Sparse Recovery With Integrality Constraints

Jan-Hendrik Lange

Max-Planck-Institute for Informatics, Campus E1 4, 66123 Saarbrücken, Germany

Marc E. Pfetsch, Bianca M. Seib[1]

Research Group Optimization, TU Darmstadt, Dolivostr. 15, 64293 Darmstadt, Germany

Andreas M. Tillmann*

Visual Computing Institute & Operations Research Group, RWTH Aachen University, Lehrstuhl für Informatik 8, Ahornstr. 55, 52074 Aachen, Germany

Abstract

In this paper, we investigate conditions for the unique recoverability of sparse integer-valued signals from few linear measurements. Both the objective of minimizing the number of nonzero components, the so-called ℓ_0-norm, as well as its popular substitute, the ℓ_1-norm, are covered. Furthermore, integrality constraints and possible bounds on the variables are investigated. Our results show that the additional prior knowledge of signal integrality allows for recovering more signals than what can be guaranteed by the established recovery conditions from (continuous) compressed sensing. Moreover, even though the considered problems are NP-hard in general (even with an ℓ_1-objective), we investigate testing the ℓ_0-recovery conditions via some numerical experiments; it turns out that the corresponding problems are quite hard to solve in practice using black-box software. However, medium-sized instances of ℓ_0- and ℓ_1-minimization with binary variables can be solved exactly within reasonable time.

Keywords: Sparse recovery, compressed sensing, integrality constraints, nullspace conditions

1. Introduction

The recovery of sparse signals has received a tremendous interest in recent years. The basic setting without noise is as follows: Under the prior knowledge that

*Corresponding author

Email addresses: jlange@mpi-inf.mpg.de (Jan-Hendrik Lange), pfetsch@mathematik.tu-darmstadt.de (Marc E. Pfetsch), bianca.seib@gmail.com (Bianca M. Seib), tillmann@or.rwth-aachen.de (Andreas M. Tillmann)

[1]Present address: Wiesbaden, Germany

a measurement vector $\boldsymbol{b} \in \mathbb{R}^m \setminus \{0\}$ is generated by a sparse signal $\boldsymbol{x} \in \mathbb{R}^n$ via $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(\boldsymbol{A}) = m < n$ is the sensing matrix, the question is whether $\boldsymbol{x}$ can be uniquely recovered, given $\boldsymbol{A}$ and $\boldsymbol{b}$. Thus, one approach is to find the sparsest $\boldsymbol{x}$ that explains the measurements, i.e., one minimizes $\|\boldsymbol{x}\|_0 \coloneqq |\{i \in \{1, \dots, n\} : x_i \neq 0\}|$ under the constraint $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. However, this problem is NP-hard, see Garey and Johnson [1]. The crucial idea in this context (see, e.g., Chen et al. [2]) is to replace $\|\boldsymbol{x}\|_0$ by the ℓ_1-norm $\|\boldsymbol{x}\|_1 \coloneqq |x_1| + \cdots + |x_n|$, which results in a convex problem that can even be cast as a linear program (LP) and is therefore tractable. The literature contains an abundance of conditions under which minimizers of $\|\boldsymbol{x}\|_1$ subject to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ are unique and equal to the sparsest solution; at this place, we refer to the book by Foucart and Rauhut [3] for more information and an overview of selected specialized algorithms to solve the ℓ_1-minimization problem.

The key point for the mentioned series of striking results is the prior knowledge that $\boldsymbol{b}$ can be sparsely represented or approximated. A natural question is whether further knowledge about the structure of the representations $\boldsymbol{x}$ can lead to stronger results about the recoverability. In general terms, the two problems from above can be written as

$$\min \{\|\boldsymbol{x}\|_0 : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in X\}, \tag{$\mathrm{P}_0(X)$}$$

$$\min \{\|\boldsymbol{x}\|_1 : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in X\}, \tag{$\mathrm{P}_1(X)$}$$

where $X \subseteq \mathbb{R}^n$ is a constraint set representing further restrictions on the representations.

The "classical" results in the literature refer to the case $X = \mathbb{R}^n$. One main example in which $X \neq \mathbb{R}^n$ is the case in which $\boldsymbol{x}$ has to be nonnegative, i.e., $X = \mathbb{R}^n_+$, see, for instance, Donoho and Tanner [4], Bruckstein et al. [5] and Khajehnejad et al. [6].

In this paper, we investigate the case in which $\boldsymbol{x}$ is required to be *integral*, i.e., $X \subseteq \mathbb{Z}^n$. This is motivated by various applications in which signals are composed from finite symbol alphabets, such as machine-to-machine communication (see Knopp et al. [7]), spectrum sensing for cognitive radio (cf. Axell et al. [8]), detection, localization or interference cancellation in multiple-input/output (MIMO) systems (see, e.g., Zhu and Giannakis [9], Knopp et al. [10] and Rossi et al. [11]), or discrete tomography tasks as described by Batenburg and Sijbers [12] or Kuske et al. [13], to name but a few.

One particular application arises when constellation signals are used in massive MIMO; we briefly describe the real-valued case, for simplicity. Here, the components of the signal are chosen from a small set of constellation signals $\{C_1, \dots, C_M\}$. One class of examples is given by the M phase-shift keying (M-PSK); several different types of such configurations with different values of M exist. Then, one can multiply the columns of the sensing matrix with all constellation signals, and binary variables $\boldsymbol{x}$ can be used to select the corresponding signal for each component. If one searches for the sparsest signal vector (e.g., in the context of low-activity multi-user detection as in [7]), one arrives at an instance of problem ($\mathrm{P}_0(X)$) in the noise-free case. Hegde et al. [14] discuss an optimization approach to relax such signals, [15] propose an exact method, and, e.g., [9] treats binary PSK.

Example 1. *To illustrate some of the issues investigated in this paper, consider* $\boldsymbol{A} = (2, 3, 6) \in \mathbb{R}^{1\times 3}$ *and* $\boldsymbol{b} = (11)$*. Then* $(\mathrm{P}_0(\mathbb{Z}^3_+))$ *has the two optimal solutions* $(4, 1, 0)^\top$ *and* $(1, 3, 0)^\top$*. Furthermore,* $(\mathrm{P}_1(\mathbb{Z}^3_+))$ *has optimal solution* $(1, 1, 1)^\top$*. Finally,* $(\mathrm{P}_0(\mathbb{R}^3_+))$ *has three optimal solutions, each with one nonzero entry, while* $(\mathrm{P}_1(\mathbb{R}^3_+))$ *has the unique optimal solution* $(0, 0, \frac{11}{6})^\top$*. This shows that requiring integrality affects the optimal solutions of* $(\mathrm{P}_0(X))$ *and* $(\mathrm{P}_1(X))$*. Moreover, these problems may yield different solutions.*

Despite the apparent practical interest, there are only a few articles in the literature that deal with integral signal recovery, both theoretically and algorithmically. For instance, Sparrer and Fischer [16] present a heuristic approach based on orthogonal matching pursuit. A further heuristic was proposed by Flinth and Kutyniok [17] based on a combination of projection and orthogonal matching pursuit ideas. The binary case—particularly prominent in the context of digital/wireless communication systems, where transmitted signals can often be represented as simple bit sequences—has been treated, for instance, by Nakarmi and Rahnavard [18] and Wu et al. [19]. Mangasarian and Recht [20] gave conditions for uniqueness of vectors in $X = \{-1, 1\}^n$ as solutions of $\boldsymbol{Ax} = \boldsymbol{b}$, $-\mathbb{1} \le \boldsymbol{x} \le \mathbb{1}$, while Stojnic [21] determined empirical and probabilistic theoretical recovery thresholds for binary signals via $(\mathrm{P}_1([0, 1]^n))$. Swoboda et al [22] presented a Lagrangian relaxation based heuristic for solving problems with integer variables.

The above-mentioned works mostly exhibit a clear focus on the algorithmic side, proposing relaxations or heuristic methods and empirical results on their success. Rigorous theoretical conditions for the recovery of sparse integral signals by ℓ_1-minimization with relaxed integrality requirements were presented in Keiper et al. [23]. They investigate binary and ternary signals by means of $(\mathrm{P}_1([0, 1]^n))$ and $(\mathrm{P}_1([-1, 1]^n))$. One of their main contributions is the investigation of phase transitions of unique recovery. They showed that recovery exploiting the bounds $[0, 1]$ or $[-1, 1]$ takes place earlier. Furthermore, Flinth and Keiper [24] provided a more detailed investigation of binary signal recovery by providing probabilistic conditions for unique recovery.

Our work complements those results and generalizes some of them: We consider more general integral sets and their continuous relaxations along with both ℓ_0- and ℓ_1-objectives. For instance, one of our results shows that explicitly treating integrality constraints can allow for the recovery of essentially arbitrarily many more integral signals than could be recovered by the associated relaxed (integrality-oblivious) problem:

Example 2. *Let* $-\boldsymbol{\ell} = \boldsymbol{u} = \mathbb{1} \in \mathbb{Z}^n$ *with* $n \ge 6$*, let* $\boldsymbol{v} = (1, -1, (\boldsymbol{v}')^\top)^\top$ *and* $\boldsymbol{w} = (\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor, (\boldsymbol{w}')^\top)^\top$ *with* $\boldsymbol{v}', \boldsymbol{w}' \in \{-1, 1\}^{n-2}$ *arbitrary. Let* $\boldsymbol{A}$ *be such that its nullspace is* $\operatorname{span}\{\boldsymbol{v}, \boldsymbol{w}\}$*. Then Theorem 30 below yields recoverability of all* $\boldsymbol{x} \in X \coloneqq \{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{\ell} \le \boldsymbol{x} \le \boldsymbol{u}\}$ *with at most* $(\lfloor \frac{n}{2} \rfloor - 1)$ *nonzeros by means of* $(\mathrm{P}_1(X))$*. However, ignoring integrality, which amount to solving*

($\mathrm{P}_1(\mathrm{conv}(X))$), *cannot recover all sparse signals from* X *with as few as* 2 *nonzero components; see Example 33 in Section 4.2 for the details.*

The computational price one has to pay for such strong results is that the ℓ_0- and ℓ_1-problems are NP-hard if integrality of the signals is enforced, see Section 2. Therefore, the main motivation for this paper is to provide a theoretical characterization of cases in which it is worth investing additional computational resources to take integrality into account. This might happen via heuristic methods or exact algorithms. Our results can also provide a motivation for developing such techniques.

It is important to note that we only treat the noise-free case, i.e., we consider equations $\boldsymbol{Ax} = \boldsymbol{b}$ instead of an error bound like $\|\boldsymbol{Ax} - \boldsymbol{b}\|_2 \leq \delta$. Thus, our results mark a first step towards an understanding of the underlying structure and should be extended to the noise-aware case to be relevant for real-world applications in the future.

The remainder of the paper is organized as follows. Apart from NP-hardness, in Section 2, we show that choosing rational $\boldsymbol{A}$ has a crucial impact on the recoverability properties. In Section 3, we derive recoverability characterizations for the ℓ_0-problem. In Section 4, we turn to the ℓ_1-case and start by investigating cases in which an integral solution can be guaranteed when solving the continuous relaxation. We then derive characterizations for uniform (Section 4.2) and individual (Section 4.3) recoverability in the ℓ_1-case. In Section 5, we report on computational experiments, and close with some final remarks in Section 6.

Remark 3. *Many of the main results in compressed sensing also hold with respect to complex data and signals, see, e.g., the survey of real and complex nullspace conditions characterizing* ℓ_0-ℓ_1*-equivalence in [3]. Nevertheless, for the sake of simplicity, we only consider the real-valued case in this paper. For instance, extensions to complex signals with (say) integral real and imaginary parts are not treated here.*

We use the following notation: We use $\mathbb{N} = \{1, 2, \dots\}$ and define $[n] \coloneqq \{1, \dots, n\}$ for $n \in \mathbb{N}$. Furthermore, for $s \in [n]$, the vector $\boldsymbol{x} \in \mathbb{R}^n$ is *s-sparse*, if $\|\boldsymbol{x}\|_0 \leq s$. The *support* of $\boldsymbol{x}$ is defined as $\mathrm{supp}(\boldsymbol{x}) \coloneqq \{i \in [n] : x_i \neq 0\}$. Moreover, for $S \subseteq [n]$, $\boldsymbol{x}_S \in \mathbb{R}^n$ denotes the vector which equals $\boldsymbol{x}$ for all components indexed by S and is 0 otherwise. The complement of a set $S \subseteq [n]$ is denoted by $S^c \coloneqq [n] \setminus S$. The nullspace (kernel) of a matrix $\boldsymbol{A}$ is defined as $\mathcal{N}(\boldsymbol{A}) \coloneqq \{\boldsymbol{x} : \boldsymbol{Ax} = \boldsymbol{0}\}$. By $\mathbb{1}$, we denote the all-ones vector of appropriate dimension.

2. Basic Problems and Results

In this paper, we investigate the following five basic integrality requirements for ($\mathrm{P}_0(X)$) and ($\mathrm{P}_1(X)$):

$$X = \mathbb{Z}^n, \quad X = \mathbb{Z}^n_+, \quad X = [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}, \quad X = [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}, \quad X = [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, \quad (1)$$

where $\boldsymbol{\ell} \leq \mathbf{0} \leq \boldsymbol{u} \in \mathbb{R}^n$ and $[\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}} \coloneqq \{\boldsymbol{x} \in \mathbb{Z}^n \,:\, \boldsymbol{\ell} \leq \boldsymbol{x} \leq \boldsymbol{u}\}$. Note that we have to make sure that $\mathbf{0} \in X$ in order to allow sparse solutions; moreover, throughout the paper, we assume w.l.o.g. that $\boldsymbol{\ell} < \boldsymbol{u}$ (otherwise, $\ell_i = u_i = 0$ and $x_i = 0$ can be eliminated from the problem a priori). When considering $[\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$, we can round the components of $\boldsymbol{\ell}$ and $\boldsymbol{u}$ up and down, respectively; thus, in this case, we may assume that $\boldsymbol{\ell}$, $\boldsymbol{u} \in \mathbb{Z}^n$. However, in particular cases, we also deal with boxes $[\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{R}} \coloneqq \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \boldsymbol{\ell} \leq \boldsymbol{x} \leq \boldsymbol{u}\}$ for which $\boldsymbol{\ell}$ and $\boldsymbol{u}$ can be real-valued. Clearly, $[\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ is the most general (integral) case; the others can be written in this form (if $\boldsymbol{\ell}$ and $\boldsymbol{u}$ are allowed to take $\mp\infty$ values, respectively).

The first observation is that all of the considered problems are NP-hard.

Proposition 4. *The problems* ($\mathrm{P}_0(X)$) *and* ($\mathrm{P}_1(X)$) *are NP-hard in the strong sense for each of the sets* X *in* (1), *even if* $\boldsymbol{A}$ *is binary and* $\boldsymbol{b} = \mathbb{1}$.

Proof. Garey and Johnson [1] proved that ($\mathrm{P}_0(\mathbb{R}^n)$) is strongly NP-hard using a reduction from "exact cover by 3-sets" (this proof is reproduced in [3]). The proof shows that, given an instance of this problem, one can construct a binary matrix $\boldsymbol{A}$ such that solutions $\boldsymbol{x}$ of $\boldsymbol{A}\boldsymbol{x} = \mathbb{1}$ minimizing $\|\boldsymbol{x}\|_0$ are necessarily 0/1, i.e., $\boldsymbol{x} \in \{0,1\}^n$. These solutions are feasible for any of the considered problems and furthermore satisfy $\|\boldsymbol{x}\|_0 = \|\boldsymbol{x}\|_1$. □

This proposition carries an unfortunate negative message: Changing $\|\boldsymbol{x}\|_0$ to $\|\boldsymbol{x}\|_1$ does not change the complexity status of the problem, and all considered problems are hard to solve. On the other hand, modern integer optimization technology allows to solve small to medium sized instances of these problems. Moreover, empirically, the ℓ_1-case is often slightly easier.

In any case, it is a fundamental question to what extent integrality requirements allow to increase the amount of cases in which a signal can be uniquely recovered. Such a solution might then be found efficiently in practice, e.g., by heuristics such as that in Flinth and Kutyniok [17].

When considering integrality requirements, it is of fundamental importance whether the matrix $\boldsymbol{A}$ is rational:

Theorem 5. *For any* $n \in \mathbb{N}$, *there exists a single-row matrix* $\boldsymbol{A} \in \mathbb{R}^{1\times n}$ *such that for every* $\boldsymbol{b} \in \mathrm{range}_{\mathbb{Z}}(\boldsymbol{A}) \coloneqq \{\boldsymbol{A}\boldsymbol{z} \,:\, \boldsymbol{z} \in \mathbb{Z}^n\}$, *there exists a unique* $\boldsymbol{x} \in \mathbb{Z}^n$ *such that* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

Proof. Since $\mathbb{R}$ is an infinite-dimensional vector space over $\mathbb{Q}$, choosing n real numbers that are linearly independent over $\mathbb{Q}$ as the components of $\boldsymbol{A}$ suffices. For instance, taking the nth roots of pairwise different prime numbers ≥ 2 will do, see, e.g., Besicovitch [25]. Thus, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a unique integral solution for every $\boldsymbol{b} \in \mathrm{range}_{\mathbb{Z}}(\boldsymbol{A})$. □

As a consequence, for such a matrix $\boldsymbol{A}$, the recovery problem with integral $\boldsymbol{x}$ is always uniquely solvable and thus, ideal recovery is possible. However, in general, such matrices cannot be stored in a computer. Moreover, the complexity of solving $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{x} \in \mathbb{Z}^n$ is unclear; in particular, one needs to use a "non-standard" model of computation, cf., e.g., Blum et al. [26].

In the following, we will often consider *rational* $\boldsymbol{A}$. Note that in this case, finding *some* integral solution $\boldsymbol{x}$ of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ can be done in polynomial time using the Hermite normal form, see, e.g., Schrijver [27]. However, as Proposition 4 shows, minimizing $\|\boldsymbol{x}\|_0$ or $\|\boldsymbol{x}\|_1$ is still NP-hard.

3. The ℓ_0-case

In this section, we provide conditions on the uniform recoverability via ($\mathrm{P}_0(X)$). For this, we define the set

$$S(s, X; \boldsymbol{b}) \coloneqq \{\boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \|\boldsymbol{x}\|_0 \leq s,\ \boldsymbol{x} \in X\}.$$

The key point is uniqueness of sparse solutions, i.e., whether $|S(s, X; \boldsymbol{A}\hat{\boldsymbol{x}})| = 1$ for s-sparse $\hat{\boldsymbol{x}} \in X$. Inspired by the terminology of Juditsky and Nemirovski [28], we define the following.

Definition 6. *Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $s \in [n]$, and $X \subseteq \mathbb{R}^n$. The matrix $\boldsymbol{A}$ is* $(s, X, 0)$-good, *if for* every *s-sparse vector $\hat{\boldsymbol{x}} \in X$, it holds that $|S(s, X; \boldsymbol{A}\hat{\boldsymbol{x}})| = 1$.*

We first state some obvious results for $(s, X, 0)$-good matrices.

Lemma 7. *Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $s \in [n]$, and $X \subseteq \mathbb{R}^n$.*

1. *If $\boldsymbol{A}$ is $(s, X, 0)$-good, it is $(s, X', 0)$-good for every $X' \subseteq X$. Therefore,*

$$(s, \mathbb{R}^n, 0)\text{-good} \quad \Rightarrow \quad (s, \mathbb{Z}^n, 0)\text{-good} \quad \Rightarrow \quad (s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 0)\text{-good.}$$

   Moreover, if $\boldsymbol{A}$ is $(s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 0)$-good, it is $(s, [\boldsymbol{\ell}', \boldsymbol{u}']_{\mathbb{Z}}, 0)$-good for every $\boldsymbol{\ell} \leq \boldsymbol{\ell}' \leq \boldsymbol{0} \leq \boldsymbol{u}' \leq \boldsymbol{u}$ ($\boldsymbol{\ell}, \boldsymbol{u} \in \mathbb{R}^n$).

2. *If $\boldsymbol{A}$ is $(s, X, 0)$-good, it is $(s', X, 0)$-good for every $s' \leq s$, $s' \in \mathbb{N}$.*

Furthermore, we recall the following well-known result from the literature.

Theorem 8 ([29], [3])**.** *A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is $(s, \mathbb{R}^n, 0)$-good for $s \in [n]$ if and only if $\mathcal{N}(\boldsymbol{A}) \cap \{\boldsymbol{z} \in \mathbb{R}^n \,:\, \|\boldsymbol{z}\|_0 \leq 2s\} = \{\boldsymbol{0}\}$.*

The statement of this theorem can be rephrased by using $\operatorname{spark}(\boldsymbol{A}) \coloneqq \min\{\|\boldsymbol{x}\|_0 \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},\ \boldsymbol{x} \neq \boldsymbol{0}\}$, which refers to the smallest number of linearly dependent columns of $\boldsymbol{A}$. Then, $\boldsymbol{A}$ is $(s, \mathbb{R}^n, 0)$-good if and only if $\operatorname{spark}(\boldsymbol{A}) > 2s$. Since the decision problem "$\operatorname{spark}(\boldsymbol{A}) \leq k$?" is NP-complete (cf. [30]) and a $\boldsymbol{z} \in \mathbb{Q}^n$ with $1 \leq \|\boldsymbol{z}\|_0 \leq 2s$ serves as a certificate for $\boldsymbol{A} \in \mathbb{Q}^{m \times n}$ *not* being $(s, \mathbb{R}^n, 0)$-good, this shows that checking the condition in Theorem 8 is coNP-complete.

By a completely analogous proof, Theorem 8 carries over to the integral case by requiring $\boldsymbol{z} \in \mathbb{Z}^n$. Moreover, if $\boldsymbol{A}$ is rational, we can always scale vectors in the nullspace $\mathcal{N}(\boldsymbol{A})$ to be integral. This yields:

Theorem 9. *A matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is $(s, \mathbb{Z}^n, 0)$-good if and only if $\mathcal{N}(\boldsymbol{A}) \cap \{\boldsymbol{z} \in \mathbb{Z}^n : \|\boldsymbol{z}\|_0 \le 2s\} = \{\boldsymbol{0}\}$. Thus, if $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$, then $\boldsymbol{A}$ is $(s, \mathbb{Z}^n, 0)$-good if and only if it is $(s, \mathbb{R}^n, 0)$-good.*

Again using NP-completeness for the spark, checking the condition in Theorem 9 is also coNP-complete. Moreover, Theorem 9 has the following interesting consequence, compare with Theorem 5.

Theorem 10. *Let $s \in [n]$. The minimal number of rows m for which a* rational *matrix $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$ can be $(s, \mathbb{Z}^n, 0)$-good is $2s$.*

Proof. If $\boldsymbol{A}$ is rational, the condition in Theorem 9 is equivalent to that of Theorem 8. Moreover, for continuous settings, [3, Theorems 2.13 and 2.14] (see also Cohen et al. [29]) show that $m \ge 2\,s$ is necessary in general and equality can be achieved using a Vandermonde matrix. □

On the other hand, when additionally considering bounds on the variables, we get a similar behavior as in Theorem 5 even for rational matrices:

Theorem 11. *Let $X = [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$ for $\boldsymbol{u} \in \mathbb{Z}^n_{>0}$. Then for any $n \in \mathbb{N}$ there exists a rational matrix $\boldsymbol{A} \in \mathbb{Q}^{1\times n}$ such that for every $\boldsymbol{b} \in \operatorname{range}_X(\boldsymbol{A}) \coloneqq \{\boldsymbol{A}\boldsymbol{z} : \boldsymbol{z} \in X\}$ there exists a unique $\boldsymbol{x} \in X$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.*

Proof. Define $\delta \coloneqq \max\{u_1, \dots, u_n\}+1$ and let $A = (a_{1k})$ be defined by $a_{1k} \coloneqq \delta^k$ for $k = 0, \dots, n-1$. Then, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ for $\boldsymbol{x} \in X$ amounts to a δ-ary expansion of $\boldsymbol{b}$, which is unique. □

Note that this result is of theoretical interest only, since the large coefficients from the proof of Theorem 11 will produce numerical problems for larger n.

3.1. Recovery Conditions for the ℓ_0-Case

To treat the case of $(\mathrm{P}_0([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$, we need the following notation. For $\boldsymbol{a} \le \boldsymbol{b} \in \mathbb{R}^n$, we consider closed boxes $[\boldsymbol{a}, \boldsymbol{b}] \coloneqq \{\boldsymbol{v} : \boldsymbol{a} \le \boldsymbol{v} \le \boldsymbol{b}\} = [a_1, b_1] \times \cdots \times [a_n, b_n]$; similarly, half-open boxes are defined in the obvious way. For $\boldsymbol{z} \in \mathbb{R}^n$ and one of these boxes $B = B_1 \times \cdots \times B_n \subseteq \mathbb{R}^n$, we define $\operatorname{supp}(\boldsymbol{z}; B) \coloneqq \{i \in [n] : z_i \in B_i\}$.

Theorem 12. *Let $X = [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$, $\delta_i^{\min} \coloneqq \min\{-\ell_i, u_i\}$ and $\delta_i^{\max} \coloneqq \max\{-\ell_i, u_i\}$ for all $i \in [n]$. Furthermore, define the following sets depending on a vector $\boldsymbol{z} \in \mathbb{R}^n$*

$$
\begin{aligned}
S_1^+ &\coloneqq \operatorname{supp}(\boldsymbol{z}; (\boldsymbol{0}, \boldsymbol{\delta}^{\min}]), & S_1^- &\coloneqq \operatorname{supp}(\boldsymbol{z}; [-\boldsymbol{\delta}^{\min}, \boldsymbol{0})),\\
S_2^+ &\coloneqq \operatorname{supp}(\boldsymbol{z}; (\boldsymbol{\delta}^{\min}, \boldsymbol{u}]), & S_2^- &\coloneqq \operatorname{supp}(\boldsymbol{z}; [\boldsymbol{\ell}, -\boldsymbol{\delta}^{\min})),\\
S_3^+ &\coloneqq \operatorname{supp}(\boldsymbol{z}; (\boldsymbol{u}, \boldsymbol{\delta}^{\max}]), & S_3^- &\coloneqq \operatorname{supp}(\boldsymbol{z}; [-\boldsymbol{\delta}^{\max}, \boldsymbol{\ell})),\\
S_4^+ &\coloneqq \operatorname{supp}(\boldsymbol{z}; (\boldsymbol{\delta}^{\max}, \boldsymbol{u} - \boldsymbol{\ell}]), & S_4^- &\coloneqq \operatorname{supp}(\boldsymbol{z}; [\boldsymbol{\ell} - \boldsymbol{u},\, -\boldsymbol{\delta}^{\max}))
\end{aligned}
$$

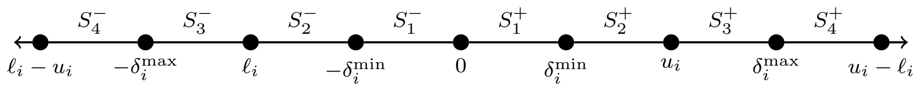

Figure 1: Illustration of the intervals in Theorem 12.

and let

$$\begin{aligned} C(\boldsymbol{\ell}, \boldsymbol{u}) \coloneqq \big\{ \boldsymbol{z} \in [\boldsymbol{\ell} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{\ell}]_{\mathbb{Z}} \,:\, & |S_4^+| + |S_4^-| + |S_3^+| + |S_3^-| \leq s, \\ & |S_4^+| + |S_4^-| + |S_2^+| + |S_2^-| \leq s, \\ 2\left(|S_4^+| + |S_4^-|\right) + |S_3^+| + |S_3^-| + |S_2^+| + |S_2^-| + & |S_1^+| + |S_1^-| \leq 2s \big\}. \end{aligned}$$

Then, an $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *is* $(s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 0)$*-good if and only if* $\mathcal{N}(\boldsymbol{A}) \cap C(\boldsymbol{\ell}, \boldsymbol{u}) = \{\boldsymbol{0}\}$.

Proof. W.l.o.g., we assume that $\boldsymbol{\ell}, \boldsymbol{u} \in \mathbb{Z}^n$. We first observe that

$$\boldsymbol{\ell} - \boldsymbol{u} \leq -\boldsymbol{\delta}^{\max} \leq \boldsymbol{\ell} \leq -\boldsymbol{\delta}^{min} \leq \boldsymbol{0} \leq \boldsymbol{\delta}^{\min} \leq \boldsymbol{u} \leq \boldsymbol{\delta}^{\max} \leq \boldsymbol{u} - \boldsymbol{\ell},$$

see also Figure 1. This shows that the boxes on which the sets S_1^+ to S_4^- are based are well-defined, although they may be empty.

Let $\boldsymbol{A}$ be $(s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 0)$-good and let $\boldsymbol{z} \in \mathcal{N}(\boldsymbol{A}) \cap C(\boldsymbol{\ell}, \boldsymbol{u})$. Define $k \coloneqq |S_4^+| + |S_4^-| + |S_3^+| + |S_3^-| \leq s$ and $r \coloneqq |S_1^+| + |S_1^-|$. Let $\tilde{S}$ be composed of $\min\{r, s-k\}$ arbitrary indices of $S_1^+ \cup S_1^-$ and $\tilde{S}^c \coloneqq (S_1^+ \cup S_1^-) \setminus \tilde{S}$ be its complement (w.r.t. $S_1^+ \cup S_1^-$). Now, we define

$$x_i \coloneqq \begin{cases} \ell_i & \text{if } i \in S_4^-, \\ 0 & \text{if } i \in S_3^-, \\ z_i & \text{if } i \in S_2^-, \\ 0 & \text{if } i \in \tilde{S}, \\ z_i & \text{if } i \in \tilde{S}^c, \\ z_i & \text{if } i \in S_2^+, \\ 0 & \text{if } i \in S_3^+, \\ u_i & \text{if } i \in S_4^+, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad y_i \coloneqq \begin{cases} \ell_i - z_i & \text{if } i \in S_4^-, \\ -z_i & \text{if } i \in S_3^-, \\ 0 & \text{if } i \in S_2^-, \\ -z_i & \text{if } i \in \tilde{S}, \\ 0 & \text{if } i \in \tilde{S}^c, \\ 0 & \text{if } i \in S_2^+, \\ -z_i & \text{if } i \in S_3^+, \\ u_i - z_i & \text{if } i \in S_4^+, \\ 0 & \text{otherwise.} \end{cases}$$

These two vectors satisfy $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$. Considering each case, one can see that $\boldsymbol{x}$, $\boldsymbol{y} \in X$. Moreover,

$$|\mathrm{supp}(\boldsymbol{y})| = |S_4^+| + |S_4^-| + |S_3^+| + |S_3^-| + |\tilde{S}| = k + \min\{r, s-k\} \leq k + s - k = s.$$

Furthermore, assume first that $r < s - k$, i.e., $\tilde{S} = S_1^+ \cup S_1^-$ and $\tilde{S}^c = \varnothing$. Then, by assumption,

$$|\mathrm{supp}(\boldsymbol{x})| = |S_4^+| + |S_4^-| + |S_2^+| + |S_2^-| \leq s.$$

On the other hand, if $r \geq s-k$, then $|\tilde{S}| = s-k$ and $|\tilde{S}^c| = r-s+k$, which yields

$$
\begin{aligned}
&|\mathrm{supp}(\boldsymbol{x})| = |S_4^+| + |S_4^-| + |S_2^+| + |S_2^-| + |\tilde{S}^c| \\
&= |S_4^+| + |S_4^-| + |S_2^+| + |S_2^-| + (|S_1^+| + |S_1^-| - s + k) \\
&= |S_4^+| + |S_4^-| + |S_2^+| + |S_2^-| + |S_1^+| + |S_1^-| - s + |S_4^+| + |S_4^-| + |S_3^+| + |S_3^-| \\
&= 2\,(|S_4^+| + |S_4^-|) + |S_3^+| + |S_3^-| + |S_2^+| + |S_2^-| + |S_1^+| + |S_1^-| - s \leq s.
\end{aligned}
$$

Since $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$, it follows that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}$ and consequently, by $(s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 0)$-goodness of $\boldsymbol{A}$, that $\boldsymbol{x} = \boldsymbol{y}$, i.e., $\boldsymbol{z} = \boldsymbol{0}$.

Conversely, assume that $\mathcal{N}(\boldsymbol{A}) \cap C(\boldsymbol{\ell}, \boldsymbol{u}) = \{\boldsymbol{0}\}$. Consider $\boldsymbol{x}, \tilde{\boldsymbol{x}} \in X$ with $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\tilde{\boldsymbol{x}}$, $\|\boldsymbol{x}\|_0 \leq s$, and $\|\tilde{\boldsymbol{x}}\|_0 \leq s$. By construction, $\boldsymbol{z} \coloneqq \boldsymbol{x} - \tilde{\boldsymbol{x}} \in \mathcal{N}(\boldsymbol{A}) \cap [\boldsymbol{\ell} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{\ell}]_{\mathbb{Z}}$.

Now observe that if $i \in S_4^+$ then $x_i > 0$ and $\tilde{x}_i < 0$ and similarly, if $i \in S_4^-$ then $x_i < 0$ and $\tilde{x}_i > 0$. This implies that $S_4^+ \subseteq \mathrm{supp}(\boldsymbol{x}) \cap \mathrm{supp}(\tilde{\boldsymbol{x}})$ and $S_4^- \subseteq \mathrm{supp}(\boldsymbol{x}) \cap \mathrm{supp}(\tilde{\boldsymbol{x}})$. For $S_4^c \coloneqq S_3^+ \cup S_3^- \cup S_2^+ \cup S_2^- \cup S_1^+ \cup S_1^-$, we thus obtain

$$
\begin{aligned}
2s &\geq |\mathrm{supp}(\boldsymbol{x})| + |\mathrm{supp}(\tilde{\boldsymbol{x}})| = |\mathrm{supp}(\boldsymbol{x}) \cup \mathrm{supp}(\tilde{\boldsymbol{x}})| + |\mathrm{supp}(\boldsymbol{x}) \cap \mathrm{supp}(\tilde{\boldsymbol{x}})| \\
&\geq |\big(\mathrm{supp}(\boldsymbol{x}) \cup \mathrm{supp}(\tilde{\boldsymbol{x}})\big) \cap (S_4^+ \cup S_4^-)| \\
&\qquad\qquad + |\big(\mathrm{supp}(\boldsymbol{x}) \cup \mathrm{supp}(\tilde{\boldsymbol{x}})\big) \cap S_4^c| + |S_4^+ \cup S_4^-| \\
&\geq 2\,|S_4^+ \cup S_4^-| + |S_4^c|.
\end{aligned}
$$

(The last inequality follows because, by construction, $S_i^{\pm} \subseteq \mathrm{supp}(\boldsymbol{x}) \cup \mathrm{supp}(\tilde{\boldsymbol{x}})$ for all $i \in [4]$.)

Observe that for $i \in S_3^+ \cup S_4^+$, necessarily $\tilde{x}_i < 0$. Similarly, if $i \in S_3^- \cup S_4^-$, then $\tilde{x}_i > 0$. This shows that $|S_4^+| + |S_4^-| + |S_3^+| + |S_3^-| \leq |\mathrm{supp}(\tilde{\boldsymbol{x}})| \leq s$.

Moreover, if $i \in S_2^+$ then $u_i > \delta_i^{\min} = -\ell_i$; furthermore, $-\tilde{x}_i \leq -\ell_i = \delta_i^{\min}$ (because $\tilde{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$), which implies $x_i > 0$ (since $-\ell_i < z_i = x_i - \tilde{x}_i \leq x_i - \ell_i$). Similarly, if $i \in S_2^-$ then $\ell_i < -\delta_i^{\min} = -u_i$; thus, $-\tilde{x}_i \geq -u_i$, which shows that $x_i > 0$. Moreover, if $i \in S_4^+ \cup S_4^-$ then $x_i \neq 0$. In total, this shows that $|S_4^+| + |S_4^-| + |S_2^+| + |S_2^-| \leq |\mathrm{supp}(\boldsymbol{x})| \leq s$. Since all sets $S_i^{\pm}$ are disjoint or empty, this concludes the proof. $\square$

Corollary 13.

1. *Let $X = [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$. A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is $(s, [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}, 0)$-good if and only if*

$$\mathcal{N}(\boldsymbol{A}) \cap \{\boldsymbol{z} \in [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}} \,:\, |\mathrm{supp}(\boldsymbol{z}; (\boldsymbol{0}, \boldsymbol{u}])| \leq s,\ |\mathrm{supp}(\boldsymbol{z}; [-\boldsymbol{u}, \boldsymbol{0}))| \leq s\} = \{\boldsymbol{0}\}. \tag{2}$$

2. *Let $X = \mathbb{Z}_+^n$. A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is $(s, \mathbb{Z}_+^n, 0)$-good if and only if*

$$\mathcal{N}(\boldsymbol{A}) \cap \{\boldsymbol{z} \in \mathbb{Z}^n \,:\, |\mathrm{supp}(\boldsymbol{z}; \mathbb{Z}_{>0}^n)| \leq s,\ |\mathrm{supp}(\boldsymbol{z}; \mathbb{Z}_{<0}^n)| \leq s\} = \{\boldsymbol{0}\}. \tag{3}$$

3. *Let* $X = [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$. *A matrix* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *is* $(s, [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}, 0)$*-good if and only if*

$$\mathcal{N}(\boldsymbol{A}) \cap \{\boldsymbol{z} \in [-2\cdot \boldsymbol{u}, 2\cdot \boldsymbol{u}]_{\mathbb{Z}} \,:\, 2\,|\mathrm{supp}(\boldsymbol{z}; [-2\cdot\boldsymbol{u}, -\boldsymbol{u}))| \\ +2|\mathrm{supp}(\boldsymbol{z}; (\boldsymbol{u}, 2\cdot\boldsymbol{u}])| + |\mathrm{supp}(\boldsymbol{z}; [-\boldsymbol{u}, \boldsymbol{0}) \cup (\boldsymbol{0}, \boldsymbol{u}])| \leq 2s\} = \{\boldsymbol{0}\}. \quad (4)$$

Proof. 1. We set $\boldsymbol{\ell} = \boldsymbol{0}$ in Theorem 12. In this case, $\boldsymbol{\delta}^{\min} = \boldsymbol{0}$ and $\boldsymbol{\delta}^{\max} = \boldsymbol{u}$. Thus, only S_2^+ and S_3^- can be nonempty, which yields

$$C(\boldsymbol{0}, \boldsymbol{u}) := \{\boldsymbol{z} \in [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}} \,:\, |S_3^-| \leq s,\ |S_2^+| \leq s,\ |S_3^-| + |S_2^+| \leq 2s\}.$$

Since $|S_3^-| + |S_2^+| \leq 2s$ is redundant, $S_3^+ = \{i \,:\, z_i \in (0, u_i]\}$ and $S_2^- = \{i \,:\, z_i \in [-u_i, 0)\}$, the condition from Theorem 12 is equivalent to that stated in (2).

2. This follows from the previous part by letting the components of $\boldsymbol{u}$ tend to infinity.

3. We set $\boldsymbol{\ell} = -\boldsymbol{u}$ in Theorem 12. In this case, $\boldsymbol{\delta}^{\min} = \boldsymbol{u}$ and $\boldsymbol{\delta}^{\max} = \boldsymbol{u}$. Thus, only S_1^+, S_1^-, S_4^+, and S_4^- can be nonempty, which yields

$$C(-\boldsymbol{u}, \boldsymbol{u}) := \{\boldsymbol{z} \in [-2\cdot\boldsymbol{u}, 2\cdot\boldsymbol{u}]_{\mathbb{Z}} \,:\, |S_4^-| + |S_4^+| \leq s,\ |S_4^-| + |S_4^+| \leq s, \\ 2\left(|S_4^+| + |S_4^-|\right) + |S_1^-| + |S_1^+| \leq 2s\}.$$

Since the first two constraints are identical and implied by the third one, we obtain from Theorem 12 the equivalent condition (4). ☐

Remark 14. *In particular, in the binary case, i.e., if* $X = [\boldsymbol{0}, \mathbb{1}]_{\mathbb{Z}}$*, then* $A \in \mathbb{R}^{m\times n}$ *is* $(s, [\boldsymbol{0}, \mathbb{1}]_{\mathbb{Z}}, 0)$*-good if and only if* $\mathcal{N}(\boldsymbol{A}) \cap \{\boldsymbol{z} \in \{-1, 0, +1\}^n \,:\, |\{i \,:\, z_i = -1\}| \leq s,\ |\{i \,:\, z_i = 1\}| \leq s\} = \{\boldsymbol{0}\}$.

Remark 15. *Note that Theorem 9 also follows from Theorem 12: We let the components of* $\boldsymbol{\ell}$ *and* $\boldsymbol{u}$ *simultaneously tend to* $-\infty$ *and* ∞*, respectively. Then* $\delta_i^{\min} = -\ell_i \to \infty$ *and* $\delta_i^{\max} = u_i \to \infty$ *for all* i*. In this case, only* S_1^+ *and* S_1^- *can be nonempty. Thus, the conditions of Theorem 12 reduce to* $|S_1^+| + |S_1^-| \leq 2s$ *and hence yield Theorem 9.*

Remark 16. *In the case of real-valued vectors, the proof of Theorem 12 carries over directly and yields an analogous statement for* $(\mathrm{P}_0([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{R}}))$ *in which the vectors in* $\mathcal{N}(\boldsymbol{A})$ *are allowed to be real. The same holds for results analogous to Corollary 13; however, note that* $\boldsymbol{A}$ *is* $(s, [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{R}}, 0)$*-good if and only if it is* $(s, \mathbb{R}^n_+, 0)$*-good, due to the scalability of (real) nullspace vectors.*

4. The ℓ_1-case

As noted earlier, if it were not for the integrality constraints, ($\mathrm{P}_1(\mathbb{Z}^n)$), ($\mathrm{P}_1(\mathbb{Z}^n_+)$), ($\mathrm{P}_1([-\boldsymbol{u},\boldsymbol{u}]_\mathbb{Z})$), ($\mathrm{P}_1([\boldsymbol{0},\boldsymbol{u}]_\mathbb{Z})$) and ($\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_\mathbb{Z})$) could all be reformulated as linear programs (LPs). Hence, from the viewpoint of integer programming, it is natural to ask under which conditions the LP relaxations of these problems (which will be denoted by ($\mathrm{P}_1^{\mathrm{LP}}(X)$)) are guaranteed to have integral optimal solutions themselves. To that end, we can resort to some well-established polyhedral results often encountered in discrete and combinatorial optimization which build on the concepts of (total) unimodularity and total dual integrality; we provide several results obtained by this approach in Subsection 4.1 below. A different viewpoint is taken by Keiper et al. [23], who consider recovery conditions and phase transitions, but restrict to solutions in $[\boldsymbol{0},\mathbb{1}]_\mathbb{R}$, $[-\mathbb{1},\mathbb{1}]_\mathbb{R}$ or $[\boldsymbol{0},2\cdot\mathbb{1}]_\mathbb{R}$.

We will briefly recall some terminology from polyhedral theory and refer to the books by Schrijver [27] and Korte and Vygen [31] for broad overviews and collections of classical results. Since the above-mentioned general polyhedral integrality results do not involve the aspect of solution sparsity, in Subsection 4.2, we further give characterizations of unique recoverability of sparse integral vectors by ℓ_1-minimization, based on extensions of the well-known nullspace condition.

4.1. Integral LP Relaxations

We begin by considering the following question: When does the LP relaxation ($\mathrm{P}_1^{\mathrm{LP}}(X)$) of ($\mathrm{P}_1(X)$) have integral optimal solutions for every right hand side vector $\boldsymbol{b}$? A first answer can be obtained for unimodular matrices $\boldsymbol{A}\in\mathbb{Z}^{m\times n}$, i.e., those for which every regular $m\times m$ submatrix has determinant ±1:

Proposition 17. *Let $\boldsymbol{A}\in\mathbb{Z}^{m\times n}$ be unimodular with* $\mathrm{rank}(\boldsymbol{A})=m\leq n$ *and let $X=\mathbb{Z}^n$ or $X=\mathbb{Z}^n_+$. Then, for every $\boldsymbol{b}\in\mathbb{Z}^m$, the LP relaxation ($\mathrm{P}_1^{\mathrm{LP}}(X)$) has an integral optimal (vertex) solution, i.e., it lies in X.*

Proof. First, we consider $X=\mathbb{Z}^n_+$ and the LP

$$\min\,\left\{\|\boldsymbol{x}\|_1\,:\,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\ \boldsymbol{x}\geq\boldsymbol{0}\right\}=\min\,\left\{\mathbb{1}^\top\boldsymbol{x}\,:\,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\ \boldsymbol{x}\geq\boldsymbol{0}\right\}.\quad(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n_+))$$

Standard results (cf., e.g., [27]) show that unimodularity of $\boldsymbol{A}$ is equivalent to the integrality of the polyhedron $\{\boldsymbol{x}\,:\,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq\boldsymbol{0}\}$ for every $\boldsymbol{b}\in\mathbb{Z}^m$. In particular, there exists an integral optimal vertex solution for ($\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n_+)$), since ($\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n_+)$) always has a finite value; this solution is also optimal for ($\mathrm{P}_1(\mathbb{Z}^n_+)$).

Now consider $X=\mathbb{Z}^n$. By means of the standard variable split $\boldsymbol{x}=\boldsymbol{x}^+-\boldsymbol{x}^-$ with $\boldsymbol{x}^+:=\max\{\boldsymbol{0},\boldsymbol{x}\}$, $\boldsymbol{x}^-:=\max\{\boldsymbol{0},-\boldsymbol{x}\}$ (component-wise), we transform ($\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n)$) into the LP

$$\min\,\left\{\mathbb{1}^\top\boldsymbol{x}^++\mathbb{1}^\top\boldsymbol{x}^-\,:\,\boldsymbol{A}\boldsymbol{x}^+-\boldsymbol{A}\boldsymbol{x}^-=\boldsymbol{b},\ \boldsymbol{x}^+\geq\boldsymbol{0},\ \boldsymbol{x}^-\geq\boldsymbol{0}\right\}.\qquad(5)$$

Clearly, if $\boldsymbol{A}$ is unimodular, then so is $(\boldsymbol{A}, -\boldsymbol{A})$, and since (5) is of the same form as ($\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}_+^n)$), the conclusion carries over. It remains to note that every integral optimal vertex solution $(\bar{\boldsymbol{x}}^+, \bar{\boldsymbol{x}}^-)$ of (5) yields a corresponding integral optimal solution $\bar{\boldsymbol{x}} \coloneqq \bar{\boldsymbol{x}}^+ - \bar{\boldsymbol{x}}^-$ for ($\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n)$), which also solves ($\mathrm{P}_1(\mathbb{Z}^n)$). □

Strengthening the structural assumption on $\boldsymbol{A}$, we obtain analogous results for the remaining cases of integral sets considered in this paper:

Proposition 18. *Let $\boldsymbol{A} \in \mathbb{Z}^{m\times n}$ be totally unimodular (i.e., every square submatrix has determinant 0 or ± 1) with $\operatorname{rank}(\boldsymbol{A}) = m \le n$ and let $X = [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$, $X = [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$ or $X = [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ (with $\boldsymbol{\ell}, \boldsymbol{u} \in \mathbb{Z}^n$). Then, for every $\boldsymbol{b} \in \mathbb{Z}^m$, the LP relaxation ($\mathrm{P}_1^{\mathrm{LP}}(X)$) has an integral optimal (vertex) solution (in X).*

Proof. The results follow along the same lines as in the proof of Prop. 17 by rewriting ($\mathrm{P}_1^{\mathrm{LP}}(X)$) as LPs whose feasible sets are polyhedra which are integral for every $\boldsymbol{b} \in \mathbb{Z}^m$ and $\boldsymbol{\ell}, \boldsymbol{u} \in \mathbb{Z}^n$ if and only if $\boldsymbol{A}$ is totally unimodular. (The latter well-known characterizations can be found, e.g., in [27].) We omit the details to avoid repetition. □

Remark 19. *Note that, while Propositions 17 and 18 do not assert* unique *recoverability of $\hat{\boldsymbol{x}} \in X$ by solving ($\mathrm{P}_1^{\mathrm{LP}}(X)$) with $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$, they nevertheless guarantee that a feasible integral vector with the same ℓ_1-norm as $\hat{\boldsymbol{x}}$ can be found efficiently. Also, the requirement of (total) unimodularity can be checked in polynomial time, cf. Seymour [32], Truemper [33] and Walter and Truemper [34]. (Furthermore, total unimodularity naturally implies unimodularity, so the statement of Proposition 17 also holds for totally unimodular $\boldsymbol{A}$.)*

If one is interested in solution integrality of ($\mathrm{P}_1^{\mathrm{LP}}(X)$) for a *specific* $\boldsymbol{b}$ only, (total) unimodularity can be weakened to requiring *total dual integrality*: For $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$, $\boldsymbol{b} \in \mathbb{Q}^m$, the system $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$ is totally dual integral (TDI) if for every $\boldsymbol{c} \in \mathbb{Q}^n$ such that the LP $\min\{\boldsymbol{b}^\top \boldsymbol{y} \,:\, \boldsymbol{A}^\top \boldsymbol{y} = \boldsymbol{c},\ \boldsymbol{y} \ge \boldsymbol{0}\}$ is finite, it has an integral optimal solution. Then, a well-known result (see, e.g., [31, Corollary 5.14]) states that if $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$ is TDI and $\boldsymbol{b} \in \mathbb{Z}^m$, all vertices of $\{\boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}\}$ are integral. This and related results can be combined with the LP relaxation ($\mathrm{P}_1^{\mathrm{LP}}(X)$) to obtain sufficient conditions for integrality of optimal solutions similar to those presented in Propositions 17 and 18; for the sake of brevity, we do not state this explicitly here. However, testing whether an (in-)equality system is TDI is NP-hard (see Ding et al. [35]), so these weaker conditions are harder to verify (for a given instance).

Note also that the LP relaxations have (possibly after a variable split) objective function coefficients $\boldsymbol{c} = \mathbb{1}$. Thus, TDI is a stronger requirement, since it pertains to essentially *all* $\boldsymbol{c}$, not just one specific one. Nevertheless, we can make use of total dual integrality to obtain characterizations of relaxation solution integrality for every $\boldsymbol{b} \in \mathbb{Z}^m$ by requiring the description of the *dual* polyhedron (in which $\boldsymbol{c} = \mathbb{1}$ acts as the right hand side vector) to be TDI. For the sake of exposition, we do not go into full generality, but will consider only binary matrices $\boldsymbol{A}$ in the remainder of this subsection.

Let us start by considering $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}_+^n)) = (\mathrm{P}_1(\mathbb{R}_+^n))$. We need some more terminology (see [31] and Cornuéjols [36] for more details): Given a (simple, undirected) graph $G = (V, E)$, the *clique-node (incidence) matrix* $\boldsymbol{A}_G$ of G has one column per node and one row per clique (i.e., complete subgraph) of G, with the (i, j)-entry equal to 1 if clique i contains node j, and zero otherwise. Further, recall that a graph G is called *perfect* if for every node-induced subgraph H of G, the chromatic number $\chi(H)$ equals the clique number $\omega(H)$ (i.e., the minimal number of colors needed to color the nodes of H such that no neighbors have the same color coincides with the cardinality of a maximum clique in H).

Theorem 20. *Let $\boldsymbol{A}^\top \in \{0,1\}^{n\times m}$ be the clique-node matrix of a perfect graph. Then, for every $\boldsymbol{b} \in \mathbb{Z}^m$, $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}_+^n))$ has integral optimal solutions.*

For the proof, we need the following well-known result; in the Appendix, we provide a proof that will be used later.

Proposition 21 ([36, Exercise 3.6])**.** *Let $G = (V, E)$ be a perfect graph with clique-node incidence matrix $\boldsymbol{A}_G$. Then, the system $\boldsymbol{A}_G\,\boldsymbol{y} \leq \mathbb{1}$, $\boldsymbol{y} \geq \boldsymbol{0}$ is TDI.*

Proof of Theorem 20. Let $\boldsymbol{A}^\top$ be the clique-node matrix of a perfect graph. Then, the dual problem of $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}_+^n))$ can be written as

$$\max\ \left\{\boldsymbol{b}^\top\boldsymbol{y} \,:\, \boldsymbol{A}^\top\boldsymbol{y} \leq \mathbb{1}\right\} \quad \Leftrightarrow \quad \max\ \Big\{\boldsymbol{b}^\top\boldsymbol{y} \,:\, \sum_{i\in C} y_i \leq 1\ \forall\,\text{cliques } C \text{ of } G\Big\}.$$

Tracing the proof of Proposition 21 in the Appendix, it is easy to see that the solution $\boldsymbol{x}$ for (A.1) constructed there satisfies all inequality constraints of that problem with equality; consequently, the system $\boldsymbol{A}^\top\boldsymbol{y} \leq \mathbb{1}$ (without nonnegativity of $\boldsymbol{y}$) is also TDI. Hence, by definition of total dual integrality, $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}_+^n))$ has integral optimal solutions for every $\boldsymbol{b} \in \mathbb{Z}^m$. ☐

If $\boldsymbol{A}^\top$ is neither (totally) unimodular nor the clique-node matrix of a perfect graph, integrality of the LP relaxation solutions is indeed not ensured in general, as the following example shows.

Example 22. *Consider*

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

and $\boldsymbol{b} = \mathbb{1} \in \mathbb{R}^3$. In this case, $\{\boldsymbol{x} \in \mathbb{Z}_+^3 \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ is empty, while there exists a (unique) continuous solution $\boldsymbol{x} = \frac{1}{2}\cdot\mathbb{1}$. Moreover, consider

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

and $\boldsymbol{b} = \mathbb{1} \in \mathbb{R}^6$*. Here,* $\boldsymbol{x} = (1,0,0,1,1,1)^\top$ *is an optimal solution to* $(\mathrm{P}_1(\mathbb{Z}^6_+))$*. However,* $\boldsymbol{x} = (\frac{1}{2},\frac{1}{2},\frac{1}{2},0,0,0)^\top$ *is an optimal solution to* $(\mathrm{P}_1(\mathbb{R}^6_+))$ *(with fewer nonzeros).*

Remark 23. *Note that in Theorem 20,* $\boldsymbol{A}^\top$ *is the clique-node matrix with respect to* all *cliques in a perfect graph. This differs from the usual theory, which allows for restricting to (inclusion-wise) maximal cliques (see the already accordingly restricted definition of clique-node matrices in [36]). Indeed, Proposition 21 remains true under such a restriction, but this does not carry over to Theorem 20. Also, one can easily find examples which show that one does not necessarily need to include all cliques to achieve total dual integrality of* $\boldsymbol{A}^\top\boldsymbol{y} \leq \mathbb{1}$*, which in turn (with binary* $\boldsymbol{A}$*) does not imply that* $\boldsymbol{A}^\top$ *is the clique-node matrix of a perfect graph.*

We can extend the results from Theorem 20 to a sufficient condition for solution integrality for $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n))$.

Theorem 24. *Let* $\boldsymbol{A}^\top \in \{0,1\}^{n\times m}$ *be the clique-node matrix of a perfect graph. Then, for every* $\boldsymbol{b} \in \mathbb{Z}^m$*,* $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n))$ *has integral optimal solutions.*

Proof. By means of a standard variable split $\boldsymbol{x} = \boldsymbol{x}^+ - \boldsymbol{x}^-$ with $\boldsymbol{x}^\pm := \max\{\boldsymbol{0}, \pm\boldsymbol{x}\}$ (component-wise), we can rewrite $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n))$ as the LP

$$\min\ \left\{\mathbb{1}^\top\boldsymbol{x}^+ + \mathbb{1}^\top\boldsymbol{x}^- \,:\, \boldsymbol{A}\boldsymbol{x}^+ - \boldsymbol{A}\boldsymbol{x}^- = \boldsymbol{b},\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}. \tag{6}$$

W.l.o.g., we may assume that $\boldsymbol{b} = ((\boldsymbol{b}^+)^\top, (\boldsymbol{b}^-)^\top)^\top$ with $\boldsymbol{b}^+ \geq \boldsymbol{0}$, $\boldsymbol{b}^- \leq \boldsymbol{0}$ (permuting rows, if necessary). Let $\boldsymbol{B} = (\boldsymbol{A}, -\boldsymbol{A})$, and denote by $\boldsymbol{B}^+$ and $\boldsymbol{B}^-$ the submatrices corresponding to the rows associated with $\boldsymbol{b}^+$ and $\boldsymbol{b}^-$, respectively. Furthermore, we write $\boldsymbol{B}^+ = (\boldsymbol{B}^+_+, \boldsymbol{B}^+_-)$ and $\boldsymbol{B}^- = (\boldsymbol{B}^-_+, \boldsymbol{B}^-_-)$ to distinguish the respective columns corresponding to $\boldsymbol{x}^+$ and $\boldsymbol{x}^-$. We can now rewrite (6) and relax its constraints as follows:

$$\begin{aligned}
(6) = \min\ &\left\{\mathbb{1}^\top\boldsymbol{x}^+ + \mathbb{1}^\top\boldsymbol{x}^- \,:\, \begin{pmatrix}\boldsymbol{B}^+_+ & \boldsymbol{B}^+_-\\ \boldsymbol{B}^-_+ & \boldsymbol{B}^-_-\end{pmatrix}\begin{pmatrix}\boldsymbol{x}^+\\ \boldsymbol{x}^-\end{pmatrix} = \begin{pmatrix}\boldsymbol{b}^+\\ \boldsymbol{b}^-\end{pmatrix},\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}\\
\geq \min\ &\left\{\mathbb{1}^\top\boldsymbol{x}^+ \,:\, \boldsymbol{B}^+_+\boldsymbol{x}^+ + \boldsymbol{B}^+_-\boldsymbol{x}^- = \boldsymbol{b}^+,\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}\\
&+ \min\ \left\{\mathbb{1}^\top\boldsymbol{x}^- \,:\, \boldsymbol{B}^-_+\boldsymbol{x}^+ + \boldsymbol{B}^-_-\boldsymbol{x}^- = \boldsymbol{b}^-,\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}\\
\geq \min\ &\left\{\mathbb{1}^\top\boldsymbol{x}^+ \,:\, \boldsymbol{B}^+_+\boldsymbol{x}^+ \geq \boldsymbol{b}^+ - \boldsymbol{B}^+_-\boldsymbol{x}^-,\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}\\
&+ \min\ \left\{\mathbb{1}^\top\boldsymbol{x}^- \,:\, \boldsymbol{B}^-_+\boldsymbol{x}^+ \leq \boldsymbol{b}^- - \boldsymbol{B}^-_-\boldsymbol{x}^-,\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}.
\end{aligned}$$

Observing that $\boldsymbol{B}^+_-\boldsymbol{x}^- \leq \boldsymbol{0} \leq \boldsymbol{B}^-_+\boldsymbol{x}^+$ (since $\boldsymbol{A}$ is binary), we can further relax the last two programs as

$$\begin{aligned}
&\min\ \left\{\mathbb{1}^\top\boldsymbol{x}^+ \,:\, \boldsymbol{B}^+_+\boldsymbol{x}^+ \geq \boldsymbol{b}^+ - \boldsymbol{B}^+_-\boldsymbol{x}^-,\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}\\
\geq\ &\min\ \left\{\mathbb{1}^\top\boldsymbol{x}^+ \,:\, \boldsymbol{B}^+_+\boldsymbol{x}^+ \geq \boldsymbol{b}^+,\, \boldsymbol{x}^+ \geq \boldsymbol{0}\right\}
\end{aligned}$$

and

$$\begin{aligned}
&\min\ \left\{\mathbb{1}^\top\boldsymbol{x}^- \,:\, \boldsymbol{B}^-_+\boldsymbol{x}^+ \leq \boldsymbol{b}^- - \boldsymbol{B}^-_-\boldsymbol{x}^-,\, \boldsymbol{x}^+ \geq \boldsymbol{0},\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}\\
\geq\ &\min\ \left\{\mathbb{1}^\top\boldsymbol{x}^- \,:\, -\boldsymbol{B}^-_-\boldsymbol{x}^- \geq -\boldsymbol{b}^-,\, \boldsymbol{x}^- \geq \boldsymbol{0}\right\}.
\end{aligned}$$

Thus, (6) is bounded from below by the sum of two linear programs, each of which can easily be seen to be associated with a (generalized) set covering problem.

By definition, $(\boldsymbol{B}_{+}^{+})^{\top}$ is a matrix whose rows are the incidence vectors of all cliques of a perfect graph (a node-induced subgraph of the graph represented by $\boldsymbol{A}^{\top}$); the same holds for $(-\boldsymbol{B}_{-}^{-})^{\top}$. Hence, by Proposition 21 and the proof of Theorem 20, both LPs have optimal integral solutions—say, $\boldsymbol{x}_{*}^{+}$ and $\boldsymbol{x}_{*}^{-}$—that satisfy all inequality constraints with equality.

Moreover, for every column of $\boldsymbol{A}$ representing a clique (and thus for any column of $\boldsymbol{B}_{+}^{+}$ or $\boldsymbol{B}_{-}^{-}$ representing a clique), there is another column for every subclique. Therefore, the solutions $\boldsymbol{x}_{*}^{+}$ and $\boldsymbol{x}_{*}^{-}$ may be chosen such that $\boldsymbol{B}_{+}^{-}\boldsymbol{x}_{*}^{+} = \boldsymbol{0} = \boldsymbol{B}_{-}^{+}\boldsymbol{x}_{*}^{-}$.

It remains to observe that $\boldsymbol{x}_{*} \coloneqq \boldsymbol{x}_{*}^{+} - \boldsymbol{x}_{*}^{-}$ is an integral feasible solution for $(\mathrm{P}_1^{\mathrm{LP}}(\mathbb{Z}^n))$, and since it achieves the lower bound given by the two LPs derived above, it is, in fact, optimal. □

4.2. Uniform Sparse Recovery Conditions

The polyhedral ideas discussed in the previous section do not take solution sparsity into account. To obtain more succinct results for recovery of *sparse* integral vectors by ℓ_1-norm minimization, we will now turn to conditions on the nullspace of the sensing matrix. The goal is to investigate the following property, similar to Definition 6:

Definition 25. *Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $s \in [n]$, and $X \subseteq \mathbb{R}^n$. The matrix $\boldsymbol{A}$ is* $(s, X, 1)$-good*, if* every *s-sparse vector $\hat{\boldsymbol{x}} \in X$ is the unique solution of* $(\mathrm{P}_1(X))$ *with $\boldsymbol{b} = \boldsymbol{A}\hat{\boldsymbol{x}}$.*

In fact, if all s-sparse ℓ_1-minimizers are unique, they also solve the respective ℓ_0-minimization problems:

Proposition 26. *If $\boldsymbol{A}$ is $(s, X, 1)$-good, then it is $(s, X, 0)$-good.*

Proof. Let $\boldsymbol{A}$ be $(s, X, 1)$-good. Assume there is a minimizer $\boldsymbol{z}$ of $(\mathrm{P}_0(X))$ with $\boldsymbol{b} = \boldsymbol{A}\hat{\boldsymbol{x}}$ for some s-sparse $\hat{\boldsymbol{x}}$. Then, $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\hat{\boldsymbol{x}}$ and $\|\boldsymbol{z}\|_0 \leq \|\hat{\boldsymbol{x}}\|_0 \leq s$, so that $\boldsymbol{z} = \hat{\boldsymbol{x}}$ must hold, since $\hat{\boldsymbol{x}}$ is (by definition of $(s, X, 1)$-goodness) the unique minimizer of $(\mathrm{P}_1(X))$ with $\boldsymbol{b} = \boldsymbol{A}\hat{\boldsymbol{x}}$. Thus, $|S(s, X, \boldsymbol{b})| = 1$, i.e., $\boldsymbol{A}$ is $(s, X, 0)$-good. □

For the sake of brevity, we will not explicitly repeat the corresponding inferences regarding $(s, X, 0)$-goodness in all the following results pertaining to $(s, X, 1)$-goodness, as they simply follow from Proposition 26.

For $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, a set $S \subseteq [n]$ and some $V \subseteq \mathbb{R}^n$, we define the following two *nullspace properties* (NSPs):

$$\mathrm{NSP}(V): \qquad \|\boldsymbol{v}_S\|_1 < \|\boldsymbol{v}_{S^c}\|_1 \qquad \forall\, \boldsymbol{v} \in (V \cap \mathcal{N}(\boldsymbol{A})) \setminus \{\boldsymbol{0}\},$$
$$\mathrm{NSP}_{+}(V): \qquad \boldsymbol{v}_{S^c} \geq 0 \;\Rightarrow\; \mathbb{1}^{\top}\boldsymbol{v} > 0 \qquad \forall\, \boldsymbol{v} \in (V \cap \mathcal{N}(\boldsymbol{A})) \setminus \{\boldsymbol{0}\}.$$

If a matrix $\boldsymbol{A}$ satisfies one of these conditions for *all* sets S of cardinality $|S| \leq s$, we say that the respective NSP *of order* s is satisfied.

In the continuous setting, nullspace properties are well-known to yield the strongest results relating ℓ_1-minimization to the recovery of sparse vectors; we summarize the fundamental results in the following theorem.

Theorem 27. *Let* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *and* $S \subseteq [n]$.

1. *Every vector* $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique solution of* $(\mathrm{P}_1(\mathbb{R}^n))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}(\mathbb{R}^n)$ *w.r.t. the set* S. *Moreover,* $\boldsymbol{A}$ *is* $(s, \mathbb{R}^n, 1)$*-good if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}(\mathbb{R}^n)$ *of order* s.

2. *Every vector* $\hat{\boldsymbol{x}} \in \mathbb{R}^n_+$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique solution of* $(\mathrm{P}_1(\mathbb{R}^n_+))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}_+(\mathbb{R}^n)$ *w.r.t. the set* S. *Moreover,* $\boldsymbol{A}$ *is* $(s, \mathbb{R}^n_+, 1)$*-good if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}_+(\mathbb{R}^n)$ *of order* s.

Proof. For a proof of statement 1, see, e.g., [3], and for statement 2, see [6]. □

In fact, the proofs that yield Theorem 27 can almost literally be translated to the case of $(\mathrm{P}_1(\mathbb{Z}^n))$ and $(\mathrm{P}_1(\mathbb{Z}^n_+))$ by additionally requiring integrality of the nullspace vectors. Thus, we immediately obtain the following result.

Theorem 28. *Let* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *and* $S \subseteq [n]$.

1. *Every vector* $\hat{\boldsymbol{x}} \in \mathbb{Z}^n$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique solution of* $(\mathrm{P}_1(\mathbb{Z}^n))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}(\mathbb{Z}^n)$ *w.r.t. the set* S. *Moreover,* $\boldsymbol{A}$ *is* $(s, \mathbb{Z}^n, 1)$*-good if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}(\mathbb{Z}^n)$ *of order* s.

2. *Every vector* $\hat{\boldsymbol{x}} \in \mathbb{Z}^n_+$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique solution of* $(\mathrm{P}_1(\mathbb{Z}^n_+))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}_+(\mathbb{Z}^n)$ *w.r.t. the set* S. *Moreover,* $\boldsymbol{A}$ *is* $(s, \mathbb{Z}^n_+, 1)$*-good if and only if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}_+(\mathbb{Z}^n)$ *of order* s.

Similarly to Theorem 9, for rational matrices there is no difference between the standard (continuous) NSPs and their integral counterparts, since rational kernel vectors can always be rescaled to integrality:

Corollary 29. *Let* $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$. *Then* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}(\mathbb{Z}^n)$ *if and only if it satisfies* $\mathrm{NSP}(\mathbb{R}^n)$*, and it satisfies* $\mathrm{NSP}_+(\mathbb{Z}^n)$ *if and only if it satisfies* $\mathrm{NSP}_+(\mathbb{R}^n)$.

As a consequence, for rational data, signal integrality does not lead to recoverability (by ℓ_1-norm minimization) of lower sparsity levels—i.e., larger number of nonzeros—than in the continuous case. However, this situation again changes once the signal is bounded.

As a first criterion for $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$, we consider $\mathrm{NSP}([\boldsymbol{\ell} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{\ell}]_{\mathbb{Z}})$, which leads to the following result.

Theorem 30. *If* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *satisfies* $\mathrm{NSP}([\boldsymbol{\ell} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{\ell}]_{\mathbb{Z}})$ *w.r.t. a set* $S \subseteq [n]$*, then every vector* $\hat{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique solution of* $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$. *Moreover, if* $\boldsymbol{A}$ *satisfies* $\mathrm{NSP}([\boldsymbol{\ell} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{\ell}]_{\mathbb{Z}})$ *of order* s*, then* $\boldsymbol{A}$ *is* $(s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 1)$*-good.*

Proof. Assume $\boldsymbol{A}$ satisfies $\mathrm{NSP}([\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}})$ w.r.t. S. Suppose $\hat{\boldsymbol{x}} \in [\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}$ has $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$, and let $\boldsymbol{z} \in [\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}} \setminus \{\hat{\boldsymbol{x}}\}$ satisfy $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\hat{\boldsymbol{x}}$. Then, $\boldsymbol{v} \coloneqq \hat{\boldsymbol{x}} - \boldsymbol{z} \in (\mathcal{N}(\boldsymbol{A}) \cap [\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}}) \setminus \{\boldsymbol{0}\}$ and thus,

$$\begin{aligned}\|\hat{\boldsymbol{x}}\|_1 &\leq \|\hat{\boldsymbol{x}} - \boldsymbol{z}_S\|_1 + \|\boldsymbol{z}_S\|_1 = \|\boldsymbol{v}_S\|_1 + \|\boldsymbol{z}_S\|_1 \\ &< \|\boldsymbol{v}_{S^c}\|_1 + \|\boldsymbol{z}_S\|_1 = \|\boldsymbol{z}_{S^c}\|_1 + \|\boldsymbol{z}_S\|_1 = \|\boldsymbol{z}\|_1.\end{aligned}$$

It follows that $\hat{\boldsymbol{x}}$ is the unique optimal solution of $(\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}))$ with $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$. Furthermore, by letting the set S vary, we immediately obtain the claim about uniqueness of all s-sparse solutions of $(\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}))$, i.e., $(s, X, 1)$-goodness of $\boldsymbol{A}$. □

The above proof is a straightforward adaptation of the sufficiency part of the proof of the original results for $(\mathrm{P}_1(\mathbb{R}^n))$ and $\mathrm{NSP}(\mathbb{R}^n)$ to the bounded-integers setting. Unfortunately, the condition of Theorem 30 (i.e., $\mathrm{NSP}([\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}})$) is no longer necessary in the present case, as the following toy example shows:

Example 31. *Let $\boldsymbol{A} = (1,2)$, $-\boldsymbol{\ell} = \boldsymbol{u} = \mathbb{1}$ and $S = \{1\}$. Clearly, every vector in $[\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}$ supported on S (i.e., either $(0,0)^\top$, $(1,0)^\top$ or $(-1,0)^\top$) is the unique minimizer of $(\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}))$ with the associated $\boldsymbol{b}$. However, it holds that $(\mathcal{N}(\boldsymbol{A}) \cap [-2\cdot\mathbb{1}, 2\cdot\mathbb{1}]_{\mathbb{Z}}) \setminus \{\boldsymbol{0}\} = \{(-2,1)^\top, (2,-1)^\top\}$. Both of these vectors violate the condition of $\mathrm{NSP}([\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}})$, which here simply amounts to $|v_1| < |v_2|$ for all $\boldsymbol{v} = (v_1, v_2)^\top$ in the above nullspace subset.*

We will give a complete characterization for sparse recovery via $(\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}))$ later (see Theorem 37 below), but first point out a few more observations. The first one is again due to the scalability of nullspace vectors:

Corollary 32. *For $(\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{R}}))$ (with $\boldsymbol{\ell} \leq \boldsymbol{0} \leq \boldsymbol{u}$, $\boldsymbol{\ell} < \boldsymbol{u}$ both in $\mathbb{R}^n$), the analogous $\mathrm{NSP}([\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{R}})$ is equivalent to the standard $\mathrm{NSP}(\mathbb{R}^n)$.*

Moreover, even though Theorem 30 only provides a sufficient condition for integral sparse recovery by means of $(\mathrm{P}_1([\boldsymbol{\ell},\boldsymbol{u}]_{\mathbb{Z}}))$, one can easily find examples which demonstrate that it is already strictly weaker than its continuous analogon. Trivial examples are obtained in cases in which there are no integral kernel vectors satisfying the bounds $[\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}}$ other than $\boldsymbol{0}$ itself. More interestingly, consider the following case:

Example 33. *We revisit Example 2: Let $-\boldsymbol{\ell} = \boldsymbol{u} = \mathbb{1} \in \mathbb{Z}^n$ with $n \geq 6$, let $\boldsymbol{v} = (1, -1, (\boldsymbol{v}')^\top)^\top$ and $\boldsymbol{w} = (\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor, (\boldsymbol{w}')^\top)^\top$ with $\boldsymbol{v}', \boldsymbol{w}' \in \{-1,1\}^{n-2}$ arbitrary, and let $\boldsymbol{A}$ be such that $\mathcal{N}(\boldsymbol{A}) = \mathrm{span}\{\boldsymbol{v}, \boldsymbol{w}\}$. By construction, $\pm\boldsymbol{v}$ are the only nonzero vectors in $\mathcal{N}(\boldsymbol{A}) \cap [\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}}$. Moreover, for any $S \subset [n]$ with $|S| \leq s \coloneqq \lfloor \frac{n}{2} \rfloor - 1$, it holds that $\|\pm\boldsymbol{v}_S\|_1 \leq s < \lceil \frac{n}{2} \rceil \leq \|\pm\boldsymbol{v}_{S^c}\|_1$, which means that $\boldsymbol{A}$ satisfies $\mathrm{NSP}([\boldsymbol{\ell}-\boldsymbol{u},\boldsymbol{u}-\boldsymbol{\ell}]_{\mathbb{Z}})$ of order s. On the other hand, for $T = \{1,2\}$ we have $\|\boldsymbol{w}_T\|_1 = 2 \cdot \lfloor \frac{n}{2} \rfloor \geq n-1 > n-2 = \|\boldsymbol{w}_{S^c}\|_1$, which reveals that $\boldsymbol{A}$ violates the (standard continuous) $\mathrm{NSP}(\mathbb{R}^n)$ for all orders $t \geq 2$. In conclusion, the difference of recoverable sparsity orders $s - t$ can grow arbitrarily large—i.e., integral basis pursuit with bounds is able to reconstruct integral signals up to* *much larger numbers of nonzeros than what could be guaranteed by integrality-oblivious previous results.*

Furthermore, note that the previous example also shows that in the presence of bounds, the integral and continuous nullspace properties no longer coincide for rational matrices $\boldsymbol{A}$.

Finally, by setting $\boldsymbol{\ell} = -\boldsymbol{u}$, we obtain results analogous to Theorem 30 for the case $X = [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$:

Corollary 34. *If $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ satisfies $\mathrm{NSP}([-2\cdot\boldsymbol{u}, 2\cdot\boldsymbol{u}]_{\mathbb{Z}})$ w.r.t. a set $S \subseteq [n]$, then every vector $\hat{\boldsymbol{x}} \in [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$ with $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ is the unique solution of $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ with $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$. Moreover, if $\boldsymbol{A}$ satisfies $\mathrm{NSP}([-2\cdot\boldsymbol{u}, 2\cdot\boldsymbol{u}]_{\mathbb{Z}})$ of order s, then $\boldsymbol{A}$ is $(s, [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}, 1)$-good.*

We now turn our attention to $(\mathrm{P}_1([\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}))$. The following result provides a characterization of recoverability for sparse nonnegative and upper-bounded integral signals.

Theorem 35. *Every vector $\hat{\boldsymbol{x}} \in [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$ with $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ is the unique optimal solution of $(\mathrm{P}_1([\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}))$ with $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ if and only if $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ satisfies $\mathrm{NSP}_+([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}})$ w.r.t. S. Moreover, $\boldsymbol{A}$ is $(s, [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}, 1)$-good if and only if $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ satisfies $\mathrm{NSP}_+([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}})$ of order s.*

Proof. We only prove the first statement, since the second one is again obtained immediately by letting the set S vary. We modify the proof of Theorem 27 part 2) (see [6]) to suit our setting: Suppose every $\hat{\boldsymbol{x}} \in [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$ with $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S \subseteq [n]$ is the unique minimizer of $(\mathrm{P}_1([\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}))$ with $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$. Let $\boldsymbol{0} \neq \boldsymbol{v} \in \mathcal{N}(\boldsymbol{A}) \cap [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$ and suppose $\boldsymbol{v}_{S^c} \geq \boldsymbol{0}$. Then,

$$\begin{aligned} \boldsymbol{A}\boldsymbol{v} = \boldsymbol{0} \quad &\Leftrightarrow \quad \boldsymbol{A}\boldsymbol{v}_S = \boldsymbol{A}(-\boldsymbol{v}_{S^c}) \\ &\Leftrightarrow \quad \boldsymbol{A}\boldsymbol{v}_S^+ - \boldsymbol{A}\boldsymbol{v}_S^- = -\boldsymbol{A}\boldsymbol{v}_{S^c} \quad \Leftrightarrow \quad \boldsymbol{A}\boldsymbol{v}_S^- = \boldsymbol{A}(\boldsymbol{v}_{S^c} + \boldsymbol{v}_S^+), \end{aligned}$$

where $\boldsymbol{v}^\pm \coloneqq \max\{\boldsymbol{0}, \pm\boldsymbol{v}\}$ (component-wise), so that $\boldsymbol{v} = \boldsymbol{v}^+ - \boldsymbol{v}^-$ and $\boldsymbol{v}^\pm \in [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$. Obviously, $\boldsymbol{v}_S^-$ is supported on S and $\boldsymbol{v}_{S^c} + \boldsymbol{v}_S^+ \geq \boldsymbol{0}$. By construction, $\boldsymbol{v}_S^-$ uniquely solves $(\mathrm{P}_1([\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}))$ with $\boldsymbol{b}_v \coloneqq \boldsymbol{A}\boldsymbol{v}_S^-$, so that $\|\boldsymbol{v}_S^-\|_1 < \|\boldsymbol{v}_{S^c} + \boldsymbol{v}_S^+\|_1$. In fact, we obtain

$$\begin{aligned} &\|\boldsymbol{v}_S^-\|_1 < \|\boldsymbol{v}_{S^c} + \boldsymbol{v}_S^+\|_1 = \|\boldsymbol{v}_{S^c}\|_1 + \|\boldsymbol{v}_S^+\|_1 \\ \Rightarrow \quad &\mathbb{1}^\top \boldsymbol{v} = \|\boldsymbol{v}_{S^c}\|_1 + \|\boldsymbol{v}_S^+\|_1 - \|\boldsymbol{v}_S^-\|_1 > 0. \end{aligned}$$

For the converse direction, suppose $\boldsymbol{A}$ satisfies $\mathrm{NSP}_+([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}})$ w.r.t. a set $S \subseteq [n]$. Let $\hat{\boldsymbol{x}} \in [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}$ with $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ and let $\boldsymbol{z} \in [\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}} \setminus \{\hat{\boldsymbol{x}}\}$ with $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\hat{\boldsymbol{x}}$. Consider $\boldsymbol{v} \coloneqq \boldsymbol{z} - \hat{\boldsymbol{x}}$; clearly, $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}) \cap [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}) \setminus \{\boldsymbol{0}\}$, and by construction, it holds that $\boldsymbol{v}_{S^c} = \boldsymbol{z}_{S^c} \geq \boldsymbol{0}$. By $\mathrm{NSP}_+([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}})$, this implies

$$\begin{aligned} &0 < \mathbb{1}^\top \boldsymbol{v} = \|\boldsymbol{v}_{S^c}\|_1 + \|\boldsymbol{v}_S^+\|_1 - \|\boldsymbol{v}_S^-\|_1 \\ \Leftrightarrow \quad &0 < \|\boldsymbol{z}_{S^c}\|_1 + \|(\boldsymbol{z} - \hat{\boldsymbol{x}})_S^+\|_1 - \|(\boldsymbol{z} - \hat{\boldsymbol{x}})_S^-\|_1 = \|\boldsymbol{z}\|_1 - \|\hat{\boldsymbol{x}}\|_1. \end{aligned}$$

To see the last equation, note that for $i \in S$, since $\hat{\boldsymbol{x}}, \boldsymbol{z} \geq \mathbf{0}$, either $0 \leq z_i - \hat{x}_i$ so that $|(z_i - \hat{x}_i)^+| = |z_i - \hat{x}_i| = z_i - \hat{x}_i = |z_i| - |\hat{x}_i|$ and $|(z_i - \hat{x}_i)^-| = 0$, or $0 < \hat{x}_i - z_i$ so that $|(z_i - \hat{x}_i)^+| = 0$ and $|(z_i - \hat{x}_i)^-| = |\hat{x}_i - z_i| = \hat{x}_i - z_i = |\hat{x}_i| - |z_i|$. Thus, $\|\hat{\boldsymbol{x}}\|_1 < \|\boldsymbol{z}\|_1$, which concludes the proof. □

For the continuous case, there is again no difference between the NSP_+ with bounds and the standard NSP_+, since we can always scale the kernel vectors accordingly:

Corollary 36. *For* $(\mathrm{P}_1([\mathbf{0}, \boldsymbol{u}]_{\mathbb{R}}))$*, the analogous* $\mathrm{NSP}_+([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{R}})$ *is equivalent to the standard* $\mathrm{NSP}_+(\mathbb{R}^n)$*.*

As mentioned earlier, the result from Theorem 35 can be transferred to the previously considered problem $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ by utilizing a standard variable split. We obtain the following recoverability characterizations:

Theorem 37. *Let* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *and* $S \subseteq [n]$*. Every vector* $\hat{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique optimal solution of* $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} := \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if* $(\boldsymbol{A}, -\boldsymbol{A})$ *satisfies* $\mathrm{NSP}_+([\binom{-\boldsymbol{u}}{\boldsymbol{\ell}}, \binom{\boldsymbol{u}}{-\boldsymbol{\ell}}]_{\mathbb{Z}})$ *w.r.t.* S*. Moreover,* $\boldsymbol{A}$ *is* $(s, [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}, 1)$*-good if and only if* $(\boldsymbol{A}, -\boldsymbol{A})$ *satisfies* $\mathrm{NSP}_+([\binom{-\boldsymbol{u}}{\boldsymbol{\ell}}, \binom{\boldsymbol{u}}{-\boldsymbol{\ell}}]_{\mathbb{Z}})$ *of order* s*.*

Proof. We split $\boldsymbol{x} = \boldsymbol{x}^+ - \boldsymbol{x}^-$ with $\boldsymbol{x}^\pm := \max\{\mathbf{0}, \pm\boldsymbol{x}\}$ (component-wise). Thus, $\boldsymbol{x}^+ \in [\mathbf{0}, \boldsymbol{u}]_{\mathbb{Z}}$ and $\boldsymbol{x}^- \in [\mathbf{0}, -\boldsymbol{\ell}]_{\mathbb{Z}}$ when $\boldsymbol{x} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$. Then, we can rewrite $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ as a problem in the form of nonnegative integral basis pursuit with upper bounds:

$$\min \left\{ \left\| \begin{pmatrix} \boldsymbol{x}^+ \\ \boldsymbol{x}^- \end{pmatrix} \right\|_1 : (\boldsymbol{A}, -\boldsymbol{A}) \begin{pmatrix} \boldsymbol{x}^+ \\ \boldsymbol{x}^- \end{pmatrix} = \boldsymbol{b}, \ \begin{pmatrix} \boldsymbol{x}^+ \\ \boldsymbol{x}^- \end{pmatrix} \in \left[\mathbf{0}, \begin{pmatrix} \boldsymbol{u} \\ -\boldsymbol{\ell} \end{pmatrix} \right]_{\mathbb{Z}} \right\}. \tag{7}$$

Note that we may assume w.l.o.g. complementarity of $\boldsymbol{x}^+$ and $\boldsymbol{x}^-$, i.e., that $x_i^+ \cdot x_i^- = 0$ for all i; otherwise, we could subtract $\min\{x_i^+, x_i^-\}$ from both values and thus reduce the objective, which can be written equivalently as $\mathbb{1}^\top \boldsymbol{x}^+ + \mathbb{1}^\top \boldsymbol{x}^-$. Hence, $\hat{\boldsymbol{x}}$ is the unique optimal solution of $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ if and only if $\hat{\boldsymbol{x}}^+$ and $\hat{\boldsymbol{x}}^-$ form the unique minimizer of (7). The claims now follow from Theorem 35 applied to the reformulation (7) of $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$. □

Remark 38. *The condition from Theorem 37 can be rephrased as follows: If* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$*, then* $((\hat{\boldsymbol{x}}^+)^\top, (\hat{\boldsymbol{x}}^-)^\top)^\top$ *is supported on* $T := S_+ \cup (n + S_-) \subseteq [2n]$*, where* $S_+ := \{i \in S : \hat{x}_i > 0\}$ *and* $S_- := \{i \in S : \hat{x}_i < 0\}$*. Since* $\mathcal{N}(\boldsymbol{A}, -\boldsymbol{A}) = \{\binom{\boldsymbol{v}}{\boldsymbol{w}} \in \mathbb{R}^{2n} : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}\boldsymbol{w}\}$ *and* $T^c = S_+^c \cup (n + S_-^c)$*, the variable split therefore yields that* $(\boldsymbol{A}, -\boldsymbol{A})$ *satisfies* $\mathrm{NSP}_+([\binom{-\boldsymbol{u}}{\boldsymbol{\ell}}, \binom{\boldsymbol{u}}{-\boldsymbol{\ell}}]_{\mathbb{Z}})$ *w.r.t.* S *if and only if for all* $\boldsymbol{v} \in [\mathbf{0}, \boldsymbol{u}]_{\mathbb{Z}}$ *and* $\boldsymbol{w} \in [\mathbf{0}, -\boldsymbol{\ell}]_{\mathbb{Z}}$ *with* $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}\boldsymbol{w}$ *and* $\|\boldsymbol{v}\|_0 + \|\boldsymbol{w}\|_0 \geq 1$*, the following implication holds true:*

$$\boldsymbol{v}_{S_+^c}, \boldsymbol{w}_{S_-^c} \geq \mathbf{0} \quad \Rightarrow \quad \mathbb{1}^\top (\boldsymbol{v} + \boldsymbol{w}) > 0.$$

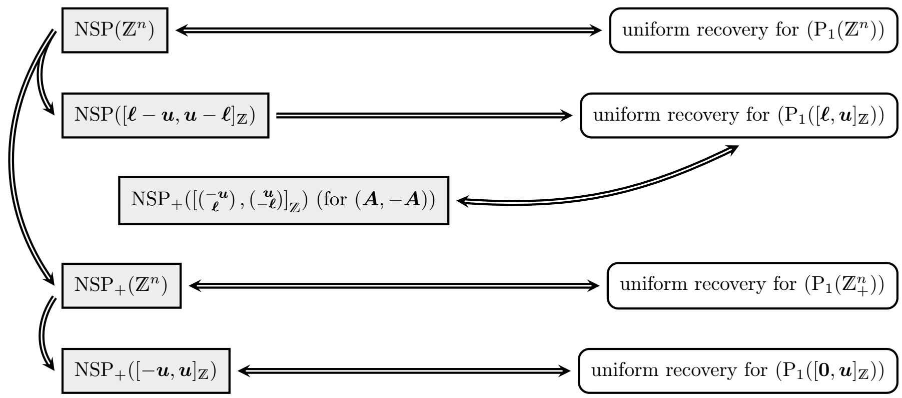

Figure 2: NSP-based recovery conditions for $(\mathrm{P}_1(X))$ for different $X \subseteq \mathbb{Z}^n$ and their relationship to each other. Arrows correspond to implications, whereas directions that are not depicted do not hold in general. The shorthand "uniform recovery" refers to guaranteed recovery of all vectors with a specific support S (by NSPs w.r.t. S) and also to that of all s-sparse vectors (by NSPs of order s). Results pertaining to $(\mathrm{P}_1([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}))$ are not shown since these are simple special cases of those for $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$.

Note also that a similar condition could be derived for $(\mathrm{P}_1(\mathbb{Z}^n))$ *by applying Theorem 28 part 2) to the corresponding split formulation (which is of the form* $(\mathrm{P}_1(\mathbb{Z}^n_+))$*), but of course Theorem 28 part 1) already provides a full (and simpler) characterization for sparse recovery by* $(\mathrm{P}_1(\mathbb{Z}^n))$*.*

Corollary 39. *Let* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *and* $S \subseteq [n]$*. Every vector* $\hat{\boldsymbol{x}} \in [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$ *with* $\mathrm{supp}(\hat{\boldsymbol{x}}) \subseteq S$ *is the unique optimal solution of* $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} := \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if* $(\boldsymbol{A}, -\boldsymbol{A})$ *satisfies* $\mathrm{NSP}_+([-\binom{\boldsymbol{u}}{\boldsymbol{u}}, \binom{\boldsymbol{u}}{\boldsymbol{u}}]_{\mathbb{Z}})$ *w.r.t.* S*. Moreover,* $\boldsymbol{A}$ *is* $(s, [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}, 1)$*-good if and only if* $(\boldsymbol{A}, -\boldsymbol{A})$ *satisfies* $\mathrm{NSP}_+([-\binom{\boldsymbol{u}}{\boldsymbol{u}}, \binom{\boldsymbol{u}}{\boldsymbol{u}}]_{\mathbb{Z}})$ *of order* s*.*

Proof. Set $\boldsymbol{\ell} = -\boldsymbol{u}$ and apply Theorem 37. □

Naturally, the NSPs for larger integral sets imply those for smaller sets. It is not hard to find examples that show that the converse directions are false in general; for brevity, we do not list such examples here, but provide an overview diagram to summarize our results and display the implications, see Figure 2.

4.3. Recovery of Individual Vectors

The focus so far was on conditions that guarantee the recovery of all vectors with a certain support or given sparsity level. In this section, we consider similar conditions for the recovery of individual integral signals.

In the continuous setting, when $(\mathrm{P}_1(X))$ with $X \subseteq \mathbb{R}^n$ can be rewritten as a linear program, there are well-known characterizations for recoverability of a specific vector $\hat{\boldsymbol{x}}$ as the unique ℓ_1-minimizer, see, e.g., [3, Theorem 4.26]. Attempting to directly transfer these results to $(\mathrm{P}_1(\mathbb{Z}^n))$ gives the following sufficient condition.

Theorem 40. *A vector* $\hat{\boldsymbol{x}} \in \mathbb{Z}^n$ *with* $\operatorname{supp}(\hat{\boldsymbol{x}}) \subseteq S \subseteq [n]$ *is the unique optimal solution of* $(\mathrm{P}_1(\mathbb{Z}^n))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if for all* $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}) \cap \mathbb{Z}^n) \setminus \{\boldsymbol{0}\}$*, it holds that*

$$\Big|\sum_{i\in S} \operatorname{sign}(\hat{x}_i)\, v_i\Big| < \|\boldsymbol{v}_{S^c}\|_1. \tag{8}$$

The proof is completely analogous to (the sufficiency part in) that of the above-cited theorem from [3] and therefore omitted for the sake of brevity.

Clearly, the condition from Theorem 40 is implied by $\mathrm{NSP}(\mathbb{Z}^n)$, since $\|\boldsymbol{v}_S\|_1 \geq |\sum_{i\in S} \operatorname{sign}(\hat{x}_i)\, v_i|$. Furthermore, note that the result also shows that it still makes no difference whether we require (8) to hold for all integral or all rational vectors in the kernel of $\boldsymbol{A}$ (the inequality is obviously scalable by $\alpha \in \mathbb{N}$), i.e., for $\boldsymbol{A} \in \mathbb{Q}^{m\times n}$ the condition is equivalent to its continuous analogon. However, the condition loses necessity in the integral setting and indeed, it is not hard to construct a simple counterexample.

For nonnegative vectors, a simple characterization of unique recoverability is given next.

Theorem 41. *A vector* $\hat{\boldsymbol{x}} \in \mathbb{Z}^n_+$ *is the unique optimal solution of* $(\mathrm{P}_1(\mathbb{Z}^n_+))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if for all* $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}) \cap \mathbb{Z}^n) \setminus \{\boldsymbol{0}\}$*, the following implication holds:*

$$\boldsymbol{v} + \hat{\boldsymbol{x}} \geq \boldsymbol{0} \quad \Rightarrow \quad \mathbb{1}^\top \boldsymbol{v} > 0.$$

Proof. Since any other feasible solution can be written as the sum of $\hat{\boldsymbol{x}}$ and an integral nullspace vector $\boldsymbol{v}$, $\hat{\boldsymbol{x}}$ is the unique point with smallest ℓ_1-norm if and only if the objective contribution of every such nullspace vector is strictly positive. □

Note that $\mathrm{NSP}_+(\mathbb{Z}^n)$ implies the condition from Theorem 41, since for $\hat{\boldsymbol{x}}$ supported on S, $\boldsymbol{v} + \hat{\boldsymbol{x}} \geq \boldsymbol{0}$ yields $\boldsymbol{v}_{S^c} \geq \boldsymbol{0}$ and $\mathrm{NSP}_+(\mathbb{Z}^n)$ implies that $\mathbb{1}^\top \boldsymbol{v} > 0$.

Finally, the two previous results can be extended directly to the remaining cases $(\mathrm{P}_1([\boldsymbol{0}, \boldsymbol{u}]_{\mathbb{Z}}))$, $(\mathrm{P}_1([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}))$ and $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$, respectively. We omit the completely analogous proofs.

Theorem 42. *A vector* $\hat{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ *with* $\operatorname{supp}(\hat{\boldsymbol{x}}) \subseteq S \subseteq [n]$ *is the unique optimal solution of* $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if for all* $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}) \cap [\boldsymbol{\ell} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{\ell}]_{\mathbb{Z}}) \setminus \{\boldsymbol{0}\}$*, it holds that*

$$\Big|\sum_{i\in S} \operatorname{sign}(\hat{x}_i)\, v_i\Big| < \|\boldsymbol{v}_{S^c}\|_1.$$

Corollary 43. *A vector* $\hat{\boldsymbol{x}} \in [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$ *with* $\operatorname{supp}(\hat{\boldsymbol{x}}) \subseteq S \subseteq [n]$ *is the unique optimal solution of* $(\mathrm{P}_1([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if for all* $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}) \cap [-2 \cdot \boldsymbol{u}, 2 \cdot \boldsymbol{u}]_{\mathbb{Z}}) \setminus \{\boldsymbol{0}\}$*, it holds that*

$$\Big|\sum_{i\in S} \operatorname{sign}(\hat{x}_i)\, v_i\Big| < \|\boldsymbol{v}_{S^c}\|_1.$$

Theorem 44. *A vector* $\hat{\boldsymbol{x}} \in [\mathbf{0}, \boldsymbol{u}]_{\mathbb{Z}}$ *is the unique optimal solution of* $(\mathrm{P}_1([\mathbf{0}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if for all* $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}) \cap [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}) \setminus \{\mathbf{0}\}$*, the following implication holds:*

$$\boldsymbol{v} + \hat{\boldsymbol{x}} \in [\mathbf{0}, \boldsymbol{u}]_{\mathbb{Z}} \quad \Rightarrow \quad \mathbb{1}^\top \boldsymbol{v} > 0.$$

The conditions for $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ and $(\mathrm{P}_1([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}))$ are again only sufficient, whereas that for $(\mathrm{P}_1([\mathbf{0}, \boldsymbol{u}]_{\mathbb{Z}}))$ gives a characterization of solution uniqueness. It is worth mentioning that the conditions from Theorem 42 (and Corollary 43) and Theorem 44 are strictly weaker than those from Theorems 40 and 41, respectively, as can easily be validated by finding toy examples confirming that "bounds on the variables matter".

Finally, by employing the usual variable split, unique recoverability w.r.t. $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ and $(\mathrm{P}_1([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}))$ can also be characterized: (Since we have seen all the arguments before, we skip the proofs for brevity.)

Theorem 45. *A vector* $\hat{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ *is the unique optimal solution of* $(\mathrm{P}_1([\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if for all* $\boldsymbol{v} \in \big(\mathcal{N}(\boldsymbol{A}, -\boldsymbol{A}) \cap [\binom{-\boldsymbol{u}}{\boldsymbol{\ell}}, \binom{\boldsymbol{u}}{-\boldsymbol{\ell}}]_{\mathbb{Z}}\big) \setminus \{\mathbf{0}\}$*, the following implication holds:*

$$\boldsymbol{v} + \begin{pmatrix} \hat{\boldsymbol{x}}^+ \\ \hat{\boldsymbol{x}}^- \end{pmatrix} \in [\mathbf{0}, \binom{\boldsymbol{u}}{-\boldsymbol{\ell}}]_{\mathbb{Z}} \quad \Rightarrow \quad \mathbb{1}^\top \boldsymbol{v} > 0.$$

Corollary 46. *A vector* $\hat{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ *is the unique optimal solution of problem* $(\mathrm{P}_1([-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}))$ *with* $\boldsymbol{b} \coloneqq \boldsymbol{A}\hat{\boldsymbol{x}}$ *if and only if the following implication holds for all* $\boldsymbol{v} \in (\mathcal{N}(\boldsymbol{A}, -\boldsymbol{A}) \cap [-\binom{\boldsymbol{u}}{\boldsymbol{u}}, \binom{\boldsymbol{u}}{\boldsymbol{u}}]_{\mathbb{Z}}) \setminus \{\mathbf{0}\}$*:*

$$\boldsymbol{v} + \begin{pmatrix} \hat{\boldsymbol{x}}^+ \\ \hat{\boldsymbol{x}}^- \end{pmatrix} \in [\mathbf{0}, \binom{\boldsymbol{u}}{\boldsymbol{u}}]_{\mathbb{Z}} \quad \Rightarrow \quad \mathbb{1}^\top \boldsymbol{v} > 0.$$

Note that, since $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}^+ - \hat{\boldsymbol{x}}^-$, the condition in Theorem 45 can be expressed equivalently as: For all $\boldsymbol{v} \in \mathcal{N}(\boldsymbol{A}) \cap [-\boldsymbol{u}, \boldsymbol{u}]_{\mathbb{Z}}$, $\boldsymbol{w} \in \mathcal{N}(\boldsymbol{A}) \cap [\boldsymbol{\ell}, -\boldsymbol{\ell}]_{\mathbb{Z}}$ with $\|\boldsymbol{v}\|_0 + \|\boldsymbol{w}\|_0 \geq 1$, it holds that $\boldsymbol{v} - \boldsymbol{w} + \hat{\boldsymbol{x}} \in [\boldsymbol{\ell}, \boldsymbol{u}]_{\mathbb{Z}}$ implies $\mathbb{1}^\top(\boldsymbol{v} + \boldsymbol{w}) > 0$. An analogous reformulation is, of course, also possible for the condition in Corollary 46.

5. Numerical Experiments

In this section, we present some computational experiments with the recovery of integer signals as a proof-of-concept. The (mixed-)integer problems were solved using Gurobi 7.5.2 on a linux cluster with Intel Xeon E5-1620 quad core CPUs with 3.5 GHz, 10 MB cache size, and 32 GB main memory. (For a primer on LP-based branch-and-bound, see, e.g., [27].)

5.1. *Solving the* ℓ_0*-Problem for Binary Signals*

We begin with the case of binary signals, i.e., $X = \{0, 1\}^n$. In this case, $(\mathrm{P}_0(X))$ equals $(\mathrm{P}_1(X))$ and can be written as

$$\min\, \{\mathbb{1}^\top \boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \{0, 1\}^n\}. \tag{9}$$

Table 1: Solving ($\mathrm{P}_0(\{0,1\}^n)$) via (10) for a random 64×256 matrix with entries in $\{0,\dots,99\}$ and vector $\boldsymbol{b} = \boldsymbol{A}\tilde{\boldsymbol{x}}$, where the random vector $\tilde{\boldsymbol{x}} \in \{0,1\}^n$ has s nonzeros. (BKZ times were obtained on an Intel i7-3770 CPU with 3.4 GHz.)

s	opt.	nodes	solver time [s]	BKZ time [s]	s	opt.	nodes	solver time [s]	BKZ time [s]
8	8	1	0.5	17.94	136	136	1335	16.2	18.16
16	16	1	0.5	19.05	144	144	1451	16.1	18.83
24	24	1	0.6	18.88	152	152	1219	12.1	18.28
32	32	1	1.1	18.13	160	160	1400	14.7	18.35
40	40	57	4.5	17.96	168	168	3309	23.8	18.13
48	48	47	4.0	18.09	176	176	1178	13.0	18.60
56	56	140	4.7	18.57	184	184	1614	15.1	18.48
64	64	1089	8.2	18.44	192	192	2229	15.9	18.02
72	72	1349	15.8	18.64	200	200	210	4.5	18.27
80	80	1799	21.2	18.40	208	208	132	3.9	18.07
88	88	2951	24.7	18.62	216	216	39	3.1	18.84
96	96	1408	16.9	18.47	224	224	50	3.3	18.50
104	104	1983	20.6	17.97	232	232	1	0.6	18.56
112	112	2050	24.1	18.55	240	240	1	0.6	18.60
120	120	16692	132.3	18.33	248	248	1	0.5	18.69
128	128	2034	21.7	18.51	256	256	1	0.5	18.33

In the first experiment, we deal with the solution of (9). We generated a 64×256 matrix $\boldsymbol{A}$ with random entries from $\{0,\dots,99\}$. (Note that here, we avoid the situation that, for some families of random matrices, all binary signals can be reconstructed by solving ($\mathrm{P}_1([0,1]^n)$), i.e., the LP relaxation of (9), if the number of measurements $m \geq n/2$, cf. [21].) Then for $s = 8, 16, \dots, 256$ we generated a $\{0,1\}$-vector $\boldsymbol{x}^s$ with 1s at s random positions. The right handside is then $\boldsymbol{b}^s \coloneqq \boldsymbol{A}\boldsymbol{x}^s$.

It turns out that directly solving problem (9) is quite hard – only 12 instances can be solved to optimality within four hours; for the solved instances the optimal solution was found right away and optimality is proved fast. The remaining instances seem to be hopeless to solve. Indeed, it is known that the so-called market-split instances, which have a quite similar structure, are very challenging, see Cornuéjols and Dawande [37]. In fact, for such instances it has been known to be hard to find a feasible solution or decide that none exists in practice using standard solution techniques. Then, Aardal et al. [38] observed that using basis reduction techniques, the market-split instances can be transformed such that they can be solved easily for medium-sized instances. We tested this transformation, but it turned out to be inefficient, possibly because, to retain the objective function, one needs to keep the original variables. Even branching on the transformed variables first does not help here. However, we also tested the so-called rangespace formulation of Krishnamoorthy and Pataki [39]. Here, we computed a unimodular matrix $\boldsymbol{U} \in \mathbb{Z}^{n \times n}$ such that the columns of $\binom{\boldsymbol{A}}{\boldsymbol{I}}\boldsymbol{U}$ are almost orthogonal and of similar length. The resulting model is then

$$\min\,\{\mathbb{1}^\top \boldsymbol{U}\boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{U}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{0} \leq \boldsymbol{U}\boldsymbol{x} \leq \mathbb{1}\}. \tag{10}$$

The intuition is that the corresponding polytope is transformed to be more "round". This makes it easier for variable-branching based branch-and-bound solvers to find feasible solutions and then prove optimality.

The results for solving (10) is shown in Table 1. The columns provide the sparsity level s, the optimal value of (10) and therefore of (9), the number of nodes in the branch-and-bound tree, and the running time in seconds. We use the library `fplll` [40] and its python interface `fpylll` [41] using the Block Korkin-Zolotarev (BKZ) basis reduction technique, see Schnorr [42]. The corresponding time in seconds is given in the last column.

The results show that the optimal solution always coincides with the sparsity level s used to construct the instances. Moreover, all instances can be solved quite fast after performing basis reduction. We have to note, however, that this approach will not scale well, since the basis reduction algorithms will take a significant time for larger instances and the solution of the instances as well. Nevertheless our results hopefully motivate research to improve the presented techniques.

The results in Table 1 also suggest that the vectors $\boldsymbol{x}^s$ used for the construction of the instances are in fact the only feasible integer points. This can be tested by adding the constraint

$$\sum_{i \in [n]:\, x_i^s = 1} (1 - x_i) + \sum_{i \in [n]:\, x_i^s = 0} x_i \geq 1.$$

to the model (9). If the problem turns out to be infeasible, $\boldsymbol{x}^s$ is the unqiue solution. In fact, infeasibility of all these instances together can be proved within one 0.1 seconds. Note, however, that $\boldsymbol{x}^s$ is usually not known.

Finally, note that the solution performance also depends on the size of the coefficients in the matrix $\boldsymbol{A}$. If we use a random 64×256 binary matrix, i.e., with entries from $\{0, 1\}$, the problems become harder to solve. In this case, (10) can only be solved for 12 instances within four hours. Nevertheless, the optimal values agree with s in all these instances. Again for all instances, uniqueness of $\boldsymbol{x}^s$ is easily proven.

5.2. Recoverability Test for Binary Signals

As a next step, we check whether the recoverability test of Remark 14 allows to guarantee unique solutions. For $X = \{0, 1\}^n$, this test can be modeled as

$$\max \{\mathbb{1}^\top \boldsymbol{v} + \mathbb{1}^\top \boldsymbol{w} \,:\, \boldsymbol{A}\boldsymbol{v} - \boldsymbol{A}\boldsymbol{w} = \boldsymbol{0},\ \mathbb{1}^\top \boldsymbol{v} \leq s,\ \mathbb{1}^\top \boldsymbol{w} \leq s, \\ v_i + w_i \leq 1\ \forall\, i \in [n], \mathbb{1}^\top \boldsymbol{v} \geq \mathbb{1}^\top \boldsymbol{w};\ \boldsymbol{v}, \boldsymbol{w} \in \{0, 1\}^n\}, \quad (11)$$

where the last inequality removes symmetry w.r.t. sign flips (scaling by -1). This formulation yields that the matrix $\boldsymbol{A}$ is $(s, X, 0)$-good if and only if the optimal objective is 0. Thus, one can use a cutoff value and stop the computation as soon as we found a solution with positive objective.

Table 2: Results of the recoverability formulation (11) for a 32×64 random binary matrix and binary signals.

s	10	11	12	13	14	15	16	17	18
best obj.	0	0	0	0	0	0	0	33	33
time [s]	284.1	834.5	1014.6	8051.6	6270.9	11632.6	8610.2	14k	14k

Alternatively, the following model can be used:

$$\min \{\mathbb{1}^\top \boldsymbol{v} + \mathbb{1}^\top \boldsymbol{w} \,:\boldsymbol{A}\boldsymbol{v} - \boldsymbol{A}\boldsymbol{w} = \boldsymbol{0},\ \mathbb{1}^\top \boldsymbol{v} \leq s,\ \mathbb{1}^\top \boldsymbol{w} \leq s,\ \mathbb{1}^\top(\boldsymbol{v}+\boldsymbol{w}) \geq 1,\\ v_i + w_i \leq 1\ \forall\, i \in [n],\ \mathbb{1}^\top \boldsymbol{v} \geq \mathbb{1}^\top \boldsymbol{w};\ \boldsymbol{v}, \boldsymbol{w} \in \{0,1\}^n\}. \quad (12)$$

Here, $\boldsymbol{A}$ is $(s, X, 0)$-good if and only if this problem is infeasible. In practice, (12) performed worse than (11).

To illustrate the behavior of (11), we first consider a matrix $\boldsymbol{A}$ of size 32×64 with random entries from $\{0, \dots, 99\}$. We then transformed the problem using basis reduction, as described above. In this case, all instances are solved within a few seconds and recoverability is proven. We also performed a similar test using a random 32×96 matrix with random entries from $\{0, \dots, 99\}$. However, only instances up to $s = 6$ could be solved using the transformed problem with basis reduction within four hours; recoverability could be proven for each of these cases. No instance could be solved for the original formulation.

Not surprisingly, the recoverability behavior depends on the sizes of the coefficients in the matrix $\boldsymbol{A}$. If we consider a 32×64 binary matrix, recoverability can be proven for $s \leq 16$. Starting from $s = 17$, the optimal value is positive, i.e., no universal recovery holds, see Table 2. For 32×96 matrices the picture is similar. One can prove recoverability up to $s = 14$. In both cases, applying basis reduction does not help to speed up the solution process.

5.3. Continuous Signals in the Unit Interval

In a next step, we consider the relaxed version ($\mathrm{P}_0([\boldsymbol{0}, \mathbb{1}]_{\mathbb{R}})$) and compare it to the integral problem ($\mathrm{P}_0([\boldsymbol{0}, \mathbb{1}]_{\mathbb{Z}})$). The goal is to quantify the effect of requiring the signals to be integer on recovery guarantees.

In the first experiment for continuous signals, we consider the recovery test of Remark 16 specialized to $X = [\boldsymbol{0}, \mathbb{1}]_{\mathbb{R}}$, which can be modeled as

$$\max \{\mathbb{1}^\top(\boldsymbol{v}+\boldsymbol{w}) \,:\boldsymbol{A}\boldsymbol{v} - \boldsymbol{A}\boldsymbol{w} = \boldsymbol{0},\ \boldsymbol{0} \leq \boldsymbol{v} \leq \boldsymbol{y},\ \boldsymbol{0} \leq \boldsymbol{w} \leq \boldsymbol{z},\ y_i + z_i \leq 1\ \forall\, i \in [n],\\ \mathbb{1}^\top \boldsymbol{y} \leq s,\ \mathbb{1}^\top \boldsymbol{z} \leq s,\ \mathbb{1}^\top \boldsymbol{v} \geq \mathbb{1}^\top \boldsymbol{w},\ \boldsymbol{y}, \boldsymbol{z} \in \{0,1\}^n\}. \quad (13)$$

As for (11), the matrix $\boldsymbol{A}$ is $(s, X, 0)$-good if and only if the optimal objective is 0. In fact, for the same 32×64 random matrix with entries from $\{0, \dots, 99\}$ from above, formulation (13) yields solutions with positive value for every $s \in [64]$, i.e., ($\mathrm{P}_0([0, \mathbb{1}]_{\mathbb{R}})$) is never universally unique. For the 32×64 *binary* matrix from above, it turns out that for $s = 10, \dots, 16$ the instances cannot be solved within a time limit of one hour, and the results for these sparsity levels remain

Table 3: Solving ($P_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}})$), ($P_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}})$), ($P_1([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}})$), and ($P_1([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}})$) for a random 32×64 binary matrix and vectors $\boldsymbol{b}$ generated by vectors of support size s.

	($P_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}})$)		($P_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}})$)		($P_1([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}})$)			($P_1([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}})$)	
s	$\lVert\cdot\rVert_0$	time [s]	$\lVert\cdot\rVert_0$	time [s]	$\lVert\cdot\rVert_1$	$\lVert\cdot\rVert_0$	time [s]	$\lVert\cdot\rVert_1$	$\lVert\cdot\rVert_0$
12	12	0.0	12	0.0	18	12	0.0	18.00	12
16	16	0.1	16	0.1	23	16	0.0	22.61	34
20	20	0.1	20	0.1	29	20	0.0	28.10	34
24	24	0.1	24	0.1	37	24	0.0	35.94	36
28	28	0.1	28	0.1	44	28	0.1	41.73	38
32	32	0.0	32	0.1	48	32	0.0	47.18	41
36	36	22.6	35	181.3	57	36	3.8	51.02	40
40	40	18.7	38	54.6	60	40	0.6	58.44	46
44	44	8.6	41	65.0	66	44	0.3	63.44	47
48	48	51.4	41	101.5	70	48	13.3	65.09	48
52	52	68.0	46	51.8	80	52	6.8	75.92	56
56	56	1459.1	44	277.5	78	56	170.9	72.09	52
60	60	5.6	51	3.7	91	60	1.3	87.40	59
64	64	3.5	53	2.8	95	64	0.4	91.06	59

inconclusive. For $s = 17$ and $s = 18$, a solution with positive objective could be found, as in the integral case (cf. Table 2).

Next, we directly consider ($P_0([\mathbf{0}, \mathbb{1}]_{\mathbb{R}})$), which can be written as

$$\min \{\mathbb{1}^\top \boldsymbol{y} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ 0 \leq \boldsymbol{x} \leq \boldsymbol{y},\ \boldsymbol{y} \in \{0,1\}^n\}. \tag{14}$$

Note that solving (14) is NP-hard by the same arguments as used to prove Proposition 4.

Using the same 64×256 matrix and instances as shown in Table 1, it is possible to solve the instances with $s = 8$, 16, 24, 32, 40, 48 (and $s = 216$, 224, 232, 240, 248, 256) within four hours. In all these cases, the (integral) solution $\tilde{\boldsymbol{x}}$ used to generate $\boldsymbol{b}$ is recovered. This shows that adding bounds on the variables can result in quite strong (individual) recovery guarantees, even if no integrality requirements are imposed. For the binary 64×256 matrix from above, one can solve $s = 8$, 16, 24, 32, 224, 232, 240, 248, 256 with the same conclusions.

In conclusion, the computations in this section do not show a difference in the recovery properties between $X = [\mathbf{0}, \mathbb{1}]_{\mathbb{Z}}$ and $X = [\mathbf{0}, \mathbb{1}]_{\mathbb{R}}$. Nevertheless, the solution performance might be different. Moreover, this property does not hold in general, as shown by Example 22 earlier—there, the given solutions are also ℓ_0-minimizers of the respective problems and the constraint $\boldsymbol{x} \leq \mathbb{1}$ can be added since it is already implied by the data and nonnegativity.

5.4. Signals With Values 0, 1 *and* 2

In a next experiment, we consider ($P_0(X)$) and ($P_1(X)$) for the cases $X = [\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}}$ and $X = [\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}}$. For the latter, we work with a formulation similar to (14). It turns out that these problems are much harder to solve than in the binary case; therefore, we again use a binary matrix of size 32×64.

The results are given in Table 3 and demonstrate that the integer program $(\mathrm{P}_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}}))$ again recovers all sparsity values of $\tilde{\boldsymbol{x}}$. However, the continuous problem $(\mathrm{P}_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}}))$ now produces solutions of smaller support, beginning with $s = 36$. This shows the larger potential of the integral version compared to the continuous version regarding recovery of a given solution with bounds.

The results for using an ℓ_1-objective are also shown in Table 3. Interestingly, solving $(\mathrm{P}_1([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}}))$ instead of $(\mathrm{P}_0([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{Z}}))$ also recovers all solutions $\tilde{\boldsymbol{x}}$ used to generate $\boldsymbol{b}$ and is notably faster. Moreover, solving the continuous counterpart $(\mathrm{P}_1([\mathbf{0}, 2\cdot\mathbb{1}]_{\mathbb{R}}))$ amounts to the solution of one LP and is very fast in comparison to the other approaches; we therefore do not list running times for this variant in Table 3. However, it never recovers the generating solutions $\tilde{\boldsymbol{x}}$ and produces significantly denser solutions, demonstrating again the stronger reconstructability properties using integer variables.

6. Concluding Remarks

Various aspects of integral sparse recovery are yet unexplored. For instance, while the results obtained in the present paper pertain to quite fundamental problems, it is also very important to explore the stability and robustness of the recovery problems if the measurements are corrupted by noise. In the continuous case, many explicit bounds on recovery errors are known (i.e., estimates on how far away from the sought true signal the solution of a recovery problem may be), but it seems no such investigations have so far been carried out assuming signal integrality. Similarly, it will be of interest to see how integrality constraints influence (probabilistic) bounds on the minimum number of measurements needed to ensure unique recoverability under certain matrix conditions. Also, one could consider integrality in the context of the so-called cosparse (analysis) model. Finally, the practical solution of all associated optimization problems involving integrality remains challenging, similar to the exact solution of $(\mathrm{P}_0(\mathbb{R}^n))$, cf. [43]. Thus, the development of further heuristics or approximation schemes as well as exact solution algorithms for sparse recovery problems with integrality constraints remains an important task; the same can be said about the actual practical evaluation of sparse recovery conditions such as the various NSPs.

Acknowledgments

This work was partially supported by the EXPRESS project within the DFG priority program CoSIP (DFG-SPP 1798).

References

[1] M. R. Garey, D. S. Johnson, Computers and intractability. A guide to the theory of NP-completeness, W. H. Freeman and Company, 1979.

[2] S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit, SIAM J. Sci. Comput. 20 (1) (1998) 33–61.

[3] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, Basel, 2013.

[4] D. Donoho, J. Tanner, Sparse nonnegative solution of underdetermined linear equations by linear programming, Proc. of the National Academy of Sciences of the United States of America – PNAS 102 (27) (2005) 9446–9451. doi:10.1073/pnas.0502269102.

[5] A. Bruckstein, M. Elad, M. Zibulevsky, On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations, IEEE Trans. Inf. Theory 54 (11) (2008) 4813–4820. doi:10.1109/TIT.2008.929920.

[6] M. A. Khajehnejad, A. G. Dimakis, W. Xu, B. Hassibi, Sparse recovery of nonnegative signals with minimal expansion, IEEE Trans. Signal Process. 59 (1) (2011) 196–208. doi:10.1109/TSP.2010.2082536.

[7] B. Knoop, F. Monsees, C. Bockelmann, D. Peters-Drolshagen, S. Paul, A. Dekorsy, Compressed sensing k-best detection for sparse multi-user communications, in: Proc. EUSIPCO, 2014, pp. 1726–1730.

[8] E. Axell, G. Leus, E. G. Larsson, H. V. Poor, Spectrum sensing for cognitive radio: State-of-the-art and recent advances, IEEE Signal Processing Magazine 29 (3) (2012) 101–116.

[9] H. Zhu, G. B. Giannakis, Exploiting sparse user activity in multiuser detection, IEEE Transactions on Communications 59 (2) (2011) 454–465.

[10] B. Knoop, F. Monsees, C. Bockelmann, D. Wübben, S. Paul, A. Dekorsy, Sparsity-aware successive interference cancellation with practical constraints, in: Proc. WSA, 2013.

[11] M. Rossi, A. M. Haimovich, Y. C. Eldar, Spatial compressive sensing for mimo radar, IEEE Transactions on Signal Processing 62 (2) (2014) 419–430.

[12] K. J. Batenburg, J. Sijbers, Dart: A practical reconstruction algorithm for discrete tomography, IEEE Transactions on Image Processing 20 (9) (2011) 2542–2553.

[13] J. Kuske, P. Swoboda, S. Petra, A novel convex relaxation for non-binary discrete tomography, in: F. Lauze, Y. Dong, A. B. Dahl (Eds.), Scale Space and Variational Methods in Computer Vision, Springer, Cham, 2017, pp. 235–246.

[14] G. Hegde, Y. Yang, C. Steffens, M. Pesavento, Parallel low-complexity M-PSK detector for large-scale MIMO systems, in: Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016.

[15] G. Hegde, M. Pesavento, M. E. Pfetsch, Joint active device identification and symbol detection using sparse constraints in massive MIMO systems, in: Proc. 25th European Signal Processing Conference (EUSIPCO), 2017. doi:10.23919/EUSIPCO.2017.8081298.

[16] S. Sparrer, R. F. H. Fischer, MMSE-based version of OMP for recovery of discrete-valued sparse signals, Electronics Letters 52 (1) (2016) 75–77. doi:10.1049/el.2015.0924.

[17] A. Flinth, G. Kutyniok, PROMP: A sparse recovery approach to lattice-valued signals, Applied and Computational Harmonic AnalysisTo appear. doi:10.1016/j.acha.2016.12.004.

[18] U. Nakarmi, N. Rahnavard, BCS: Compressive sensing for binary sparse signals, in: Proc. Military Communications Conference (MILCOM), IEEE, 2012.

[19] F. Wu, J. Fu, Z. Lin, B. Zeng, Analysis on rate-distortion performance of compressive sensing for binary sparse source, in: Data Compression Conference, 2009, pp. 113–122. doi:10.1109/DCC.2009.24.

[20] O. L. Mangasarian, B. Recht, Probability of unique integer solution to a system of linear equations, Eur. J. Oper. Res. 214 (1) (2011) 27–30.

[21] M. Stojnic, Recovery thresholds for ℓ_1 optimization in binary compressed sensing, in: Proc. ISIT, 2010, pp. 1593–1597.

[22] P. Swoboda, J. Kuske, B. Savchynskyy, A dual ascent framework for lagrangean decomposition of combinatorial problems, preprint abs/1612.05460, arXiv (2016). URL http://arxiv.org/abs/1612.05460

[23] S. Keiper, G. Kutyniok, D. G. Lee, G. E. Pfander, Compressed sensing for finite-valued signals, Linear Algebra and its Applications 532 (2017) 570–613. doi:http://dx.doi.org/10.1016/j.laa.2017.07.006.

[24] A. Flinth, S. Keiper, Recovery of binary sparse signals with biased measurement matrices, preprint math.OC/1801.03381, arXiv (2018).

[25] A. S. Besicovitch, On the linear independence of fractional powers of integers, Journal of the London Mathematical Society s1-15 (1) (1940) 3–6. doi:10.1112/jlms/s1-15.1.3.

[26] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer, 1997.

[27] A. Schrijver, Theory of Linear and Integer Programming, John Wiley & Sons, Chichester, UK, 1986.

[28] A. Juditsky, A. Nemirovski, On verifiable sufficient conditions for sparse signal recovery via ℓ_1 minimization, Math. Program., Ser. B 127 (2011) 57–88. doi:10.1007/s10107-010-0417-z.

[29] A. Cohen, W. Dahmen, R. DeVore, Compressed sensing and best k-term approximation, J. Amer. Math. Soc. 22 (1) (2009) 211–231.

[30] A. M. Tillmann, M. E. Pfetsch, The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing, IEEE Trans. Inf. Theory 60 (2) (2014) 1248–1259.

[31] B. Korte, J. Vygen, Combinatorial Optimization. Theory and Algorithms, 5th Edition, Vol. 21 of Algorithms and Combinatorics, Springer, Heidelberg, 2012.

[32] P. D. Seymour, Decomposition of regular matroids, J. Comb. Theory, Ser. B 28 (1980) 305–359.

[33] K. Truemper, A decomposition theory for matroids. V. Testing of matrix total unimodularity, J. Comb. Theory, Ser. B 49 (1990) 241–281.

[34] M. Walter, K. Truemper, Implementation of a unimodularity test, Math. Prog. Comput. 5 (1) (2013) 57–73.

[35] G. Ding, L. Feng, W. Zang, The complexity of recognizing linear systems with certain integrality properties, Math. Prog., Ser. A 114 (2) (2008) 321–334.

[36] G. Cornuéjols, Combinatorial Optimization: Packing and Covering, Vol. 74 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 2001.

[37] G. Cornuéjols, M. Dawande, A class of hard small 0-1 programs, INFORMS J. Computing 11 (2) (1999) 205–210.

[38] K. Aardal, R. E. Bixby, C. Hurkens, A. K. Lenstra, J. Smeltink, Market split and basis reduction: towards a solution of the Cornuéjols-Dawande instances, INFORMS J. Comput. 12 (3) (2000) 192–202.

[39] B. Krishnamoorthy, G. Pataki, Column basis reduction and decomposable knapsack problems, Discrete Optimization 6 (2009) 242–270.

[40] fplll, Lattice algorithms using floating-point arithmetic, https://github.com/fplll/fplll.

[41] fpylll, A python interface for fplll, https://github.com/fplll/fpylll.

[42] C. Schnorr, A hierarchy of polynomial time lattice basis reduction algorithms, Theor. Comput. Sci. 53 (1987) 201–224. doi:10.1016/0304-3975(87)90064-8.
[43] S. Jokar, M. E. Pfetsch, Exact and approximate sparse solutions of underdetermined linear equations, SIAM J. Sci. Comput 31 (1) (2008) 23–44.

Appendix A. Proof of Proposition 21

Recall that G is perfect if and only if its complement graph $\overline{G}$ is perfect (see, e.g., [36, Theorem 3.4]) and that cliques in G correspond exactly to stable sets in $\overline{G}$. Hence, the system $\boldsymbol{A}_G\,\boldsymbol{y} \leq \mathbb{1}$, $\boldsymbol{y} \geq \boldsymbol{0}$ is equivalent to

$$\sum_{i\in S} y_i \leq 1 \quad \forall\, \text{stable sets } S \text{ of } G, \quad \boldsymbol{y} \geq \boldsymbol{0}.$$

Suppose $\boldsymbol{w} \in \mathbb{Z}^V$ and consider the linear program

$$\max\,\{\boldsymbol{w}^\top \boldsymbol{y} \,:\, \boldsymbol{A}_G\,\boldsymbol{y} \leq \mathbb{1},\, \boldsymbol{y} \geq \boldsymbol{0}\},$$

whose dual program is given by

$$\min\,\Big\{\mathbb{1}^\top \boldsymbol{x} \,:\, \sum_{S\ni i} x_S \geq w_i \quad \forall\, i \in V,\ \boldsymbol{x} \geq \boldsymbol{0}\Big\}. \tag{A.1}$$

(Note that here, x_S is the component of $\boldsymbol{x}$ associated with S.)

We proceed to show that (A.1) has an integral optimal solution for every $\boldsymbol{w} \in \mathbb{Z}^V$. W.l.o.g., we may assume that $\boldsymbol{w} \geq \mathbb{1}$ (if $w_i \leq 0$, the corresponding constraint is automatically satisfied and can be omitted).

Let $G' = (V', E')$ be the graph obtained from G by adding $w_i - 1$ copies of each node $i \in V$ along with edges connecting each node copy to the respective original node and all its neighbors (including the other node copies). By the Replication Lemma (see, e.g., [36, Lemma 3.3]), G' is perfect. With every node $i \in V$, we thus associate the set W_i with $|W_i| = w_i$ of nodes in V'; indeed, $V' = \bigcup_{i\in V} W_i$. Now, the LP

$$\min\,\Big\{\mathbb{1}^\top \boldsymbol{x}' \,:\, \sum_{S'\ni j} x'_{S'} \geq 1 \quad \forall\, j \in V',\ \boldsymbol{x}' \geq \boldsymbol{0}\Big\}, \tag{A.2}$$

where the sum is over the stable sets of G', is equivalent to (A.1) in the sense that feasible solutions of one problem can be transferred directly to feasible solutions of the other: To see this, note that we may identify a stable set S in G with all stable sets S' in G' whose nodes lie in $\bigcup_{i\in S} W_i$, and vice versa. More precisely, define

$$\rho(S) \coloneqq \{S' \text{ stable set in } G' \,:\, |S'| = |S|,\, S' \cap W_i \neq \emptyset \quad \forall\, i \in S\}$$

as the set of stable sets in G' corresponding to a stable set S in G. Conversely, for any stable set S' in G' there exists a unique stable set S in G such that $S' \in \rho(S)$. Moreover, it holds that

$$|\rho(S)| = \prod_{i\in S} |W_i| = \prod_{i\in S} w_i.$$

If $\boldsymbol{x}$ is feasible for (A.1), then $\boldsymbol{x}'$ defined by $x'_{S'} \coloneqq \frac{1}{|\rho(S)|} x_S$, where $S' \in \rho(S)$, is feasible for (A.2) and has the same objective value. Indeed, it holds that for every $j \in W$,

$$\sum_{S' \ni j} x'_{S'} = \sum_{S \ni i \colon j \in W_i} \prod_{k \in S,\, k \neq i} w_k \, \frac{x_S}{|\rho(S)|} = \sum_{S \ni i \colon j \in W_i} \frac{x_S}{w_i} \geq \frac{w_i}{w_i} = 1$$

$$\text{and} \quad \sum_{S'} x'_{S'} = \sum_{S} \sum_{S' \in \rho(S)} x'_{S'} = \sum_{S} \sum_{S' \in \rho(S)} \frac{x_S}{|\rho(S)|} = \sum_{S} x_S. \tag{A.3}$$

Similarly, if $\boldsymbol{x}'$ is feasible for (A.2), then $\boldsymbol{x}$ given by $x_S \coloneqq \sum_{S' \in \rho(S)} x'_{S'}$ is feasible for (A.1) with the same objective value: Feasibility follows from

$$\sum_{S \ni i} x_S = \sum_{S \ni i} \sum_{S' \in \rho(S)} x'_{S'} = \sum_{j \in W_i} \sum_{S' \ni j} x'_{S'} \geq \sum_{j \in W_i} 1 = w_i,$$

while (A.3) shows equality of the objective values.

Now, consider a $\chi(G')$-coloring of G', i.e., a partition of the node set V' into disjoint stable sets $S'_1, \ldots, S'_{\chi(G')}$. Setting $x'_{S'_t} = 1$ for all $t \in [\chi(G')]$ (and $x'_{S'} = 0$ for all other S') yields a feasible solution $\boldsymbol{x}'$ of (A.2). Because G' is perfect, $\boldsymbol{x}'$ is actually optimal: Any incidence vector of a clique C' in G' is feasible for the dual program of (A.2) with objective value equal to the number of elements in C'. Since $\chi(G') = \omega(G')$ by definition of perfectness, the objective values for the coloring above and any maximum clique coincide. Thus, by strong duality, $\boldsymbol{x}'$ is optimal for (A.2) and consequently, so is the corresponding solution $\boldsymbol{x}$ for (A.1), which concludes the proof. □