Mathematical Programming
https://doi.org/10.1007/s10107-022-01837-0

FULL LENGTH PAPER

Series B

The double exponential runtime is tight for 2-stage stochastic ILPs

Klaus Jansen[1] · Kim-Manuel Klein[1] · Alexandra Lassota[2]

Received: 19 May 2021 / Accepted: 2 May 2022

Abstract

We consider fundamental algorithmic number theoretic problems and their relation to a class of block structured Integer Linear Programs (ILPs) called 2-stage stochastic. A 2-stage stochastic ILP is an integer program of the form $\min\{c^T x \mid \mathcal{A}x = b, \ell \leq x \leq u, x \in \mathbb{Z}^{r+ns}\}$ where the constraint matrix $\mathcal{A} \in \mathbb{Z}^{nt \times r+ns}$ consists of n matrices $A_i \in \mathbb{Z}^{t \times r}$ on the vertical line and n matrices $B_i \in \mathbb{Z}^{t \times s}$ on the diagonal line aside. We show a stronger hardness result for a number theoretic problem called QUADRATIC CONGRUENCES where the objective is to compute a number $z \leq \gamma$ satisfying $z^2 \equiv \alpha \bmod \beta$ for given $\alpha, \beta, \gamma \in \mathbb{Z}$. This problem was proven to be NP-hard already in 1978 by Manders and Adleman. However, this hardness only applies for instances where the prime factorization of β admits large multiplicities of each prime number. We circumvent this necessity proving that the problem remains NP-hard, even if each prime number only occurs constantly often. Using this new hardness result for the QUADRATIC CONGRUENCES problem, we prove a lower bound of $2^{2^{\delta(s+t)}}|I|^{O(1)}$ for some $\delta > 0$ for the running time of any algorithm solving 2-stage stochastic ILPs assuming the Exponential Time Hypothesis (ETH). Here, $|I|$ is the encoding length of the instance. This result even holds if r, $||b||_\infty$, $||c||_\infty$, $||\ell||_\infty$ and the largest absolute value Δ in the constraint matrix $\mathcal{A}$ are constant. This shows that the state-of-the-art algorithms are nearly tight. Further, it proves the suspicion that these ILPs are indeed

An extended abstract of this work appeared in the proceedings of the 22nd Conference on Integer Programming and Combinatorial Optimization (IPCO 2021).

✉ Alexandra Lassota
alexandra.lassota@epfl.ch

Klaus Jansen
kj@informatik.uni-kiel.de

Kim-Manuel Klein
kmk@informatik.uni-kiel.de

[1] Faculty of Engineering, Department of Computer Science, Kiel University, Kiel, Germany

[2] Institute of Mathematics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Published online: 28 May 2022

harder to solve than the closely related n-fold ILPs where the constraint matrix is the transpose of $\mathcal{A}$.

Keywords 2-Stage stochastic ILPs · Quadratic congruences · Lower bound · Exponential time hypothesis

Mathematics Subject Classification 90C10

1 Introduction

One of the most fundamental problems in algorithm theory and optimization is the INTEGER LINEAR PROGRAMMING problem. Many theoretical and practical problems can be modeled as integer linear programs (ILPs) and thus, they serve as a very general but powerful framework for tackling various questions. Formally, the INTEGER LINEAR PROGRAMMING problem is defined as

$$\min\{c^\top x \mid \mathcal{A}x = b, \ell \leq x \leq u, x \in \mathbb{Z}^{d_2}\}$$

for some matrix $\mathcal{A} \in \mathbb{Z}^{d_1 \times d_2}$, a right-hand side $b \in \mathbb{Z}^{d_1}$, an objective function $c \in \mathbb{Z}^{d_2}$ and some lower and upper bounds $\ell, u \in \mathbb{Z}^{d_2}$. The goal is to find a solution x such that the value of the objective function $c^\top x$ is minimized. In general, this problem is NP-hard. Thus, it is of great interest to find structures to these ILPs which make them solvable more efficiently. This work considers 2-stage stochastic integer linear programs where the constraint matrix admits a specific block structure. Namely, the constraint matrix $\mathcal{A}$ only contains non-zero entries in the first few columns and block-wise along the the diagonal aside. This yields a constraint matrix $\mathcal{A}$ of 2-stage stochastic form:

$$\mathcal{A} = \begin{pmatrix} A_1 & B_1 & 0 & \dots & 0 \\ A_2 & 0 & B_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ A_n & 0 & \dots & 0 & B_n \end{pmatrix}.$$

Thereby $A_1, \dots, A_n \in \mathbb{Z}^{t \times r}$ and $B_1, \dots, B_n \in \mathbb{Z}^{t \times s}$ are integer matrices themselves. The complete constraint matrix $\mathcal{A}$ has size $nt \times r + ns$. Let Δ denote the largest absolute entry in $\mathcal{A}$.

Such 2-stage stochastic ILPs are a common tool in stochastic programming and they are often used in practice to model uncertainty of decision making over time [1, 11, 22, 28]. In particular, each block of the second stage encodes a different scenario, i. e., its restrictions and behaviour. The first stage is used to encode the probability that the respective scenario occurs. Due to the applicability, a lot of research has been done in order to solve these (mixed) ILPs efficiently in practice. Since we focus on the theoretical aspects of 2-stage stochastic ILPs in this chapter, we only refer the reader to the surveys [13, 26, 32] and the references therein regarding the practical methods.

The current state-of-the-art algorithms to solve 2-stage stochastic ILPs admits a running time of $3^{(r+s)s^s(2r\Delta+1)^{rs}} n \log^3(n) \cdot |I|$ where $|I|$ is the binary encoding length of the input [12, 20] or respectively of $n \log^{O(rs)}(n) 2^{(2\Delta)^{O(r^2+rs)}}$ [7] by a recent result. The first result improves upon the result in [23] due to Klein where the dependence on n was quadratic. The dependencies on the block dimensions and $|I|$ were similar. The first result in that respect was by Hemmecke and Schulz [16] who provided an algorithm with a running time of $f(r, s, t, \Delta) \cdot \text{poly}(n)$ for some computable function f. However, due to the use of an existential result from commutative algebra, no explicit bound could be stated for f.

Let us turn our attention to the n-fold ILPs for a moment, which where first introduced in [9]. These ILPs admit a constraint matrix which is the transpose of the 2-stage stochastic constraint matrix. Despite being so closely related, n-fold ILPs can be solved in time near linear in the number of blocks and only single exponentially in the block-dimensions of A_i^T, B_i^T [6, 21].

Thus, it is an intrinsic questions whether we can solve 2-stage stochastic ILPs more efficient or – as the latest algorithms suggest – whether 2-stage stochastic ILPs are indeed harder to solve than the closely related n-fold ILPs. We answer this question by showing a double-exponential lower bound in the running time for any algorithm solving the 2- STAGE STOCHASTIC INTEGER LINEAR PROGRAMMING (2- STAGE ILP) problem. Here, the 2- STAGE ILP problem is the corresponding decision variant which asks whether the ILP admits a feasible solution. We summarize this problem formally as follows:

2- STAGE STOCHASTIC INTEGER LINEAR PROGRAMMING problem

Input: Constraint matrix $\mathcal{A}$ of 2-stage stochastic form, $b \in \mathbb{Z}^{r+ns}$, an objective function $c \in \mathbb{Z}^{nt}$, and lower and upper bounds $\ell, u \in \mathbb{Z}^{nt}$.

Objective: Decide whether the ILP $\min\{c^\top x \mid \mathcal{A}x = b, \ell \leq x \leq u, x \in \mathbb{Z}^{nt}\}$ is feasible.

To prove this hardness, we reduce from the QUADRATIC CONGRUENCES problem. This problem asks whether there exists a $z \leq \gamma$ such that $z^2 \equiv \alpha \bmod \beta$ for some $\gamma, \alpha, \beta \in \mathbb{N}$. Formally, we get:

QUADRATIC CONGRUENCES problem

Input: Numbers $\alpha, \beta, \gamma \in \mathbb{N}$, prime factorization $b_1^{\beta_1}, \ldots, b_{n_{\text{QC}}}^{\beta_{n_{\text{QC}}}}$ of β with $b_1, \ldots, b_{n_{\text{QC}}}$ denoting the different prime factors of β and β_i the occurrence of b_i.

Objective: Decide whether there exists a $z \leq \gamma$ such that $z^2 \equiv \alpha \bmod \beta$.

This problem was proven to be NP-hard by Manders and Adleman [29] already in 1978 by showing a reduction from 3-SAT. This hardness even persists if the prime

factorization of β is given [29]. By this result, Manders and Adleman prove that it is NP-complete to compute the solutions of diophantine equations of degree 2. However, their reduction yields large parameters: the occurrences of each prime factor in the prime factorization of β is too large to obtain the desired lower bound for the 2- STAGE ILP problem. In particular, the occurrence of each prime factor is at least linear in the number of variables and clauses of the underlying 3-SAT problem. As the reduction generates block dimensions of size logarithmic in the largest prime factor to the power of its occurrence, the dependence on its occurrence and thus, on n_3, is linear, whereas we aim for a logarithmic one to show the desired hardness.

We give a new reduction yielding a stronger statement: The QUADRATIC CONGRUENCES problem is NP-hard even if the prime factorization of β is given and each prime factor occurs at most once (except 2 which occurs four times). Beside being useful to prove the lower bounds for solving the 2-stage stochastic ILPs, we think this results is of independent interest. We obtain a neat structure which may be helpful in various related problems or may yield stronger statements of past results which use the QUADRATIC CONGRUENCES problem.

Along the way to reduce to the 2-stage stochastic ILPs, based on this new reduction, we show strong NP-hardness for another problem we call the NON- UNIQUE REMAINDER problem. In this algorithmic number theoretic problem, we are given $x_1, \ldots, x_{n_{\mathrm{NR}}}, y_1, \ldots, y_{n_{\mathrm{NR}}}, \zeta \in \mathbb{N}$ and pairwise coprime numbers $q_1, \ldots, q_{n_{\mathrm{NR}}}$. The question is to decide whether there exists a number $z \in \mathbb{Z}_{>0}$ with $z \leq \zeta$ satisfying $z \bmod q_i \in \{x_i, y_i\}$ for all $i \in [n_{\mathrm{NR}}]$. In other words, either the residue x_i or y_i should be met for each equation. We summarize this problem as follows:

NON- UNIQUE REMAINDER problem

Input: Numbers $x_1, \ldots, x_{n_{\mathrm{NR}}}, y_1, \ldots, y_{n_{\mathrm{NR}}}, \zeta \in \mathbb{N}$ and pairwise coprime numbers $q_1, \ldots, q_{n_{\mathrm{NR}}} \in \mathbb{N}$.

Objective: Decide whether there exists a number $z \in \mathbb{Z}_{>0}$ with $z \leq \zeta$ satisfying $z \bmod q_i \in \{x_i, y_i\}$ for all $i \in [n_{\mathrm{NR}}]$.

This problem is a natural generalization of the Chinese Remainder problem where $x_i = y_i$ for all i. In that case, however, the problem can be solved using the Extended Euclidean algorithm. To the best of our knowledge the NON- UNIQUE REMAINDER problem has not been considered in the literature so far.

In order to finally achieve the desired lower bounds on the running time for the 2-stage stochastic ILP problem, we make use of the Exponential Time Hypothesis (ETH) – a widely believed conjecture stating that the 3-SAT problem cannot be solved in subexponentially time with respect to the number of variables:

Conjecture 1 (ETH [17]) *The* 3-SAT *problem cannot be solved in time less than* $O(2^{\delta_3 n_3})$ *for some constant* $\delta_3 > 0$ *where* n_3 *is the number of variables in the instance.*

Note that we use the index 3 for all variables of the 3-SAT problem.

Using the ETH, plenty lower bounds for various problems are shown, for an overview on the techniques and results see e.g. [8]. So far, the best algorithm runs in time $O(2^{0.387n_3})$, i.e., it follows that $\delta_3 \leq 0.387$ [8].

In the following, we also need the Chinese Remainder Theorem (CRT) for some of the proofs, which states the following:

Proposition 1 (CRT [19]) *Let $n_1, \ldots, n_k$ be pairwise co-prime. Further, let $i_1, \ldots, i_k$ be some integers. Then there exists integers x satisfying $x \equiv i_j \bmod n_j$ for all j. Further, any two solutions x_1, x_2 are congruent modulo $\prod_{j=1}^{k} n_j$.*

Summary of Results

- We give a new reduction from the 3-SAT problem to the QUADRATIC CONGRUENCES problem which proves a stronger NP-hardness result: The QUADRATIC CONGRUENCES problem remains NP-hard, even if the prime factorization of β is given and each prime number greater than 2 occurs at most once and the prime number 2 occurs four times. This does not follow from the original proof. In contrast, the original proof generates each prime factor at least $O(n_3 + m_3)$ times, where m_3 is the number of clauses in the formula. Our reduction circumvents this necessity, yet neither introduces noteworthily more nor larger prime factors. The proof is based on the original one. We believe this result is of independent interest.
- Based on this new reduction, we show strong NP-hardness for the NON- UNIQUE REMAINDER problem. This problem is a natural generalization of the Chinese Remainder problem where $x_i = y_i$ for all i. To the best of our knowledge the NON- UNIQUE REMAINDER problem has not been considered in the literature so far.
- Finally, we show that the NON- UNIQUE REMAINDER problem can be modeled by a 2-stage stochastic ILP. Assuming the ETH, we can then conclude a doubly exponential lower bound of $2^{2^{\delta(s+t)}}|I|^{O(1)}$ on the running time for any algorithm solving 2-stage stochastic ILPs. The double exponential lower bound even holds if the number of first stage variables $r = 1$, and the largest entries in the constraint matrix, the right-hand side and the objective function are constant, i.e., $\Delta, ||b||_\infty, ||c||_\infty \in O(1)$. This proves the suspicion that 2-stage stochastic ILPs are significantly harder to solve than n-fold ILPs with respect to the dimensions of the block matrices and Δ. Furthermore, it implies that the current state-of-the-art algorithms for solving 2-stage stochastic ILPs is indeed (nearly) optimal.

Further Related Work In recent years, there was significant progress in the development of algorithms for n-fold ILPs and lower bounds on the other hand. Assume the parameters as of the transpose of the 2-stage stochastic constraint matrix, i.e., the blocks A_i^T in the first few rows have dimension $r \times t$ and the blocks B_i^T along the diagonal beneath admit a dimension of $s \times t$. The best known algorithms to solve these ILPs have a running time of $2^{O(rs^2)}(rs\Delta)^{O(r^2s+s^2)}(nt)^{1+o(1)}$ [6] or respectively a running time of $(rs\Delta)^{r^2s+s^2}L^2(nt)^{1+o(1)}$ [21] where L denotes the encoding length of the largest number in the input. The best known lower bound is $\Delta^{\delta_{n\text{-fold}}(r+s)^2}$ for some $\delta_{n\text{-fold}} > 0$ [12].

Despite their similarity, it seems that 2-stage stochastic ILPs are significantly harder to solve than n-fold ILPs. Yet, no superexponential lower bound for the running time of

any algorithm solving the 2- STAGE ILP problem was shown. There is a lower bound for a more general class of ILPs in [12] that contain 2-stage stochastic ILPs showing that the running time is double-exponential parameterized by the topological height of the treedepth decomposition of the primal or dual graph. However, the topological height of 2-stage stochastic ILPs is constant and thus, no strong lower bound can be derived for this case.

If we relax the necessity of an integral solution, the 2-stage stochastic LP problem becomes solvable in time $2^{2\Delta^{O(t^3)}} n \log^3(n) \log(||u-\ell||_\infty) \log(||c||_\infty)$ [3]. For the case of mixed integer linear programs, there exists an algorithm solving 2-stage stochastic MILPs in time $2^{\Delta^{\Delta^{t^{O(t^2)}}}} n \log^3(n) \log(||u-\ell||_\infty) \log(||c||_\infty)$ [3]. Note that t can be replaced by $r+s$ as this value corresponds to the size of a submatrix with full rang derived from any block, see [3]. Both results rely on the fractionality of a solution whose size is only dependent on the parameters. This allows us to scale the problem such that it becomes an ILP (as the solution has to be integral) and thus, state-of-the-art algorithms for 2-stage stochastic ILPs can be applied.

There are also studies for a more general case called 4-Block ILPs where the constraint matrix consists of non-zero entries in the first few columns, the first few rows and block-wise along the diagonal. This may be seen as the combination of n-fold and 2-stage stochastic ILPs. Only little is known about them: They are in XP [15]. Further, a lower and upper bound on the Graver Basis elements (inclusion-wise minimal kernel elements) of $O(n^r f(k, \Delta))$ was shown recently [4], where r is the number of rows in the submatrix appearing repeatedly in the first few rows and k denotes the sum of the remaining block dimensions. There are also various results for recursive block structures, for an overview see [12, 24].

Structure of this Chapter Sect. 2 presents the stronger hardness result for the QUADRATIC CONGRUENCES problem we derive by enhancing the original reduction by Manders and Adleman from the 3-SAT problem. Due to the technical depth and length, however, the formal proof it postponed to Sect. 5. In Sect. 3, we show that the QUADRATIC CONGRUENCES problem can be modeled as a 2-stage stochastic ILP. To do so, we utilize the NON- UNIQUE REMAINDER problem as an intermediate step during the reduction. Finally, in Sect. 4, we bring the reductions together to prove the desired lower bound. This involves a construction which lowers the absolute value of Δ at the cost of slightly larger block dimensions.

2 Advanced hardness for QUADRATIC CONGRUENCES

This section presents that every instance of the 3-SAT problem can be transformed into an equivalent instance of the QUADRATIC CONGRUENCES problem in polynomial time. Recall that the QUADRATIC CONGRUENCES problem asks whether there exists a number $z \leq \gamma$ such that $z^2 \equiv \alpha \mod \beta$ holds. This problem was proven to be NP-hard by Manders and Adleman [29] showing a reduction from 3-SAT. This hardness even persists when the prime factorization of β is given [29]. However, we aim for an even stronger statement: The QUADRATIC CONGRUENCES problem remains NP-hard even

if the prime factorization of β is given and each prime number greater than 2 occurs at most once and the prime number 2 occurs four times. This does not follow from the original hardness proof. In contrast, if n_3 is the number of variables and m_3 the number of clauses in the 3-SAT formula then β admits a prime factorization with $O(n_3 + m_3)$ different prime numbers each with a multiplicity of at least $O(n_3 + m_3)$. Even though our new reduction lowers the occurrence of each prime factor greatly, the number of prime factors as well as their size do not enlarge notably.

While the idea and thus, the structure follows the proof of the original one from [29], adapting it to our needs requires various new observations concerning the behaviour of the newly generated prime factors and the functions we introduce. The original proof heavily depends on the numbers being high powers of the prime factors whereas we employ careful combinations of (new) prime factors. This requires us to introduce other number theoretical results as for example Lemma 1 into the arguments to estimate the bounds and function appropriately.

In the following, we want to give an idea of the hardness proof. The reduction may seem non-intuitive at first as it only shows the final result of equivalent transformations between various problems until we reach the QUADRATIC CONGRUENCES one. We list all these problems in order of their appearance whose strong NP-hardness is shown implicitly along the way. Afterwards, we give short ideas of their respective equivalence, which are proven in separate claims in the full proof, see Sect. 5. Note that not all variables are declared at this point, but also not necessary to understand the proof sketch.

- (3-SAT) Is there a function $\eta\colon x_i \to \{0, 1\}$ assigning a truth value to each variable that satisfies all clauses σ_k of the 3-SAT formula Φ simultaneously?
- (P2) Are there values $y_k \in \{0, 1, 2, 3\}$ and a truth assignment η such that $0 = y_k - \sum_{x_i \in \sigma_k} \eta(x_i) - \sum_{\bar{x_i} \in \sigma_k} (1 - \eta(x_i)) + 1$ for all k?
- (P3) Are there values $\alpha_j \in \{-1, +1\}$ such that $\sum_{j=0}^{\nu} \theta_j \alpha_j \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ for some θ_j and τ specified in dependence on the formula later on, and some prime numbers p_i and p^*?
- (P5) Is there an $x \in \mathbb{Z}$ satisfying

$$0 \leq |x| \leq H \tag{P5.1}$$

$$x \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i \tag{P5.2}$$

$$(H + x)(H - x) \equiv 0 \bmod K\,? \tag{P5.3}$$

  for some H dependent on the θ_j and K being a product of primes?
- (P6) Is there an $x \in \mathbb{Z}$ satisfying

$$0 \leq |x| \leq H \tag{P6.1}$$

$$(\tau - x)(\tau + x) \equiv 0 \bmod 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \tag{P6.2}$$

$$(H + x)(H - x) \equiv 0 \bmod K? \tag{P6.3}$$

– (QUADRATIC CONGRUENCES) Is there a number $x \leq H$ such that $(2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i + K)x^2 \equiv K\tau^2 + 2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i H^2 \bmod 2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i \cdot K$?

The 3-SAT problem is transformed to Problem (P2) by using the straight-forward interpretation of truth values as numbers 0 and 1 and the satisfiability of a clause as the sum of its literals being larger zero. Introducing slack variables y_k yields the above form, see Claim 3.

Multiplying each equation of (P2) with exponentially growing factors and then forming their sum preserves the equivalence of these systems. Introducing some modulo consisting of unique prime factors larger than the outcome of the largest possible sum obviously does not influence the system. Replacing the variables $\eta(x_i)$ and y_k by variables α_j with domain $\{-1, +1\}$, re-arranging the term and defining parts of the formula as the variables θ_j and τ yields Problem (P3), see Claim 4.

We then introduce some Problem (P4) to integrate the condition $x \leq H$. The problem asks whether there exists some $x \in \mathbb{Z}$ such that

$$0 \leq |x| \leq H \tag{P4.1}$$

$$(H + x)(H - x) \equiv 0 \bmod K? \tag{P4.2}$$

By showing that each solution to the system (P4) is of form $\sum_{j=0}^{v} \theta_j \alpha_j$, we can combine (P3) and (P4) yielding (P5), see Claim 5.

Using some observations about the form of solutions for the second constraint of Problem (P5), we can re-formulate it as Problem (P6), see Claim 6.

Next, we use the fact that $p^* \prod_{i=1}^{m'} p_i$ and K are co-prime per definition and thus, we can combine (P6.2) and (P6.3) to one equivalent equation. To do so, we take each left-hand side of (P6.2) and (P6.3) and multiply the modulo of the respective other equation and form their overall sum. Using a little re-arranging, this finally yields the desired QUADRATIC CONGRUENCES problem, see Claim 7. Note that by re-arranging the factor to the other side of the equivalence, this form is exactly the same as in the problem statement. Overall, we get the following Theorem.

Theorem 1 *An instance of the* 3-SAT *problem with n_3 variables and m_3 clauses is reducible to an instance of the* QUADRATIC CONGRUENCES *problem in polynomial time with the properties that $\alpha, \beta, \gamma \in 2^{O((n_3+m_3)^2 \log(n_3+m_3))}$, $n_{QC} \in O((n_3+m_3)^2)$, $max_i\{b_i\} \in O((n_3+m_3)^2 \log(n_3+m_3))$, and each prime factor in β occurs at most once except the prime factor 2 which occurs four times.*

Due to the technical depth and length, the actual proof it postponed to Sect. 5.

3 Reduction from the quadratic congruences problem

This section presents the reduction from the QUADRATIC CONGRUENCES problem to the 2- STAGE ILP problem. First, we present a transformation of an instance of the QUADRATIC CONGRUENCES problem to an instance of the NON- UNIQUE REMAINDER problem. This problem was not considered in the literature so far and serves as an intermediate step in this chapter. However, it might be of independent interest as it generalizes the prominent Chinese Remainder theorem. Secondly, we show how an instance of the NON- UNIQUE REMAINDER problem can be modelled as a 2-stage stochastic ILP. Recall that in the NON- UNIQUE REMAINDER problem, we are given numbers $x_1, \ldots, x_{n_{\text{NR}}}, y_1, \ldots, y_{n_{\text{NR}}}, q_1, \ldots, q_{n_{\text{NR}}}, \zeta \in \mathbb{N}$ where the q_is are pairwise co-prime. The question is to decide whether there exists a natural number z satisfying $z \bmod q_i \in \{x_i, y_i\}$ simultaneously for all $i \in \{1, 2, \ldots, n_{\text{NR}}\}$ and which is smaller or equal to ζ.

In other words, we either should meet the residue x_i or y_i. Thus, we can re-write the equation as $z \equiv x_i \bmod q_i$ or $z \equiv y_i \bmod q_i$ for all i.

Indeed, this problem becomes easy if $x_i = y_i$ for all i, i. e., we know the remainder we want to satisfy for each equation [33]: First, compute s_i and r_i with $r_i \cdot q_i + s_i \cdot \prod_{j=1, j\neq i}^{n_{\text{NR}}} q_j = 1$ for all i using the Extended Euclidean algorithm. Now it holds that $s_i \cdot \prod_{j=1, j\neq i}^{n_{\text{NR}}} q_j \equiv 1 \bmod q_i$ as q_i and $\prod_{j=1, j\neq i}^{n_{\text{NR}}} q_j$ are coprime, and $s_i \cdot \prod_{j=1, j\neq i}^{n_{\text{NR}}} q_j \equiv 0 \bmod q_j$ for $j \neq i$. Thus, the smallest solution corresponds to $z = \sum_{i=1}^{n_{\text{NR}}} x_i \cdot s_i \cdot \prod_{j=1, j\neq i}^{n_{\text{NR}}} q_j$ due to the Chinese Remainder theorem [33]. Comparing z to the bound ζ finally yields the answer. Also note that if n_{NR} is constant, we can solve the problem by testing all possible vectors $(v_1, \ldots, v_{n_{\text{NR}}})$ with $v_i \in \{x_i, y_i\}$ and then use the Chinese Remainder theorem as explained above.

Theorem 2 *The* QUADRATIC CONGRUENCES *problem is reducible to the* NON- UNIQUE REMAINDER *problem in polynomial time with the properties that* $n_{NR} \in O(n_{QC})$, $\max_{i\in\{1,\ldots,n_{NR}\}}\{q_i, x_i, y_i\} = O(\max_{j\in\{1,\ldots,n_{QC}\}}\{b_j^{\beta_j}\})$, *and* $\zeta \in O(\gamma)$.

Proof *Transformation:* Set $q_1 = b_1^{\beta_1}, \ldots, q_{n_{\text{NR}}} = b_{n_{\text{QC}}}^{\beta_{\text{QC}}}$ and $\zeta = \gamma$ where β_i denotes the occurrence of the prime factor b_i in the prime factorization of β. Compute $\alpha_i \equiv \alpha \bmod q_i$. Set $x_i^2 = \alpha_i$ if there exists such an $x_i \in \mathbb{Z}_{q_i}$. Further, compute $y_i = -x_i + q_i$. If there is no such number x_i and thus, y_i, produce a trivial no-instance.
Instance size: The numbers we generate in the reduction equal the prime numbers of the QUADRATIC CONGRUENCES problem including their occurrence. Hence, it holds that $\max_{i\in\{1,\ldots,n_{\text{NR}}\}}\{q_i\} = O(\max_{j\in\{1,\ldots,n_{\text{QC}}\}}\{b_j^{\beta_j}\})$. Due to the modulo, this value also bounds x_i and y_i. The upper bound on a solution equals the ones from the instance of the QUADRATIC CONGRUENCES problem, i. e., $\zeta \in O(\gamma)$, and $n_{\text{NR}} = n_{\text{QC}}$ holds.
Correctness: First, let us verify that producing a trivial no-instance is correct if we cannot find some x_i. Indeed, this can be traced back to the Chinese Remainder theorem: If and only if there is an x with $x^2 \equiv \alpha \bmod \beta$ and $q_1, \ldots, q_{n_{\text{NR}}}$ (i. e., the equivalences to $b_i^{\beta_i}$) is the prime factorization of β, then $x^2 \equiv \alpha_i \bmod q_i$, $\alpha_i \in \mathbb{Z}_{q_i}$ for all i. In other words, it is has to be dividable by all $b_i^{\beta_i}$ yielding the same remainder α (modulo

$b_i^{\beta_i}$). Hence, if there does not exists a square root of α in one of the systems then $x^2 \equiv \alpha \bmod \beta$ has no solution.

But if there exists x_i and y_i, these values are in $\mathbb{Z}_{q_i}$ as $x_i \leq \alpha_i < q_i$ per definition of x_i and α_i. Further, both values solve the problem $x_i^2, y_i^2 \equiv \alpha \bmod q_i$ as $x_i^2 \equiv \alpha_i \bmod q_i \equiv \alpha_i + \lambda \cdot q_i \bmod q_i \equiv \alpha \bmod q_i$ for some $\lambda \in \mathbb{N}$. Moreover,

$$\begin{aligned} y_i^2 &\equiv (-x_i + q_i)^2 \bmod q_i = q_i^2 - 2x_i q_i + x_i^2 \bmod q_i \\ &\equiv x_i^2 \bmod q_i \equiv \alpha \bmod q_i . \end{aligned}$$

The third equation holds as each summand except the last one is a multiple of q_i. The last transformation is true due to the computation above.

Note that for all prime numbers greater than 2 it holds that $x_i \neq y_i$. This can easily be seen as we already argued that x_i and y_i are in $\mathbb{Z}_{p_i}$. Let us suppose both values are equal, i. e.,

$$\begin{aligned} & x_i^2 = y_i^2 \\ \Leftrightarrow\ & \alpha_i = (-x_i + q_i)^2 \\ \Leftrightarrow\ & \alpha_i = q_i^2 - 2q_i x_i + x_i^2 \\ \Leftrightarrow\ & \alpha_i = q_i^2 - 2q_i x_i + \alpha_i \\ \Leftrightarrow\ & 2q_i x_i = q_i^2 \\ \Leftrightarrow\ & 2x_i = q_i . \end{aligned}$$

The factor q_i is a product of some prime number greater than 2 by the assumption above. Thus, there is no x_i satisfying the formula.

Let us now prove the equivalence of the reduction.

$\Rightarrow$ Let the instance of the QUADRATIC CONGRUENCES problem be a yes-instance. Then there exists a z satisfying $z^2 \equiv \alpha \bmod \beta$ with $0 < z \leq \gamma$. This solution directly corresponds to a solution of the generated instance of the NON- UNIQUE REMAINDER problem. First, $z \leq \gamma = \zeta$. Secondly, z satisfies all equations as it holds that

$$z^2 \equiv \alpha \bmod \beta \equiv \alpha \bmod \prod_{i=1}^{n_{\mathrm{NR}}} b_i^{\beta_i} \equiv \alpha \bmod b_i^{\beta_i} \text{ for all } i .$$

The first equivalence holds as the $b_i^{\beta_i}$s are the prime factorization of β. The second equivalence is true as we can decompose the solution as follows: $z^2 = \lambda \cdot \prod_{i=1}^{n} b_i^{\beta_i} + \alpha$ for some $\lambda \in \mathbb{N}$. Thus, the first summand is not only divided without remainder by $\prod_{i=1}^{n_{\mathrm{NR}}} b_i^{\beta_i}$ but also by all primes along with their occurrences alone, leaving only the second summand α as the remainder. Further, since $x_i^2, y_i^2 \equiv \alpha \bmod q_i$ as shown before, it holds that

$$z^2 \equiv \alpha \bmod b_i^{\beta_i} \equiv \alpha \bmod q_i \equiv x_i^2 \equiv y_i^2 \text{ for all } i .$$

Hence, this satisfies all equations of the generated instance of the NON- UNIQUE REMAINDER problem making it a yes-instance.

$\Leftarrow$ Let the instance of the NON- UNIQUE REMAINDER problem be a yes-instance. Hence, we could verify that there exists a solution to the given equations smaller than ζ. Let this solution be denoted as z^*. It holds that $z^* \equiv x_i \bmod q_i$ or $z \equiv y_i \bmod q_i$. Let v_i correspond to the residue that was satisfied, i. e., $v_i = x_i$ or $v_i = y_i$. The solution z^* also solves the QUADRATIC CONGRUENCES problem. First, $z^* \leq \zeta = \gamma$. Further, it holds per definition of the numbers that

$$(z^*)^2 \equiv (v_i)^2 \equiv \alpha \bmod q_i \text{ for all } i.$$

As it satisfies all equations simultaneously and the b_i are pairwise co-prime, it follows from the Chinese Remainder theorem that

$$\begin{aligned}(z^*)^2 &\equiv (v_i)^2 \equiv \alpha \bmod q_i \text{ for all } i \\ &\equiv (z^*)^2 \equiv \alpha \bmod \prod_{i=1}^{n_{\text{NR}}} q_i \equiv \alpha \bmod \prod_{i=1}^{n_{\text{QC}}} b_i^{\beta_i} \equiv \alpha \bmod \beta\end{aligned}$$

as the $b_i^{\beta_i}$s are the prime factorization of β.

Running time: Setting the variables accordingly can be done in time polynomial in n_{QC}. Further, computing each x_i, y_i can be done in poly-logarithmic time regarding the largest absolute number for each $i \in \{1, \ldots, n_{\text{NR}}\}$ [5]. □

Finally, we reduce the NON- UNIQUE REMAINDER problem to the 2- STAGE ILP problem. Note that the considered 2- STAGE ILP problem is a decision problem. In other words, we only seek to determine whether there exists a feasible solution. We neither optimize a solution vector nor are we interested in the solution vector itself.

Theorem 3 *The* NON- UNIQUE REMAINDER *problem is reducible to the* 2- STAGE ILP *problem in polynomial time with the properties that* $n \in O(n_{NR})$, $r, s, t, ||c||_\infty, ||b||_\infty, ||\ell||_\infty \in O(1)$, $||u||_\infty \in O(\zeta)$, *and* $\Delta \in O(\max_i\{q_i\})$.

Proof *Transformation:* Having the instance for the NON- UNIQUE REMAINDER problem at hand we construct our ILP as follows with $n = n_{\text{NR}}$:

$$\mathcal{A} \cdot x = \begin{pmatrix} -1 & q_1 & x_1 & y_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ -1 & 0 & \ldots & 0 & 0 & \ldots & 0 & q_n & x_n & y_n \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & 1 & 1 \end{pmatrix} \cdot x = b = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

All variables get a lower bound of 0 and an upper bound of ζ. We can set the objective function arbitrarily as we are just searching for a feasible solution, hence we set it to $c = (0, 0, \ldots, 0)^\top$.

Instance size: Due to our construction, it holds that $t = 2, r = 1, s = 3$. The number n of repeated blocks equals the number n_{NR} of equations in the instance of the NON- UNIQUE REMAINDER problem. The largest entry Δ can be bounded by $\max_i\{q_i\}$. The lower and upper bounds are at most $||u||_\infty = O(\zeta)$, $||\ell||_\infty = O(1)$. The objective function c is set to zero and is thus of constant size. The largest value in the right-hand side is $||b||_\infty = 1$.

Correctness: $\Rightarrow$ Let the given instance of the NON- UNIQUE REMAINDER problem be a yes-instance. Thus, there exists a solution $z^* < \zeta$ satisfying all equations. As before, let v_i correspond to the remainder that was satisfied in each equation i, i. e., $v_i = x_i$ or $v_i = y_i$. A solution to our integer linear program now looks as follows: Set the first variable to z^*. Let the columns corresponding to x_i and y_i be set as follows for each i: If $v_i = x_i$ then set this variable occurrence in the solution vector to 1. Set the occurrence to the corresponding variable of y_i to zero. Otherwise, set the variables the other way round. Finally, the variable corresponding to the columns of the q_i are computed as $(z^* - v_i)/q_i$. It is easy to see that this solution is feasible and satisfies the bounds on the variable sizes.

$\Leftarrow$ Let the given instance of the 2- STAGE ILP problem be a yes-instance. By definition of the constraint matrix we have for every $1 \leq i \leq n$ that there exists a multiple $\lambda_i \geq 0$ such that $z = x_i + \lambda_i q_i$ or $z = y_i + \lambda_i q_i$. Hence $z \equiv x_i \mod q_i$ or $z \equiv y_i \mod q_i$ for every $1 \leq i \leq n$. Further, $z \leq u$. Thus, the solution z is a solution of the NON- UNIQUE REMAINDER problem.

Running time: Mapping the variables and computing the values for the q_is can all be done in polynomial time regarding the largest occurring number and n. □

4 Runtime bounds for 2-stage stochastic ILPs under ETH

This sections presents the proof that the double exponential running time in the current state-of-the-art algorithms is nearly tight assuming the Exponential Time Hypothesis (ETH). To do so, we make use of the reductions above showing that we can transform an instance of the 3-SAT problem to an instance of the 2- STAGE ILP problem.

Corollary 1 *The* 2- STAGE ILP *problem cannot be solved in time less than* $2^{\delta\sqrt{n}}$ *for some* $\delta > 0$ *assuming ETH.*

Proof Suppose the opposite. That is, there is an algorithm solving the 2- STAGE ILP problem in time less than $2^{\delta\sqrt{n}}$. Let an instance of the 3-SAT problem with n_3 variables and m_3 clauses be given. Due to the Sparsification lemma, we may assume that $m_3 \in O(n_3)$ [18]. The Sparsification lemma states that any 3-SAT formula can be replaced by subexponentially many 3-SAT formulas, each with a linear number of clauses with respect to the number of variables. The original formula is satisfiable if at least one of the new formulas is. This yields that if we cannot decide a 3-SAT problem in subexponential time, we can also not do so for a 3-SAT problem where $m_3 \in O(n_3)$.

We can reduce such an instance to an instance of the QUADRATIC CONGRUENCES problem in polynomial time regarding n_3 such that $n_{\text{QC}} \in O(n_3^2)$, $max_i\{b_i\} \in O(n_3^2 \log(n_3))$, $\alpha, \beta, \gamma = 2^{O(n_3^2 \log(n_3))}$, see Theorem 1.

Next, we reduce this instance to an instance of the NON- UNIQUE REMAINDER problem. Using Theorem 2, this yields the parameter sizes $n_{\text{NR}} \in O(n_3^2)$, $\max_{i\in\{1,\ldots,n_{\text{NR}}\}}\{q_i, x_i, y_i\} = O(n_3^2 \log(n_3))$, and finally $\zeta \in 2^{O(n_3^2 \log(n_3))}$. Note that all prime numbers greater than 2 appear at most once in the prime factorization of β and 2 appears 4 times. Thus, the largest q_i, which corresponds to $\max_i\{b_i^{\beta_i}\}$ equals the largest prime number in the QUADRATIC CONGRUENCES problem: The largest prime number is at least the $(\nu^2 + 2\nu + 2m' + 13) \geq$ 13th prime number by a rough estimation. The 13th prime number is 41 and thus, larger than $2^4 = 16$.

Finally, we reduce that instance to an instance of the 2- STAGE ILP problem with parameters $r, s, t, ||c||_\infty, ||b||_\infty, ||\ell||_\infty \in O(1)$, $||u||_\infty \in 2^{O(n_3^2 \log(n_3))}$, $n \in O(n_3^2)$, and $\Delta \in O(n_3^2 \log(n_3))$, see Theorem 3.

Hence, if there is an algorithm solving the 2- STAGE ILP problem in time less than $2^{\delta\sqrt{n}}$ this would result in the 3-SAT problem to be solved in time less than $2^{\delta\sqrt{n}} = 2^{\delta\sqrt{C_1 n_3^2}} = 2^{\delta(C_2 n_3))}$ for some constants C_1, C_2. Setting $\delta_3 \leq \delta/C_2$, this would violate the ETH. □

To prove our main result, we still have to reduce the size of the coefficients in the constraint matrix. To do so, we encode large coefficients into submatrices. This reduces the size of the entries greatly while just extending the matrix dimensions slightly. A similar approach was used for example in [10, 23] to prove a lower bound for the size of inclusion minimal kern-elements of 2-stage stochastic ILPs or in [25] to decrease the value of Δ in the matrices.

Theorem 4 *The* 2- STAGE ILP *problem cannot be solved in time less than* $2^{2^{\delta(s+t)}}|I|^{O(1)}$ *for some constant* $\delta > 0$*, even if* $r = 1$*,* $\Delta, ||b||_\infty, ||c||_\infty, ||b||_\infty \in O(1)$*, assuming ETH. Here* $|I|$ *denotes the encoding length of the total input.*

Proof First, we show that we can alter the resulting integer linear program such that we reduce the size of Δ to $O(1)$. We do so by encoding large coefficients with base 2, which comes at the cost of enlarged dimensions of the constraint matrix. Let $\text{enc}(x)$ be the encoding of a number x with base 2. Further, let $\text{enc}_i(x)$ be the ith number of $\text{enc}(x)$. Finally, $\text{enc}_0(x)$ denotes the last significant number of the encoding. Hence, the encoding of a number x is $\text{enc}(x) = \text{enc}_0(x)\text{enc}_1(x)\ldots\text{enc}_{\lfloor\log(\Delta)\rfloor}(x)$ and x can be reconstructed by $x = \sum_{i=0}^{\lfloor\log(\Delta)\rfloor} \text{enc}_i(x) \cdot 2^i$.

Let a matrix E be defined as,

$$E = \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 \\ 0 & 2 & -1 & 0\ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \ldots & 0 & 2 & -1 \end{pmatrix}.$$

We re-write the constraint matrix as follows: For each coefficient $a > 1$, we insert its encoding $\text{enc}(a)$ and beneath we put the matrix E. Furthermore, we have to fix the dimensions for the first row in the constraint matrix, the columns without great coefficients and the right-hand side b by filling the matrix at the corresponding positions with zeros. The altered integer linear program $\mathcal{A} \cdot x = b$ is displayed in Fig. 1.

Note that the ones beneath the sub-matrices $\text{enc}(x_i)$ and $\text{enc}(y_i)$ correspond to $\text{enc}_0(x_i)$ and $\text{enc}_0(y_i)$. The independent blocks consisting of $\text{enc}(a)$ and the matrix E beneath correctly encodes the number $a > 1$, i. e., it preserves the solution space: Let x_a be the number in the solution corresponding to the column with entry a of the original instance. The solution for the altered column (i. e., the sub-matrix) is $(x_a \cdot 2^0, x_a \cdot 2^1, \ldots, x_a \cdot 2^{\lfloor \log(\Delta) \rfloor})$. The additional factor of 2 for each subsequent entry is due to the diagonal of E. It is easy to see that $a \cdot x_a = \sum_{i=0}^{\lfloor \log(\Delta) \rfloor} \text{enc}_i(a) \cdot x_a \cdot 2^i$ as we can extract x_a on the right-hand side and solely the encoding of a remains. Thus, the solutions of the original matrix and the altered one directly transfer to each other. Hence, the solution space is preserved.

Regarding the dimensions, each coefficient $a > 1$ is replaced by a $(O(\log(\Delta)) \times O(\log(\Delta)))$ matrix. Thus, the dimension expands to $t' = t \cdot O(\log(\Delta)) = O(\log(\Delta))$, $s' = s \cdot O(\log(\Delta)) = O(\log(\Delta))$, while r and n stay the same. Further, we have to adjust the bounds. The lower bound for all new variables is also zero. For the upper bounds we allow an additional factor of 2^i for the ith value of the encoding. Thus, $||u'||_\infty = 2^{\lfloor \log(\Delta) \rfloor} ||u||_\infty$. Further, we get that the largest coefficient is bounded by $\Delta' = O(1)$. The right-hand side b enlarges to a vector b' with $O(n \log(\Delta))$ entries.

Now suppose there is an algorithm solving the 2- STAGE ILP problem in time less than $2^{2^{\delta(s+t)}} |I|^{O(1)}$. The Proof of Corollary 1 shows that we can transform an instance of the 3-SAT problem with n_3 variables and m_3 clauses to an 2-stage stochastic ILP with parameters $r, s, t, ||c||_\infty, ||b||_\infty, ||\ell||_\infty \in O(1)$, $||u||_\infty \in 2^{O(n_3^2 \log(n_3))}$, $n \in O(n_3^2)$, and $\Delta \in O(n_3^2 \log(n_3))$. Further, we explained above that we can transform this ILP to an equivalent one where

$$
\begin{aligned}
t' &= O(\log(\Delta)) = O(\log(n_3^2 \log(n_3))) = O(\log(n_3)),\\
s' &= O(\log(\Delta)) = O(\log(n_3^2 \log(n_3))) = O(\log(n_3)),\\
\Delta' &= O(1),\\
b' &\in \mathbb{Z}^{O(n_3^2 \log(n_3))},\\
||u'||_\infty &= 2^{\lfloor \log(\Delta) \rfloor} ||u||_\infty = 2^{\lfloor \log(n_3^2 \log(n_3)) \rfloor} 2^{O(n_3^2 \log(n_3))} = 2^{O(n_3^2 \log(n_3))},
\end{aligned}
$$

while r, and n stay the same. The encoding length $|I|$ is then given by

$$
\begin{aligned}
|I| &= (nt'(r + ns')) \log(\Delta') + (r + ns') \log(||\ell||_\infty)\\
&\quad + (r + ns') \log(||u'||_\infty) + nt' \log(||b'||_\infty) + (r + ns') \log(||c||_\infty)\\
&= 2^{O(n_3^2)}.
\end{aligned}
$$

Hence, if there is an algorithm solving the 2- STAGE ILP problem in time less than $2^{2^{\delta(s+t)}} |I|^{O(1)}$ this would result in the 3-SAT problem to be solved in time less than

$$
\begin{aligned}
2^{2^{\delta(s+t)}} |I|^{O(1)} &= 2^{2^{\delta(C_1 \log(n_3) + C_2 \log(n_3))}} 2^{n_3^{O(1)}} = 2^{2^{\delta C_3 \log(n_3)}} 2^{n_3^{O(1)}}\\
&= 2^{n_3^{\delta \cdot C_3}} 2^{n_3^{O(1)}} = 2^{n_3^{\delta \cdot C_4}}
\end{aligned}
$$

$$\begin{pmatrix}
-1 & \mathrm{enc}(q_1) & \mathrm{enc}(x_1) & \mathrm{enc}(y_1) & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & E & 0\dots0 & 0\dots0 & 0 & \dots & 0 & 0 & \dots & 0 \\
\vdots & & & & & & & & & \\
\vdots & 0\dots0 & E & 0\dots0 & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0\dots0 & 0\dots0 & E & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0\dots0 & 10\dots0 & 10\dots0 & 0 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
-1 & 0 & \dots & 0 & 0 & \dots & 0 & \mathrm{enc}(q_n) & \mathrm{enc}(x_n) & \mathrm{enc}(y_n) \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & E & 0\dots0 & 0\dots0 \\
\vdots & & & & & & & & & \\
\vdots & 0 & \dots & 0 & 0 & \dots & 0 & 0\dots0 & E & 0\dots0 \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & 0\dots0 & 0\dots0 & E \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & 0\dots0 & 10\dots0 & 10\dots0
\end{pmatrix} \cdot x = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Fig. 1 The displayed ILP is the altered ILP after encoding large entries with basis 2

for some constants C_1, C_2, C_3, C_4. Setting $\delta = \delta'/C_4$ we get $2^{n_3^{\delta C_4}} = 2^{n_3^{\delta'}}$. As it holds for sufficient large x and $\epsilon < 1$ that $x^\epsilon < \epsilon x$ it follows that $2^{n_3^{\delta'}} < 2^{\delta' n_3}$. This violates the ETH. Note that this result even holds if $r = 1$, $\Delta, ||c||_\infty, ||b||_\infty, ||\ell||_\infty \in O(1)$ as constructed by our reductions. □

5 Full proof of Theorem 1

This section presents the full Proof of Theorem 1. For an intuition and road map of the proof, we refer to Sect. 2.

First, let us prove a lemma about the size of the product of prime numbers, which comes in handy in the respective theorem.

Lemma 1 *Denote by q_i the ith prime number. The product of the first k prime numbers $\prod_{i=1}^{k} q_i$ is bounded by $2^{2k\log(k)}$ for all $k \geq 2$.*

Proof Denote by $\pi(x)$ the number of prime numbers of size at most x. It holds that $\pi(x) > x/\log(x)$ for $x \geq 17$ [30]. Note that the original statement uses the natural logarithm. But due to the division, the estimation also holds for the logarithm with base 2. Setting $x = y^2$, it holds that $\pi(y^2) > y^2/\log(y^2)$ for $y \geq 5$. As $y^2/\log(y^2) = y^2/(2\log(y)) \geq y^2/y = y$ for $y \geq 5$, it also holds that $\pi(y^2) > y$ for $y \geq 5$. Thus $p_i < i^2$ for $i \geq 5$, as we have at least i many prime numbers in the interval $[1, i^2]$.

Manually checking the values for the first four prime numbers shows that the equation $p_i \leq i^2$ even holds for all prime numbers greater 2. For $p_1 = 2 > 1^2$, we can simply multiply an additional factor of 2. Altogether, we can thus estimate the product of the first k prime numbers for $k \geq 2$ as

$$\begin{aligned}\prod_{i=1}^{k} q_i &\leq \prod_{i=1}^{k} (i^2) \cdot 2 = (\prod_{i=1}^{k} i)^2 \cdot 2 = (k!)^2 \cdot 2 \leq (2(k/2)^k)^2 \cdot 2 \\ &= 2^2((k/2)^k)^2 \cdot 2 = 2^3(k/2)^{2k} = 2^3 2^{2k\log(k/2)} \leq 2^{2k\log(k)}\end{aligned}$$

proving the statement. We use the estimation $k! = 2(k/2)^k$ which can easily be proved using induction. Further, note that $k \geq 2$ has to hold for the last estimation. □

Theorem 5 *The* QUADRATIC CONGRUENCES *problem is NP-hard even if the prime factorization of β is given and each prime factor greater than* 2 *occurs at most once and the prime factor* 2 *occurs* 4 *times.*

Proof We show a reduction from the well-known NP-hard problem 3-SAT where we are given a 3-SAT formula Φ with n_3 variables and m_3 clauses.
Transformation: First, eliminate duplicate clauses from Φ and those where some variable x_i and its negation $\bar{x_i}$ appear together. Call the resulting formula Φ', the number of occurring variables n' and denote by m' the number of appearing clauses respectively. Let $\Sigma = (\sigma_1, \ldots, \sigma_{m'})$ be some enumeration of the clauses. Denote by $p_1, \ldots, p_{2m'}$ the first $2m'$ prime numbers greater 2. Compute

$$\tau_{\Phi'} = -\sum_{k=1}^{m'} \prod_{i=1}^{k} p_i.$$

Further, compute for each $i \in 1, 2, \ldots, n'$:

$$f_j^+ = \sum_{x_i \in \sigma_k} \prod_{i=1}^{k} p_i \text{ and } f_j^- = \sum_{\bar{x_i} \in \sigma_k} \prod_{i=1}^{k} p_i.$$

Set $\nu = 2m' + n'$. Compute the coefficients c_j for all $j = 0, 1, \ldots, \nu$ as follows: Set $c_0 = 0$. For $j = 1, \ldots, 2m'$ set

$$c_j = -\frac{1}{2} \prod_{i=1}^{k} p_i \text{ if } j = 2k-1 \text{ and } c_j = -\prod_{i=1}^{k} p_i \text{ if } j = 2k, \text{ for some } k \in \mathbb{N}.$$

Compute the remaining coefficients for $j = 1, \ldots, n'$ as $c_{2m'+j} = \frac{1}{2} \cdot (f_j^+ - f_j^-)$. Further, set $\tau = \tau_{\Phi'} + \sum_{j=0}^{\nu} c_j + \sum_{j=1}^{n'} f_j^-$.

Denote by $q_1, \ldots, q_{\nu^2+2\nu+1}$ the first $\nu^2+2\nu+1$ prime numbers. Let $p_{0,0}$, $p_{0,1}, \ldots,$ $p_{0,\nu}$, $p_{1,0}, \ldots, p_{\nu,\nu}$ be the first $(\nu+1)^2 = \nu^2 + 2\nu + 1$ prime numbers greater than $(4(\nu+1)2^3 \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$ and greater than $p_{2m'}$. Define p^* as the $(\nu^2 + 2\nu + 2m' + 13)$th prime number.

Determine the parameters θ_j for $j = 0, 1, \ldots, \nu$ as the least θ_j satisfying:

$$\theta_j \equiv c_j \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i,$$

$$\theta_j \equiv 0 \bmod \prod_{i=0, i \neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k},$$

$$\theta_j \not\equiv 0 \bmod p_{j,1}.$$

Set the following parameters:

$$H = \sum_{j=0}^{\nu} \theta_j \text{ and } K = \prod_{i=0}^{\nu} \prod_{k=0}^{\nu} p_{i,k}.$$

Finally, set

$$\begin{aligned}
\alpha &= (2^4 \cdot p^* \prod_{i=1}^{m'} p_i + K)^{-1} \cdot (K\tau^2 + 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot H^2),\\
\beta &= 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot K,\\
\gamma &= H,
\end{aligned}$$

where $(2^4 \cdot p^* \prod_{i=1}^{m'} p_i + K)^{-1}$ is the inverse of $(2^4 \cdot p^* \prod_{i=1}^{m'} p_i + K) \bmod 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot K$.

Correctness: We show that the satisfiability of the formula Φ is equivalent to a line of (systems of) equations, i. e., the formula has a satisfying truth assignment on the variables if and only if the (systems of) equations admit a solution. By this, we prove the hardness for various problems along the way. These are listed above with their respective equivalence sketched. In the following, we separate each of these steps by claims. We do not state the formula for each variable repeatedly and refer to the transformation section for an overview.

However, before we start with the transformations of the formula, we first observe two properties about the generated prime factors. These come in handy for the estimations later on we need to prove the equivalence of the systems.

Claim 1 *Choosing p^* as the $(\nu^2 + 2\nu + 2m' + 13)$th prime factor satisfies $p^* > p_{\nu,\nu}$.*

Proof of Claim Denote by q_i the ith prime number: Suppose $p_{2m'} \geq (4(\nu + 1)2^3 \cdot \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$. Then $p_{\nu,\nu}$ is the $(\nu^2 + 2\nu + 1 + 2m' + 1)$th prime number and thus, $p^* > p_{\nu,\nu}$. Otherwise, if $p_{2m'} < (4(\nu + 1)2^3 \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$, we bound the function values as follows:

$$\begin{aligned}
&(4(\nu+1)2^3 \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}\\
&= 4^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))} (\nu+1)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}\\
&\quad \cdot (2^3)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}
\end{aligned}$$

$$\cdot \Big(\prod_{i=1}^{\nu^2+2\nu+1} q_i \Big)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$$
$$\leq 2 \cdot 2 \cdot 2 \cdot (2^{2(\nu^2+2\nu+1)\log(\nu^2+2\nu+1)})^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$$
$$\leq 8 \cdot (4^{(\nu^2+2\nu+1)\log(\nu^2+2\nu+1)})^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$$
$$= 8 \cdot 4 = 32.$$

The second transformation holds as the product of the first k prime numbers is bounded by $2^{2k\log(k)}$ (for $k \geq 2$, which obviously holds here), see Lemma 1. There are 11 prime numbers in the interval [1, 32]. Hence, $p_{\nu,\nu}$ is at most the $(11 + \nu^2 + 2\nu + 1)$th prime number and thus, $p^* > p_{\nu,\nu}$. □

Claim 2 *It holds that* $p^* \leq \prod_{i=m'+1}^{\nu^2+2\nu} q_i$.

Proof of Claim We can bound the value of the product from beneath as $\prod_{i=m'+1}^{\nu^2+2\nu} q_i \geq q_{m'+1}^{\nu^2+\nu}$. Estimating the value for p^*, we use that the value of the next prime number after a number ρ is at most 2ρ [2]. Thus, as there are $\nu^2 + 2\nu + m' + 10$ prime numbers between $p_{m'+1}$ and p^*, we get $p^* \leq p_{m'+1} \cdot 2^{\nu^2+2\nu+m'+10} \leq q_{m'+1} \cdot 2^{\nu^2+2\nu+m'+11} \leq q_{m'+1} \cdot 2^{\nu^2+3\nu+11}$ since per definition $p_i \leq 2 \cdot q_i$ and $\nu \geq m'$ holds. Dividing both sides of the estimation by $q_{m'+1}$, it thus remains to show that $2^{\nu^2+3\nu+11} \leq q_{m'+1}^{\nu^2+\nu-1}$. Obviously, $q_{m'+1}^{\nu^2+\nu-1}$ grows for larger values of m'. The smallest reasonable value for $m' = 2$ and thus, $q_{m'+1} \geq 5$. By that, we get that

$$q_{m'+1}^{\nu^2+\nu-1} \geq 5^{\nu^2+\nu-1} \geq 2^{2(\nu^2+\nu-1)} = 2^{2\nu^2+2\nu-2} \geq 2^{\nu^2+3\nu+11}$$

for all $\nu \geq 5$ and thus, for all reasonable values of ν, showing the statement. □

Let us now focus on the transformations of the formula Φ yielding the equivalence of the first two above-mentioned problems:

Claim 3 *The* 3-SAT *problem asking whether there is a function* $\eta\colon x_i \to \{0, 1\}$ *assigning a truth value to each variable that satisfies all clauses* σ_k *of the* 3*-SAT formula* Φ *simultaneously is a yes-instance if and only if Problem (P2) asking whether there are values* $y_k \in \{0, 1, 2, 3\}$ *and a truth assignment* η *such that* $0 = R_k = y_k - \sum_{x_i \in \sigma_k} \eta(x_i) - \sum_{\bar{x_i} \in \sigma_k} (1 - \eta(x_i)) + 1$ *for all k is a yes-instance.*

Proof of Claim Obviously, the reduced formula Φ' is satisfiable if and only if Φ is. The formula Φ' is satisfiable if there exists a truth assignment $\eta\colon \{x_1, \ldots, x_{n'}\} \to \{0, 1\}$ assigning a logical value to each variable $x_1, \ldots, x_{n'}$ which satisfies all clauses $\sigma_1, \ldots, \sigma_{m'}$ simultaneously. This can be re-written to the following equation for each clause $\sigma_k \in \Phi_k$ interpreting the truth values as numbers:

$$0 = R_k = y_k - \sum_{x_i \in \sigma_k} \eta(x_i) - \sum_{\bar{x_i} \in \sigma_k} (1 - \eta(x_i)) + 1, \quad y_k \in \{0, 1, 2, 3\}.$$

For a clause σ_k, this equation is only satisfiable if at least one variable $x_i \in \sigma_k$ has value $\eta(x_i) = 1$ or one variable occurring in its negation $\bar{x_i} \in \sigma_k$ has value $\eta(x_i) = 0$. Otherwise, we have to set $y_k = -1$ which is not allowed. □

Note that we never have to set $y_k = 3$ to satisfy the formula. However, we allow this value as it will come in handy later on when transforming the equation. Further, set $0 = R_0 = \alpha_0 + 1$ for $\alpha_0 \in \{-1, +1\}$ for later convenience. Clearly, the new equation is satisfiable.

Claim 4 *The Problem (P2) asking whether there are values $y_k \in \{0, 1, 2, 3\}$ and a truth assignment η such that $0 = R_k = y_k - \sum_{x_i \in \sigma_k} \eta(x_i) - \sum_{\bar{x_i} \in \sigma_k} (1 - \eta(x_i)) + 1$ for all k is a yes-instance if and only if Problem (P3) asking whether there are values $\alpha_j \in \{-1, +1\}$ such that $\sum_{j=0}^{\nu} \theta_j \alpha_j \equiv \tau \mod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ is a yes-instance.*

Proof of Claim We can bound the values of R_k for $k \in \{0, 1, \ldots, m'\}$ by $-2 \leq R_k \leq 4$. For the lower bound, the values are given by $y_k = 0$, all $x_i \in \sigma_k$ have value $\eta(x_i) = 1$ and all $\bar{x_i} \in \sigma_k$ have value $\eta(x_i) = 0$. For the upper bound we set $y_k = 3$, all $x_i \in \sigma_k$ to $\eta(x_i) = 0$ and $\bar{x_i} \in \sigma_k$ to $\eta(x_i) = 1$. For R_0 obviously $0 \leq R_0 \leq 2$ holds. Thus,

$$R_k = 0, \forall k \in \{0, 1, \ldots, m'\} \Leftrightarrow \sum_{k=0}^{m'} R_k \prod_{i=0}^{k} p_i = 0$$

as the sum is zero if all $R_k = 0$. For the opposite direction, if the sum is zero, then no $R_k \neq 0$ as the product of the prime numbers grows too fast. Thus, the other summands cannot compensate for some $R_k \neq 0$. We can bound the expression further by

$$\left| \sum_{k=0}^{m'} R_k \prod_{i=0}^{k} p_i \right| \leq 4 \sum_{k=0}^{m'} \prod_{i=0}^{k} p_i \leq 4(m' + 1) \prod_{i=0}^{m'} p_i < 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$$

as $p^* > p_{\nu,\nu}$, see Claim 1, and as $p_{\nu,\nu} > p_{m'} > m' + 1$. This yields

$$R_k = 0, \forall k \in \{0, 1, \ldots, m'\} \Leftrightarrow \sum_{k=0}^{m'} R_k \prod_{i=0}^{k} p_i \equiv 0 \mod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i \qquad \text{(I)}$$

as the modulo has no impact on the satisfiability of the equation.

Next, we aim to re-write R_k by replacing the variables y_k and $\eta(x_i)$ with new variables admitting a domain of $\{-1, 1\}$:

$$\begin{aligned} y_k &= 1/2 \cdot [(1 - \alpha_{2k-1}) + 2 \cdot (1 - \alpha_{2k})], \ k \in \{1, \ldots, m'\}, \\ \eta(x_i) &= 1/2 \cdot (1 - \alpha_{2m'+i}), \ i \in \{1, \ldots, n'\}. \end{aligned}$$

Obviously the value domains of y_k and $\eta(x_i)$ are preserved. Substituting the variables and re-arranging the Eq. (I) yields

$$\sum_{j=0}^{\nu} c_j \alpha_j \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i, \alpha_j \in \{-1, +1\}.$$

Intuitively, the α_js corresponding to the truth assignment each appear c_j times, a number which captures how often a variable occurs positive and, respectively, negative in the original formula. Additionally, we get some α_j variables due to the y_k variables. Their additional occurrences introduced by the corresponding c_j are cancelled out by τ. By definition of θ_j this is equivalent to

$$\sum_{j=0}^{\nu} \theta_j \alpha_j \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i, \alpha_j \in \{-1, +1\}$$

proving the claim. □

Let $H = \sum_{j=0}^{\nu} \theta_j$ and $K = \prod_{i=0}^{\nu} \prod_{j=0}^{\nu} p_{i,j}$ be defined as before. Consider the following system asking whether there is an $x \in \mathbb{Z}$ such that:

$$0 \leq |x| \leq H \tag{P4.1}$$

$$(H + x)(H - x) \equiv 0 \bmod K \tag{P4.2}$$

We use this system to integrate the condition $x \leq H$ into the transformations. In the following, we prove that each solution of this system is of form $x = \sum_{j=0}^{\nu} \alpha_j \theta_j$ and thus, Problem (P4) can be combined with Problem (P3) yielding Problem (P5).

Claim 5 *The Problem (P3) asking whether there are values $\alpha_j \in \{-1, +1\}$ such that $\sum_{j=0}^{\nu} \theta_j \alpha_j \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ is a yes-instance if and only if the Problem (P5) is a yes-instance.*

Proof of Claim The unique solutions x to the given system (P4) are of form

$$x = \sum_{j=0}^{\nu} \alpha_j \theta_j, \alpha \in \{-1, +1\}, j = 0, 1, \ldots, \nu.$$

Let us first verify that an x of such form solves the system. First

$$|x| = \left| \sum_{j=0}^{\nu} \alpha_j \theta_j \right| \leq \sum_{j=0}^{\nu} \theta_j = H$$

satisfies (P4.1). Further, we have that each summand in the expanded formula $(H + x)(H - x)$ has to contain all prime factors $p_{i,j}$ for $i = 0, 1, \ldots, \nu$ and $j = 0, 1, \ldots, \nu$

in its prime factorization to satisfy (P4.2). For $(H + x) = (\sum_{j=0}^{n} \theta_j + \sum_{j=0}^{n} \theta_j \alpha_j)$ it holds that each θ_j where $\alpha_j = +1$ occurs twice while each θ_j where $\alpha_j = -1$ is canceled out by H. The other way round holds for $(H - x)$. Thus, expanding the brackets yields that each summand is a product of some θ_j and θ_k where $\alpha_j = +1$ and $\alpha_k = -1$. This implies that $j \neq k$. As each θ_j contains all prime factors of K except $p_{j,0}, \ldots, p_{j,\nu}$, the product of two different θ_j and θ_k contains each prime factor occurring in K satisfying (P4.2).

Regarding the uniqueness, we first prove that each solution x' to the given system satisfies $x' \equiv x \bmod K$. Then we show that the distance of two solutions is at most $2H$. Finally, proving that $2H < K$ yields the desired statement.

Observe that

$$(H + x)(H - x) \equiv 0 \bmod \prod_{j=0}^{\nu} p_{i,j}, \forall i = 0, 1, \ldots, \nu.$$

Assume there exists some number $\tilde{p} = \prod_{j=0}^{\nu} p_{i,j}$ for some $i \in \{0, 1, \ldots, \nu\}$ which divides $(H + x)$ and $(H - x)$ (without remainder). Thus, $(H + x) + (H - x) \equiv 0 \bmod \tilde{p} \Leftrightarrow 2H \equiv 0 \bmod \tilde{p}$. As $\tilde{p}$ is a product of prime numbers greater than 2, it follows that $H \equiv 0 \bmod \tilde{p} \Leftrightarrow \sum_{j=0}^{\nu} \theta_j \equiv 0 \bmod \tilde{p}$. However, from the definition of θ_j (third condition) it follows that for each j there exist different prime numbers not present in the prime factorization of θ_j contradicting the assumption. Thus, $\tilde{p}$ divides either $(H + x)$ or $(H - x)$ (without remainder). Define

$$\alpha_j = \begin{cases} +1 & \text{if } (H - x) \equiv 0 \bmod \prod_{i=0}^{\nu} p_{j,i} \\ -1 & \text{if } (H + x) \equiv 0 \bmod \prod_{i=0}^{\nu} p_{j,i} \end{cases}$$

$$x' = \sum_{j=0}^{\nu} \alpha_j \theta_j.$$

In the following, we show that $x' \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k}$ holds for all $j \in \{0, 1, \ldots, \nu\}$:

$$\begin{aligned} x' &\equiv x \bmod \prod_{k=0}^{\nu} p_{j,k} \\ &\Leftrightarrow \sum_{j=0}^{\nu} \alpha_j \theta_j \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k} \\ &\Leftrightarrow \alpha_j \theta_j \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k} \\ &\Leftrightarrow \sum_{i=0}^{\nu} \alpha_j \theta_i \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k} \\ &\Leftrightarrow \alpha_j \sum_{i=0}^{\nu} \theta_i \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k} \end{aligned}$$

$$\Leftrightarrow \alpha_j H \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k}.$$

The first transformation simply inserts the definition of x'. Due to the definition of the θ_i, only the summand θ_j remains after calculating the modulo. Thus, we can sum up all θ_i with arbitrary sign as they equal zero after calculating the modulo. In the last step, we insert the definition of H. Now we either have $\alpha_j = +1$. Then $H \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k}$, i. e., $H - x \equiv 0 \bmod \prod_{k=0}^{\nu} p_{j,k}$, which is true by definition of $\alpha_j = +1$. Otherwise, $\alpha_j = -1$. Then $-H \equiv x \bmod \prod_{k=0}^{\nu} p_{j,k}$, i. e., $H + x \equiv 0 \bmod \prod_{k=0}^{\nu} p_{j,k}$, which is again true by the definition of α_j. Thus, the initial statement is correct. As it holds for all j, we can conclude that $x' \equiv x \bmod K$.

As $\alpha_j \in \{-1, +1\}$ for all $j \in \{0, 1, \ldots, \nu\}$, it holds that $-H \leq x \leq H$. Since the same holds for x' it follows that $|x - x'| \leq 2H$.

We can bound the value of θ_j as $\theta_j < 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot \prod_{i=0, i\neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k}$, as $2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ and $\prod_{i=0, i\neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k}$ are coprime and thus, the least θ_j satisfying the equivalence conditions in the definition of θ_j is at most their product [31]. The additional factor of 2 is introduced by the inequality constraint $\theta_j \not\equiv 0 \bmod p_{j,1}$, as if the calculated θ_j for the equality constraints does not satisfy that condition, we can extend it to $\theta'_j = \theta_j + 2^3 \cdot p^* \prod_{i=1}^{m'} p_i \cdot \prod_{i=0, i\neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k}$. This doubles the size estimation and as $p_{j,1}$ is coprime to $2^3 \cdot p^* \prod_{i=1}^{m'} p_i \cdot \prod_{i=0, i\neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k}$, it holds that θ'_j is not equivalent to 0 mod $p_{j,1}$.

As $p^* \leq \prod_{i=m'+1}^{\nu^2+2\nu} q_i$, see Claim 2, we get

$$2^4 \cdot p^* \prod_{i=1}^{m'} p_i \leq 2^4 \prod_{i=m'+1}^{\nu^2+2\nu} q_i \prod_{i=1}^{m'} p_i = 2^4 \prod_{i=1}^{\nu^2+2\nu+1} q_i.$$

Using this and our choice for the prime factors to satisfy $p_{0,0} > (4(\nu + 1)2^3 \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$, we estimate:

$$\begin{aligned}
&2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot \prod_{i=0, i\neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k} \\
&= \frac{2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot K}{\prod_{k=0}^{\nu} p_{j,k}} \\
&\leq \frac{2^4 \prod_{i=1}^{\nu^2+2\nu+1} q_i \cdot K}{\prod_{k=0}^{\nu} (4(\nu+1)2^3 \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}} \\
&\leq \frac{K}{2(\nu+1)}.
\end{aligned}$$

This term bounds each value of θ_j. It follows that $2H = 2\sum_{j=0}^{\nu} \theta_j < 2 \cdot (\nu + 1) \cdot K/(2(\nu + 1)) = K$. Thus, $x = x'$, as each solution x' to the given system satisfies

$x' \equiv x \bmod K$ and the distance of two solutions is at most $2H < K$. Hence, we conclude that solutions of the form $x = \sum_{j=0}^{v} \theta_j \alpha_j$ are the unique solutions to the system (P4.1) and (P4.2).

Thus, we can re-write

$$\sum_{j=0}^{v} \theta_j \alpha_j \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i, \alpha_j \in \{-1, +1\}$$

using the system (P4.1) and (P4.2) to the following one:

$$0 \leq |x| \leq H, x \in \mathbb{Z} \tag{P5.1}$$

$$x \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i \tag{P5.2}$$

$$(H + x)(H - x) \equiv 0 \bmod K \tag{P5.3}$$

proving their equivalence. □

Next, we re-write the system (P5) to:

$$0 \leq |x| \leq H, x \in \mathbb{Z} \tag{P6.1}$$

$$(\tau - x)(\tau + x) \equiv 0 \bmod 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \tag{P6.2}$$

$$(H + x)(H - x) \equiv 0 \bmod K. \tag{P6.3}$$

Claim 6 *The Problem (P5) is a yes-instance if and only if the Problem (P6) is a yes-instance.*

Proof of Claim As only the second conditions differ, we focus on their equivalence in the following. First, we prove that if (P5.2) holds, i. e., $x \equiv \tau \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$, then (P6.2) holds, i. e., $(\tau - x)(\tau + x) \equiv 0 \bmod 2^4 \cdot p^* \prod_{i=1}^{m'} p_i$. We can re-write (P5.2) to $x = \lambda 2^3 \cdot p^* \prod_{i=1}^{m'} p_i + \tau$ for some $\lambda \in \mathbb{Z}$. Inserting this in (P6.2) yields:

$$\begin{aligned}
&(\tau + \lambda 2^3 \cdot p^* \prod_{i=1}^{m'} p_i + \tau)(\tau - \lambda 2^3 \cdot p^* \prod_{i=1}^{m'} p_i - \tau) \\
&= (2\tau + \lambda 2^3 \cdot p^* \prod_{i=1}^{m'} p_i)(\lambda 2^3 \cdot p^* \prod_{i=1}^{m'} p_i) \equiv 0 \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i
\end{aligned}$$

as each factor is multiplied with $\lambda 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$.

Next, we prove the opposite direction. First, observe that if $(\tau - x)(\tau + x) \equiv 0 \bmod 2^4 \cdot p^* \prod_{i=1}^{m'} p_i$ then either $(\tau - x) \equiv 0 \bmod 2^3$ or $(\tau + x) \equiv 0 \mod 2^3$: As (P6.2) holds, $(\tau + x) = \lambda_i \cdot 2^i$ and $(\tau - x) = \lambda_j \cdot 2^j$ for some $i, j \in \mathbb{Z}$ and $\lambda_i, \lambda_j \not\equiv 0 \bmod 2$. It follows that

$$\begin{aligned}(\tau + x) + (\tau - x) &= \lambda_i \cdot 2^i + \lambda_j \cdot 2^j \\ \Leftrightarrow 2\tau &= \lambda_i \cdot 2^i + \lambda_j \cdot 2^j \\ \Leftrightarrow \tau &= \lambda_i \cdot 2^{i-1} + \lambda_j \cdot 2^{j-1}.\end{aligned}$$

As τ is odd per definition, either i or j has to be 1 and thus, the other parameter has to be 3. Using this, we know that if x satisfies (P6.2), then $(\tau - x) \equiv 0 \bmod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ or $(\tau + x) \equiv 0 \mod 2^3 \cdot p^* \prod_{i=1}^{m'} p_i$. In the first case, x directly corresponds to a solution of (P5.2) as $x - \tau$ is a multiple of $2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ and thus, x is a multiple of $2^3 \cdot p^* \prod_{i=1}^{m'} p_i$ with a residue of τ. Otherwise $-x$ satisfies the condition using the same argument. Obviously the other conditions are also satisfied in both systems. □

Lastly, we re-write the system one final time to:

$$0 \leq x \leq H, x \in \mathbb{Z} \tag{QC.1}$$

$$2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i (H^2 - x^2) + K(\tau^2 - x^2)$$

$$\equiv 0 \bmod 2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i \cdot K. \tag{QC.2}$$

Claim 7 *The Problem (P6) is a yes-instance if and only if the* QUADRATIC CONGRUENCES *problem is a yes-instance.*

Proof of Claim First, as we only consider x^2, we can suppose $x \geq 0$ and thus, re-writting (P6.1) to (QC.1) is correct. Further, (P6.2) and (P6.3) merge into (QC.2). Recall that $2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i$ and K are co-prime. The first summand obviously always contains the factor $2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i$, thus we have to find an x such that $(H^2 - x^2) \equiv 0 \bmod K$ which corresponds to (P6.3). The second summand clearly is a multiple of K, thus we have to assure that $(\tau^2 - x^2) \equiv 0 \bmod 2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i$. This matches (P6.2).

Dissolving the brackets and rearranging the term (QC.2) we get

$$\begin{aligned}&(2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i + K)x^2 \\ &\equiv K\tau^2 + 2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i H^2 \bmod 2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i \cdot K.\end{aligned}$$

As $2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i + K$ is relatively prime to $2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i \cdot K$ it has an inverse modulo $2^4 \cdot p^* \cdot \prod_{i=1}^{m'} p_i \cdot K$ [27]. Thus, multiplying by the inverse we get the values for α, β and γ as in the transformation above. □

Overall, this proves that satisfying the formula Φ is equivalent to an instance of the QUADRATIC CONGRUENCES problem admitting a feasible solution.

Running time: All steps, numbers and their computation can be bounded in a polynomial dependent of n_3, i. e., the number of variables in the 3-Sat formula, and m_3, i. e., the number of clauses in the formula. First, we eliminate unnecessary clauses from the formula. Hence, we have to go through all clauses once. The first $2m' + 1$ prime numbers have a value of at most $O(m' \log(m'))$ and can thus be found in polynomial time via sieving. The function $(4(\nu+1)2^3 \prod_{i=1}^{\nu^2+2\nu+1} q_i)^{1/((\nu^2+2\nu+1)\log(\nu^2+2\nu+1))}$ is at most 32 as shown before. Hence, we can also bound the value of the next $\nu^2 + 2\nu + 1$ prime numbers larger than 32 and $p_{2m'}$ by a polynomial in n_3 and m_3 and we can compute them efficiently by sieving. All other numbers calculated in the transformation are a product or sum over these prime numbers (each occurring at most once in the calculation) and thus, their values are also in poly(n_3, m_3). We can compute the inverse $(2^4 \cdot p^* \prod_{i=1}^{m'} p_i + K)^{-1}$ in polynomial time [27]. □

Now we have proved that the QUADRATIC CONGRUENCES problem is NP-hard even in the restricted case where all prime factors in β only appear at most once (except 2). To apply the ETH, however, we also have to estimate the dimensions of the generated instance. The above reduction yields the following parameters:

Theorem 1 *An instance of the* 3-SAT *problem with n_3 variables and m_3 clauses is reducible to an instance of the* QUADRATIC CONGRUENCES *problem in polynomial time with the properties that $\alpha, \beta, \gamma \in 2^{O((n_3+m_3)^2 \log(n_3+m_3))}$, $n_{\mathrm{QC}} \in O((n_3+m_3)^2)$, $max_i\{b_i\} \in O((n_3+m_3)^2 \log(n_3+m_3))$, and each prime factor in β occurs at most once except the prime factor 2 which occurs four times.*

Proof In Theorem 5, we already showed and proved a reduction from the 3-SAT problem to the QUADRATIC CONGRUENCES problem and argued the running time. It remains to bound the parameters. To do so, we bound the numbers occurring in the reduction above in order of their appearance. Again, for an overview of the generated numbers and variables, and their respective formulas, we refer to the transformation section of Theorem 5.

After eliminating the trivial clauses it obviously holds that $m' \leq m_3$ and $n' \leq n_3$. Next, we calculate $\tau_{\Phi'}$. Its absolute value can be bounded as

$$|\tau_{\Phi'}| = |-\sum_{k=1}^{m'} \prod_{i=1}^{k} p_i| = \sum_{k=1}^{m'} \prod_{i=1}^{k} p_i$$
$$\leq m_3 \prod_{i=1}^{m_3} p_i \leq m_3 2^{2m_3 \log(m_3)} \leq 2^{O(m_3 \log(m_3))}$$

since the product of the first k prime numbers is bounded by $2^{2k\log(k)}$ for all $k \geq 2$, see Lemma 1. Similarly, $\max_i\{|f_i^+|, |f_i^-|\} \leq \sum_{x_i \in \sigma_j} \prod_{k=1}^{j} p_k + \sum_{\bar{x}_i \in \sigma_j} \prod_{k=1}^{j} p_k \leq$

$2m_3 \cdot 2^{2m_3 \log(m_3)} \leq 2^{O(m_3 \log(m_3))}$ and also $\max_j\{c_j\} = \max_j\{\prod_{i=1}^{j} p_i, f_j^+ + f_j^-\} \leq 2^{O(m_3 \log(m_3))}$. Per definition, $\nu = 2m' + n' = O(n_3 + m_3)$. The largest prime number $max_i\{b_i\}$ we generate in the reduction is p^*, which is the $(\nu^2 + 2\nu + 2m' + 13)$th prime number. Thus, its value is bounded by $p^* \leq O(\nu^2 \log(\nu)) = O((n_3 + m_3)^2 \log(n_3 + m_3))$ [14]. Due to the modulo, we can bound $\max_j\{\theta_j\}$ as

$$\begin{aligned}\max_j\{\theta_j\} &\leq 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot \prod_{i=0, i\neq j}^{\nu} \prod_{k=0}^{\nu} p_{i,k} \\ &\leq 2^4 2^{O((n_3+m_3)^2 \log(n_3+m_3))} = 2^{O((n_3+m_3)^2 \log(n_3+m_3))}.\end{aligned}$$

Thus, $H = \sum_{j=0}^{\nu} \theta_j \leq \nu \cdot 2^{O((n_3+m_3)^2 \log(n_3+m_3))} = 2^{O((n_3+m_3)^2 \log(n_3+m_3))}$ and $K = \prod_{i=0}^{\nu} \prod_{k=0}^{\nu} p_{i,k} \leq 2^{O((n_3+m_3)^2 \log(n_3+m_3))}$. Finally, we can bound the main parameters. As α is bounded by the modulo of β is follows that $\alpha \leq \beta$. Further, $\beta = 2^4 \cdot p^* \prod_{i=1}^{m'} p_i \cdot K \leq 2^{O((n_3+m_3)^2 \log(n_3+m_3))}$. Per definition $\gamma = H$ and thus, $\gamma \leq 2^{O((n_3+m_3)^2 \log(n_3+m_3))}$, which finalizes the estimation of the numbers. □

Funding Open access funding provided by EPFL Lausanne. This work was supported by the DFG project JA 612/20-1.

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of interest Not applicable.

References

1. Albareda-Sambola, M., van der Vlerk, M.H., Fernández, E.: Exact solutions to a class of stochastic generalized assignment problems. Eur. J. Oper. Res. **173**(2), 465–487 (2006)
2. Bertrand, J.: Bertrand's postulate chapter 2. Proofs from THE BOOK, page 9, (2018)
3. Brand, C., Koutecký, M., Ordyniak, S.: Parameterized algorithms for MILPs with small treedepth. CoRR, arXiv:1912.03501, (2019)
4. Chen, L., Koutecký, M., Xu, L., Shi, W.: New bounds on augmenting steps of block-structured integer programs. In ESA, volume 173 of LIPIcs, pages 33:1–33:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, (2020)
5. Crandall, R., Pomerance, C.B.: Prime Numbers: a Computational Perspective. Springer, Berlin (2006)

6. Cslovjecsek, J., Eisenbrand, F., Hunkenschröder, C., Rohwedder, L., Weismantel, R.: Block-structured integer and linear programming in strongly polynomial and near linear time. In SODA, pages 1666–1681. SIAM, (2021)
7. Cslovjecsek, J., Eisenbrand, F., Pilipczuk, M., Venzin, M., Weismantel, R.: Efficient sequential and parallel algorithms for multistage stochastic integer programming using proximity. CoRR, arXiv:2012.11742, (2020)
8. Cygan, M., Fomin, F.V., Kowalik, L., Lokshtanov, D., Marx, D., Pilipczuk, M., Saurabh, S.: Parameterized Algorithms. Springer, Marcin Pilipczuk (2015)
9. De Loera, J.A., Hemmecke, R., Onn, S., Weismantel, R.: N-fold integer programming. Discret. Optim. **5**(2), 231–241 (2008)
10. De Loera, J.A., Onn, S.: All linear and integer programs are slim 3-way transportation programs. SIAM J. Optim. **17**(3), 806–821 (2006)
11. Dempster, M.A.H., Fisher, M.L., Jansen, L., Lageweg, B.J., Lenstra, J.K., Rinnooy Kan, A.H.G.: Analysis of heuristics for stochastic programming: results for hierarchical scheduling problems. Math. Oper. Res. **8**(4), 525–537 (1983)
12. Eisenbrand, F., Hunkenschröder, C., Klein, K.-M., Koutecký, M., Levin, A., Onn, S.: An algorithmic theory of integer programming. CoRR, arXiv:1904.01361, (2019)
13. Gavenčiak, Tomáš, Koutecký, Martin, Knop, Dušan: Integer programming in parameterized complexity: Five miniatures. Discret. Optim. page 100596, (2020)
14. Hardy, G.H., Littlewood, J.E.: Contributions to the theory of the riemann zeta-function and the theory of the distribution of primes. Acta Math. **41**, 119–196 (1916)
15. Hemmecke, R., Köppe, M., Weismantel, R.: A polynomial-time algorithm for optimizing over N-fold 4-block decomposable integer programs. In IPCO, volume 6080 of Lecture Notes in Computer Science, pages 219–229. Springer, (2010)
16. Hemmecke, R., Schultz, R.: Decomposition of test sets in stochastic integer programming. Math. Program. **94**(2–3), 323–341 (2003)
17. Impagliazzo, R., Paturi, R.: On the complexity of k-SAT. J. Comput. System Sci. **62**(2), 367–375 (2001)
18. Impagliazzo, R., Paturi, R., Zane, F.: Which problems have strongly exponential complexity? J. Comput. System Sci. **63**(4), 512–530 (2001)
19. Ireland, K., Rosen, M.: A classical introduction to modern number theory, volume 84 of Graduate texts in mathematics. Springer, Berlin (1982)
20. Jansen, K., Klein, K.-M., Reute, J.: Complexity bounds for block-ips. Technical report, Department of Computer Science, Kiel University (2021)
21. Jansen, K., Lassota, A., Rohwedder, L.: Near-linear time algorithm for n-fold ilps via color coding. SIAM J. Discret. Math. **34**(4), 2282–2299 (2020)
22. Kall, P., Wallace, S.W.: Stochastic programming. Springer, Berlin (1994)
23. Klein, K.-M.: About the complexity of two-stage stochastic IPs. In IPCO, volume 12125 of Lecture Notes in Computer Science, pages 252–265. Springer, (2020)
24. Klein, K.-M., Reuter, J.: Collapsing the tower - on the complexity of multistage stochastic ips. CoRR, arXiv:2110.12743, (2021). To appear in SODA 22
25. Knop, D., Pilipczuk, M., Wrochna, M.: Tight complexity lower bounds for integer linear programming with few constraints. ACM Trans. Comput. Theory. **12**(3), 19:1-19:19 (2020)
26. Küçükyavuz, S., Sen, S.: An introduction to two-stage stochastic mixed-integer programming. In Leading Developments from INFORMS Communities, pages 1–27. INFORMS, (2017)
27. Lamé, G.: Note sur la limite du nombre des divisions dans la recherche du plus grand commun diviseur entre deux nombres entiers. (1844)
28. Laporte, G., Louveaux, F.V., Mercure, H.: A priori optimization of the probabilistic traveling salesman problem. Oper. Res. **42**(3), 543–549 (1994)
29. Manders, K.L., Adleman, L.M.: NP-complete decision problems for binary quadratics. J. Comput. Syst. Sci. **16**(2), 168–184 (1978)
30. Rosser, J.B., Schoenfeld, L.: Approximate formulas for some functions of prime numbers. Ill. J. Math. **6**, 64–94 (1962)
31. Schroeder, M.: The chinese remainder theorem and simultaneous congruences. In Number Theory in Science and Communication, pages 235–243. Springer, (2009)
32. Schultz, R., Stougie, L., Van Der Vlerk, M.H.: Two-stage stochastic integer programming: a survey. Stat. Neerl. **50**(3), 404–416 (1996)

33. Wagon, S.: Mathematica in action. Springer Science & Business Media, Berlin (1999)