columns. Trivially

$$\frac{d^2}{2N} \le V.$$

If $V > N^4/2$ then the lower bound of Theorem I.3 is satisfied, and there is nothing to worry about. Otherwise, we can deduce that

$$d < N^{5/2}.$$

[10] More precisely, assign $2n$ to the position (n, n) and $2n + 1$ to position $(n + 1, n)$.

We say that the distance of two numbers in the assigment is 1 if they are in the same row or column. They are in distance 2 from each other if they are not in the same line and there is a third number which is in distance 1 from both numbers. In generally, two numbers are at distance $k+1$ from each other if their distance is not $\leq k$ and there is a third number which is at distance k from one of the numbers, and 1 from the other number. Denote $B(k,x)$ the set of numbers which are located within distance k from x, and let $S(k,x) = B(k,x) \setminus B(k-1,x)$. By definition of $B(k,x)$, we have

$$\max\{y|y \in B(k,x)\} \leq x + kd$$

and

$$\min\{y|y \in B(k,x)\} \geq x - kd,$$

hence

$$\#B(k,x) \leq 2kd + 1.$$

Since

$$\cup_{i=0}^{k} S(i,x) = B(k,x)$$

and the sets $S(i,x)$ are disjoint from each other, there exist arbitrarily large integers k such that

$$\#S(k,x) \leq 2d. \tag{19}$$

We shall need the following lemma.

Lemma IV.10: Let $x_1, \dots, x_N, x, y, m$ be real numbers and N a positive integer. Then

(i)

$$\operatorname{Var}(x_1, \dots, x_N) = \frac{N-1}{N^2} \sum x_i^2 - \frac{2}{N^2} \sum_{i<j} x_i \cdot x_j.$$

(ii)

$$\sum_{i<j} (x_i - x_j)^2 = (N-1) \sum x_i^2 - 2 \sum_{i<j} x_i \cdot x_j.$$

(iii)

$$\operatorname{Var}(x_1, \dots, x_N) = \frac{1}{N^2} \sum_{i<j} (x_i - x_j)^2.$$

(iv)

$$(x-y)^2 = 2(m-x)^2 + 2(m-y)^2 - 4\left(m - \frac{x+y}{2}\right)^2.$$

(v)

$$\sum_{i=1}^{N^2} (x-i)^2 \geq \frac{N^6 - N^2}{12}.$$

(vi) Given $A_1, \dots, A_n, B_1, \dots, B_n$. Then

$$\sum_{i,j} (A_i - B_j)^2 = \sum_{i<j} (A_i - A_j)^2 + \sum_{i<j} (B_i - B_j)^2 + (A_1 + \dots + A_n - B_1 - \dots - B_n)^2.$$

(vii) Fix a row R of the assigment. Then

$$\sum_{x \in R, y \in V(R)} (x-y)^2 \geq \frac{N^7 - N^3}{12} + N^3 \operatorname{Var}(R).$$

Proof: (i) follows from definition. (ii), (iv) and (vi) are easy identities. (iii) follows from (i) and (ii). (v) follows from $\operatorname{Var}(1, \dots, N^2) = (N^4 - 1)/12$. To prove (vii), we shall use (v) with the remark that $N(R)$ contains N^2 distinct numbers. Let

$$m = \frac{1}{N^2} \sum_{y \in N(R)} y.$$

Then we have

$$\sum_{x \in R, y \in V(R)} (x-y)^2 = \sum_{x \in R} \sum_{y \in V(R)} (x-y)^2 = \sum_{x \in R} \sum_{y \in V(R)} ((x-m) + (m-y))^2 =$$

$$= N^2 \sum_{x \in R} (x-m)^2 + N \sum_{y \in V(R)} (m-y)^2 \geq N^3 \mathrm{Var}(R) + \frac{N^7 - N^3}{12}.$$

■

Now we are ready to start to prove the main result of the section.

Proof: (Proof of Theorem I.3, lower bound.) Fix a row R. Recall that $C(z)$ denotes the column of the number z. Let

$$\mathrm{Var}_x(R) = \frac{1}{N^2} \sum_{x_i \in R} (x - x_i)^2. \tag{20}$$

By Lemma IV.10(iii),

$$\sum_{x \in R} \mathrm{Var}_x(R) = 2\mathrm{Var}(R). \tag{21}$$

Consider the sum

$$S(R) = \sum_{x \in R, y \in V(R)} (x-y)^2. \tag{22}$$

By Lemma IV.10(vii) we have

$$S(R) \geq \frac{N^7 - N^3}{12} + N^3 \mathrm{Var}(R) \tag{23}$$

For the right hand side of (22) we apply Lemma IV.10(iv), with $m = R \cap C(y)$ (recall that $V(R)$ denotes the numbers within distance 1 from some element of R):

$$S(R) = \sum_{x \in R, y \in V(R), z = R \cap C(y)} \left(2(x-z)^2 + 2(y-z)^2 - 4\left(z - \frac{x+y}{2}\right)^2 \right). \tag{24}$$

For the first term of the summation from (24) we apply Lemma IV.10(iii):

$$2N \sum_{x, z \in R} (x-z)^2 = 4N^3 \cdot \mathrm{Var}(R).$$

For the second term, from (20) we obtain

$$2N \sum_{z \in R, y \in C(z)} (y-z)^2 = 2N^3 \sum_{z \in R} \mathrm{Var}_z(C(z)).$$

To handle the third term, we use Lemma IV.10(vi) with $A_i = z - x_i$ and $B_i = -z + y_j$:

$$\sum_{x \in R} \sum_{z \in R} \sum_{y \in C(z)} 4\left(z - \frac{x+y}{2}\right)^2 = \sum_{x \in R} \sum_{z \in R} \sum_{y \in C(z)} (2z - x - y)^2$$

$$\geq \sum_{z \in R} \left(\sum_{x_i < x_j \in R} (x_i - x_j)^2 + \sum_{y_i < y_j \in C(z)} (y_i - y_j)^2 \right) = N^3 \mathrm{Var}(R) + N^2 \sum_{z \in R} \mathrm{Var}(C(z)).$$

Combining the last three displayed statements, we get that

$$\frac{N^7 - N^3}{12} + N^3 \mathrm{Var}(R) \leq S(R) \leq 3N^3 \mathrm{Var}(R) + 2N^3 \sum_{z \in R} \mathrm{Var}_z(C(z)) - N^2 \sum_{z \in R} \mathrm{Var}(C(z)).$$

We can conclude that

$$\frac{N^4 - 1}{12} \leq 2\mathrm{Var}(R) + 2\sum_{z \in R} \mathrm{Var}_z(C(z)) - \frac{1}{N} \sum_{z \in R} \mathrm{Var}(C(z)). \tag{25}$$

Similarly, if R is any column, we get

$$\frac{N^4 - 1}{12} \leq 2\mathrm{Var}(R) + 2\sum_{z \in R} \mathrm{Var}_z(R(z)) - \frac{1}{N} \sum_{z \in R} \mathrm{Var}(R(z)), \tag{26}$$

where $R(z)$ of course stands for the row of the number z.

Let us now use the word *line* to mean either a row or column. Let k be a large integer with

$$\#S(k,x) \leq 2d. \tag{27}$$

Consider the complete set of lines containing an element of $B(k-1,x)$. Let call the cardinality of this set t. For each of these rows (resp. columns), add up the statement of equation (25) (resp. equation (26)) for that row (resp. column). The number of lines which are not entirely contained in $B(k-1,x)$ is at most $4d$, because such lines must each contain an element of $S(k,x)$, and any element of $S(k,x)$ can only be contained in two such lines (a row and a column). In the large sum, on the left hand side we get $t(1/12)(N^4-1)$.

In order to calculate the right hand side of the large sum, we consider the lines contained in $B(k-1,x)$ separately from the lines that only intersect $B(k-1,x)$ (the latter category only contains at most $4d$ lines). If R is a row contained in $B(k-1,x)$, $\mathrm{Var}(R)$ will be contained with multiplicity 5 on the right hand side of the large sum, namely,

- the first term of the right hand side of (25) yields $\mathrm{Var}(R)$ with multiplicity $+2$;
- the second term of the right hand side of (26) yields $\mathrm{Var}(R)$ with multiplicity $+4$ (we get a term from each of the columns intersecting R and combine them using (21);
- the third term on the right hand side of (26) yields $\mathrm{Var}(R)$ with multiplicity -1.

Similarly, for any column R contained in $B(k-1,x)$, we get $\mathrm{Var}(R)$ on the right hand side of the large sum with multiplicity 5.

Each of the at most $4d$ lines not entirely contained in $B(k-1,x)$ contributes at most $2N+1$ terms to the right hand side of the large sum, and each of these terms is bounded in absolute value independently of k. Thus, letting k tend to infinity[11], we obtain that the lim inf of the average variance of the lines intersecting $B(k-1,x)$ is at least $(N^4-1)/60$. Thus, there must be at least one row or column whose variance is at least $(N^4-1)/60$. This concludes the proof of our theorem. ■

Acknowledgments

The authors thank Vinay Vaishampayan, Suhas Diggavi, Jim Reeds, Gábor Pete, Jeff Lagarias, David Applegate and Henry Landau for helpful discussions.

References

[1] A.A. El Gamal and T.M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 851–857, Nov. 1982.
[2] C.E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, vol. 4, pp. 142–163, Mar. 1959.
[3] J.G. Proakis, *Digital Communications*, Macgraw–Hill, New York, 1983.
[4] V.A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 93, pp. 821–834, 1993.
[5] Tanya Y. Berger-Wolf and Edward M. Reingold, "Index assignment for multichannel communication under failure," *IEEE Trans. Inform. Theory*, 2002, to appear.
[6] L.H. Harper Tanya Y. Berger-Wolf and Mitchell A. Harris, "Bandwidth of clique products of cliques," preprint, 2002.
[7] V.A. Vaishampayan, ," personal communication, 2001.
[8] Imre Ruzsa and Gábor Székely, *Algebraic Probability Theory*, Wiley, New York, 1988.
[9] T. Lewis, "The factorization of the rectangular distribution," *J. Appl. Probability*, vol. 4, pp. 529–542, 1967.
[10] A. Tortrat, "Sur un théorème de lewis et la décomposition en facteurs premiers de la loi rectangulaire," *J. Appl. Probability*, vol. 6, pp. 177–185, 1969.

[11] We must only use k satisfying (27), but this is possible by the argument leading up to (19).