Generating Randomized Roundings with Cardinality Constraints and Derandomizations

Benjamin Doerr

Max–Planck–Institut für Informatik, Saarbrücken, Germany.

Abstract. We provide a general method to generate randomized roundings that satisfy cardinality constraints. Our approach is different from the one taken by Srinivasan (FOCS 2001) and Gandhi et al. (FOCS 2002) for one global constraint and the bipartite edge weight rounding problem.
Also for these special cases, our approach is the first that can be derandomized. For the bipartite edge weight rounding problem, in addition, we gain an $\tilde{O}(|V|)$ factor run-time improvement for generating the randomized solution.
We also improve the current best result on the general problem of derandomizing randomized roundings. Here we obtain a simple $O(mn \log n)$ time algorithm that works in the RAM model for arbitrary matrices with entries in $\mathbb{Q}_{\geq 0}$. This improves over the $O(mn^2 \log(mn))$ time solution of Srivastav and Stangier.

1 Introduction and Results

Many combinatorial optimization problems can easily be formulated as integer linear programs (ILPs). Unfortunately, solving ILPs is NP–hard, whereas solving linear programs (without integrality constraints) is easy, both in theory and practice. Therefore, a natural and widely used technique is to solve the linear relaxation of the ILP and then transform its solution into an integer one.

Typically, this requires rounding a vector x to an integer one y in such a way that the rounding errors $|(Ax)_i - (Ay)_i|$, $i \in [m] := \{1, \ldots, m\}$, are small for some given $m \times n$ matrix A.

1.1 Randomized Rounding

A very successful approach to such rounding problems is the one of *randomized rounding* introduced by Raghavan and Thompson [RT87,Rag88]. Here the integer vector y is obtained from x by rounding each component j independently with probabilities derived from the fractional part of x_j. In particular, if $x \in [0,1]^n$, we have $\Pr(y_j = 1) = x_j$ and $\Pr(y_j = 0) = 1 - x_j$ independently for all $j \in [n]$.

Since the components are rounded independently, the rounding error $|(Ax)_i - (Ay)_i|$ in constraint i is a sum of independent random variables. Thus it is highly concentrated around its mean, which by choice of the probabilities is zero. Large deviation bounds like the Chernoff inequality allow to quantify such violations and thus yield performance guarantees. The derandomization problem is to transform this randomized approach into deterministic rounding algorithms that keep the rounding errors $|(Ax)_i - (Ay)_i|$ below some threshold.

1.2 Hard Constraints

Whereas the independence in rounding the variables ensures that the rounding errors $|(Ax)_i - (Ay)_i|$ are small, it is very weak in guaranteeing that a constraint is satisfied without error. We call a constraint *hard constraint* if we require our solution to satisfy it without violation. In this paper, we are mainly concerned with *cardinality constraints*. These are constraints on unweighted sums of variables. Let us give an example where such hard constraints naturally occur.

The *integer splittable flow problem* is the following routing problem. Given an undirected graph and several source–sink pairs (s_i, t_i) together with integral demands d_i, we are looking for integer flows f_i from s_i to t_i having flow value d_i such that the maximum edge congestion is minimized. Solving the non-integral relaxation and applying path stripping (cf. [GKR$^+$99]), we end up with this rounding problem: Round a solution $(x_P)_P$ of the linear system

$$\begin{aligned} \text{Minimize } W \text{ s. t. } \quad & \sum_{P \ni e} x_P \leq W, \ \forall e \\ & \sum_{P \in \mathcal{P}_i} x_P = d_i, \ \forall i \\ & x_P \geq 0, \ \forall P \end{aligned}$$

to an integer one such that the first set of constraints is violated not too much and the second one is satisfied without any violation (hard constraints).

Further examples of rounding problems with hard constraints include other routing applications ([RT91,Sri01]), many flow problems ([RT87,RT91,GKR$^+$99]), partial and capacitated covering problems ([GKPS02,GHK$^+$03]) and the assignment problem with extra constraints ([AFK02]).

For the special case of the above problem where all d_i are one, Raghavan and Thompson [RT87] presented an easy solution: For each i, they pick one $P \in \mathcal{P}_i$ with probability x_P and then set $y_P = 1$ and $y_{P'} = 0$ for all $P' \in \mathcal{P}_i \setminus \{P\}$. For the general case, however, this idea and all promising looking extensions fail. Guruswami et al. [GKR$^+$99] state on the integral splittable flow problem (ISF) in comparison to the unsplittable flow problem that "standard roundings techniques are not as easily applied to ISF".

At FOCS 2001, Srinivasan [Sri01] presented a way to compute randomized roundings that respect the constraint that the sum of all variables remains unchanged (one *global cardinality constraint*) and fulfill some negative correlation properties.

Gandhi, Khuller, Parthasarathy and Srinivasan [GKPS02] combined the deterministic "pipage rounding" algorithm of Ageev and Sviridenko [AS04] with Srinivasan's approach to obtain randomized roundings of edge weights in bipartite graphs that are degree preserving. By this we mean that the sum of weights of all edges incident with some vertex is not changed by the rounding. The roundings of Gandhi et al. also fulfill negative correlation properties, but only on sets of edges incident with a common vertex.

Both Srinivasan [Sri01] and Gandhi et al. [GKPS02] do not consider the derandomization problem. A first derandomization of Srinivasan's [Sri01] roundings was given in [Doe05]. For the bipartite edge weight rounding problem, Ageev and Sviridenko [AS04] state that any randomized rounding algorithm "will be too sophisticated to admit derandomization".

1.3 Our Contribution

In this paper, we extend the work of [Doe05] in several directions.

Randomized roundings with constraints. We show that for all sets of cardinality constraints, the general problem of generating randomized roundings can be reduced to the one for $\{0, \frac{1}{2}\}$ vectors. This immediately yields a simpler way to generate the randomized roundings used in Srinivasan [Sri01], Gandhi et al. [GKPS02], Sadakane, Takki-Chebihi and Tokuyama [STT01] and [Doe04]. For the rounding problem of [GKPS02], we even gain an $\tilde{O}(|V|)$ factor in the run-time.

Derandomizations with constraints. Since our approach is structurally simpler than the earlier ones, we do obtain the corresponding derandomizations. In fact, we may even use classical derandomizations like Raghavan's. In consequence, derandomizing randomized rounding approaches for the bipartite edge weight rounding problem is much easier than what is conjectured in Ageev and Sviridenko [AS04]. Note that this derandomization is more than a re-invention of the original algorithm of Ageev and Sviridenko. It also keeps those rounding errors small for which the randomized approach allowed the use of Chernoff type large deviation bounds.

Counter-examples. We also show that a number of natural properties of independent randomized roundings may not hold in the presence of constraints. For example, let $f : [0,1]^n \to \mathbb{R}$ be non-decreasing, $x, x' \in [0,1]^n$ and y, y' be independent randomized roundings of x, x' respectively. Then $x \leq x'$ (component-wise) implies $E(f(y)) \leq E(f(y'))$. We show that already a single cardinality constraint may inflict that no randomized roundings respecting this constraint have the above property.

General randomized rounding derandomization. Our final result is not to be overlooked due to its simple less than a page proof. Here we give an easy $O(mn \log n)$ time derandomization for arbitrary constraint matrices $A \in ([0,1] \cap \mathbb{Q})^{m \times n}$ that works in the RAM model. This improves over the $O(mn^2 \log(mn))$ time (and 30 pages) landmark solution of Srivastav and Stangier [SS96]. Note that Raghavan's derandomization [Rag88] needs to compute the exponential function and in consequence in the RAM model only works for binary matrices (as pointed out in Section 2.2 of his paper).

2 Randomized Rounding, Constraints and Correlation

For a number r write $[r] = \{n \in \mathbb{N} \,|\, n \leq r\}$, $\lfloor r \rfloor = \max\{z \in \mathbb{Z} \,|\, z \leq r\}$, $\lceil r \rceil = \min\{z \in \mathbb{Z} \,|\, z \geq r\}$ and $\{r\} = r - \lfloor r \rfloor$. We write $z \approx r$ if $z \in \{\lfloor r \rfloor, \lceil r \rceil\}$. We use these notations for vectors as well (component-wise).

Let $x \in \mathbb{R}$. A real-valued random variable y is called *randomized rounding of* x if $\Pr(y = \lfloor x \rfloor + 1) = \{x\}$ and $\Pr(y = \lfloor x \rfloor) = 1 - \{x\}$. Since only the fractional parts of x and y are relevant, we usually have $x \in [0,1]$. In this case, we have

$$\Pr(y = 1) = x,$$
$$\Pr(y = 0) = 1 - x.$$

For $x \in \mathbb{R}^n$, we call $y = (y_1, \ldots, y_n)$ randomized rounding of x if y_j is a randomized rounding of x_j for all $j \in [n]$.

The algorithmic concept of randomized rounding can be formulated as follows: Fix a number $n \in \mathbb{N}$, the number of variables to be rounded. Let $X \subseteq [0,1]^n$. This is the set of vectors for which we allow randomized rounding. Typically, this will be $[0,1]^n$ or a suitably rich subset thereof. A family $(\Pr_x)_{x \in X}$ of probability distributions on $\{0,1\}^n$

is called *randomized rounding*, if for all $x \in X$, a sample y from $\Pr_x$ is a randomized rounding of x.

As described in the introduction, we are interested in roundings that satisfy some hard constraints. Though usually we will only regard cardinality constraints (requiring the sum of some variable to be unchanged), it will be convenient to encode hard constraints in a matrix B. Our aim then is that a rounding y of x satisfies $By = Bx$. Of course, if Bx is not integral, this can never be satisfied. We therefore relax the condition to $By \approx Bx$. In a randomized setting, we often obtain the slightly stronger condition that By is a randomized rounding of Bx.

Besides satisfying hard constraints we still want to keep other rounding errors small (as does independent randomized rounding). A useful concept here is the one of negative correlation, which implies Chernoff type large deviation inequalities.

We call a set $\{X_j \mid j \in S\}$ of binary random variables *negatively correlated* if for all $S_0 \subseteq S$, $b \in \{0,1\}$, we have $\Pr(\forall j \in S_0 : X_j = b) \le \prod_{j \in S_0} \Pr(y_j = b)$. As shown in [PS97], this implies the usual Chernoff-Hoeffding bounds on large deviations. The following version is not strongest possible, but sufficient for most purposes.

Lemma 1. *Let $\{X_j \mid j \in S\}$ be a set of negatively correlated binary random variables and $a_j \in [0,1]$, $j \in S$. Put $X = \sum_{j \in S} a_j X_j$ and $\mu = E(X)$. Then for all $\delta \in [0,1]$,*

$$\Pr(X \ge (1+\delta)\mu) \le \exp(-\tfrac{1}{3}\mu\delta^2),$$
$$\Pr(X \le (1-\delta)\mu) \le \exp(-\tfrac{1}{2}\mu\delta^2).$$

It turns out that hard constraints and negative correlation cannot always be achieved simultaneously. We therefore restrict ourselves to negative correlation on certain sets of variables. Let $\mathcal{S} \subseteq 2^{[n]}$ be closed under taking subsets, that is, $S_0 \subseteq S \in \mathcal{S}$ implies $S_0 \in \mathcal{S}$.

Definition 1. *We call* $(\Pr_x)$ randomized rounding with respect to B and $\mathcal{S}$, *if for all x a sample y from $\Pr_x$ satisfies the following.*

(A1) *y is a randomized rounding of x.*
(A2) *By is a randomized rounding of Bx.*
(A3) *For all $S \in \mathcal{S}$, $\forall b \in \{0,1\} : \Pr(\forall j \in S : y_j = b) \le \prod_{j \in S} \Pr(y_j = b)$.*

In this language, we know the following. Clearly, independent randomized rounding is a randomized rounding with respect to the empty matrix B and $\mathcal{S} = 2^{[n]}$. Srinivasan [Sri01] showed that for the $1 \times n$ matrix $B = (1 \ldots 1)$, randomized roundings with respect to B and $\mathcal{S} = 2^{[n]}$ exist and can be generated in time $O(n)$. Let $G = (V,E)$ be a bipartite graph and $B = (b_{ij})_{\substack{i \in V \\ j \in E}}$ its vertex-edge-incidence matrix. For $v \in V$ let $E_v = \{e \in E \mid v \in e\}$. Gandhi et al. [GKPS02] showed that there are randomized roundings with respect to B and $\mathcal{S} = \{E_0 \mid \exists v \in V : E_0 \subseteq E_v\}$. They can be generated in time $O(mn)$. From [Doe03,Doe04], we have that if B is totally unimodular, then randomized roundings with respect to B and $\mathcal{S} = \emptyset$ exist. Recall that a matrix is totally unimodular if each square submatrix has determinant -1, 0 or 1. If B is not totally unimodular, then not even for $X = \{0, \frac{1}{2}\}^n$ a randomized rounding $(\Pr_x)_{x \in X}$ with respect to B and $\mathcal{S} = \emptyset$ exists.

Throughout the paper let $A \in [0,1]^{m_A \times n}$ and $x \in [0,1]^n$. Let B be a totally unimodular $m_B \times n$ matrix.

3 Binary Reductions

A central step of our method is a reduction to the problem of rounding $\{0, \frac{1}{2}\}$ vectors, similar as in Beck and Spencer [BS84]. This reduced rounding problem turns out to be structurally and computationally much simpler than the general one. We start by describing the connection between the reduced and the general problem.

3.1 Randomized Roundings

Let $\mathcal{S} \subseteq 2^{[n]}$ be closed under taking subsets. Let $(\mathrm{Pr}_x)_{x \in \{0,\frac{1}{2}\}^n}$ be a family of probability distributions on $\{0,1\}^n$. We call the family (Pr_x) *basic randomized rounding with respect to B and $\mathcal{S}$*, if for all $x \in \{0, \frac{1}{2}\}^n$ a sample y from Pr_x satisfies (A1) to (A3) and

(A4) $\mathrm{Pr}_x(y) = \mathrm{Pr}_x(2x - y)$.

The key result of this subsection is that any basic randomized rounding can be extended to a randomized rounding $(\overline{\mathrm{Pr}}_x)$, where x ranges over all vectors in $[0,1]^n$ having finite binary length. The simple idea is to iterate basic randomized rounding digit by digit:

Digit by digit rounding: Let $x \in [0,1]^n$ having binary length ℓ (that is, all x_i can be written as $x_i = \sum_{j=0}^{\ell} d_j 2^{-j}$ with $d_j \in \{0,1\}$). There is nothing to show for $\ell = 0$, so assume $\ell \geq 1$. Write $x = x' + 2^{-\ell+1}x''$ with $x'' \in \{0, \frac{1}{2}\}^n$ and $x' \in [0,1]^n$ having binary length at most $\ell - 1$. Let y'' be a sample from the basic randomized rounding $\mathrm{Pr}_{x''}$. Set $\tilde{x} := x' + 2^{-\ell+1}y''$. Note that $\tilde{x}$ has binary length at most $\ell - 1$. Repeat this procedure until a binary vector is obtained. For each x having finite binary expansion, this defines a probability distribution $\overline{\mathrm{Pr}}_x$ on $\{0,1\}^n$.

Theorem 1. *Let (Pr_x) be a basic randomized rounding with respect to B and $\mathcal{S}$. Then $(\overline{\mathrm{Pr}}_x)$ with x ranging over all $[0,1]$ vectors having finite binary length is a randomized rounding with respect to B and $\mathcal{S}$.*

Proof. We proceed by induction. If $x \in \{0, \frac{1}{2}, 1\}^n$, we simply have $\mathrm{Pr}_x = \overline{\mathrm{Pr}}_x$. Let x therefore have binary length $\ell > 1$. Let $x = x' + 2^{-\ell+1}x''$ with $x'' \in \{0, \frac{1}{2}\}^n$ and $x' \in [0,1]^n$ having binary length at most $\ell - 1$. Let y'' be a sample from $\mathrm{Pr}_{x''}$. Set $\tilde{x} := x' + 2^{-\ell+1}y''$. Let y be a sample from $\overline{\mathrm{Pr}}_{\tilde{x}}$. By construction, y has distribution $\overline{\mathrm{Pr}}_x$.
(A1): Let $j \in [n]$. By induction,

$$\begin{aligned}\mathrm{Pr}(y_j = 1) &= \sum_{\varepsilon \in \{0,1\}} \mathrm{Pr}(y''_j = \varepsilon)\, \mathrm{Pr}(y_j = 1 \,|\, y''_j = \varepsilon) \\ &= \sum_{\varepsilon \in \{0,1\}} \mathrm{Pr}(y''_j = \varepsilon)(x'_j + 2^{-\ell+1}\varepsilon) = x_j.\end{aligned}$$

(A2): Let $i \in [m_B]$. If $Bx'' \in \mathbb{Z}$, then $Bx = B\tilde{x}$ with probability one and By is a randomized rounding of Bx for both y being a sample from $\overline{\mathrm{Pr}}_x$ and $\overline{\mathrm{Pr}}_{\tilde{x}}$. If $Bx'' \notin \mathbb{Z}$, then $\mathrm{Pr}((Bx'')_i = (Bx'')_i + \frac{1}{2}) = \mathrm{Pr}((Bx'')_i = (Bx'')_i - \frac{1}{2}) = \frac{1}{2}$ by (A2).

By induction, we have

$$\begin{aligned}
&\Pr((By)_i = \lfloor (Bx)_i \rfloor + 1) \\
&= \Pr\left((B\tilde{x})_i = (Bx)_i + 2^{-\ell}\right) \Pr\left((By)_i = \lfloor (Bx)_i \rfloor + 1 \,\middle|\, (B\tilde{x})_i = (Bx)_i + 2^{-\ell}\right) + \\
&\quad \Pr\left((B\tilde{x})_i = (Bx)_i - 2^{-\ell}\right) \Pr\left((By)_i = \lfloor (Bx)_i \rfloor + 1 \,\middle|\, (B\tilde{x})_i = (Bx)_i - 2^{-\ell}\right) \\
&= \tfrac{1}{2}(\{(Bx)_i\} + 2^{-\ell}) + \tfrac{1}{2}(\{(Bx)_i\} - 2^{-\ell}) = \{(Bx)_i\}.
\end{aligned}$$

(A3): Let $S \in \mathcal{S}$. Note that $\prod_{j \in S}(x'_j + 2^{-\ell+1}\varepsilon_j) + \prod_{j \in S}(x'_j + 2^{-\ell+1}(2x''_j - \varepsilon_j)) \leq 2\prod_{j \in S} x_j$ holds for all roundings ε of x''. Hence by induction and (A4),

$$\begin{aligned}
&\Pr(\forall j \in S : y_j = 1) \\
&= \sum_{\varepsilon \in \{0,1\}^n} \Pr(y'' = \varepsilon) \Pr((\forall j \in S : y_j = 1) \,|\, y'' = \varepsilon) \\
&\leq \sum_{\varepsilon \in \{0,1\}^n} \Pr(y'' = \varepsilon) \prod_{j \in S}(x'_j + 2^{-\ell+1}\varepsilon_j) \\
&= \tfrac{1}{2} \sum_{\varepsilon \in \{0,1\}^n} \Pr(y'' = \varepsilon) \Big(\prod_{j \in S}(x'_j + 2^{-\ell+1}\varepsilon_j) + \prod_{j \in S}(x'_j + 2^{-\ell+1}(2x''_j - \varepsilon_j)) \Big) \\
&\leq \tfrac{1}{2} \sum_{\varepsilon \in \{0,1\}^n} \Pr(y'' = \varepsilon) \cdot 2 \prod_{j \in S} x_j = \prod_{j \in S} x_j.
\end{aligned}$$

A similar argument shows the claim for $b = 0$. □

3.2 Derandomizations

In this subsection, we extend the binary expansion method to the derandomization problem. A *randomized rounding derandomization* (with constant c) is an algorithm that computes for given $A \in [0,1]^{m_A \times n}$ and $x \in [0,1]^n$ a $y \in \{0,1\}^n$ such that for all $i \in [m_A]$,

$$|(Ax)_i - (Ay)_i| \leq c\sqrt{\max\{(Ax)_i, \ln(2m_A)\} \ln(2m_A)}.$$

It thus achieves (with minor loss) the existential bounds given by randomized rounding.

The following derandomizations are known.

(i) If $A \in \{0,1\}^{m_A \times n}$ and $x \in \{0, \frac{1}{2}, 1\}^n$, then Spencer's [Spe94] method of conditional probabilities yields a straight-forward $O(m_A n)$–time derandomization with constant $c = \sqrt{\frac{1}{2}}$. Note that the conditional probabilities in this special case are easy to compute via binomial coefficients.

(ii) Raghavan's derandomization [Rag88] via so-called pessimistic estimators is more complicated, but allows a wider range of vectors. Still in time $O(m_A n)$, it achieves the constant $c = e - 1$. In the RAM model, it works for all $A \in \{0,1\}^{m_A \times n}$ and $x \in ([0,1] \cap \mathbb{Q})^n$. If one allows precise computations with real numbers in constant time (in particular exponential functions), then this extends to arbitrary $A \in [0,1]^{m_A \times n}$.

(iii) Srivastav and Stangier [SS96] give a derandomization for all $A \in ([0,1] \cap \mathbb{Q})^{m_A \times n}$ in the RAM model, though at the price of an increased run-time of $O(m_A n^2 \log(m_A n))$. Also, it is quite complicated from the view-point of implementation. The constant c is not explicitly stated in the paper, but by plugging in the inequality of Angluin-Valiant given there, one achieves $c = \sqrt{3}$.

(iv) In the last section of this paper, we show how to use the binary expansion ideas to obtain a relatively simple derandomization that works for all $A \in ([0,1] \cap \mathbb{Q})^{m_A \times n}$ and $x \in ([0,1] \cap \mathbb{Q})^n$ in time $O(m_A n \log n)$ in the RAM model. The constant in this case is $4(e-1)(1+o(1))$.

For $\ell \in \mathbb{N}$ and $c \in \mathbb{R}_{\geq 0}$ let $f(\ell, c) = c \sum_{i=1}^{\ell} 2^{-(i-1)/2} \prod_{j=i+1}^{\ell} (1 + 2^{-(j-1)/2} c)^{1/2}$.

Theorem 2 (Digit by digit derandomization). *Let $\mathcal{A}$ be an algorithm which for some matrix A and any $x \in \{0, \frac{1}{2}, 1\}^n$ computes a rounding y of x such that $By \approx Bx$ and*

$$\forall i \in [m_A] : |(Ax)_i - (Ay)_i| \leq c\sqrt{\max\{(Ax)_i, \ln(2m_A)\} \ln(2m_A)}.$$

Then for each $x \in [0,1]^n$ having binary length ℓ, a rounding y such that $By \approx Bx$ and

$$\forall i \in [m_A] : |(Ax)_i - (Ay)_i| \leq f(\ell, c)\sqrt{\max\{(Ax)_i, \ln(2m_A)\} \ln(2m_A)}$$

can be computed by ℓ times invoking $\mathcal{A}$.

We omit the proof for reasons of space. Similar as in the proof of Theorem 1, we use induction over the length of the binary expansion. Some care has to be taken to control the size of $(A\tilde{x})_i$ for the intermediate roundings $\tilde{x}$.

We end this section with some rough estimates of the constants $f(c, \ell)$.

Lemma 2. *$f(c) := \lim_{\ell \to \infty} f(\ell, c)$ exists for all c and satisfies $f(c) = c^{O(\log c)}$. We have $f(\sqrt{\frac{1}{2}}) \leq 4$, $f(e-1) \leq 18$, and $f(\sqrt{3}) \leq 19$.*

Let us remark that the increase in the constant in most cases in not as bad as $f(c) = c^{\Theta(\log c)}$ suggests. If $\log m_A = o((Ax)_i)$, then a closer look at the proof of Theorem 2 yields $|(Ax)_i - (Ay)_i| \leq 2(\sqrt{2}+1)(1+o(1))c\sqrt{(Ax)_i \ln(2m_A)}$. Hence asymptotically we only lose a factor of less than 5 in the large deviation bound. In fact, already if $(Ax)_i \geq c^2 \ln(2m_A)$, we obtain a bound of $|(Ax)_i - (Ay)_i| \leq 7c\sqrt{(Ax)_i \ln(2m_A)}$.

4 Randomized Roundings with Disjoint Constraints

We now use the binary expansion method developed in the previous section to generate randomized roundings that satisfy disjoint cardinality constraints. Hence throughout this section ;et $B \in \{0,1\}^{m_B \times n}$ and $\|B\|_1 := \max_j \sum_i |b_{ij}| = 1$. For the generation of the roundings, this is a microscopic extension of Srinivasan's [Sri01] result. The reader's focus should therefore be on the simplicity of our approach.

As should be clear by now, all we have to do is analyze the $\{0, \frac{1}{2}\}$ case. Let us assume that B is stored in some $O(n)$ space datastructure allowing amortized linear time enumerations of the sets $\{j \in [n] \mid b_{ij} = 1\}$ for all $i \in [m_B]$.

Lemma 3. *There are basic randomized roundings $(\Pr_{x,B})$ with respect to B and $2^{[n]}$. A sample from $(\Pr_{x,B})$ can be generated in time $O(n)$.*

Proof. Let $x \in \{0, \frac{1}{2}\}^n$. For $i \in [m_B]$ let $E_i := \{j \in [n] \mid x_j = \frac{1}{2}, b_{ij} = 1\}$. Choose a set $\mathcal{M}$ of disjoint 2–subsets of $[n]$ such that $|E_i \setminus \bigcup \mathcal{M}| \leq 1$ and $|M \cap E_i| \neq 1$ hold for all $i \in [m_B]$ and $M \in \mathcal{M}$. In other words, $\mathcal{M}$ is a maximal collection of disjoint 2–sets of $[n]$ that all intersect all E_i in a trivial way[1].

[1] As we will see, the particular choice of $\mathcal{M}$ is completely irrelevant. Assume therefore that we have fixed some deterministic way to choose it (e.g., greedily in the natural order of $[n]$).

For each $\{j_1, j_2\}$ independently we flip a coin to decide whether $(y_{j_1}, y_{j_2}) = (1, 0)$ or $(y_{j_1}, y_{j_2}) = (0, 1)$. For all $j \in [n] \setminus \bigcup \mathcal{M}$ let y_j be a randomized rounding of x_j independent from all other random choices. The above defines a basic randomized rounding $(\Pr_{x,B})$ with respect to B and $2^{[n]}$. □

From Theorem 1 and 3, the following is immediate.

Theorem 3. *There are randomized roundings* $(\overline{\Pr}_{x,B})$ *with respect to* B *and* $2^{[n]}$*. A sample from* $(\overline{\Pr}_{x,B})$ *can be generated in time* $O(n\ell)$*, where* ℓ *is the binary length of* x*.*

We now derandomize the construction above. Here the simpler, compared to previous work more sequential construction proves to be advantageous. As before, we only have to analyze the $0, \frac{1}{2}$ case.

Lemma 4. *Let* A *be an* $m_A \times n$ *matrix. Let* $x \in \{0, \frac{1}{2}\}^n$*. Then a binary vector* y *such that* $By \approx Bx$ *and*

$$\forall i \in [m_A] : |(Ax)_i - (Ay)_i| \leq 2c\sqrt{\max\{(Ax)_i, \ln(4m_A)\} \ln(4m_A)}$$

can be computed by applying a derandomization to a $2m_A \times n$ *matrix with entries from* $\{a_{ij} \mid i \in [m_A], j \in [m]\}$*.*

The proof is again omitted for reasons of space. The main idea is to note that the rounding errors inflicted by the matching rounding of Lemma 3 can be written as a weighted sum of binary random variables representing the coin flips.

Combining Theorem 2 and Lemma 4 with the derandomizations cited in Section 3.2, we obtain the following derandomized version of Srinivasan's results.

Theorem 4. *Let* $A \in [0, 1]^{m_A \times n}$*. Let* $x \in [0, 1]^n$*. Then for all* $\ell \in \mathbb{N}$*, a binary vector* y *can be computed such that* $By \approx Bx$ *and*

$$\forall i \in [m_A] : |(Ax)_i - (Ay)_i| \leq f(\ell, 2c)\sqrt{\max\{(Ax)_i, \ln(4m_A)\} \ln(4m_A)} + n2^{-\ell}$$

This has a time complexity of ℓ *times the one of applying a derandomization to a* $2m_A \times n$ *matrix with entries in* $\{a_{ij} \mid i \in [m], j \in [n]\}$*.*

Some bounds on constants that are relevant in connection with the derandomizations mentioned in Subsection 3.2 are $f(2\sqrt{\frac{1}{2}}) \leq 13$, $f(2e - 2) \leq 90$, and $f(2\sqrt{3}) \leq 92$. However, the remark following Lemma 2 also applies to the theorem above, i.e., for $(Ax)_i$ large compared to $\ln(4m_A)$, the increase in the constants become less significant.

5 Bipartite Edge Weight Rounding

In this section, we consider sets of cardinality constraints where each variable may be contained in up to two constraints. Throughout this section, let $B = \binom{B_1}{B_2}$, where the B_i are $0, 1$ matrices such that $\|B_i\|_1 = 1$. We assume that B is represented by a datastructure allowing constant time queries of type "given i, find j such that $b_{ij} = 1$" and "given j, find i such that $b_{ij} = 1$".

For such constraints, negative correlation on $\mathcal{S} = 2^{[n]}$ is too much to ask for. We restrict ourselves to $\mathcal{S}_B = \{S \subseteq [n] \mid \exists i \in [m_B]\, \forall j \in S : b_{ij} = 1\}$.

Problems of this type have been regarded in Gandhi et al. [GKPS02]. They used a formulation as rounding problem for edge weights in bipartite graphs. We briefly fix the connection.

Bipartite edge weight rounding problem: Given a bipartite graph $G = (U \dot\cup V, E)$ and edge weights $x \in [0,1]^E$, find $y \in \{0,1\}^E$ such that (B1) y_e is a randomized rounding of x_e for all $e \in E$, (B2) $\sum_{e \ni v} y_e \approx \sum_{e \ni v} x_e$ for all $v \in U \cup V$ and (B3) for all $v \in U \cup V$, $S \subseteq \{e \in E \,|\, v \in E\}$ and $b \in \{0,1\}$, we have $\Pr(\forall i \in S : y_e = b) \leq \prod_{e \in S} \Pr(y_e = b)$.

The bipartite edge weight rounding problem is easily seem to be captured by our setting: Define $B_1 = (b_{ue}) \in \{0,1\}^{U \times E}$ through $b_{ue} = 1$ if and only if $u \in e$ as well as $B_2 = (b_{ve}) \in \{0,1\}^{V \times E}$ through $b_{ve} = 1$ if and only if $v \in e$. Then $By \approx Bx$ for some $x \in [0,1]^E$, $y \in \{0,1\}^E$ is just the degree preservation condition (B2). Also, negative correlation on $\mathcal{S}_B$ is equivalent to (B3).

The bipartite edge weight rounding problem for edge weights 0 ('no edge', if you like) and $\frac{1}{2}$ is easily solved. Here the pipage rounding idea of [AS04] fixes each variable to an integer value in amortized constant time. This saves an $O(|V|)$ run-time factor compared to the general case.

Lemma 5. *There are basic randomized roundings with respect to B and $\mathcal{S}$. They can be sampled in time $O(n)$.*

The lemma above together with the general reduction of Theorem 1 yields the following version of the bipartite edge weight rounding result of Gandhi et al. Note that the time complexity here is superior to the $O(|E||V|)$ bound of Gandhi et al. [GKPS02] (unless we are working with an overly high precision ℓ).

Theorem 5. *There are randomized roundings $(\overline{\Pr}_{x,B})$ with respect to B and $2^{[n]}$. A sample from $(\overline{\Pr}_{x,B})$ can be generated in time $O(n\ell)$, where ℓ is the binary length of x.*

Again, the randomized algorithm above can be derandomized.

Lemma 6. *Let $A \in [0,1]^{m_A \times n}$. Assume that for each $i_A \in [m_A]$ there is an $i_B \in [m_B]$ such that for all $j \in [n]$, $b_{i_B j} = 1$ whenever $a_{i_A j} \neq 0$. Let $x \in \{0, \frac{1}{2}\}^n$. Then a binary vector y such that $By \approx Bx$ and*

$$|(Ax)_i - (Ay)_i| \leq 2c\sqrt{\max\{(Ax)_i, \ln(4m_A)\}\ln(4m_A)}$$

for all $i \in [m_A]$ can be computed by applying a derandomization to a matrix of dimension at most $2m_A \times n$ with entries from $\{a_{ij} \,|\, i \in [m_A], j \in [n]\}$.

Combining the Lemma 6 with Theorem 2, we obtain the following derandomization of the result of Gandhi et al.

Theorem 6. *Let $A \in [0,1]^{m_A \times n}$ such that for each $i_A \in [m_A]$ there is an $i_B \in [m_B]$ such that for all $j \in [n]$, $b_{i_B j} = 1$ whenever $a_{i_A j} \neq 0$. Then for all $\ell \in \mathbb{N}$, a binary vector y can be computed such that $By \approx Bx$ and*

$$\forall i \in [m_A] : |(Ax)_i - (Ay)_i| \leq f(\ell, 2c)\sqrt{\max\{(Ax)_i, \ln(4m_A)\}\ln(4m_A)} + n2^{-\ell}.$$

The time complexity is ℓ times the one of a derandomization for $2m \times n$ matrices with entries from $\{a_{ij} \,|\, i \in [m_A], j \in [m]\}$.

6 Other Constraints

It is relatively easy to see that Theorems 3, 4, 5 and 6 can be extended to include hard constraint matrices $B \in \{-1, 0, 1\}$ as long as B is totally unimodular. An extension to further values, however, is not possible. Also, Theorems 3 and 4 can be extended to other sparsely intersecting constraints than the ones of Section 5. We now give two examples involving substantially different hard constraints.

Sequence Rounding. In connection with an image processing application, Sadakane, Takki-Chebihi and Tokuyama [STT01] compute roundings of sequences such that the rounding errors in all intervals are less than one. This is in fact a classical problem, but the new aspect in their work is that they need a randomized solution as this is less likely to produce unwanted structures in the images. The approach taken in [STT01] is via efficiently computing several roundings and then taking a random one.

A simpler way using the framework of this paper is to compute a randomized rounding y of $x \in [0,1]^n$ with the additional constraints that for each interval $I \subseteq [n]$, $\sum_{i \in I} y_i$ is a randomized rounding of $\sum_{i \in I} x_i$. To do so, we have to understand this problem for $0, \frac{1}{2}$ sequences, which is trivial.

Matrix Rounding. Asano et al. [AKOT03] model the digital halftoning problem as matrix rounding problem. For $X \in [0,1]^{m \times n}$ and $Y \in \{0,1\}^{m \times n}$, they set

$$d(X,Y) := \sum_{\substack{i \in [m-1] \\ j \in [n-1]}} \Big| \sum_{k,\ell \in \{0,1\}} (x_{i+k,j+\ell} - y_{i+k,j+\ell}) \Big|.$$

They claim that the image represented by Y is a good halftoning of the image represented by X, if $d(X,Y)$ is small. The current best solution for computing good roundings with respect to this error measure uses dependent randomized roundings [Doe04]. Let Y be a randomized rounding of X with respect to the constraints [2]

$$\begin{aligned}
&y_{i,j} + y_{i,j+1} \approx x_{i,j} + x_{i,j+1}, i \in [m],\ j \in [n-1],\\
&y_{i,j} + y_{i+1,j} \approx x_{i,j} + x_{i+1,j}, i \in [m-1]\,\text{odd},\ j \in [n],\\
&y_{i,j} + y_{i,j+1} + y_{i+1,j} + y_{i+1,j+1} \approx x_{i,j} + x_{i,j+1} + x_{i+1,j} + x_{i+1,j+1},\\
&\quad i \in [m-1]\,\text{odd},\ j \in [n-1].
\end{aligned}$$

Then $E(d(X,Y)) \leq 0.55$ holds for all X. The existence of such dependent roundings easily follows from the totally unimodularity condition. However, actually computing them in linear time involves tedious case distinctions.

With the reduction of Theorem 1, life is much easier since we only have to regard $X \in \{0, \frac{1}{2}\}^{m \times n}$. In this case, a constraint of the third type is either implied by constraints of the first two kinds, or it contains exactly two non-integral variables. All constraints thus yield a bipartite graph $G = ([m] \times [n], E)$ with $\{(i_1, j_1), (i_2, j_2)\} \in E$ telling us that exactly one of y_{i_1,j_1} and y_{i_2,j_2} has to become one, and these are all constraints. This makes it easy to compute such a rounding: For each connected component of G, flip a fair coin to decide which of the two classes of the bipartition shall be rounded to one, and set the other variables to zero.

---

[2] Here we use the notation $y \approx x$ to denote that y is a randomized rounding of x.

7 A Word of Warning

We have to note that dependencies like cardinality constraints inflict that some natural properties are unexpectedly not satisfied. Call a function $f : \{0,1\}^n \to \mathbb{R}$ non-decreasing if $y \le z$ (component-wise) implies $f(y) \le f(z)$.

(i) There are $S \subseteq [n]$ such that the roundings of Section 4 and 5 make $x \mapsto \Pr_x(\forall i \in S : y_i = 1)$ *not* non-decreasing.

(ii) The compared to (A3) stronger property that for all disjoint $S, T \subseteq [n]$, one has $\Pr_x(y_{|S} \equiv 1 \,|\, y_{|T} \equiv 1) \le \Pr_x(y_{|S} \equiv 1)$ also does not hold for the roundings of Section 4 and 5. This is the reason why in [GKPS02] this property could only be proven for a single cardinality constraint and only by prescribing a particular order for the individual roundings.

(iii) There are non-decreasing functions f such that *any* randomized rounding with respect to a single cardinality constraint makes $x \mapsto E_x(f)$ not non-decreasing.

All these phenomena, of course, are not possible for independent randomized roundings.

Let us also mention the following. A distribution on $\{0,1\}^n$ is negatively associated (NA), if for all non-decreasing $f, g : \{0,1\}^n \to \mathbb{R}_{\ge 0}$ we have $E(fg) \le E(f)E(g)$. The distributions of Section 4 and 5 are not (NA).

8 General Derandomization

In this section, we improve and simplify the derandomization result of Stangier and Srivastav [SS96]. Recall from Section 3.2 that Raghavan's derandomization in the RAM model only works for binary matrices. This problem was solved in [SS96], though at the price of a significantly higher time complexity of $O(mn^2 \log(mn))$. Also, this approach is hard to implement due to its technical demands. We overcome these difficulties by reducing the general problem to Raghavan's setting and obtain the following result.

Theorem 7. *Let $A \in ([0,1] \cap \mathbb{Q})^{m \times n}$, $x \in ([0,1] \cap \mathbb{Q})^n$ and $\ell \in \mathbb{N}$. Then a $y \in \{0,1\}^n$ such that*

$$|(Ax)_i - (Ay)_i| \le 2(e-1)\sqrt{\max\left\{(Ax)_i, \ln(2\ell m)\right\}\ln(2\ell m)} + 2^{-\ell} n$$

holds for all $i \in [m]$ can be computed in time $O(mn\ell)$ in the RAM model.

Proof (Sketch). Use the binary expansion $A = \sum_{k=1}^{\ell} 2^{-k} A^{(k)}$ of A and apply Raghavan's derandomization to the $\ell m \times n$ matrix obtained from stacking the $A^{(i)}$. □

Note that if we choose $\ell = \lceil \log_2 n \rceil$, the additive extra term is at most one. Note also, that then the $\ln(2\ell m)$ term is just a factor away from the usual $\ln(2m)$: We may assume $m \ge \log n$. Otherwise using linear algebra we may transform x into a vector x' such that $Ax = Ax'$ and at most m components of x' are not 0 or 1. But if $\ell = \lceil \log_2 n \rceil \le 2m$, then $\ln(2\ell m) \le 2\ln(2m)$.

Finally, note the following. By combining Lemma 4 with the elementary derandomization for $\{0, \frac{1}{2}, 1\}$ vectors in Section 3.2, we obtain a very elementary and simple to implement algorithm for arbitrary vectors and binary matrices.

References

[AFK02] S. Arora, A. Frieze, and H. Kaplan. A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. *Math. Program.*, 92:1–36, 2002.

[AKOT03] T. Asano, N. Katoh, K. Obokata, and T. Tokuyama. Matrix rounding under the L_p-discrepancy measure and its application to digital halftoning. *SIAM J. Comput.*, 32:1423–1435, 2003.

[AS04] A. A. Ageev and M. I. Sviridenko. Pipage rounding: a new method of constructing algorithms with proven performance guarantee. *J. Comb. Optim.*, 8:307–328, 2004.

[BS84] J. Beck and J. Spencer. Integral approximation sequences. *Math. Programming*, 30:88–98, 1984.

[Doe03] B. Doerr. Non-independent randomized rounding. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 506–507, 2003.

[Doe04] B. Doerr. Nonindependent randomized rounding and an application to digital halftoning. *SIAM Journal on Computing*, 34:299–317, 2004.

[Doe05] B. Doerr. Roundings respecting hard constraints. In *Proceedings of the 22nd Annual Symposium on Theoretical Aspects of Computer Science (STACS'05)*, volume 3404 of *Lecture Notes in Computer Science*, pages 617–628, Berlin–Heidelberg, 2005. Springer-Verlag.

[GHK+03] R. Gandhi, E. Halperin, S. Khuller, G. Kortsarz, and A. Srinivasan. An improved approximation algorithm for vertex cover with hard capacities. In J. Baeten, J. Lenstra, J. Parrow, and G. Woeginger, editors, *Automata, Languages and Programming. 30th International Colloquium, ICALP 2003*, volume 2719 of *Lecture Notes in Computer Science*, pages 164–175, Berlin–Heidelberg, 2003. Springer Verlag.

[GKPS02] R. Gandhi, S. Khuller, S. Parthasarathy, and A. Srinivasan. Dependent rounding in bipartite graphs. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 323–332, 2002.

[GKR+99] V. Guruswami, S. Khanna, R. Rajaraman, B. Shepherd, and M. Yannakakis. Near-optimal hardness results and approximation algorithms for edge-disjoint paths and related problems. In *Annual ACM Symposium on Theory of Computing (STOC)*, pages 19–28, New York, 1999. ACM.

[PS97] A. Panconesi and A. Srinivasan. Randomized distributed edge coloring via an extension of the Chernoff-Hoeffding bounds. *SIAM J. Comput.*, 26:350–368, 1997.

[Rag88] P. Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *J. Comput. Syst. Sci.*, 37:130–143, 1988.

[RT87] P. Raghavan and C. D. Thompson. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7:365–374, 1987.

[RT91] P. Raghavan and C. D. Thompson. Multiterminal global routing: a deterministic approximation scheme. *Algorithmica*, 6:73–82, 1991.

[Spe94] J. Spencer. *Ten lectures on the probabilistic method*, volume 64 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[Sri01] A. Srinivasan. Distributions on level-sets with applications to approximations algorithms. In *Proc. 41th Ann. IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 588–597, 2001.

[SS96] A. Srivastav and P. Stangier. Algorithmic Chernoff-Hoeffding inequalities in integer programming. *Random Structures & Algorithms*, 8:27–58, 1996.

[STT01] K. Sadakane, N. Takki-Chebihi, and T. Tokuyama. Combinatorics and algorithms on low-discrepancy roundings of a real sequence. In *ICALP 2001*, volume 2076 of *Lecture Notes in Computer Science*, pages 166–177, Berlin Heidelberg, 2001. Springer-Verlag.