

Non-independent Randomized Rounding, Linear Discrepancy and an Application to Digital Halftoning

Benjamin Doerr*

Abstract

Motivated by an application from image processing (Asano et al., SODA 2002), we investigate the problem to round a given $[0, 1]$ -valued matrix to a $\{0, 1\}$ matrix such that the L_1 rounding error with respect to all 2×2 boxes is small. We present a randomized algorithm computing roundings with expected error at most 0.5463 per box. Our algorithm is the first one solving this problem fast enough for practical application, namely in linear time. We use a modification of randomized rounding. Instead of independently rounding the variables, we impose a number of suitable dependencies. This reduces the rounding error significantly compared to independent randomized rounding, which leads to an expected error of 0.8294 per box. Finally, we give a characterization of realizable dependencies.

1 Introduction

In this paper, we are concerned with rounding problems. In general form, these problems are of the following type: Given some numbers x_1, \dots, x_n , one is looking for roundings y_1, \dots, y_n such that some given error measures are small. By rounding we always mean that $y_i = \lfloor x_i \rfloor$ or $y_i = \lceil x_i \rceil$. Since there are 2^n possibilities, such rounding problems are good candidates for hard problems. In fact, even several restricted versions like the discrepancy problem are known to be NP -hard.

On the other hand, there are cases that can be solved optimally in polynomial time. Knuth [4] for example has shown that given a permutation π there exists a rounding such that the errors $|\sum_{i=1}^k (y_i - x_i)|$ and $|\sum_{i=1}^k (y_{\pi(i)} - x_{\pi(i)})|$ are at most $\frac{n}{n+1}$ for all $1 \leq k \leq n$. Such roundings can be obtained by computing a maximum flow in a network. A recent generalization to arbitrary totally unimodular rounding problems can be found in [2].

*Address: Mathematisches Seminar II, Christian-Albrechts-Universität zu Kiel, 24098 Kiel, Germany. Email: bed@numerik.uni-kiel.de. Partially supported (associate member) by the graduate school 'Effiziente Algorithmen und Multiskalenmethoden', Deutsche Forschungsgemeinschaft.

2 The Digital Halftoning Problem: A Matrix Rounding Problem

The rounding problem regarded in this paper is motivated by an application from image processing, namely the digital halftoning problem. The digital halftoning problem is to convert a continuous-tone intensity image (each pixel may have an arbitrary ‘color’ on the white-to-black scale) into a binary image (only black and white dots are allowed). An intensity image can be represented by a $[0, 1]$ -valued $m \times n$ matrix A . Each entry a_{ij} corresponds to the brightness level of the pixel with coordinates (i, j) . Since many devices, e.g., laser printers, can only output white and black dots, we have to round A towards a $0, 1$ matrix. Naturally, this has to be done in a way that the resulting image looks similar to the original one.

This notion of similarity is a crucial point. From the viewpoint of application, similarity is defined via the human visual system: A rounding is good, if an average human being can retrieve most of the original information from the rounded image. What was missing so far from the theoretical point of view is a good mathematical formulation of similarity. So far, the most widely accepted criterion for a good halftoning algorithm is that it has the ‘blue noise’ property (cf. the surveys Ulichney [9] and Lau and Arce [5]). This refers not to a similarity measure comparing two images, but an analysis on how the algorithm performs on constant grey level areas. Thus, on the other hand, this criterion gives little information on how changing intensities, in particular, shapes, are reproduced. On last year’s SODA conference, Asano et al. [1] reported that they made progress towards finding a similarity measure. Their experimental results indicate that good digital halftonings have small error with respect to all 2×2 subregions. More formally, we end up with this problem:

Let $A \in [0, 1]^{m \times n}$ denote our input matrix. The set $R_{ij} := \{(i, j), (i+1, j), (i, j+1), (i+1, j+1)\}$ for some $i \in [m-1], j \in [n-1]$ is called a 2×2 subregion (or box) in $[m] \times [n]$.¹ Denote by \mathcal{R} the set of all these boxes. We write $A_{R_{ij}}$ for the 2×2 matrix $\begin{pmatrix} a_{i,j} & a_{i,j+1} \\ a_{i+1,j} & a_{i+1,j+1} \end{pmatrix}$ induced by R_{ij} . For any matrix A put $\Sigma A := \sum_{i,j} a_{ij}$.

For a matrix $B \in \{0, 1\}^{m \times n}$ — which by definition is a rounding of A — we define the rounding error of A with respect to B by

$$d_{\mathcal{R}}(A, B) := \sum_{R \in \mathcal{R}} |\Sigma A_R - \Sigma B_R|.$$

We usually omit the subscript \mathcal{R} when there is no danger of confusion.

¹For an arbitrary number r we denote by $[r]$ the set of positive integers not exceeding r .

3 Results

Asano et al. [1] exhibited that roundings B such that $d(A, B)$ is small, yield good digital halftonings. They showed that for any A an optimal rounding B^* satisfies $d(A, B^*) \leq \frac{3}{4}|\mathcal{R}|$. They also gave a polynomial time algorithm computing a rounding B such that $d(A, B) \leq d(A, B^*) + 0.5625|\mathcal{R}|$. It is easy to see that there are matrices A such that all roundings (and in fact all integral matrices) B have $d(A, B) \geq \frac{1}{2}|\mathcal{R}|$.

A major draw-back of the algorithm given in [1] is that it is not very practical. It requires the solution of an integer linear program with totally unimodular constraint matrix. This leads to a run-time bound which is at least quadratic in the number nm of pixels. As pointed out by the authors, this is too slow for digital halftoning applications. We obtain the following result.

Theorem 1. *There is a randomized algorithm computing in linear time a rounding B of $A \in [0, 1]^{m \times n}$ such that*

- (i) *the expected rounding error satisfies $E(d(A, B)) \leq 0.5463|\mathcal{R}|$;*
- (ii) *if B^* is an optimal rounding, then $E(d(A, B)) \leq d(A, B^*) + 0.3125|\mathcal{R}|$;*
- (iii) *for all $\varepsilon > 0$, $\Pr(d(A, B) > E(d(A, B)) + \varepsilon|\mathcal{R}|) < 3 \exp(-\frac{1}{6}\varepsilon^2(m - 3))$.*

Our algorithm can be derandomized. This yields a deterministic linear-time algorithm computing roundings with error guarantee at most the expected one of the randomized version. Using the language of discrepancy theory, our result states that the L_1 linear discrepancy of a set \mathcal{R} of 2×2 boxes in the grid is at most $0.5463|\mathcal{R}|$. For an experimental study of how well our algorithm performs in practise, we refer to Schnieder [8]. Some of these results are also contained in [3].

4 Non-independent Randomized Rounding

The key idea of our algorithm might also be of a broader interest. We develop a randomized rounding scheme where the individual roundings are not independent. The classical approach of randomized rounding due to Raghavan and Thompson [6, 7] is to round each variable x_i independently with probabilities given by the fractional part $\{x_i\}$ of its value, i.e., $P(y_i = \lfloor x_i \rfloor + 1) = \{x_i\} = 1 - \Pr(y_i = \lfloor x_i \rfloor)$. Independence of the roundings allows to use Chernoff-type large deviation inequalities showing that a sum of independent random variables is highly concentrated around its expectation.

In addition to the randomized rounding condition for single variables, we impose dependencies of the type

$$P\left(\sum_{i \in I_k} y_i = \left\lfloor \sum_{i \in I_k} x_i \right\rfloor + 1\right) = \left\{ \sum_{i \in I_k} x_i \right\} = 1 - P\left(\sum_{i \in I_k} y_i = \left\lfloor \sum_{i \in I_k} x_i \right\rfloor\right) \quad (1)$$

for some sets I_k . Hence our roundings not only are randomized roundings with respect to single variables, but also with respect to the sums $\sum_{i \in I_k} x_i$. The roundings B computed in Theorem 1 satisfy

- (i) for all $i \in [\frac{m}{2}]$, $j \in [n-1]$, $R := \{2i-1, 2i\} \times \{j, j+1\}$, $\Sigma B|_R$ is a randomized rounding of $\Sigma A|_R$,
- (ii) for all $i \in [\frac{m}{2}]$, $j \in [n]$, $b_{2i-1,j} + b_{2i,j}$ is a randomized rounding of $a_{2i-1,j} + a_{2i,j}$,
- (iii) for all $i \in [m]$, $j \in [n-1]$, $b_{i,j} + b_{i,j+1}$ is a randomized rounding of $a_{i,j} + a_{i,j+1}$,
- (iv) for all $i \in [m]$, $j \in [n]$, $b_{i,j}$ is a randomized rounding of $a_{i,j}$.
- (v) $b_{i,j}$ is mutually independent from all $b_{i',j'}$ such that $\lceil i/2 \rceil \neq \lceil i'/2 \rceil$.

Such roundings can be computed in linear time (in mn) and yield the bounds claimed in Theorem 1.

We also regard the question what dependencies can be realized. More formally, we ask how the sets I_k have to be chosen, such that a randomized rounding satisfying (1) exists for all values of the variables. Surprisingly, there is a simple characterization:

Theorem 2. *Let $\mathcal{H} = (I, \mathcal{E})$ be a hypergraph. The following two properties are equivalent:*

- (i) *For all $(x_i)_{i \in I}$ there is a randomized rounding $(y_i)_{i \in I}$ satisfying (1) for all $E \in \mathcal{E}$.*
- (ii) *\mathcal{H} is totally unimodular.*

Thus, in particular, roundings B such that $\Sigma B|_R$ is a randomized rounding of $\Sigma A|_R$ for all $R \in \mathcal{R}$ — which would be perfect for our purposes — do not exist for all A .

5 Open Problems

From the work of Asano et al. [1] and this work, several open problems and areas for future research arise. Both motivate a closer investigation of L_p , $p < \infty$, discrepancy problems, which attracted less attention than the L_∞ discrepancy. They also suggest the study of

discrepancy problems where the rounding error depends on small regions (like here, 2×2 boxes). The latter is particularly interesting since many of the commonly used methods like randomization together with large deviation bounds fail.

A problem simple to state but seemingly hard to solve in this context is the one of the L_∞ linear discrepancy of the 2×2 boxes, i.e., what is the smallest k such that for any $A \in [0, 1]^{m \times n}$ there is a $B \in \{0, 1\}^{m \times n}$ such that $|\Sigma A_R - \Sigma B_R| \leq k$ holds for all 2×2 boxes R . Asano (private communication) has shown $1 \leq k \leq 5/3$. Since greedy rounding trivially yields an upper bound of 2, one has the feeling that this problem is not too well understood.

Of course, the same problem for the L_1 discrepancy, i.e., the approximation problem regarded in this paper, is also not completely understood, and it seems to be an interesting question whether the lower bound of $0.5|\mathcal{R}|$ is sharp or not.

References

- [1] T. Asano, N. Katoh, K. Obokata, and T. Tokuyama. Matrix rounding under the L_p -discrepancy measure and its application to digital halftoning. In *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 896–904, 2002.
- [2] B. Doerr. Lattice approximation and linear discrepancy of totally unimodular matrices. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 119–125, 2001. Journal version to appear in *Combinatorica*.
- [3] B. Doerr and H. Schnieder. Non-independent randomized rounding and an application to digital halftoning. In R. Möhring and R. Raman, editors, *Proceedings of ESA 2002*, volume 2461 of *Lecture Notes in Computer Science*, pages 399–410, Berlin–Heidelberg, 2002. Springer Verlag.
- [4] D. E. Knuth. Two-way rounding. *SIAM J. Discrete Math.*, 8:281–290, 1995.
- [5] D. L. Lau and G. R. Acre. *Modern Digital Halftoning*, volume 8 of *Signal Processing and Communication*. Marcel Dekker, Inc., New York, Basel, 2001.
- [6] P. Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *J. Comput. Syst. Sci.*, 37:130–143, 1988.
- [7] P. Raghavan and C. D. Thompson. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7:365–374, 1987.
- [8] H. Schnieder. *Digital Halftoning through Non-independent Randomized Rounding*. Diploma Thesis, Christian–Albrechts–Universität zu Kiel, 2003.
- [9] R. Ulichney. *Digital Halftoning*. MIT Press, 1987.