Linear Algebra and its Applications 504 (2016) 64–107

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

The complexity of divisibility

Johannes Bausch [a,*], Toby Cubitt [a,b]

[a] *DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK*
[b] *Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK*

ARTICLE INFO

Article history:
Received 17 April 2015
Accepted 24 March 2016
Available online xxxx
Submitted by V. Mehrmann

MSC:
60-08
81-08
68Q30

Keywords:
Stochastic matrices
CPTP maps
Probability distributions
Divisibility
Decomposability
Complexity theory

ABSTRACT

We address two sets of long-standing open questions in linear algebra and probability theory, from a computational complexity perspective: stochastic matrix divisibility, and divisibility and decomposability of probability distributions. We prove that finite divisibility of stochastic matrices is an NP-complete problem, and extend this result to nonnegative matrices, and completely-positive trace-preserving maps, i.e. the quantum analogue of stochastic matrices. We further prove a complexity hierarchy for the divisibility and decomposability of probability distributions, showing that finite distribution divisibility is in P, but decomposability is NP-hard. For the former, we give an explicit polynomial-time algorithm. All results on distributions extend to weak-membership formulations, proving that the complexity of these problems is robust to perturbations.

* Corresponding author.
E-mail addresses: jkrb2@cam.ac.uk (J. Bausch), t.cubitt@ucl.ac.uk (T. Cubitt).

http://dx.doi.org/10.1016/j.laa.2016.03.041

1. Introduction and overview

People have pondered divisibility questions throughout most of Western science and philosophy. Perhaps the earliest written mention of divisibility is in *Aristotle's Physics* in 350 BC, in the form of the *Arrow paradox*—one of *Zeno of Elea's* paradoxes (ca. 490–430 BC). Aristotle's lengthy discussion of divisibility (he devotes an entire chapter to the topic) was motivated by the same basic question as more modern divisibility problems in mathematics: can the behaviour of an object—physical or mathematical—be subdivided into smaller parts?

For example, given a description of the evolution of a system over some time interval t, what can we say about its evolution over the time interval $t/2$? If the system is stochastic, this question finds a precise formulation in the *divisibility problem* for stochastic matrices [19]: given a stochastic matrix $\mathbf{P}$, can we find a stochastic matrix $\mathbf{Q}$ such that $\mathbf{P} = \mathbf{Q}^2$?

This question has many applications. For example, in information theory stochastic matrices model noisy communication channels, and divisibility becomes important in *relay coding*, when signals must be transmitted between two parties where direct end-to-end communication is not available [23]. Another direct use is in the analysis of chronic disease progression [3], where the transition matrix is based on sparse observations of patients, but finer-grained time-resolution is needed. In finance, changes in companies' credit ratings can be modelled using discrete time Markov chains, where rating agencies provide a transition matrix based on annual estimates—however, for valuation or risk analysis, a transition matrix for a much shorter time periods needs to be inferred [17].

We can also ask about the evolution of the system for *all* times up to time t, i.e. whether the system can be described by some continuous evolution. For stochastic matrices, this has a precise formulation in the *embedding problem*: given a stochastic matrix $\mathbf{P}$, can we find a generator $\mathbf{Q}$ of a continuous-time Markov process such that $\mathbf{P} = \exp(\mathbf{Q}t)$? The embedding problem seems to date back further still, and was already discussed by Elfving in 1937 [10]. Again, this problem occurs frequently in the field of systems analysis, and in analysis of experimental time-series snapshots [7,22,27].

Many generalizations of these divisibility problems have been studied in the mathematics and physics literature. For example, the question of square-roots of (entry-wise) nonnegative matrices is an old open problem in matrix analysis [24]: given an entry-wise nonnegative matrix $\mathbf{M}$, does it have an entry-wise nonnegative square-root? In quantum mechanics, the analogue of a stochastic matrix is a completely-positive trace preserving (CPTP) map, and the corresponding divisibility problem asks: when can a CPTP map $\mathbf{T}$ be decomposed as $\mathbf{T} = \mathbf{R} \circ \mathbf{R}$, where $\mathbf{R}$ is itself CPTP? The continuous version of this, whether a CPTP can be embedded into a completely-positive semi-group, is sometimes called the *Markovianity problem* in physics [8]—the latter again has applications to subdivision coding of quantum channels in quantum information theory [26].

Instead of dynamics, we can also ask whether the description of the static state of a system can be subdivided into smaller, simpler parts. Once again, probability theory provides a rich source of such problems. The most basic of these is the classic topic of

divisible distributions: given a random variable X, can it be decomposed into $X = Y + Z$ where Y, Z are some other random variables? What if Y and Z are identically distributed? If we instead ask for a decomposition into infinitely many random variables, this becomes the question of whether a distribution is infinitely divisible.

In this work, we address two of the most long-standing open problems on divisibility: divisibility of stochastic matrices, and divisibility and decomposability of probability distributions. We also extend our results to divisibility of nonnegative matrices and completely positive maps. Surprisingly little is known about the divisibility of stochastic matrices. Dating back to 1962 [19], the most complete characterization remains for the case of a 2×2 stochastic matrix [14]. The infinite divisibility problem has recently been solved [8], but the finite case remains an open problem. Divisibility of random variables, on the other hand, is a widely-studied topic. Yet, despite first results dating back as far as 1934 [5], no general method of answering whether a random variable can be written as the sum of two or more random variables—whether distributed identically, or differently—is known.

We focus on the computational complexity of these divisibility problems. In each case, we show which of the divisibility problems have efficient solutions—for these, we give an explicit efficient algorithm. For all other cases, we prove reductions to the famous P = NP-conjecture, showing that those problems are NP-hard. This essentially implies that—unless P = NP—the geometry of the corresponding divisible and non-divisible is highly complex, and these sets have no simple characterization beyond explicit enumeration. In particular, this shows that any future concrete classification of these NP-hard problems will be at least as hard as answering P = NP.

The following theorems summarize our main results on maps. Precise formulations and proofs can be found in section 2.

Theorem 1. *Given a stochastic matrix* $\mathbf{P}$*, deciding whether there exists a stochastic matrix* $\mathbf{Q}$ *such that* $\mathbf{P} = \mathbf{Q}^2$ *is* NP*-complete.*

Theorem 2. *Given a* CPTP *map* $\mathbf{B}$*, deciding whether there exists a* CPTP *map* $\mathbf{A}$ *such that* $\mathbf{B} = \mathbf{A} \circ \mathbf{A}$ *is* NP*-complete.*

In fact, the last two theorems are strengthenings of the following result.

Theorem 3. *Given a nonnegative matrix* $\mathbf{M}$*, deciding whether there exists a nonnegative matrix* $\mathbf{N}$ *such that* $\mathbf{M} = \mathbf{N}^2$ *is* NP*-complete.*

The following theorems summarize our main results on distributions. Precise formulations and proofs can be found in section 3.

Theorem 4. *Let* X *be a finite discrete random variable. Deciding whether* X *is* n*-divisible—i.e. whether there exists a random variable* Y *such that* $X = \sum_{i=1}^{n} Y$*—is in* P*.*

Theorem 5. *Let X be a finite discrete random variable, and $\epsilon > 0$. Deciding whether there exists a random variable Y ϵ-close to X such that Y is n-divisible, or that there exists such a Y that is nondivisible, is contained in* P.

Theorem 6. *Let X be a finite discrete random variable. Deciding whether X is decomposable—i.e. whether there exist random variables Y, Z such that $X = Y + Z$—is* NP*-complete.*

Theorem 7. *Let X be a finite discrete random variable, and $\epsilon > 0$. Deciding whether there exists a random variable Y ϵ-close to X such that Y is decomposable, or that there exists such an ϵ-close Y that is indecomposable, is* NP*-complete.*

It is interesting to contrast the results on maps and distributions. In the case of maps, the homogeneous 2-divisibility problems are already NP-hard, whereas finding an inhomogeneous decomposition is straightforward. For distributions, on the other hand, the homogeneous divisibility problems are efficiently solvable to all orders, but becomes NP-hard if we relax it to the inhomogeneous decomposability problem.

This difference is even more pronounced for infinite divisibility. The infinite divisibility problem for maps is NP-hard (shown in [8]), whereas the infinite divisibility and decomposability problems for distributions are computationally trivial, since indivisible and indecomposable distributions are both dense—see section 3.5.8 and 3.4.5.

The paper is divided into two parts. We first address stochastic matrix and CPTP divisibility in section 2, obtaining results on entry-wise positive matrix roots along the way. Divisibility and decomposability of probability distributions is addressed in section 3. In both sections, we first give an overview of the history of the problem, stating previous results and giving precise definitions of the problems. We introduce the necessary notation at the beginning of each section, so that each section is largely self-contained.

2. CPTP and stochastic matrix divisibility

2.1. Introduction

Mathematically, subdividing Markov chains is known as the *finite divisibility* problem. The simplest case is the question of finding a stochastic root of the transition matrix (or a CPTP root of a CPTP map in the quantum setting), which corresponds to asking for the evolution over half of the time interval. While the question of divisibility is rather simple to state mathematically, it is not clear a priori whether a stochastic matrix root for a given stochastic matrix exists at all. Historically, this has been a long-standing open question, dating back to at least 1962 [19]. Matrix roots were also suggested early on in other fields, such as economics and general trade theory, at least as far back as 1967 [31], to model businesses and the flow of goods. Despite this long history, very little is known about the existence of stochastic roots of stochastic matrices. The most

complete result to date is a full characterization of 2×2 matrices, as given for example in [14]. The authors mention that *"... it is quite possible that we have to deal with the stochastic root problem on a case-by-case basis."* This already suggests that there might not be a simple mathematical characterization of divisible stochastic matrices—meaning one that is simpler than enumerating the exponentially many roots and checking each one for stochasticity.

There are similarly few results if we relax the conditions on the matrix normalization slightly, and ask for (entry-wise) nonnegative roots of (entry-wise) nonnegative matrices—for a precise formulation, see Definition 10 and 11. An extensive overview can be found in [24]. Following this long history of classical results, quantum channel divisibility recently gained attention in the quantum information literature. The foundations were laid in [33], where the authors first introduced the notion of *channel divisibility*. A divisible quantum channel is a CPTP map that can be written as a nontrivial concatenation of two or more quantum channels.

A related question is to ask for the evolution under infinitesimal time steps, which is equivalent to the existence of a logarithm of a stochastic matrix (or CPTP map) that generates a stochastic (resp. CPTP) semi-group. Classically, the question is known as *Elfving's* problem or the *embedding problem*, and it seems to date back even further than the finite case, to 1937 [10]. In the language of Markov chains, this corresponds to determining whether a given stochastic matrix can be embedded into an underlying continuous time Markov chain. Analogously, infinite quantum channel divisibility—also known as the *Markovianity* condition for a CPTP map—asks whether the dynamics of the quantum system can be described by a *Lindblad* master equation [21,12]. The infinite divisibility problems in both the classical and quantum case were recently shown to be NP-hard [8]. Formulated as weak membership problems, these results imply that it is NP-hard to extract dynamics from experimental data [7].

However, while related, it is not at all clear that there exists a reduction of the finite divisibility question to the case of infinite divisibility. In fact, mathematically, the infinite divisibility case is a special case of finite divisibility, as a stochastic matrix is infinitely divisible if and only if it admits an n^{th} root for all $n \in \mathbb{N}$ [19].

The finite divisibility problem for stochastic matrices is still an open question, as are the nonnegative matrix and CPTP map divisibility problems. We will show that the question of existence of stochastic roots of a stochastic matrix is NP-hard. We also extend this result to (doubly) stochastic matrices, nonnegative matrices, and CPTP maps.

We start out by introducing the machinery we will use to prove Theorem 1 and 3 in section 2.2. A reduction from the quantum to the classical case can be found in section 2.4, from the nonnegative to the stochastic case in section 2.5 and the main result—in a mathematically rigorous formulation—is then presented as Theorem 20 in section 2.6.

2.2. Preliminaries

2.2.1. Roots of matrices

In our study of matrix roots we restrict ourselves to the case of square roots. The more general case of p^{th} roots of matrices remains to be discussed. We will refer to square roots simply as roots. To be explicit, we state the following definition.

Definition 8. Let $\mathbf{M} \in \mathbb{K}^{d\times d}$, $d \in \mathbb{N}$, $\mathbb{K}$ some field. Then we say that $\mathbf{R} \in \mathbb{K}^{d\times d}$ is a root of $\mathbf{M}$ if $\mathbf{R}^2 = \mathbf{M}$. We denote the set of all roots of $\mathbf{M}$ with $\sqrt{\mathbf{M}}$.

Following the theory of matrix functions—see for example [15]—we remark that in the case of nonsingular $\mathbf{M}$, $\sqrt{\mathbf{M}}$ is nonempty and can be expressed in Jordan normal form via $\sqrt{\mathbf{M}} = \mathbf{Z}\mathbf{J}\mathbf{Z}^{-1}$ for some invertible $\mathbf{Z}$, where $\mathbf{J} = \operatorname{diag}(\mathbf{J}_1^{\pm}, \ldots, \mathbf{J}_m^{\pm})$. Here $\mathbf{J}_i^{\pm}$ denotes the $\pm$-branch of the root function $f(x) = \sqrt{x}$ of the Jordan block corresponding to the ith eigenvalue λ_i,

$$\mathbf{J}_i^{\pm} = \begin{pmatrix} \pm f(\lambda_i) & \pm f'(\lambda_i)/1! & \ldots & \pm f^{(m_i-1)}(\lambda_i)/(m_i-1)! \\ 0 & \pm f(\lambda_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \pm f'(\lambda_i)/1! \\ 0 & \ldots & 0 & \pm f(\lambda_i) \end{pmatrix} .$$

If $\mathbf{M}$ is diagonalizable, $\mathbf{J}$ simply reduces to the canonical diagonal form $\mathbf{J} = \operatorname{diag}(\pm\sqrt{\lambda_1}, \ldots, \pm\sqrt{\lambda_m})$.

If $\mathbf{M}$ is derogatory—i.e. there exist multiple Jordan blocks sharing the same eigenvalue λ—it has continuous families of so-called *nonprimary* roots $\sqrt{\mathbf{M}} = \mathbf{Z}\mathbf{U}\mathbf{J}\mathbf{U}^{-1}\mathbf{Z}^{-1}$, where $\mathbf{U}$ is an arbitrary nonsingular matrix that commutes with the Jordan normal form $[\mathbf{U}, \mathbf{J}] = 0$.

We cite the following result from [16, Th. 2.6].

Theorem 9 *(Classification of roots). Let $\mathbf{M} \in \mathbb{K}^{d\times d}$ have the Jordan canonical form $\mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^{-1}$, where $\mathbf{\Lambda} = \operatorname{diag}(\mathbf{J}_0, \mathbf{J}_1)$, such that $\mathbf{J}_0$ collects all Jordan blocks corresponding to the eigenvalue 0, and $\mathbf{J}_1$ collects the remaining ones. Assume further that*

$$d_i := \dim(\ker \mathbf{M}^i) - \dim(\ker \mathbf{M}^{i-1})$$

has the property that for all $i \in \mathbb{N}_{\geq 0}$, no more than one element of the sequence satisfies $d_i \in (2i, 2(i+1))$. Then $\sqrt{\mathbf{M}} = \mathbf{Z}\sqrt{\mathbf{\Lambda}}\mathbf{Z}^{-1}$, where $\sqrt{\mathbf{\Lambda}} = \operatorname{diag}(\sqrt{\mathbf{J}_0}, \sqrt{\mathbf{J}_1})$.

For a given matrix, the classification gives the set of all roots. If $\mathbf{M}$ is a real matrix, a similar theorem holds and there exist various numerical algorithms for calculating real square roots, see for example [15].

2.2.2. Roots of stochastic matrices

Remember the following two definitions.

Definition 10. A matrix $\mathbf{M} \in \mathbb{K}^{d\times d}$ is said to be nonnegative if $0 \leq \mathbf{M}_{ij}\ \forall i,j = 0,\ldots,d$.

Definition 11. A matrix $\mathbf{Q} \in \mathbb{K}^{d\times d}$ is said to be stochastic if it is nonnegative and $\sum_{k=1}^{d} \mathbf{Q}_{ik} = 1\ \forall i = 0,\ldots,d$.

In contrast to finding a general root of a matrix, very little is known about the existence of nonnegative roots of nonnegative matrices—or stochastic roots of stochastic matrices—if $d \geq 3$. For stochastic matrices and in the case $d = 2$, a complete characterization can be given explicitly, and for $d \geq 3$, all real stochastic roots that are functions of the original matrix are known, as demonstrated in [14]. Further special classes of matrices for which a definite answer exists can be found in [16]. But even for $d = 3$, the general case is still an open question—see [20, ch. 2.3] for details.

Indeed, a stochastic matrix may have no stochastic root, a primary or nonprimary root—or both. To make things worse, if a matrix has a p^{th} stochastic root, it might or might not have a q^{th} stochastic root if $p \nmid q$—p is not a divisor of q—, $q > p$ or $q \nmid p$, $q < p$.

A related open problem is the inverse eigenspectrum problem, as described in the extensive overview in [9]. While the sets $\Omega_n \subset \mathbb{D}$—denoting all the possible valid eigenvalues of an n-dimensional stochastic matrix—can be given explicitly, and hence also Ω_n^p, almost nothing is known about the sets of valid eigenspectra. Any progress in this area might yield necessary conditions for the existence of stochastic roots.

In recent years, some approaches have been developed to approximate stochastic roots numerically, see the comments in [14, sec. 4]. Unfortunately, most algorithms are highly unstable and do not necessarily converge to a stochastic root. A direct method using nonlinear optimization techniques is difficult and depends heavily on the algorithm employed [20].

It remains an open question whether there exists an efficient algorithm that decides whether a stochastic matrix $\mathbf{Q}$ has a stochastic root.

In this paper, we will prove that this question is NP*-hard to answer.*

2.2.3. The Choi isomorphism

For the results on CPTP maps, we will need the following basic definition and results.

Definition 12. Let $\mathbf{A} : \mathcal{H} \longrightarrow \mathcal{H}$ be a linear map on $\mathcal{H} = \mathbb{C}^{d\times d}$. We say that $\mathbf{A}$ is positive if for all Hermitian and positive definite $\rho \in \mathcal{H}$, $\mathbf{A}\rho$ is Hermitian and positive definite. It is said to be completely positive if $\mathbf{A} \otimes \mathbb{1}_n$ is positive $\forall n \in \mathbb{N}$.

A map $\mathbf{A}$ which is completely positive and trace-preserving—i.e. $\text{tr}(\mathbf{A}\rho) = \text{tr}\ \rho$ $\forall \rho \in \mathcal{H}$—is called a *completely positive trace-preserving* map, or short CPTP map.

In contrast to positivity, complete positivity is easily characterized using the well-known *Choi–Jamiolkowski* isomorphism—cf. [4, Th. 2].

Remark 13. Let the notation be as in Definition 12 and pick a basis $e_1, \ldots, e_d$ of $\mathbb{C}^d$. Then $\mathbf{A}$ is completely positive if and only if the *Choi* matrix

$$\mathrm{C}_{\mathbf{A}} := (\mathbb{1}_d \otimes \mathbf{A})\Omega\Omega^{\mathrm{T}} = \sum_{i,j=1}^{d} e_i e_j^{\mathrm{T}} \otimes \mathbf{A}(e_i e_j^{\mathrm{T}})$$

is positive semidefinite, where $\Omega := \sum_{i=1}^{d} e_i \otimes e_i$.

The condition of trace-preservation then translates to the following.

Remark 14. A map $\mathbf{A}$ is trace-preserving if and only if $\mathrm{tr}_2(\mathrm{C}_{\mathbf{A}}) = \mathbb{1}_d$, where tr_2 denotes the partial trace over the second pair of indices.

2.3. Equivalence of computational questions

In the following we denote with S some arbitrary finite index set, not necessarily the same for all problems. We begin by defining the following decision problems.

Definition 15 *(*CPTP Divisibility*).*

Instance. CPTP map $\mathbf{B} \in \mathbb{Q}^{d\times d}$.
Question. Does there exist a CPTP map $\mathbf{A} : \mathbf{A}^2 = \mathbf{B}$?

Definition 16 *(*CPTP Root*).*

Instance. Family of matrices $(\mathbf{A}_s)_{s\in \mathrm{S}}$ that comprises all the roots of a matrix $\mathbf{B}$.
Question. Does there exist an $s \in S : \mathbf{A}_s$ is a CPTP map?

Definition 17 *(*Stochastic Divisibility*).*

Instance. Stochastic matrix $\mathbf{P} \in \mathbb{Q}^{d\times d}$.
Question. Does there exist a stochastic matrix $\mathbf{Q} : \mathbf{Q}^2 = \mathbf{P}$?

Definition 18 *(*Stochastic Root*).*

Instance. Family of matrices $(\mathbf{Q}_s)_{s\in \mathrm{S}}$ comprising all the roots of a matrix $\mathbf{P}$.
Question. Does there exist an $s \in \mathrm{S} : \mathbf{Q}_s$ stochastic?

Definition 19 *(*Nonnegative Root*).*

Instance. Family of matrices $(\mathbf{M}_s)_{s\in\mathrm{S}}$ comprising all the roots of a matrix $\mathbf{N}$, where all $\mathbf{M}_s$ have at least one positive entry.
Question. Does there exist an $s \in \mathrm{S} : \mathbf{M}_s$ nonnegative?

Theorem 20. *The reductions as shown in Fig. 1 hold.*

Proof. The implication Stochastic Divisibility⟵Stochastic Root needs one intermediate step. If $\mathbf{P}$ is not stochastic, the answer is negative. If it is stochastic, we can apply Stochastic Divisibility. The opposite direction holds for non-derogatory stochastic $\mathbf{P}$: in this case we can enumerate all roots of $\mathbf{P}$ as a finite family which forms a valid instance for Stochastic Root.

The reduction Stochastic Root⟵Nonnegative Root can be resolved by Lemma 25 and Lemma 26—we construct a family of matrices $(\mathbf{Q}_s)_{s\in\mathrm{S}}$ that contains a stochastic root iff $(\mathbf{M}_s)_{s\in\mathrm{S}}$ contains a nonnegative root. The result then follows from applying Stochastic Root. If our stochastic matrix $\mathbf{P}$ is irreducible, then any nonnegative root $\mathbf{Q}_{s'} : \mathbf{Q}_{s'}^2 = \mathbf{P}$ is stochastic, and in that case Stochastic Root⟷Nonnegative Root—see [16, sec. 3] for details.

The link cptp Divisibility⟵cptp Root again needs the following intermediate step. If $\mathbf{A}$ is not cptp, the answer is negative. If it is cptp, then we can apply cptp Divisibility. Similarly, if $\mathbf{A}$ is non-derogatory, the reduction works in the opposite direction as well.

The direction cptp Root⟵Stochastic Root follows from Corollary 24. We start out with a family $(\mathbf{Q}_s)_{s\in S}$ comprising all the roots of a stochastic matrix $\mathbf{P}$. Then let $(\mathbf{A}_s := \mathrm{emb}\ \mathbf{Q}_s)_{s\in S}$—this family then comprises all of the roots of $\mathbf{B} := \mathbf{A}_k^2 \equiv \mathbf{A}_s^2$ $\forall k, s$. Furthermore, by Lemma 23, there exists a cptp $\mathbf{A}_s$ if and only if there exists a stochastic $\mathbf{Q}_s$, and the reduction follows.

Finally, we can extend our reduction to the programs Doubly Stochastic Root and Doubly Stochastic Divisibility as well as Nonnegative Divisibility, defined analogously, see our comment in Corollary 27 and the complete reduction tree in Fig. 1. □

At this point, we observe the following fact.

Lemma 21. *All the above* Divisibility *and* Root *problems in Definition 15 to 19 are contained in* NP*.*

Proof. It is straightforward to come up with a witness and a verifier circuit that satisfies the definition of the decision class NP. For example in the cptp case, a witness is a matrix root that can be checked to be a cptp map using Remark 13 and squared in polynomial time, which is the verifier circuit. Both circuit and witness are clearly poly-sized and hence the claim follows. □

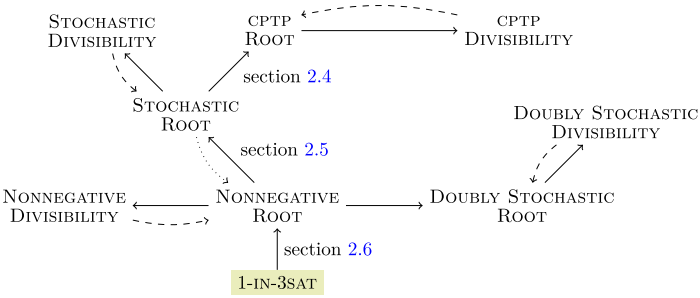


Fig. 1. Complete chain of reduction for our programs. The dashed line between the Divisibility and Root problems hold for non-derogatory matrices, respectively. The dotted line between Stochastic Root and Nonnegative Root holds only for irreducible matrices. The doubly stochastic and nonnegative branch are included for completeness but not described in detail here—see Corollary 27.

By encoding an instance of 1-in-3sat into a family of nonnegative matrices $(\mathbf{M}_s)_{s\in\mathrm{S}}$, we show the implication 1-in-3sat$\longrightarrow$Nonnegative Root and 1-in-3sat$\longrightarrow$(Doubly) Stochastic/cptp Divisibility, accordingly, from which NP-hardness of (Doubly) Stochastic/cptp Divisibility follows. The entire chain of reduction can be seen in Fig. 1.

2.4. Reduction of Stochastic Root *to* CPTP Root

This reduction is based on the following embedding.

Definition 22. Let $\{e_i\}$ be an orthonormal basis of $\mathbb{K}^d$. The embedding emb is defined as

$$\mathrm{emb} : \mathbb{K}^{d\times d} \hookrightarrow \mathbb{K}^{d^2\times d^2} \quad ,$$

$$\mathbf{A} \longmapsto \mathbf{B} := \sum_{i,j=1}^{d} \mathbf{A}_{ij}(e_i\otimes e_i)(e_j\otimes e_j)^{\mathrm{T}} = \sum_{i,j=1}^{d} \mathbf{A}_{ij}(e_ie_j^{\mathrm{T}})\otimes(e_ie_j^{\mathrm{T}}) \quad .$$

We observe the following.

Lemma 23. *We use the same notation as in Remark 13. Let* $\mathbf{A}\in\mathbb{K}^{d\times d}$ *and* $\mathbf{B} := \mathrm{emb}\,\mathbf{A}$. *Then* $\mathbf{A}$ *is positive (nonnegative) if and only if the Choi matrix* $\mathrm{C}_{\mathbf{B}}$ *is positive (semi-) definite. Furthermore, the row sums of* $\mathbf{A}$ *are* 1*—i.e.* $\sum_{j=1}^{d}\mathbf{A}_{ij} = 1\ \forall j = 1,\dots,d$*—if and only if* $\mathrm{tr}_2(\mathrm{C}_{\mathbf{B}}) = \mathbb{1}_d$. *In addition, the spectrum of* $\mathbf{B}$ *satisfies* $\sigma(\mathbf{B})\subseteq\sigma(\mathbf{A})\cup\{0\}$.

Proof. The first claim follows directly from the matrix representation of our operators. There, the *Choi* isomorphism is manifest as the reshuffling operation or partial transpose

$$\cdot^{\Gamma} : \mathbb{K}^{d^2\times d^2} \longrightarrow \mathbb{K}^{d^2\times d^2}, \left[(e_ie_j^{\mathrm{T}})\otimes(e_ie_j^{\mathrm{T}})\right]^{\Gamma} \longmapsto (e_ie_i^{\mathrm{T}})\otimes(e_je_j^{\mathrm{T}}) \quad .$$

For more details, see e.g. [2].

The second statement follows from

$$\begin{aligned}\mathrm{tr}_2(\mathrm{C}_{\mathbf{B}}) &= \mathrm{tr}_2\left(\sum_{i,j=1}^{d} \mathbf{A}_{ij}(e_i e_j^{\mathrm{T}}) \otimes (e_i e_j^{\mathrm{T}})\right) \\ &= \sum_{i,j=1}^{d} \mathbf{A}_{ij} e_i e_i^{\mathrm{T}} = \mathrm{diag}\left(\sum_{j=1}^{d} \mathbf{A}_{1j}, \ldots, \sum_{j=1}^{d} \mathbf{A}_{dj}\right) \quad .\end{aligned}$$

The final claim is trivial. □

This remark immediately yields the following consequence.

Corollary 24. *For a family of stochastic matrices $(\mathbf{Q}_s)_{s\in S}$ parametrized by the index set S, there exists a family of square matrices $(\mathbf{A}_s)_{s\in S} := (\mathrm{emb}\ \mathbf{Q}_s)_{s\in \mathrm{S}}$, such that $(\mathbf{Q}_s)_{s\in \mathrm{S}}$ contains a stochastic matrix if and only if $(\mathbf{A}_s)_{s\in \mathrm{S}}$ contains a* CPTP *matrix.*

2.5. Reduction of Nonnegative Root *to* Stochastic Root

The difference between Nonnegative Root and Stochastic Root is the extra normalization condition in the latter, see Definition 11. The following two lemmas show that this normalization does not pose an issue, so we can efficiently reduce the problem Nonnegative Root to Stochastic Root.

Lemma 25. *For a family of square matrices $(\mathbf{M}_s)_{s\in \mathrm{S}}$ parametrized by the index set* S*, all of which with at least one positive entry, there exists a family of square matrices $(\mathbf{Q}_s)_{s\in \mathrm{S}}$ such that $(\mathbf{M}_s)_{s\in \mathrm{S}}$ contains a nonnegative matrix if and only if $(\mathbf{Q}_s)_{s\in \mathrm{S}}$ contains a stochastic matrix and such that* $\mathrm{rank}\,\mathbf{Q}_s = \mathrm{rank}\,\mathbf{M}_s + 2\ \forall s \in \mathrm{S}$*. Furthermore, $(\mathbf{Q}_s)_{s\in \mathrm{S}}$ can be constructed efficiently from $(\mathbf{M}_s)_{s\in \mathrm{S}}$.*

Proof. We explicitly construct our family $(\mathbf{Q}_s)_{s\in \mathrm{S}}$ as follows. Pick an $s \in \mathrm{S}$ and denote $\mathbf{M} := \mathbf{M}_s$. Let d be the dimension of $\mathbf{M}$. We first pick $a \in \mathbb{R}^+$ such that $a \max_{ij} \mathbf{M}_{ij} = 1/2$[1] and define

$$\begin{aligned}\mathbf{Q}_s &:= \frac{1}{1764d}\begin{pmatrix} 1764a\mathbf{M} + 637 & 735 - 1260a\mathbf{M} & 392 - 504a\mathbf{M} \\ 735 - 1260a\mathbf{M} & 900a\mathbf{M} + 1029 & 360a\mathbf{M} \\ 392 - 504a\mathbf{M} & 360a\mathbf{M} & 144a\mathbf{M} + 1372 \end{pmatrix} \\ &\equiv \frac{a}{d} AA^{\mathrm{T}} \otimes \mathbf{M} + \frac{1}{d}\left(BB^{\mathrm{T}} + CC^{\mathrm{T}}\right) \otimes \mathbf{1} \quad ,\end{aligned}$$

where by sum of matrix $\mathbf{M}$ and scalar x we mean $\mathbf{M} + x\mathbf{1}$, $\mathbf{1} := (1)_{1\le i,j\le d} \in \mathbb{R}^{d\times d}$, and

[1] The exact bound is $a \max_{ij} \mathbf{M}_{ij} \le 43/81$.

$$A := \left(1, -\frac{5}{7}, -\frac{2}{7}\right)^{\mathrm{T}} \quad , \quad B := \left(\frac{1}{6}, \frac{1}{2}, -\frac{2}{3}\right)^{\mathrm{T}} \quad , \quad C := -\frac{1}{\sqrt{3}}(1,1,1)^{\mathrm{T}} \quad .$$

Observe that $\{A, B, C\}$ form an orthogonal set—if one wishes, normalizing and pulling out the constant as eigenvalue to the corresponding eigenprojectors would work equally well.

By construction, $\mathbf{Q}_s$ is nonnegative if and only if $\mathbf{M}_s$ is. Since the row sums of $\mathbf{Q}_s$ are always 1, $\mathbf{Q}_s$ is stochastic if and only if $\mathbf{M}_s$ is nonnegative, and the claim follows. □

Lemma 26. *Let the notation be as in Lemma 25 and write $\sqrt{\mathbf{N}}$ for the set of roots of $\mathbf{N}$, see Definition 8. Assume $(\mathbf{M}_s)_{s\in\mathrm{S}} = \sqrt{\mathbf{N}}$ for some $\mathbf{N} \in \mathbb{C}^{d\times d}$. Then there exists a $\mathbf{P} \in \mathbb{C}^{d\times d}$, such that $\mathbf{Q}_s^2 = \mathbf{P}\ \forall s \in S$ and $(\mathbf{Q}_s)_{s\in\mathrm{S}} \subset \sqrt{\mathbf{P}}$. Furthermore, the complement of $(\mathbf{Q}_s)_{s\in\mathrm{S}}$ in $\sqrt{\mathbf{P}}$ does not contain any stochastic roots.*

Proof. The first statement is obvious, since for all $s \in \mathrm{S}$,

$$\mathbf{Q}_s^2 = \frac{a^2}{d^2}\frac{78}{49} AA^{\mathrm{T}} \otimes \mathbf{M}_s^2 + \frac{1}{d}\left(\frac{13}{18} BB^{\mathrm{T}} + CC^{\mathrm{T}}\right) \otimes \mathbf{1} =: \mathbf{P} \quad ,$$

and hence clearly $(\mathbf{Q}_s)_{s\in\mathrm{S}} \subset \sqrt{\mathbf{P}}$.

The last statement is not quite as straightforward—it is the main reason our carefully crafted matrix $\mathbf{Q}_s$ has its slightly unusual shape. All possible roots of $\mathbf{P}$ are of the form

$$\sqrt{\mathbf{P}} = \frac{a}{d} AA^{\mathrm{T}} \otimes \sqrt{\mathbf{N}} \pm \frac{1}{d}\left(BB^{\mathrm{T}} \pm CC^{\mathrm{T}}\right) \otimes \mathbf{1} \quad .$$

It is easy to check that none of the other sign choices yields any stochastic matrix, so the claim follows.[2] □

Corollary 27. *The results of Lemma 25 and 26 also hold for doubly stochastic matrices—observe that our construction of $\mathbf{Q}_s$ is already doubly stochastic.*

2.6. Reduction of 1-IN-3SAT to NONNEGATIVE ROOT

We now embed an instance of a boolean satisfiability problem, 1-IN-3SAT—see Definition 87 for details—into a family of matrices $(\mathbf{M}_s)_{s\in\mathrm{S}}$ in a way that there exists an s such that $\mathbf{M}_s$ is nonnegative if and only if the instance of 1-IN-3SAT is satisfiable. The construction is inspired by [8].

We identify

$$\text{TRUE} \longleftrightarrow 1 \quad , \quad \text{FALSE} \longleftrightarrow -1 \quad . \tag{1}$$

[2] The reader may try to find a simpler matrix that does the trick.

Denote with $(m_{i1}, m_{i2}, m_{i3}) \in \{\pm 1\}^3$ the three boolean variables occurring in the i^{th} boolean clause, and let $m_i \in \{\pm 1\}$ stand for the single i^{th} boolean variable. Then 1-IN-3SAT translates to the inequalities

$$-\frac{3}{2} \leq m_{i1} + m_{i2} + m_{i3} \leq -\frac{1}{2} \quad \forall i = 1, \ldots, n_c \quad . \tag{2}$$

Theorem 28. *Let (n_v, n_c, m_i, m_{ij}) be a 1-IN-3SAT instance. Then there exists a family of matrices $(\mathbf{M}_s)_{s\in \mathrm{S}}$ such that $\exists s : \mathbf{M}_s$ nonnegative iff the instance is satisfiable.*

To prove this, we first need the following technical lemma.

Lemma 29. *Let (n_v, n_c, m_i, m_{ij}) be a 1-IN-3SAT instance. Then there exists a family of matrices $(\mathbf{C}_s)_{s\in \mathrm{S}}$ such that $\exists s$: the first n_c on-diagonal 4×4 blocks of $\mathbf{C}_s$ are nonnegative iff the instance is satisfiable. In addition, we have $\mathbf{C}_s^2 = \mathbf{C}_t^2\ \forall s, t$. Furthermore, $(\mathbf{C}_s)_{s\in \mathrm{S}} \subset \sqrt{\mathbf{C}_s^2}$, and the complement contains no nonnegative root.*

Proof. For every boolean variable m_k, define a vector $v_k \in \mathbb{R}^d$ such that their first n_c elements are defined via

$$(v_k)_i := \begin{cases} 1 & m_k \text{ occurs in } i^{\text{th}} \text{ clause} \\ 0 & \text{otherwise} \end{cases} .$$

We will specify the dimension d later—obviously $d \geq n_c$, and the free entries are used to orthonormalize all vectors in the end. For now, we denote the orthonormalization region with $\vec{o}$. We further define the vectors $c_1, c_2 \in \mathbb{R}^d$ to have all 1s in the first n_c entries, i.e. $c_{1,2} = (1, \ldots, 1, \vec{o}_{1,2})$. Let then

$$\mathbf{C}'_s := c_1 c_1^{\mathrm{T}} \otimes \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} + \frac{1}{2} c_2 c_2^{\mathrm{T}} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sum_{k=1}^{n_v} p_k v_k v_k^{\mathrm{T}} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad . \tag{3}$$

The variables p_k denote a specific rescaled choice of the boolean variables m_i, which—in order to avoid degeneracy—have to be distinct, i.e. via

$$p_i = \left(1 - \frac{1}{N} - \frac{i}{N n_v}\right) m_i \quad \forall i = 1, \ldots, n_v \quad . \tag{4}$$

The p_{ij} are defined accordingly from the m_{ij} and $N \in \mathbb{N}$ is large but fixed.

Let further

$$\mathbf{C}_s := \begin{pmatrix} \mathbf{C}'_s & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}^{d\times d} \quad ,$$

where we have used an obvious block notation to pad $\mathbf{C}'_s$ with zeroes, which will come into play later.

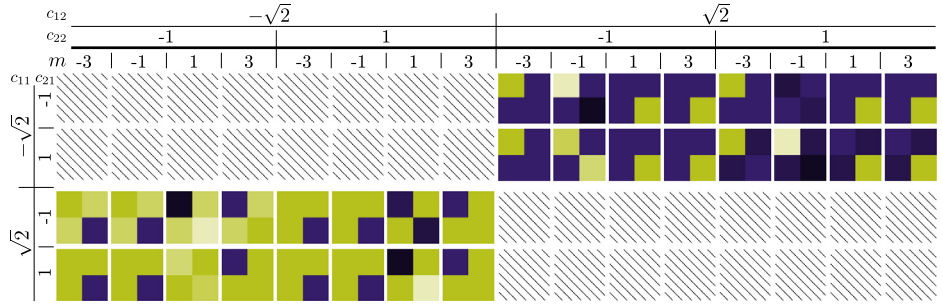


Fig. 2. $\mathbf{C}'_s$ for various sign choices of the eigenvalues c_{ij}, $i,j = 1,2$ corresponding to the eigenvectors $c_{1,2}$. Only all positive signs and $m := \sum_j m_{ij} = -1$ yields a nonnegative block (third from right in top row). Hatching signifies complex numbers, the colour scale is the same as in Fig. 3, i.e. light green denotes negative numbers, dark purple nonnegative entries.

The on-diagonal 2×2 blocks of $\mathbf{C}_s$ then encode the 1-IN-3SAT inequalities from equation (2)—demanding nonnegativity—as the set of equations

$$\frac{3}{2} + p_{i1} + p_{i2} + p_{i3} \geq 0 \quad \text{and} \quad -\frac{1}{2} - p_{i1} - p_{i2} - p_{i3} \geq 0 \ .$$

Note that we leave enough head space such that the rescaling in equation (4) does not affect any of the inequalities—see section 2.8 for details.

Observe further that the eigenvalues corresponding to each eigenprojector in the last term of equation (3) necessarily have opposite sign, otherwise we create complex entries. We will later rescale $\mathbf{C}_s$ by a positive factor, under which the inequalities are invariant, so the first claim follows.

We can always orthonormalize the vectors $c_{1,2}$ and v_k using the freedom left in $\vec{o}$, hence we can achieve that $\mathbf{C}_s^2 = \mathbf{C}_t^2 \ \forall s,t$. It is straightforward to check that no other sign choice for the eigenvalues of the first two terms yields nonnegative blocks—see Fig. 2 for details. From this, the last two claims follow. □

2.7. Orthonormalization and handling the unwanted inequalities

As in [8], we have unwanted inequalities—the off-diagonal blocks in the first $4n_c$ entries and the blocks involving the orthonormalization region $\vec{o}$. We first deal with the off-diagonal blocks in favour of enlarging the orthonormalization region, creating more—potentially negative—entries in there, and then fix the latter.

Off-diagonal blocks. We begin with the following lemma.

Lemma 30. *Let the family* $(\mathbf{C}_s)_{s\in\mathrm{S}}$ *be defined as in the proof of Lemma 29, and* (n_v, n_c, m_i, m_{ij}) *the corresponding 1-*IN-*3*SAT *instance. Then there exists a matrix* $\mathbf{E} \in \mathbb{C}^{d\times d}$ *such that the top left* $4n_c \times 4n_c$ *block of* $\mathbf{C}_s + \mathbf{E}$ *has at least one negative entry* $\forall s$ *iff the instance is not satisfiable. Furthermore,* $\mathrm{im}\mathbf{C}_s \perp \mathrm{im}\mathbf{E}\ \forall s$*, and* $\mathbf{C}_s + \mathbf{E}'$ *has negative entries* $\forall s$*,* $\forall \mathbf{E}' \in \sqrt{\mathbf{E}^2} \setminus \{\mathbf{E}\}$.

Proof. Define

$$\mathbf{E}_1 := E_1 E_1^{\mathrm{T}} \otimes \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \text{where} \quad E_1 := (1, \ldots, 1, \vec{o})^{\mathrm{T}} \quad .$$

Then $\mathbf{E}_1$ has rank 1.

From this mask, we now erase the first n_c on-diagonal 4×4-blocks, while leaving all other entries in the upper left $4n_c \times 4n_c$ block positive. Define $b_i := (e_i, \vec{o}) \in \mathbb{C}^d$ for $i = 1, \ldots, n_c$ where e_i denotes the i^{th} unit vector, and let

$$\mathbf{E} := \frac{7}{2}\mathbf{E}_1 - \frac{7}{2} \sum_{i=1}^{n_c} t_i b_i b_i^{\mathrm{T}} \otimes \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad .$$

The variables t_i are chosen close to 1 but distinct, e.g.

$$t_i := \left(1 - \frac{1}{M} - \frac{i}{M n_c}\right) \quad , \tag{5}$$

where $M \in \mathbb{N}$ large but fixed. Then $\mathbf{E}$ has rank $n_c + 1$, and adding $\mathbf{E}$ to $\mathbf{C}_s$ trivializes all unwanted inequalities in the upper left $4n_c \times 4n_c$ block. By picking M large enough, the on-diagonal inequalities are left intact.

One can check that all other possible sign choices for the roots of $\mathbf{E}$ create negative entries in parts of the upper left block where $\mathbf{C}_s$ is zero $\forall s$. Furthermore, $\mathbf{C}_s$ and $\mathbf{E}$ have distinct nonzero eigenvalues by construction—the orthogonality condition is again straightforward, hence the last two claims follow. □

Orthonormalization region.

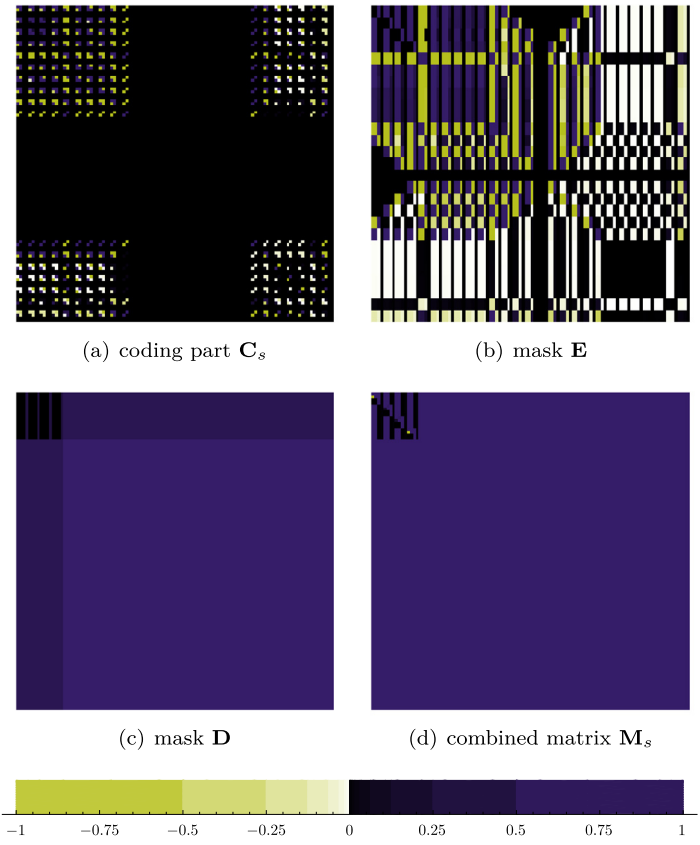
Lemma 31. *Let $4n < d$ and $\delta \gg 1$. There exists a nonnegative rank 2 matrix $\mathbf{D} \in \mathbb{C}^{d \times d}$ such that the top left $4n \times 4n$ block of $\mathbf{D}$ has entries $\mathbf{D}_{ij} \in \mathcal{O}(\delta^{-2})$ if $j \nmid 4$ and the rest of the matrix entries are $\Omega(\delta^{-1})$. If $\mathbf{D}' \in \sqrt{\mathbf{D}^2}$, either the same holds true for $\mathbf{D}'$, or $\mathbf{D}'_{ij} < 0$ $\forall j < 4n + 1, j \mid 4$.*

Proof. Define

$$E_2 := \Big(\underbrace{\frac{1}{\delta}, \ldots, \frac{1}{\delta}}_{n \text{ times}}, 1, \ldots, 1 \Big) \in \mathbb{C}^d$$

and let $\mathbf{E}_2 := E_2 E_2^{\mathrm{T}} \otimes \mathbf{1}_4$, where $\mathbf{1}_4 := (1)_{1 \leq i,j \leq 4}$. Let further

$$\Delta := \Big(\underbrace{\frac{1}{\delta}, \ldots, \frac{1}{\delta}}_{n \text{ times}}, -\frac{1}{\delta}, \ldots, -\frac{1}{\delta}, a \Big) \in \mathbb{C}^d \quad ,$$



(a) coding part $\mathbf{C}_s$

(b) mask $\mathbf{E}$

(c) mask $\mathbf{D}$

(d) combined matrix $\mathbf{M}_s$

(e) color scale, clamped

$$\mathrm{R}(m_1, m_2, m_4) \wedge \mathrm{R}(m_2, m_3, m_4) \wedge \mathrm{R}(m_1, m_3, m_4) \wedge \mathrm{R}(m_1, m_2, m_3)$$

(f) 1-IN-3SAT instance

Fig. 3. One branch of an unsatisfiable instance of 1-IN-3SAT encoded into a matrix of total rank 19. The negative entries—two bright dots—in the upper left block in the combined matrix (d) indicate that this branch does not satisfy the given instance. By looking at all other blocks, one sees that none translates to a nonnegative matrix. Observe that in this naïve implementation the orthonormalization region is suboptimally large.

where $0 < a < 1$ is used to orthonormalize Δ and E_2, which is the case if

$$a = -\frac{n}{\delta^2} + \frac{d - n - 1}{\delta} \quad .$$

By explicitly writing out the rank 2 matrix

$$\mathbf{D} := \mathbf{E}_2 \pm \Delta\Delta^{\mathrm{T}} \otimes \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad ,$$

it is straightforward to check that $\mathbf{D}$ fulfils all the claims of the lemma—see Fig. 3 for an example. □

2.8. Lifting singularities

The reader will have noted by now that even though we have orthonormalized all our eigenspaces, ensuring that the nonzero eigenvalues are all distinct, we have at the same time introduced a high-dimensional kernel in $\mathbf{C}_s$, $\mathbf{E}$ and $\mathbf{D}$. The following lemma shows that this does not pose an issue.

Lemma 32. *Let $(\mathbf{A}_s)_{s\in\mathrm{S}}$ be the family of primary rational roots of some degenerate $\mathbf{B} \in \mathbb{Q}^{d\times d}$. Then there exists a non-degenerate matrix $\mathbf{B}'$, such that for the family $(\mathbf{A}'_s)_{s\in\mathrm{S}}$ of roots of $\mathbf{B}'$, we have $\mathbf{A}_s$ positive iff $\mathbf{A}'_s$ positive. Furthermore, the entries of $\mathbf{A}'_s$ are rational with bit complexity $\mathfrak{r}(\mathbf{A}'_s) = \mathcal{O}(\mathrm{poly}(\mathfrak{r}(\mathbf{A}_s)))$.*

Proof. Take a matrix $\mathbf{A} \in (\mathbf{A}_s)_{s\in\mathrm{S}}$. We need to distort the zero eigenvalues $\{\lambda_i^{(0)}\}$ slightly away from 0. Using notation from Definition 36, a conservative estimate for the required smallness without affecting positivity would be

$$\lambda_i^{(0)} \longmapsto \lambda_i'^{(0)} \quad : \quad 0 < \lambda_i'^{(0)} \leq \frac{1}{C \cdot d^3 \cdot \max_{ij}\{|\mathbf{Z}_{ij}|, |\mathbf{Z}^{-1}_{ij}|\}} \quad ,$$

where we used the Jordan canonical form $\mathbf{A} = \mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^{-1}$ for some invertible $\mathbf{Z}$ and $\mathbf{\Lambda} = \mathrm{diag}(\mathbf{J}_0, \mathbf{J}_1)$, such that $\mathbf{J}_0$ collects all Jordan blocks corresponding to the eigenvalue 0, and $\mathbf{J}_1$ collects the remaining ones. □

This will lift all remaining degeneracies and singularities, without affecting our line of argument above. Observe that all inequalities in our construction were bounded away from 0 with enough head space independent of the problem size, so positivity in the lemma is sufficient.

We thus constructed an embedding of 1-IN-3SAT into non-derogatory and non-degenerate matrices, as desired. It is crucial to note that we do not lose anything by restricting the proof to the study of these matrices, as the following lemma shows.

Lemma 33. *There exists a Karp reduction of the* DIVISIBILITY *problems when defined for* all *matrices to the case of* non-degenerate and non-derogatory *matrices.*

Proof. As shown in Lemma 21, containment in NP for this problem is easy to see, also in the degenerate or derogatory case. Since 1-IN-3SAT is NP-complete, there has to exist a poly-time reduction of the DIVISIBILITY problems—when defined for *all* matrices—to 1-IN-3SAT. Now embed this 1-IN-3SAT-instance with our construction. This yields a poly-time reduction to the non-degenerate non-derogatory case. □

2.9. Complete embedding

We now finally come to the proof of Theorem 28.

Proof of Theorem 28. Construct the family $(\mathbf{C}_s+\mathbf{E})_{s\in\mathrm{S}}$ using Lemma 29 and Lemma 30, ensuring that all orthonormalizing is done, which preliminarily fixes the dimension d. By Lemma 31, we now construct a mask $\mathbf{D}(\delta)$ of dimension $d+d'$, where $d'>0$ is picked such that we can also orthonormalize all previous vectors with respect to E_2 and δ.

By Lemmas 29, 30, 31 and 32, the perturbed family $(\mathbf{M}'_s)_{s\in\mathrm{S}} := (\mathbf{C}_s+\mathbf{E}+N\mathbf{D}(\delta))'_s$—where N and $\delta\in\mathbb{Q}$ are chosen big enough so that all unwanted inequalities are trivially satisfied—fulfils the claims of the theorem and the proof follows. □

We finalize the construction as follows. In Theorem 28, we have embedded a given 1-IN-3SAT instance into a family of matrices $(\mathbf{M}_s)_{s\in\mathrm{S}}$, such that the instance is satisfiable if and only if at least one of those matrices is nonnegative.

By rescaling the entire matrix such that $\max_{ij}(\mathbf{M}_s)_{ij}=1/2$, we could show that this instance of 1-IN-3SAT is satisfiable if and only if the normalized matrix family $(\mathbf{Q}_s)_{s\in\mathrm{S}}$, which we construct explicitly, contains a stochastic matrix.

As shown in section 2.3, this can clearly be answered by STOCHASTIC DIVISIBILITY, as the family $(\mathbf{Q}_s)_{s\in\mathrm{S}}$ comprises all the roots of a unique matrix $\mathbf{P}$. If this matrix is *not* stochastic, our instance of 1-IN-3SAT is trivially not satisfiable. If the matrix *is* stochastic, we ask STOCHASTIC DIVISIBILITY for an answer—a positive outcome signifies satisfiability, a negative one non-satisfiability.

2.10. Bit complexity of embedding

To show that our results holds for only polynomially growing bit complexity, observe the following proposition.

Proposition 34. *The bit complexity* $\mathfrak{r}(\mathbf{M}_s)$ *of the constructed embedding of a 1-*IN-*3*SAT *instance* (n_v,n_c,m_i,m_{ij}) *equals* $\mathcal{O}(\mathrm{poly}(n_v,n_c))$.

Proof. We can ignore any construction that multiplies by a constant prefactor, for example Lemma 25 and Lemma 26. The renormalization for Lemma 25 to $\max_{ij}\mathbf{M}_{s,ij}=1/2$ does not affect $\mathfrak{r}$ either.

The rescaling in equation (4) and equation (5) yields a complexity of $\mathcal{O}(\log n_v)$, and the same thus holds true for Lemma 29 and Lemma 30.

The only other place of concern is the orthonormalization region. Let us write a_i for all vectors that need orthonormalization. In the n^{th} step, we need to make up for $\mathcal{O}(n)$ entries with our orthonormalization, using the same amount of precision to solve the linear equations $(a_i^{\mathrm{T}}a_n=0)_{1\leq i<n}$. This has to be done with a variant of the standard *Gauss* algorithm, e.g. the *Bareiss* algorithm—see for example [1]—which has nonexponential bit complexity.

Together with the lifting of our singularities, which has polynomial precision, we obtain $\mathfrak{r}(\mathbf{M}_s)=\mathcal{O}(\mathrm{poly}(n_v,n_c))$. Completing the embedding in section 2.9 changes the bit complexity by at most another polynomial factor, and hence the claim follows. □

3. Distribution divisibility

3.1. Introduction

Underlying stochastic and quantum channel divisibility, and—to some extent—a more fundamental topic, is the question of divisibility and decomposability of probability distributions and random variables. An illustrative example is the distribution of the sum of two rolls of a standard six-sided die, in contrast to the single roll of a twelve-sided die. Whereas in the first case the resulting random variable is obviously the sum of two uniformly distributed random variables on the numbers $\{1, \ldots, 6\}$, there is no way to achieve the outcome of the twelve-sided die as any sum of nontrivial "smaller" dice—in fact, there is no way of dividing *any* uniformly distributed discrete random variable into the sum of non-constant random variables. In contrast, a uniform continuous distribution can always be decomposed[3] into two *different* distributions.

To be more precise, a random variable X is said to be *divisible* if it can be written as $X = Y + Z$, where Y and Z are non-constant independent random variables that are identically distributed (iid). Analogously, *infinite divisibility* refers to the case where X can be written as an infinite sum of such iid random variables.

If we relax the condition $Y \overset{d}{=} Z$—i.e. we allow Y and Z to have different distributions—we obtain the much weaker notion of *decomposability*. This includes using other sources of randomness, not necessarily uniformly distributed.

Both divisibility and decomposability have been studied extensively in various branches of probability theory and statistics. Early examples include *Cramer's theorem* [6], proven in 1936, a result stating that a Gaussian random variable can only be decomposed into random variables which are also normally distributed. A related result on χ^2 distributions by *Cochran* [5], dating back to 1934, has important implications for the analysis of covariance.

An early overview over divisibility of distributions is given in [28]. Important applications of n-divisibility—the divisibility into n iid terms—is in modelling, for example of bug populations in entomology [18], or in financial aspects of various insurance models [30,29]. Both examples study the overall distribution and ask if it is compatible with an underlying subdivision into smaller random events. The authors also give various conditions on distributions to be infinitely divisible, and list numerous infinitely divisible distributions.

Important examples for infinite divisibility include the *Gaussian*, *Laplace*, *Gamma* and *Cauchy* distributions, and in general all normal distributions. It is clear that those distributions are also finitely divisible, and decomposable. Examples of indecomposable distributions are *Bernoulli* and discrete uniform distributions.

[3] All continuous uniform distributions decompose into the sum of a discrete *Bernoulli* distribution and another continuous uniform distribution. This decomposition is never unique.

However, there does not yet exist a straightforward way of checking whether a given discrete distribution is divisible or decomposable. We will show in this work that the question of decomposability is NP-hard, whereas divisibility is in P. In the latter case, we outline a computationally efficient algorithm for solving the divisibility question. We extend our results to weak-membership formulations (where the solution is only required to within an error ϵ in total variation distance), and argue that the continuous case is computationally trivial as the indecomposable distributions form a dense subset.

We start out in section 3.2 by introducing general notation and a rigorous formulation of divisibility and decomposability as computational problems. The foundation of all our distribution results is by showing equivalence to polynomial factorization, proven in section 3.3. This will allow us to prove our main divisibility and decomposability results in section 3.4 and 3.5, respectively.

3.2. Preliminaries

3.2.1. Discrete distributions

In our discussion of distribution divisibility and decomposability, we will use the standard notation and language as described in the following definition.

Definition 35. Let $(\Omega, \mathcal{F}, \mathrm{p})$ be a discrete *probability space*, i.e. Ω is at most countably infinite and the *probability mass function* $\mathrm{p} : \Omega \longrightarrow [0,1]$—or *pmf*, for short—fulfils $\sum_{x\in\Omega} \mathrm{p}(x) = 1$. We take the σ-*algebra* $\mathcal{F}$ to be maximal, i.e. $\mathcal{F} = 2^{\Omega}$, and without loss of generality assume that the *state space* $\Omega = \mathbb{N}$. Denote the *distribution* described by p with $\mathfrak{D}$. A *random variable* $X : \Omega \longrightarrow \mathrm{B}$ is a measurable function from the sample space to some set B, where usually $\mathrm{B} = \mathbb{R}$.

For the sake of completeness, we repeat the following well-known definition of characteristic functions.

Definition 36. Let $\mathfrak{D}$ be a discrete probability distribution with pmf p, and $X \sim \mathfrak{D}$. Then

$$\phi_X(\omega) := \mathbb{E}(\mathrm{e}^{\mathrm{i}\omega X}) = \int_{\Omega} \mathrm{e}^{\mathrm{i}\omega x} \mathrm{d}F_X(x) = \sum_{x\in\Omega} \mathrm{p}(x)\mathrm{e}^{\mathrm{i}\omega x}$$

defines the *characteristic function* of $\mathfrak{D}$.

It is well-known that two random variables with the same characteristic function have the same cumulative density function.

Definition 37. Let the notation be as in Definition 35. Then the distribution $\mathfrak{D}$ is called *finite* if $\mathrm{p}(k) = 0\ \forall k \geq N$ for some $N \in \mathbb{N}$.

Remark 38. Let $\mathfrak{D}$ be a discrete probability distribution with pmf p. We will—without loss of generality—assume that $\mathrm{p}(0) \neq 0$ and $\mathrm{p}(k) = 0\ \forall k < 0$ for the pmf p of a finite distribution. It is a straightforward shift of the origin that achieves this.

3.2.2. Continuous distributions

Definition 39. Let $(\mathcal{X}, \mathcal{A})$ be a *measurable space*, where $\mathcal{A}$ is the σ-algebra of $\mathcal{X}$. The probability distribution of a *random variable* X on $(\mathcal{X}, \mathcal{A})$ is the *Radon–Nikodym derivative* f, which is a measurable function with $\mathbb{P}(X \in \mathrm{A}) = \int_{\mathrm{A}} f \mathrm{d}\mu$, where μ is a reference measure on $(\mathcal{X}, \mathcal{A})$.

Observe that this definition is more general than Definition 35, where the reference measure is simply the counting measure over the discrete sample space Ω. Since we are only interested in real-valued univariate continuous random variables, observe the following important remark.

Remark 40. We restrict ourselves to the case of $\mathcal{X} = \mathbb{R}$ with $\mathcal{A}$ the *Borel* sets as measurable subsets and the *Lebesgue* measure μ. In particular, we only regard distributions with a *probability density function* f—or *pdf*, for short—i.e. we require the *cumulative distribution function* $\mathrm{P}(x) := \mathbb{P}(X \leq x) \equiv \int_{y \leq x} f(y) \mathrm{d}y$ to be absolutely continuous.

Corollary 41. *The cumulative distribution function* P *of a continuous random variable* X *is almost everywhere differentiable, and any piecewise continuous function* f *with* $\int_{\mathbb{R}} f(x) \mathrm{d}x = 1$ *defines a valid continuous distribution.*

3.2.3. Divisibility and decomposability of distributions

To make the terms mentioned in the introduction rigorous, note the two following definitions.

Definition 42. Let X be a random variable. It is said to be *n-decomposable* if $X = Z_1 + \ldots + Z_n$ for some $n \in \mathbb{N}$, where $Z_1, \ldots, Z_n$ are independent non-constant random variables. X is said to be *indecomposable* if it is not decomposable.

Definition 43. Let X be a random variable. It is said to be *n-divisible* if it is n-decomposable as $X = \sum_{i=1}^{n} Z_i$ and $Z_i \overset{d}{=} Z_j\ \forall i, j$. X is said to be *infinitely* divisible if $X = \sum_{i=1}^{\infty} Z_i$, with $Z_i \sim \mathfrak{D}$ for some nontrivial distribution $\mathfrak{D}$.

If we are not interested in the exact number of terms, we also simply speak of *decomposable* and *divisible*. We will show in section 3.5.7 that—in contrast to divisibility—the question of decomposability into more than two terms is not well-motivated.

Observe the following extension of Remark 38.

Lemma 44. *Let $\mathfrak{D}$ be a discrete probability distribution with pmf* p. *If* p *obeys Remark 38, then we can assume that its factors do as well. In the continuous case, we can without loss of generality assume the same.*

Proof. Obvious from positivity of convolutions in case of divisibility. For decomposability, we can achieve this by shifting the terms symmetrically. □

3.2.4. Markov chains

To establish notation, we briefly state some well-known properties of *Markov* chains.

Remark 45. Take discrete iid random variables $Y_1, \ldots, Y_n \sim \mathfrak{D}$ and write $\mathbb{P}(Y_i = k) = p_k := \mathrm{p}(k)$ for all $k \in \mathbb{N}$, independent of $i = 1, \ldots, n$. Define further

$$X_i := \begin{cases} Y_1 + \ldots + Y_i & i > 0 \\ 0 & \text{otherwise} \end{cases} .$$

Then $\{X_n, n \geq 0\}$ defines a discrete-time *Markov chain*, since

$$\begin{aligned} \mathbb{P}(X_{n+1} = k_{n+1} | X_0 = k_0 \wedge \ldots \wedge X_n = k_n) &= \mathbb{P}(Y_{n+1} = k_{n+1} - k_n) \\ &\equiv p_{k_{n+1}-k_n} . \end{aligned}$$

This last property is also called stationary independent increments, i.e. we add an iid random variable at each step.

Remark 46. Let the notation be as in Remark 45. The transition probabilities of the Markov chain are then given by

$$P_{ij} := \begin{cases} p_{j-i} & j \geq i \\ 0 & \text{otherwise} \end{cases} .$$

In matrix form, we write the *transition matrix*

$$\mathbf{P} := \begin{pmatrix} p_0 & p_1 & p_2 & \cdots \\ & p_0 & p_1 & \cdots \\ & & p_0 & \cdots \\ & & & \ddots \end{pmatrix} .$$

Working with transition matrices is straightforward—if the initial distribution is given by $\pi := (1, 0, \ldots)$, then obviously $(\pi\mathbf{P})_i = p_i$. Iterating $\mathbf{P}$ then yields the distributions of $X_2, X_3, \ldots$, respectively—e.g. $(\pi\mathbf{P}^2)_i = \mathbb{P}(X_2 = i) \equiv \mathbb{P}(Y_1 + Y_2 = i)$.

We know that X_2 is divisible—namely into $X_2 = Y_1 + Y_2$, by construction—but what if we ask this question the other way round? We will show in the next section that there exists a relatively straightforward way to calculate if an (infinite) matrix in the shape of

P has a stochastic root—i.e. if $\mathfrak{D}$ is divisible. Observe that this is not in contradiction with Theorem 1, as the theorem does not apply to infinite operators.

In contrast, the more general question of whether we can write a finite discrete random variable as a sum of nontrivial, potentially distinct random variables will be shown to be NP-hard.

3.3. Equivalence to polynomial factorization

Starting from our digression in section 3.2.4 and using the same notation, we begin with the following definition.

Definition 47. Denote with **S** the shift matrix $S_{ij} := \delta_{i+1,j}$. Then we can write

$$\mathbf{P} = p_0\mathbb{1} + p_1\mathbf{S} + p_2\mathbf{S}^2 + \ldots = \sum_{i=0}^{\infty} p_i\mathbf{S}^i \in \mathbb{R}_{[0,1]}[\mathbf{S}] \quad .$$

Since **S** just acts as a symbol, we write

$$f_{\mathfrak{D}}(x) := \sum_{i=0}^{N} p_i x^i \in \mathcal{R} \quad \text{where} \quad \mathcal{R} := \mathbb{R}_{\geq 0}[x]/\sim \quad ,$$

and $f \sim g :\Leftrightarrow f = cg$, $c > 0$. We call $f_{\mathfrak{D}}$ the *characteristic polynomial* of $\mathfrak{D}$—not to be confused with the characteristic polynomial of a matrix. The equivalence space $\mathcal{R}$ defines the set of all characteristic polynomials, and can be written as

$$\mathcal{R} = \bigcup_{i=n}^{\infty} \mathcal{R}_i \quad \text{where} \quad \mathcal{R}_n := \mathcal{R}/(x^n) \quad .$$

We mod out the overall scaling in order to keep the normalization condition $\sum_k \mathrm{p}(k) = 1$ implicit—if we write $f_{\mathfrak{D}}$, we will always assume $f_{\mathfrak{D}}(1) = 1$. An alternative way to define these characteristic polynomials is via characteristic functions, as given in Definition 36.

Definition 48. $f_{\mathfrak{D}}(\mathrm{e}^{\mathrm{i}\omega}) = \phi_X(\omega)$.

The reason for this definition is that it allows us to reduce operations on the transition matrix **P** or products of characteristic functions ϕ_X to algebraic operations on $f_{\mathfrak{D}}$. This enables us to translate the divisibility problem into a polynomial factorization problem and use algebraic methods to answer it. Observe that the $\mathcal{R}_N$ are normed vector spaces, which we will make use of later.

Definition 49. We define norms on the space of characteristic polynomials of degree N—$\mathcal{R}_N$—via $\|f_{\mathfrak{D}}\|_{N,p} := \|(p_i)_{1\leq i\leq N}\|_{\ell^p}$. If N is not explicitly specified, we usually assume $N = \deg f_{\mathfrak{D}}$.

First note the following proposition.

Proposition 50. *There is a* 1*-to-*1 *correspondence between finite distributions* $\mathfrak{D}$ *and characteristic polynomials* $f_{\mathfrak{D}}$*, as defined in Definition 47.*

Proof. Clear by Definition 48 and the uniqueness of characteristic functions. □

While this might seem obvious, it is worth clarifying, since this correspondence will allow us to directly translate results on polynomials to distributions.

The following lemma reduces the question of divisibility and decomposability—see Definition 42 and 43—to polynomial factorization.

Lemma 51. *A finite discrete distribution* $\mathfrak{D}$ *is* n*-divisible iff there exists a polynomial* $g \in \mathcal{R}$ *such that* $g^n = f_{\mathfrak{D}}$*.* $\mathfrak{D}$ *is* n*-decomposable iff there exist polynomials* $g_1, \ldots, g_n \in \mathcal{R}$ *such that* $\prod_{i=1}^n g_i = f_{\mathfrak{D}}$*.*

Proof. Assume that $\mathfrak{D}$ is n-divisible, i.e. that there exists a distribution $\mathfrak{D}'$ and random variables $Z_1, \ldots, Z_n \sim \mathfrak{D}'$ such that $X = \sum_{i=1}^n Z_i$. Denote with $\mathbf{Q}$ the transition matrix of $\mathfrak{D}'$, as defined in Remark 46, and write q for its probability mass function. Then

$$\mathbb{P}(X = j) = \mathbb{P}\left(\sum_{i=1}^{n} Z_i = j\right) = (\mathbf{Q}^n \pi)_j \quad ,$$

as before. Write $g_{\mathfrak{D}'}$ for the characteristic polynomial of $\mathfrak{D}'$. By Definition 47, $g^n(\mathbf{S}) \equiv f_{\mathfrak{D}}(\mathbf{S})$, and hence $g^n_{\mathfrak{D}'} = f_{\mathfrak{D}}$. Observe that

$$1 = \sum_i \mathrm{p}(i) = f_{\mathfrak{D}}(1) \equiv g^n_{\mathfrak{D}'}(1) = \left(\sum_i \mathrm{q}(i)\right)^n \quad ,$$

and hence $\sum_i \mathrm{q}(i) = 1$ is normalized automatically.

The other direction is similar, as well as the case of decomposability, and the claim follows. □

3.4. Divisibility

3.4.1. Computational problems

We state an exact variant of the computational formulation of the question according to Definition 43—i.e. one with an allowed margin of error—as well as a weak membership formulation.

Definition 52 *(*Distribution Divisibility$_n$*).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$.

Question. Does there exist a finite discrete distribution $\mathfrak{D}' : X = \sum_{i=1}^{n} Z_i$ for random variables $Z_i \sim \mathfrak{D}'$?

Observe that this includes the case $n = 2$, which we defined in Definition 43.

Definition 53 *(Weak Distribution Divisibility$_{n,\epsilon}$).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$ with pmf $\mathrm{p}_X(k)$.
Question. If there exists a finite discrete random variable Y with pmf $\mathrm{p}_Y(k)$, such that $\|\mathrm{p}_X - \mathrm{p}_Y\|_\infty < \epsilon$ and such that
1. Y is n-divisible—return Yes
2. Y is not n-divisible—return No.

3.4.2. Exact divisibility

Theorem 54. Distribution Divisibility$_n \in$ P.

Proof. By Lemma 51 it is enough to show that for a characteristic polynomial $f \in \mathcal{R}_N$, we can find a $g \in \mathcal{R} : g^n = f$ in polynomial time. In order to achieve this, write $(f)^{1/n}$ as a Taylor expansion with rest, i.e.

$$\sqrt[n]{f(x)} = p(x) + R(x) \quad \text{where} \quad p \in \mathcal{R}_{N/n}, R \in \mathcal{R}_N \quad .$$

If $R \equiv 0$, then $g = p$ n-divides f, and then the distribution described by f is n-divisible. Since the series expansion is constructive and can be done efficiently—see [25]—the claim follows.

If the distribution coefficients are rational numbers, another method is to completely factorize the polynomial—e.g. using the *LLL* algorithm, which is known to be easy in this setting—sort and recombine the linear factors, which is also in $\mathrm{O}(\mathrm{poly}(\mathrm{ord} f))$, see for example [13]. Then check if all the polynomial root coefficients are positive. □

We collect some further facts before we move on.

Remark 55. Let p be the probability mass function for a finite discrete distribution $\mathfrak{D}$, and write $\operatorname{supp} \mathrm{p} = \{k : \mathrm{p}(k) \neq 0\}$. If $\max \operatorname{supp} \mathrm{p} - \min \operatorname{supp} \mathrm{p} =: w$, then $\mathfrak{D}$ is obviously not n-divisible for $n > w/2$, and furthermore not for any n that do not divide $w, n < w/2$. Indeed, $\mathfrak{D}$ is not n-divisible if the latter condition holds for *either* $\max \operatorname{supp} \mathrm{p}$ or $\min \operatorname{supp} \mathrm{p}$.

Remark 56. Let $X \sim \mathfrak{D}$ be an n-divisible random variable, i.e. $\exists Z_1, \ldots, Z_n \sim \mathfrak{D}' : \sum_{i=1}^{n} Z_i = X$. Then $\mathfrak{D}'$ is unique.

Proof. This is clear, because $\mathbb{R}[x]$ is a unique factorization domain. □

3.4.3. Divisibility with variation

As an intermediate step, we need to extend Theorem 54 to allow for a margin of error ϵ, as captured by the following definition.

Definition 57 *(*Distribution Divisibility$_{n,\epsilon}$*).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$ with pmf $\mathrm{p}_X(k)$.

Question. Do there exist finite random variables $Z_1, \ldots, Z_n \sim \mathfrak{D}'$ with pmfs $\mathrm{p}_Z(k)$, such that $\| \underbrace{\mathrm{p}_Z * \ldots * \mathrm{p}_Z}_{n \text{ times}} - \mathrm{p}_X \|_\infty < \epsilon$?

Lemma 58. Distribution Divisibility$_{n,\epsilon}$ *is in* P.

Proof. Let $f(x) = \sum_{i=0}^{N} p_i x^i$ be the characteristic polynomial of a finite discrete distribution, and $\epsilon > 0$. By padding the distribution with 0s, we can assume without loss of generality that $N = \deg f$ is a multiple of n. A polynomial root—if it exists—has the form $g(x) = \sum_{i=0}^{N} a_i x^i$, where $a_i \geq 0 \ \forall i$. Then

$$\begin{aligned} g(x)^n &= (\ldots + a_3 x^3 + a_2 x^2 + a_1 x + a_0)^n \\ &= \ldots + ((n-1)a_1^2 + n a_0^{n-2} a_2) x^2 + n a_0^{n-1} a_1 x + a_0^n \quad . \end{aligned}$$

Comparing coefficients in the divisibility condition $f(x) = g(x)^n$, the latter translates to the set of inequalities

$$\begin{aligned} a_0^n &\in (p_0 - \epsilon, p_0 + \epsilon) \\ n a_0^{n-1} a_1 &\in (p_1 - \epsilon, p_1 + \epsilon) \\ (n-1)a_1^2 + n a_0^{n-2} a_2 &\in (p_2 - \epsilon, p_2 + \epsilon) \\ &\vdots \end{aligned}$$

Each term but the first one is of the form $\mathrm{h}_i(a_1, \ldots, a_{i-1}) + n a_0^{n-i} a_i \in (p_i - \epsilon, p_i + \epsilon)$, where $\mathrm{h}_i \geq 0 \ \forall i$ is monotonic. This can be rewritten as $a_i \in \mathrm{U}_{\epsilon / n a_0^{n-i}}((p - \mathrm{h}_i(a_1, \ldots, a_{i-1}))/n a_0^{n-i})$. It is now easy to solve the system iteratively, keeping track of the allowed intervals I_i for the a_i.

If $\mathrm{I}_i = \varnothing$ for some i, we return No, otherwise Yes. We have thus developed an efficient algorithm to answer Distribution Weak Divisibility$_{n,\epsilon}$, and the claim of Lemma 58 follows. □

Remark 59. Given a random variable X, the algorithm constructed in the proof of Lemma 58 allows us to calculate the closest n-divisible distribution to X in polynomial time.

Proof. Straightforward, e.g. by using binary search over ϵ. □

3.4.4. Weak divisibility

For the weak membership problem, we reduce Weak Distribution Divisibility$_{n,\epsilon}$ to Distribution Divisibility$_{n,\epsilon}$.

Theorem 60. Weak Distribution Divisibility$_{n,\epsilon} \in$ P.

Proof. Let $\mathfrak{D}$ be a finite discrete distribution. If Distribution Divisibility$_{n,\epsilon}$ answers Yes, we know that there exists an n-divisible distribution ϵ-close to $\mathfrak{D}$. In case of No, $\mathfrak{D}$ itself is not n-divisible, hence we know that there exists a non-n-divisible distribution close to $\mathfrak{D}$. □

3.4.5. Continuous distributions

Let us briefly discuss the case of continuous distributions—continuous meaning a non-discrete state space $\mathcal{X}$, as specified in section 3.2.2. Although divisibility of continuous distributions is well-defined and widely studied, formatting the continuous case as a computational problem is delicate, as the continuous distribution must be specified by a finite amount of data for the question to be computationally meaningful. The most natural formulation is the continuous analogue of Definition 43 as a weak-membership problem. However, we can show that this problem is computationally trivial.

First observe the following intermediate result.

Lemma 61. *Take* $f \in \mathcal{C}^{+}_{\mathrm{c,b}}$ *with* $\operatorname{supp} f \subset \mathrm{A} \cup \mathrm{B}$*, where* $\mathrm{A} := [0, M], \mathrm{B} := [2M, 3M]$*,* $M \in \mathbb{R}_{>0}$*. We claim that if* f *is divisible, then both* $f|_{\mathrm{A}}$ *and* $f|_{\mathrm{B}}$ *are divisible.*

Proof. Due to symmetry, it is enough to show divisibility for $f|_{\mathrm{A}}$. Assuming f is divisible, we can write $f = r * r$, i.e. $f(x) = \int_{\mathbb{R}} r(x-y)r(y)\mathrm{d}y$. It is straightforward to show that $r(x) = 0\ \forall x < 0$. Define

$$\bar{r}(x) = \begin{cases} r(x) & x \in \mathrm{A}/2 \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

where $\mathrm{A}/2 := \{a/2 : a \in \mathrm{A}\}$. Then

$$\begin{aligned} (\bar{r} * \bar{r})(x) &= \int_{\mathbb{R}} \bar{r}(x-y)\bar{r}(y)\mathrm{d}y \\ &= \int_{\mathbb{R}} \mathrm{d}y \begin{cases} r(x-y) & x-y \in \mathrm{A}/2 \\ 0 & \text{otherwise} \end{cases} \cdot \begin{cases} r(y) & y \in \mathrm{A}/2 \\ 0 & \text{otherwise} \end{cases}. \end{aligned}$$

We see that $(\bar{r}*\bar{r})(x) = 0$ for $x \notin \mathrm{A}$. For $x \in \mathrm{A}$, the support of the integrand is contained in $\{y : y \in x - \mathrm{A}/2 \wedge y \in \mathrm{A}/2\} = x - \mathrm{A}/2 \cap \mathrm{A}/2 := \mathrm{S}_x$, and hence we can write $(\bar{r}*\bar{r})(x) =$

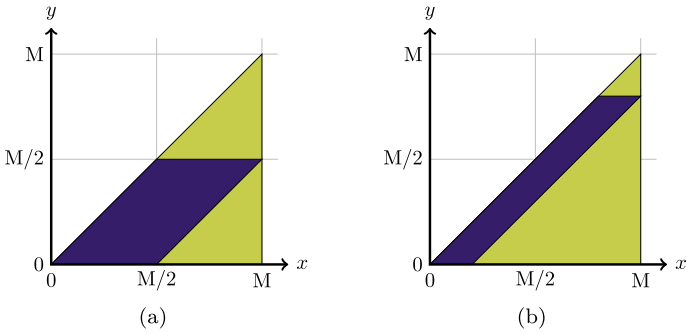


Fig. 4. (a) Integration domains in Proposition 62 for $\bar{r}*\bar{r}$ (dark purple shading) and $f|_{\mathrm{A}}$ (light green shading), respectively. (b) Example for integration domains in Proposition 85 for $\bar{r}*\bar{s}$ (dark purple shading) and $f|_{\mathrm{A}}$ (light green shading), respectively.

$\int_{\mathrm{S}_x} r(x-y)r(y)\mathrm{d}y$. It hence remains to show that $f|_{\mathrm{A}}(x) = \int_{\mathrm{S}_x} r(x-y)r(y)\mathrm{d}y\ \forall x \in \mathrm{A}$. The integrand $r(x-y)r(y) = 0\ \forall y < 0 \vee y > x$. The difference in the integration domains can be seen in Fig. 4. We get two cases.

Let $x \in \mathrm{A}$. Assume $\exists y' \in (M/2, M)$ such that $r(x-y')r(y') > 0$. Let $x' := 2y'$. We then have $r(y')^2 = r(x'-y')r(y') > 0$, and due to continuity $f(x') > 0$, contradiction, because $x' \in (M, 2M)$.

Analogously fix $x' \in (M/2, M)$. Assume $\exists y' \in (0, x'-M/2)$ such that $r(x'-y')r(y') > 0$, and thus $r(x'-y') > 0$, where $a := x'-y' > M/2$, $2a \in (M, 2M)$. Then $r(a)^2 = r(2a-a)r(a) > 0$, due to continuity $f(2a) > 0$, again contradiction. □

Proposition 62. *Let $\mathcal{C}^+_{\mathrm{c,b}}$ denote the set of piecewise continuous nonnegative functions of bounded support. Then the set of* nondivisible *functions, $\mathcal{I} := \{f : \nexists r \in \mathcal{C}_{\mathrm{c,b}} : f = r * r\}$ is dense in $\mathcal{C}_{\mathrm{c,b}}$.*

Proof. It is enough to show the claim for functions $f \in \mathcal{C}^+_{\mathrm{c,b}}$ with $\inf \operatorname{supp} f \geq 0$. Let $\epsilon > 0$, and $M := \sup \operatorname{supp} f$. Take $j \in \mathcal{C}_{\mathrm{c,b}}$ to be nondivisible with $\operatorname{supp} j \subset (2M, 3M)$, and define

$$g(x) := \begin{cases} f(x) & x < M \\ \epsilon j(x)/\|j\|_\infty & x \in (2M, 3M) \\ 0 & \text{otherwise} \end{cases}.$$

By construction, $\|f-g\|_\infty < \epsilon$, but $g|_{(2M,3M)} \equiv j$ is not divisible, hence by Lemma 61 g is not divisible, and the claim follows. □

Corollary 63. *Let $\epsilon > 0$. Let X be a continuous random variable with pdf $\mathrm{p}_X(k)$. Then there exists a nondivisible random variable Y with pdf $\mathrm{p}_Y(k)$, such that $\|\mathrm{p}_X - \mathrm{p}_Y\| < \epsilon$.*

Proof. Let $\epsilon > 0$ small. Since $\mathcal{C}_{\mathrm{c,b}} \subset \{f \text{ integrable}\} =: \mathrm{L}$, we can pick $f_M \in \mathrm{L} : \operatorname{supp} f_M \in (-M, M), \|\mathrm{p}_X - f_M\| < \epsilon/3$ and $\|f_M\| = 1 + \delta$ with $|\delta| \leq \epsilon/3$. Then

$$\left\| \mathrm{p}_X - \frac{f_M}{\|f_M\|} \right\| = \left\| \mathrm{p}_X - \frac{f_M}{1+\delta} \right\| \leq \|\mathrm{p}_X - f_M\| + \frac{\epsilon}{2}\|f_M\| \leq \epsilon \quad ,$$

and Proposition 62 finishes the claim. □

Corollary 64. *Any weak membership formulation of divisibility in the continuous setting is trivial to answer, as for all $\epsilon > 0$, there always exists a nondivisible distribution ϵ close to the one at hand. Similar considerations apply to other formulations of the continuous divisibility problem.*

3.4.6. Infinite divisibility

Let us finally and briefly discuss the case of infinite divisibility. While interesting from a mathematical point of view, the question of infinite divisibility is ill-posed computationally. Trivially, discrete distributions cannot be infinitely divisible, as follows directly from Theorem 54. A similar argument shows that neither the ϵ, nor the weak variant of the discrete problem is a useful question to ask, as can be seen from Lemma 58 and 60.

By the same arguments as in section 3.4.5, the weak membership version is thus easy to answer and therefore trivially in P.

3.5. Decomposability

3.5.1. Computational problems

We define the decomposability analogue of Definition 52 and 53 as follows.

Definition 65 *(*Distribution Decomposability*).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$.
Question. Do there exist finite discrete distributions $\mathfrak{D}', \mathfrak{D}'' : X = Z_1 + Z_2$ for random variables $Z_1 \sim \mathfrak{D}', Z_2 \sim \mathfrak{D}''$?

Definition 66 *(*Weak Distribution Decomposability$_\epsilon$*).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$ with pmf $\mathrm{p}_X(k)$.
Question. If there exists a finite discrete random variable Y with pmf $\mathrm{p}_Y(k)$, such that $\|\mathrm{p}_X - \mathrm{p}_Y\|_\infty < \epsilon$ and such that
1. Y is decomposable—return Yes
2. Y is indecomposable—return No.

In this section, we will show that Distribution Decomposability is NP-hard, for which we will need a series of intermediate results. Requiring the support of the first random variable Z_1 to have a certain size, i.e. $|\operatorname{supp}(p_{\mathfrak{D}'})| = m$, yields the following program.

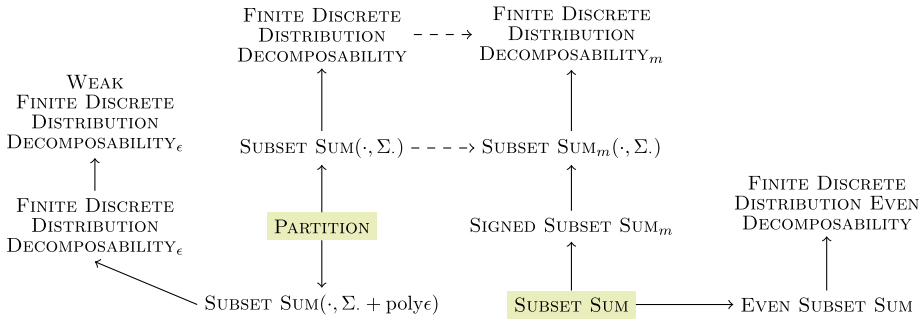


Fig. 5. Complete chain of reduction for our discrete programs. The dashed lines are obvious and not mentioned explicitly.

Definition 67 *(*Distribution Decomposability$_m$*,* $m \geq 2$*).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$ with $|\operatorname{supp}(p_{\mathfrak{D}})| > m$.

Question. Do there exist finite discrete distributions $\mathfrak{D}', \mathfrak{D}'' : X = Z_1 + Z_2$ for random variables $Z_1 \sim \mathfrak{D}', Z_2 \sim \mathfrak{D}''$ and such that $|\operatorname{supp}(p_{\mathfrak{D}'})| = m$?

We then define Distribution Even Decomposability to be the case where the two factors have equal support.

The full reduction tree can be seen in Fig. 5.

Analogous to Lemma 21, we state the following observation.

Lemma 68. *All the above* Decomposability *problems in Definition 65 to 66 are contained in* NP*.*

Proof. It is straightforward to construct a witness and a verifier that satisfies the definition of the decision class NP. For example in Definition 78, a witness is given by two tables of numbers which are easily checked to form finite discrete distributions. Convolving these lists and comparing the result to the given distribution can clearly be done in polynomial time. Both verification and witness are thus poly-sized, and the claim follows. □

3.5.2. Even decomposability

We continue by proving that Distribution Even Decomposability is NP-hard. We will make use of the following variant of the well-known Subset Sum problem, which is NP-hard—see Lemma 92 for a proof. The interested reader will find a rigorous digression in section A.1.

Definition 69 *(*Even Subset Sum*).*

Instance. Multiset S of reals with |S| even, $l \in \mathbb{R}$.

Question. Does there exist a multiset $\mathrm{T} \subsetneq \mathrm{S}$ with $|\mathrm{T}| = |\mathrm{S}|/2$ and such that $|\sum_{t\in\mathrm{T}} t - \sum_{s\in\mathrm{S}\setminus\mathrm{T}} s| < l$?

This immediately leads us to the following intermediate result.

Lemma 70. Distribution Even Decomposability *is* NP*-hard.*

Proof. Let (S, l) be an instance of Even Subset Sum. We will show that there exists a polynomial $f \in \mathcal{R}$ of degree $2|\mathrm{S}|$ such that f is divisible into $f = g \cdot h$ with $\deg g = \deg h$ iff (S, l) is a Yes instance. We will explicitly construct the polynomial $f \in \mathcal{R}$. As a first step, we transform the Even Subset Sum instance (S, l), making it suited for embedding into f.

Let $N := |\mathrm{S}|$ and denote the elements in S with $s_1, \ldots, s_N$. We perform a linear transformation on the elements s_i via

$$b_i := a\left(s_i - \frac{1}{|\mathrm{S}'|}\sum_{s\in\mathrm{S}'} s\right) + \frac{al}{2|\mathrm{S}'|} \quad \text{for} \quad i = 1, \ldots, N \ , \tag{7}$$

where $a \in \mathbb{R}_{>0}$ is a free scaling parameter chosen later such that $|b_i| < \delta \in \mathbb{R}_+$ small. Let $\mathrm{B} := \{b_1, \ldots, b_N\}$. By Lemma 93, we see that Even Subset Sum(S, l) = Even Subset Sum(B, al). Since further $\sum_i b_i = al/2 > 0$, we know that (B, al) is a Yes instance if and only if there exist two non-empty disjoint subsets $\mathrm{B}_1 \cup \mathrm{B}_2 = \mathrm{B}$ such that both

$$\sum_{i\in\mathrm{B}_1} b_i > 0 \quad \text{and} \quad \sum_{i\in\mathrm{B}_2} b_i > 0 \ . \tag{8}$$

The next step is to construct the polynomial f and prove that it is divisible into two polynomial factors $f = g \cdot h$ if and only if (B, al) is a Yes instance. We first define quadratic polynomials $g(b_i, x) := x^2 + b_i x + 1$ for $i = 1, \ldots, N$, and set $f_\mathrm{T}(x) := \prod_{b\in\mathrm{T}} g(b, x)$ for $\mathrm{T} \subset \mathrm{B}$. Observe that for suitably small δ, the $g(b_i, x)$ are irreducible over $\mathbb{R}[x]$. With this notation, we claim that $f_\mathrm{B}(x)$ has the required properties.

In order to prove this claim, we first show that for sufficiently small scaling parameter a, a generic subset $\mathrm{T} \subset \mathrm{B}$ with $n := |\mathrm{T}|$ and $f_\mathrm{T}(x) =: \sum_{i=1}^{2|\mathrm{T}|} c_i x^i$, the coefficients c_i satisfy

$$c_0 = 1 \ , \tag{9}$$

$$\mathrm{sgn}(c_1) = \mathrm{sgn}(\Sigma) \ , \tag{10}$$

$$c_{2j} > 0 \quad \text{for} \quad j = 1, \ldots, |\mathrm{T}| \ , \tag{11}$$

$$\mathrm{sgn}(c_{2j+1}) \geq \mathrm{sgn}(\Sigma) \quad \text{for} \quad j = 1, \ldots, |\mathrm{T}| - 1 \ , \tag{12}$$

where $\Sigma := \sum_{t\in\mathrm{T}} t$. Indeed, if then $f_\mathrm{B} = g \cdot h$, where $g, h \in \mathcal{R}$, then $g = f_{\mathrm{B}_1}$ and $h = f_{\mathrm{B}_2}$ for aforementioned subsets $\mathrm{B}_1, \mathrm{B}_2 \subsetneq \mathrm{B}$, and conversely if (B, al) is a Yes instance, then

$f_{\mathrm{B}} = f_{\mathrm{B}_1} \cdot f_{\mathrm{B}_2}$—remember that $\mathbb{R}[x]$ is a unique factorization domain, so all polynomials of the shape f_{T} necessarily decompose into quadratic factors.

By construction, $c_0 = 1$ and $c_1 = n\Sigma$, so the first two assertions follow immediately. To address equation (11) and (12), we further split up the even and odd coefficients into

$$c_j =: \begin{cases} c_{j,0} + c_{j,2} + \ldots + c_{j,j} & \text{if } j \text{ even} \\ c_{j,1} + c_{j,3} + \ldots + c_{j,j} & \text{if } j \text{ odd} \end{cases} , \tag{13}$$

where $c_{j,k}$ is the coefficient of $x^j b_{i_1} \cdots b_{i_k}$. We thus have $c_{j,k} = \mathcal{O}(\delta^k)$ in the limit $\delta \to 0$—we will implicitly assume the limit in this proof and drop it for brevity. Our goal is to show that the scaling in δ suppresses the combinatorial factors, i.e. that c_j is dominated by its first terms $c_{j,0}$ and $c_{j,1}$, respectively.

In order to achieve this, we need some more machinery. First regard $g(\delta, x) = x^2 + \delta x + 1$. It is immediate that for an expansion

$$g(\delta, x)^n =: \sum_{j=0}^{2n} x^j \sum_{k=0}^{n} d_{j,k} \delta^k \quad ,$$

we get coefficient-wise inequalities

$$|c_{j,k}| \leq d_{j,k} \ \forall j = 0, \ldots, 2n, k = 0, \ldots, n \quad . \tag{14}$$

We will calculate the coefficients $d_{j,k}$ of $g(\delta, x)^n$ explicitly and use them to bound the coefficients $c_{j,k}$ of $f_{\mathrm{T}}(x)$.

Using a standard *Cauchy* summation and the uniqueness of polynomial functions, we obtain

$$\begin{aligned} g(\delta, x)^n &= \sum_{j=0}^{n} \frac{1}{j!} (1 + x^2)^{n-j} x^j (n)_j \delta^j \\ &= \sum_{j=0}^{n} \frac{\delta^j}{j!} (n)_j x^j \sum_{k=0}^{n-j} \binom{n-j}{k} x^{2k} \\ &\equiv \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{\delta^j}{j!} (n)_j \binom{n-j}{k} x^{j+2k} \\ &= \sum_{j=0}^{\infty} \sum_{l=0}^{j} \frac{\delta^l}{l!} (n)_l \binom{n-l}{j-l} x^{2j-l} \\ &\equiv \sum_{j=0}^{n} \sum_{l=0}^{j} \frac{\delta^l}{l!} (n)_l \binom{n-l}{j-l} x^{2j-l} \\ &= \sum_{j=0}^{n} \sum_{l=j}^{2j} \frac{\delta^{2j-l}}{(2j-l)!} (n)_{2j-l} \binom{n-2j+l}{l-j} x^l \quad . \end{aligned}$$

With $(n)_l$ we denote the falling factorial, i.e. $(n)_l = n(n-1)(n-2)\cdots(n-l+1)$. By convention, $(n)_0 = 1$.

Regarding even and odd powers of x separately, we can thus deduce that

$$g(\delta,x)^n = \sum_{j=0}^{2n} x^j \begin{cases} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{\delta^{2k+1}}{(2k+1)!} \frac{(n)_{\lceil \frac{j}{2} \rceil + k}}{(\lfloor \frac{j}{2} \rfloor - k)!} & \text{if } j \text{ odd} \\ \sum_{k=0}^{\frac{j}{2}} \frac{\delta^{2k}}{(2k)!} \frac{(n)_{\frac{j}{2}+k}}{(\frac{j}{2}-k)!} & \text{if } j \text{ even} \end{cases}$$

$$= \sum_{j=0}^{2n} x^j \begin{cases} (n)_{\lceil \frac{j}{2} \rceil} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{\delta^{2k+1}}{(2k+1)!} \frac{(n-\lceil \frac{j}{2} \rceil)_k}{(\lfloor \frac{j}{2} \rfloor - k)!} & \text{if } j \text{ odd} \\ (n)_{\frac{j}{2}} \sum_{k=0}^{\frac{j}{2}} \frac{\delta^{2k}}{(2k)!} \frac{(n-\frac{j}{2})_k}{(\frac{j}{2}-k)!} & \text{if } j \text{ even} \end{cases} .$$

A straightforward estimate shows that for the even and odd case, we obtain the coefficient scaling

$$g(\delta,x)^n = \sum_{j=0}^{2n} x^j \begin{cases} (n)_{\lceil \frac{j}{2} \rceil} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \delta^{2k+1} \mathcal{O}(n^k) & \text{if } j \text{ odd} \\ (n)_{\frac{j}{2}} \sum_{k=0}^{\frac{j}{2}} \delta^{2k} \mathcal{O}(n^k) & \text{if } j \text{ even} \end{cases} ,$$

which means that e.g. picking $\delta = \mathcal{O}(1/n^2)$ is enough to exponentially suppress the higher order combinatorial factors.

We will now separately address the even and odd cases—equation (11) and (12).

Even case. As the constant coefficients $c_{j,0} = \mathcal{O}(1)$ in δ, it is the same as for $g(\delta,x)^n$ and by equation (14), we immediately get

$$\frac{|c_{j,2} + \ldots + c_{j,j}|}{c_{j,0}} = \mathcal{O}(\delta) .$$

Odd case. Note that if $\Sigma < 0$, we are done, so assume $\Sigma > 0$ in the following. A simple combinatorial argument gives

$$c_{j,1} = \binom{n-1}{(j-1)/2} \Sigma ,$$

so it remains to show that $c_{j,1} > -c_{j,3} - \ldots - c_{j,j}$. Analogously to the even case, by equation (14), we conclude

$$\frac{|c_{j,3} + \ldots + c_{j,j}|}{c_{j,1}} = \mathcal{O}(\delta) \quad ,$$

which finalizes our proof. □

3.5.3. m-Support decomposability

In the next two sections we will generalize the last result to Distribution Decomposability$_m$. As a first observation, we note the following.

Lemma 71. *Let $f(n)$ be such that $(f(n)\beta(f(n), n+1-f(n)))^{-1} = \mathcal{O}(\text{poly}(n))$. Then* Distribution Decomposability$_{f(|\cdot|)} \in$ P.

Proof. See proof of Theorem 54, and an easy scaling argument for $\binom{n}{f(n)}$ completes the proof. As in Remark 91, this symmetrically extends to Distribution Decomposability$_{|\cdot|-f(|\cdot|)} \in$ P. □

Observe that $f(n) = n/2$ yields exponential growth, hence the remark is consistent with the findings in section 3.5.2.

We now regard the general case. As in the last section, we need variants of the Subset Sum problem, which are given in the following two definitions.

Definition 72 *(*Subset Sum$_m$, $m \in \mathbb{Z}$*).*

Instance. Multiset S of reals with |S| even, $l \in \mathbb{R}$.
Question. Does there exist a multiset T $\subsetneq$ S with |T| $= m$ and such that $|\sum_{t\in \mathrm{T}} t - \sum_{s\in \mathrm{S}\setminus \mathrm{T}} s| < l$?

Definition 73 *(*Signed Subset Sum$_m$*).*

Instance. Multiset S of positive integers or reals, $x, y \in \mathbb{R} : x \leq y$.
Question. Does there exist a multiset T $\subset$ S with |T| $= m$ and such that $x < \sum_{t\in \mathrm{T}} t - \sum_{s\in \mathrm{S}\setminus \mathrm{T}} s < y$?

Both are shown to be NP-hard in Lemma 90 and 94, or by the following observation. In order to avoid having to take absolute values in the definition of Subset Sum$_m$, we reduce it to multiple instances of Signed Subset Sum$_m$, by using the following interval partition of the entire range $(-l, l)$.

Remark 74. For every $a > 0, l > 0$, there exists a partition of the interval $(-l-2a, l+2a) = \bigcup_{i=0}^{N-1}(x_i, x_{i+1})$ with suitable $N \in \mathbb{N}$ such that $x_{i+1} - x_i = 2a$ and

$$(-l, l) = \left(\bigcup_{i=1}^{N-2}(x_i, x_{i+1})\right) \setminus ((x_0, x_1) \cup (x_{N-1}, x_N)) \quad .$$

This finally leads us to the following result.

Lemma 75. Distribution Decomposability$_m$ *is* NP*-hard.*

Proof. We will show the reduction Distribution Decomposability$_m$ $\longleftarrow$ Subset Sum$_m$. Let m be fixed. Let (S, l) be an Subset Sum instance. For brevity, we write $\Sigma_\mathrm{S} := \sum_{s\in\mathrm{S}} s$. Without loss of generality, by Corollary 89, we again assume $\Sigma_\mathrm{S} \geq 0$. Now define $a := 2(|\mathrm{S}|l + 2m\Sigma_\mathrm{S} - |\mathrm{S}|\Sigma_\mathrm{S})/(2m - |\mathrm{S}|)$. Using Remark 74, pick a suitable subdivision of the interval $(-l - 2a, l + 2a)$, such that

$$\begin{aligned}\textsc{Subset Sum}_m(\mathrm{S}, l) = \Bigg(\bigvee_{i=1}^{N-2} \textsc{Signed Subset Sum}_m(\mathrm{S}, x_i, x_{i+1})\Bigg) \\ \wedge \neg\textsc{Signed Subset Sum}_m(\mathrm{S}, x_0, x_1) \\ \wedge \neg\textsc{Signed Subset Sum}_m(\mathrm{S}, x_{N-1}, x_N) \quad .\end{aligned}$$

One can verify that

$$\begin{aligned}&\textsc{Signed Subset Sum}_m(\mathrm{S}, x_i - a, x_i + a) \\ &\quad = \textsc{Signed Subset Sum}_m(\mathrm{S} + c(m,i), -\Sigma_{\mathrm{S}+c(m,i)}, \Sigma_{\mathrm{S}+c(m,i)}) \\ &\quad = \textsc{Subset Sum}_m(\mathrm{S} + c(m,i), \Sigma_{\mathrm{S}+c(m,i)}) \quad ,\end{aligned}$$

where we chose $c(m,i) = x_i/(2m - |\mathrm{S}|)$. The latter program we can answer using the same argument as for the proof of Lemma 70, and the claim follows. □

As a side remark, this also confirms the following well-known fact.

Corollary 76. *Let* $f(n)$ *be as in Lemma 71. Then* Subset Sum$_{f(|\cdot|)} \in$ P.

3.5.4. General decomposability

We have already invented all the necessary machinery to answer the general case.

Theorem 77. Distribution Decomposability *is* NP*-hard.*

Proof. Follows immediately from Lemma 70, where we regard the special set of Subset Sum instances for which (S, l) is such that $l = \sum_{s\in\mathrm{S}} s$. We show in Lemma 96 that Subset Sum$(\cdot, \Sigma_\cdot)$ is still NP-hard, thus the claim follows. □

3.5.5. Decomposability with variation

As a further intermediate result—and analogously to Definition 57—we need to allow for a margin of error ϵ.

Definition 78 *(*Distribution Decomposability$_\epsilon$*).*

Instance. Finite discrete random variable $X \sim \mathfrak{D}$ with pmf $\mathrm{p}_X(k)$.
Question. Do there exist finite discrete random variables $Z_1 \sim \mathfrak{D}', Z_2 \sim \mathfrak{D}''$ with pmfs $\mathrm{p}_{Z_1}(k)$, $\mathrm{p}_{Z_2}(k)$, such that $\|\mathrm{p}_{Z_1} * \mathrm{p}_{Z_2} - \mathrm{p}_X\|_\infty < \epsilon$?

This definition leads us to the following result.

Lemma 79. Distribution Decomposability$_\epsilon$ *is* NP*-hard.*

Proof. First observe that we can restate this problem in the following equivalent form. Given a finite discrete distribution $\mathfrak{D}$ with characteristic polynomial $f_{\mathfrak{D}}$, do there exist two finite discrete distributions $\mathfrak{D}', \mathfrak{D}''$ with characteristic polynomials $f_{\mathfrak{D}'}, f_{\mathfrak{D}''}$ such that $\|f_{\mathfrak{D}} - f_{\mathfrak{D}'} f_{\mathfrak{D}''}\|_d < \epsilon$? Here, we are using the maximum norm from Definition 49, and assume without loss of generality that $\deg f_{\mathfrak{D}} = \deg f_{\mathfrak{D}'} \deg f_{\mathfrak{D}''}$.

As $f_{\mathfrak{D}}$ is a polynomial, we can regard its *Viète* map $v : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, where $n = \deg f_{\mathfrak{D}}$, which continuously maps the polynomial roots to its coefficients. It is a well-known fact—see [32] for a standard reference—that v induces an isomorphism of algebraic varieties $w : \mathbb{A}_k^n / \mathrm{S}_n \xrightarrow{\sim} \mathbb{A}_k^n$, where S_n is the n^{th} symmetric group. This shows that w^{-1} is polynomial, and hence the roots of $f_{\mathfrak{D}'} f_{\mathfrak{D}''}$ lie in an $\mathcal{O}(\epsilon)$-ball around those of $f_{\mathfrak{D}}$. By a standard uniqueness argument we thus know that if $f_{\mathfrak{D}} = \prod_i f_i$ with $f_i = x^2 + b_i x + 1$ as in the proof of Lemma 70, then $f_{\mathfrak{D}'} = \prod_i g_i$ with $g_i = a_i x^2 + b_i' x + c_i$, where $a_i = c_i = 1 + \mathcal{O}(\epsilon)$, $b_i' = b_i + \mathcal{O}(\epsilon)$—we again implicitly assume the limit $\epsilon \to 0$.

We continue by proving the reduction Distribution Divisibility$_\epsilon$ $\longleftarrow$ Subset Sum$(\cdot, \Sigma., +\,\mathrm{poly}\,\epsilon)$, which is NP-hard as shown in Lemma 97. Let $\mathrm{S} = \{s_i\}_{i=1}^N$ be a Subset Sum multiset. We claim that it is satisfiable if and only if the generated characteristic function $f_{\mathrm{S}}(x)$—where we used the notation of the proof of Lemma 70—defines a finite discrete probability distribution and the corresponding random variable X is a Yes instance for Distribution Divisibility$_\epsilon$.

First assume f_{S} is such a Yes instance. Then $\sum_{s \in \mathrm{S}} s \geq 0$, and there exist two characteristic polynomials $g = \prod_i g_i$ and $h = \prod_i h_i$ as above and such that $\|f_{\mathrm{S}} - gh\|_d < \epsilon$. We also know that if $g_i = a_i x^2 + b_i x + c_i$, then $\exists \mathrm{T} \subsetneq \mathrm{S}$ such that $\{b_i\}_i \in \mathrm{B}_\epsilon(\mathrm{T}) \subseteq \mathbb{R}^{|\mathrm{T}|}$, where $\mathrm{T} \subsetneq \mathrm{S}$ and $\mathrm{B}_\epsilon(\mathrm{T})$ denotes an ϵ ball around the set T, and analogously for $h_i = a_i' x^2 + b_i' x + c_i'$, with $\{b_i'\}_i \in \mathrm{B}_\epsilon(\mathrm{S} \setminus \mathrm{T}) \subseteq \mathbb{R}^{|\mathrm{S}|-|\mathrm{T}|}$. Regarding the linear coefficients, we thus have

$$\left| \sum_{s \in \mathrm{S}} s - \sum_{t \in \mathrm{T}} t - \sum_{s \in \mathrm{S} \setminus \mathrm{T}} s \right| = \left| \sum_{s \in \mathrm{S}} s - \sum_{i=1}^{|\mathrm{T}|} b_i - \sum_{i=1}^{|\mathrm{S}\setminus\mathrm{T}|} b_i' + \mathcal{O}(\epsilon) \right|$$
$$\leq \mathcal{O}(\epsilon) \leq \sum_{s \in \mathrm{S}} s + \mathcal{O}(\epsilon) \quad . \tag{15}$$

Now the case if f_{S} is a No instance. Assume there exists a nontrivial multiset $\mathrm{T} \subsetneq \mathrm{S}$ satisfying

$$\left| \sum_{t \in \mathrm{T}} t - \sum_{s \in \mathrm{S} \setminus \mathrm{T}} s \right| < \sum_{s \in \mathrm{S}} s + \mathcal{O}(\epsilon) \quad .$$

Then by construction $\sum_{t \in \mathrm{T}} t, \sum_{s \in \mathrm{S} \setminus T} s \geq -\mathcal{O}(\epsilon)$ and $f_{\mathrm{T}} \cdot f_{\mathrm{S} \setminus \mathrm{T}} = f_{\mathrm{S}}$, contradiction, and the claim follows. □

3.5.6. Weak decomposability

Analogously to section 3.4.4, we now regard the weak membership problem of decomposability.

Theorem 80. Weak Distribution Decomposability$_\epsilon$ *is* NP*-hard.*

Proof. In order to show the claim, we prove the reduction Weak Distribution Decomposability$_\epsilon$ $\longleftarrow$ Distribution Decomposability$_{\mathrm{g}(\epsilon)}$, where the function $\mathrm{g} = \mathcal{O}(\epsilon)$. It is clear that the polynomial factor leaves the NP-hardness of the latter program intact.

We use the same notation as in the proof of Lemma 79. Let f_{S} be a Yes instance of Distribution Decomposability$_\epsilon$, and define $\mathrm{S}' := \{s + \mathcal{O}(\epsilon) : s \in \mathrm{S}\}$. From equation (15) it immediately follows that $f_{\mathrm{S}'}$ is a Yes instance of Distribution Decomposability$_{\mathrm{g}(\epsilon)}$, where we allow $\mathrm{g} = \mathcal{O}(\epsilon)$. We have hence shown that there exists an $\mathcal{O}(\epsilon)$ ball around each Yes instance that *solely* contains Yes instances.

A similar argument holds for the No instances. It is clear that these cases can be answered using Weak Distribution Decomposability$_\epsilon$, and the claim follows. □

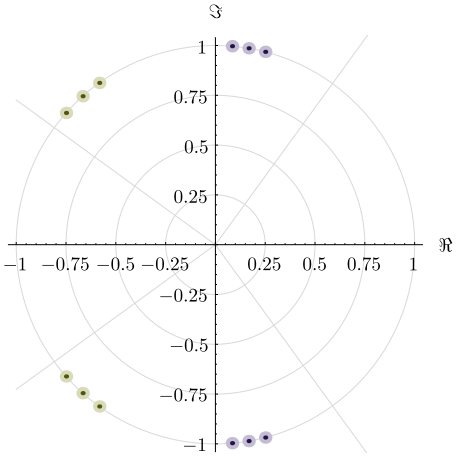
3.5.7. Complete decomposability

Another interesting question to ask is for the complete decomposition of a finite distribution $\mathfrak{D}$ into a sum of indecomposable distributions. We argue that this decomposition is not unique.

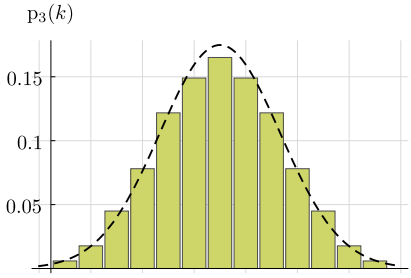
Proposition 81. *There exists a family of finite distributions* $(\mathfrak{D}_n)_{n \in \mathbb{N}}$ *with probability mass functions* $\mathrm{p}_n(k) : \max \operatorname{supp} \mathrm{p}_n(k) = 4n$ *and such that, for each* $\mathfrak{D}_n$*, there are at least* $n!$ *distinct decompositions into indecomposable distributions.*

Proof. We explicitly construct the family $(\mathfrak{D}_n)_{n \in \mathbb{N}}$. Let $n \in \mathbb{N}$. We will define a set of irreducible quadratic polynomials $\{p_k, n_k \text{ for } k = 1, \ldots, n\}$ such that n_k are *not* positive, but $p_k n_l$ are positive quartics $\forall k, l$—and thus define valid probability distributions. Since $\mathbb{R}[x]$ is a unique factorization domain the claim then follows.

Following the findings in the proof of Lemma 70, it is in fact enough to construct a set $\{a_k, b_k : 0 < |a_k| < 2, -2 < b_k < 0 \text{ for } k = 1, \ldots, n\} \subset \mathbb{R}^{2n}$ and such that



(a) complex roots of characteristic polynomial of $\mathfrak{G}_3$



(b) probability mass function for $\mathfrak{G}_3$

Fig. 6. Construction of a family of distributions $(\mathfrak{D}_n)_{n\in\mathbb{N}}$ with at least $n!$ distinct decompositions, cf. Proposition 81. Shown is $\mathfrak{D}_n$ for $n = 3$. The dashed line shows a normal distribution for comparison.

$a_k + b_l > 0\ \forall k, l$—then let $p_k := 1 + a_k x + x^2$, $n_k := 1 + b_k x + x^2$. It is straightforward to verify that e.g.

$$a_k := 1 + \frac{k}{2n} \quad \text{and} \quad b_k := -\frac{k}{2n}$$

fulfil these properties. As an illustration, $\mathfrak{D}_3$ and the roots of its characteristic polynomial are shown in Fig. 6. □

Remark 82. Observe that for $b_k := -k/2n^2$, the construction in Proposition 81 allows decompositions into m indecomposable terms, where $m = n, \ldots, 2n$.

Corollary 83. *$\mathcal{R}$ is not a unique factorization domain.*

Proposition 81 and Remark 82 show that an exponential number of complete decompositions—all of which have different distributions—do not give any further insight into the distribution of interest–indeed, as the number of positive indecomposable factors is not even unique. Asking for a non-maximal decomposition into indecomposable terms does therefore not answer more than whether the distribution is decomposable at all.

Indeed, the question whether one *can* decompose a distribution into indecomposable parts can be trivially answered with YES, but if we include the condition that the factors have to be non-trivial, or for decomposability into a certain number of terms—say $N \geq 2$ or the maximum number of terms—the problem is also obviously NP-hard by the previous results.

In short, by Theorem 77, we immediately obtain the following result.

Corollary 84. *Let $\mathfrak{D}$ be a finite discrete distribution. Deciding whether one can write $\mathfrak{D}$ as any nontrivial sum of irreducible distributions is* NP*-hard.*

3.5.8. Continuous distributions

Analogous to our discussion in section 3.4.5, the exact and ϵ variants of the decomposability question are computationally ill-posed. We again point out that answering the weak membership version is trivial, since the set of indecomposable distributions is dense, as the following proposition shows.

Proposition 85. *Let $\mathcal{C}^{+}_{\mathrm{c,b}}$ denote the piecewise linear nonnegative functions of bounded support. Then the set of* indecomposable *functions, $\mathcal{J} := \{f : \nexists r, s \in \mathcal{C}_{\mathrm{c,b}} : f = r * s\}$ is dense in $\mathcal{C}_{\mathrm{c,b}}$.*

Proof. We first extend Lemma 61, and again take $f \in \mathcal{C}^{+}_{\mathrm{c,b}} : \operatorname{supp} f \subset \mathrm{A} \cup \mathrm{B}$. While not automatically true that $r(x), s(x) = 0\ \forall x < 0$, we can assume this by shifting r and s symmetrically. We also assume $\inf \operatorname{supp} f = 0$, and hence $\inf \operatorname{supp} r = \inf \operatorname{supp} s = 0$—see Lemma 44 for details.

Since $f(x) = 0\ \forall x \in (M, 2M)$, we immediately get $r(x) = s(x) = 0\ \forall x \in (M, 2M)$. Furthermore, $\exists m \in (0, M) : r(x) = s(y) = 0\ \forall x \in (m, M], y \in (M - m, M]$. Analogously to equation (6), we define

$$\bar{r}(x) = \begin{cases} r(x) & x \in [0, m] \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \bar{s}(x) = \begin{cases} r(x) & x \in [0, M - m] \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

The integration domain difference is derived analogously, and can be seen in an example in Fig. 4. We again regard the two cases separately.

Let $x \in \mathrm{A}$. Assume $\exists y' \in (M - m, M)$ such that $r(x - y')s(y') > 0$. Then $s(y') > 0$, contradiction. Now fix $x' \in (m, M)$, and assume $\exists y' \in (0, x' - m) : r(x' - y')s(y') > 0$. Since $x' - y' > x' - x' + m = m$, $r(x' - y') > 0$ yields another contradiction.

The rest of the proof goes through analogously. □

Corollary 86. *Let $\epsilon > 0$. Let X be a continuous random variable with pmf $\mathrm{p}_X(k)$. Then there exists an indecomposable random variable Y with pmf $\mathrm{p}_Y(k)$, such that $\|\mathrm{p}_X - \mathrm{p}_Y\| < \epsilon$.*

Proof. See Corollary 63. □

4. Conclusion

In section 2, we have shown that the question of existence of a stochastic root for a given stochastic matrix is in general at least as hard as answering 1-IN-3SAT, i.e. it is NP-hard. By Corollary 27, this NP-hardness result also extends to NONNEGATIVE and

Doubly Stochastic Divisibility, which proves Theorem 1. A similar reduction goes through for cptp Divisibility in Corollary 24, proving NP-hardness of the question of existence of a cptp root for a given cptp map.

In section 3, we have shown that—in contrast to cptp and stochastic matrix divisibility—distribution divisibility is in P, proving Theorem 4. On the other hand, if we relax divisibility to the more general decomposability problem, it becomes NP-hard as shown in Theorem 6. We have also extended these results to weak membership formulations in Theorem 5 and 7—i.e. where we only require a solution to within ϵ in the appropriate metric—showing that all the complexity results are robust to perturbation.

Finally, in section 3.4.5 and 3.5.8, we point out that for continuous distributions—where the only computationally the only meaningful formulations are the weak membership problems or closely related variants—questions of divisibility and decomposability become computationally trivial, as the nondivisible and indecomposable distributions independently form dense sets.

As containment in NP for all of the NP-hard problems is easy to show (Lemma 21 and 68), these problems are also NP-complete. Thus our results imply that, apart for the distribution divisibility problem which is efficiently solvable, all other divisibility problems for maps and distributions are equivalent to the famous P = NP conjecture, in the following precise sense: A polynomial-time algorithm for answering any one of these questions—(Doubly) Stochastic, Nonnegative or cptp Divisibility, or either of the Decomposability variants—would prove P = NP. Conversely, solving P = NP would imply that there exists a polynomial-time algorithm to solve all of these Divisibility problems.

Acknowledgements

Johannes Bausch would like to thank the German National Academic Foundation and EPSRC (grant number 1600123) for financial support. Toby Cubitt is supported by the Royal Society (grant number UF110164). The authors are grateful to the Isaac Newton Institute for Mathematical Sciences, where part of this work was carried out, for their hospitality during the 2013 programme "Mathematical Challenges in Quantum Information Theory".

Appendix A

A.1. NP-*toolbox*

Boolean satisfiability problems.

Definition 87 *(1-*in*-3*sat*).*

Instance: n_v boolean variables $m_1, \ldots, m_{n_v}$ and n_c clauses $\mathrm{R}(m_{i1}, m_{i2}, m_{i3})$ where $i = 1, \ldots, n_c$, usually denoted as a 4-tuple (n_v, n_c, m_i, m_{ij}). The boolean operator R satisfies

$$\mathrm{R}(a,b,c) = \begin{cases} \textsc{True} & \text{if exactly one of } a,b \text{ or } c \text{ is } \textsc{True} \\ \textsc{False} & \text{otherwise} \quad . \end{cases}$$

Question: Does there exist a truth assignment to the boolean variables such that every clause contains exactly one true variable?

Subset sum problems. We start out with the following variant of a well-known NP-complete problem—see for example [11] for a reference.

Definition 88 *(*Subset Sum*, variant).*

Instance. Multiset S of integer or rational numbers, $l \in \mathbb{R}$.
Question. Does there exist a multiset $\mathrm{T} \subsetneq \mathrm{S}$ such that $|\sum_{t\in\mathrm{T}} t - \sum_{s\in\mathrm{S}\setminus\mathrm{T}} s| < l$?

From the definition, we immediately observe the following rescaling property.

Corollary 89. *Let* $a \in \mathbb{R}\setminus\{0\}$ *and* (S, l) *a* Subset Sum *instance. Then* Subset Sum(S, l) = Subset Sum$(a\mathrm{S}, |a|l)$.

For a special version of Subset Sum, Subset Sum$_m$—as defined in Definition 72—we observe the following.

Lemma 90. Subset Sum$_m$ $\longleftarrow$ Subset Sum.

Proof. If (S, l) is a Subset Sum instance, then

$$\textsc{Subset Sum}(\mathrm{S}, l) = \bigvee_{m=1}^{|\mathrm{S}|} \textsc{Subset Sum}_m(\mathrm{S}, l) \quad . \quad \square$$

Remark 91. It is clear that Subset Sum$_m(\mathrm{S}, l)$ = False for $|\mathrm{S}| \leq m \vee 0 \geq m$. Furthermore, Subset Sum$_m(\mathrm{S}, l)$ = Subset Sum$_{|\mathrm{S}|-m}(\mathrm{S}, l)$.

Observe that this remark indeed makes sense, as Subset Sum$_0$ should give False, which is the desired outcome for $m = |\mathrm{S}|$. We further reduce Subset Sum to Even Subset Sum, as defined in Definition 69.

Lemma 92. Even Subset Sum$\longleftarrow$Subset Sum.

Proof. Let (S, l) be an Subset Sum instance. Define $\mathrm{S}' := \mathrm{S} \cup \{0, \ldots, 0\} : |\mathrm{S}'| = 2|\mathrm{S}|$. Then if Even Subset Sum(S', l) = True, we know that there exists $\mathrm{T}' \subset \mathrm{S}' : |\sum_{t\in\mathrm{T}'} t - \sum_{s\in\mathrm{S}'\setminus\mathrm{T}'} s| < l$. Let then $\mathrm{T} := \mathrm{T}'$ without the 0s. It is obvious that then $|\sum_{t\in\mathrm{T}} t - \sum_{s\in\mathrm{S}\setminus\mathrm{T}} s| < l$. The False case reduces analogously, hence the claim follows. $\square$

For Even Subset Sum, we generalize Corollary 89 to the following scaling property.

Lemma 93. *Let $a \in \mathbb{R} \setminus \{0\}$, $c \in \mathbb{R}$, and (S, l) an* Even Subset Sum *instance. Then* Even Subset Sum(S, l) = Even Subset Sum$(a\mathrm{S} + c, |a|l)$*, where addition and multiplication is defined element-wise.*

Proof. Straightforward, since we require $|\mathrm{S}| = 2|\mathrm{T}| = 2|\mathrm{S} \setminus \mathrm{T}|$. □

For Definition 73, we finally show

Lemma 94. Signed Subset Sum$_m$ $\longleftarrow$ Subset Sum$_m$.

Proof. Immediate from Signed Subset Sum$_m(\mathrm{S}, -l, l)$ = Subset Sum$_m(\mathrm{S}, l)$. □

Partition problems. Another well-known NP-complete problem which will come into play in the proof of Theorem 77 is *set partitioning*.

Definition 95 *(*Partition*).*

Instance. Multiset A of positive integers or reals.
Question. Does there exist a multiset $\mathrm{T} \subsetneq \mathrm{A}$ with $\sum_{t\in\mathrm{T}} t = \sum_{s\in\mathrm{A}\setminus T} s$?

Lemma 96. *For the special case of* Subset Sum *with instance (S, l), where the bound $l \in \mathbb{R}$ equals the total sum of the instance numbers $l = \sum_{s\in\mathrm{S}} s$, we obtain the equivalence* Subset Sum$(\cdot, \Sigma_\cdot) \longleftrightarrow$ Partition$(\cdot)$.

Proof. Let S be the multiset of a Subset Sum instance (S, l), where we assume without loss of generality that all $\mathrm{S} \ni s \geq 0$. Now first assume $\Sigma_\mathrm{S} = 0$. In that case the claim follows immediately, since the problems are identical.

Without loss of generality, we can thus assume $\Sigma_\mathrm{S} > 0$ and regard the set $\mathrm{S}' := \mathrm{S} \cup \{-\Sigma_\mathrm{S}/2, -\Sigma_\mathrm{S}/2\}$, such that $\Sigma_{\mathrm{S}'} = 0$.

If now Subset Sum$(\mathrm{S}', 0)$ = True, we know that there exists $\mathrm{T} \subsetneq \mathrm{S}' : |\sum_{t\in\mathrm{T}} t - \sum_{s\in\mathrm{S}'\setminus T} s| = 0$. Now assume T contains both copies of $-\Sigma_\mathrm{S}/2$. Then clearly

$$|\sum_{t\in\mathrm{T}} t \; - \sum_{s\in\mathrm{S}'\setminus T} s| = |-\Sigma_\mathrm{S} + \sum_{t\in\mathrm{T}\setminus\{-\Sigma_\mathrm{S}\}} t \; - \sum_{s\in\mathrm{S}\setminus T} s| > 0 \quad ,$$

since $|\mathrm{S} \setminus \mathrm{T}| > 0$. The same argument shows that exactly one $-\Sigma_\mathrm{S}/2 \in \mathrm{T}, \mathrm{S} \setminus \mathrm{T}$, and hence Partition(S) = True.

On the other hand, if Partition(S) = True, then it immediately follows that Subset Sum$(\mathrm{S}, \Sigma_\mathrm{S})$ = True. □

Finally observe the following extension of Lemma 96.

Lemma 97. *Let $\epsilon > 0$, f a polynomial. Then* SUBSET SUM$(\cdot, \Sigma_\cdot + f(\epsilon)) \longleftrightarrow$ PARTITION$(\cdot)$.

Proof. The proof is the same as for Lemma 96, but we regard $\mathrm{S} \cup \{-\Sigma_\mathrm{S}/2 - f(\epsilon)/2, -\Sigma_\mathrm{S}/2 - f(\epsilon)/2\}$ instead. □

References

[1] E.H. Bareiss, Sylvester's identity and multistep integer-preserving Gaussian elimination, Math. Comp. 22 (103) (sep 1968) 565–578.
[2] I. Bengtsson, K. Życzkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement, Cambridge University Press, 2006.
[3] T. Charitos, P.R. de Waal, L.C. van der Gaag, Computing short-interval transition matrices of a discrete-time Markov chain from partially observed data, Stat. Med. 27 (6) (mar 2008) 905–921.
[4] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. 10 (3) (1975) 285–290.
[5] W.G. Cochran, The distribution of quadratic forms in a normal system, with applications to the analysis of covariance, Math. Proc. Cambridge Philos. Soc. 30 (02) (oct 1934) 178–191.
[6] H. Cramér, Über eine Eigenschaft der normalen Verteilungsfunktion, Math. Z. 41 (1) (dec 1936) 405–414.
[7] T.S. Cubitt, J. Eisert, M.M. Wolf, Extracting dynamical equations from experimental data is NP hard, Phys. Rev. Lett. 108 (12) (mar 2012) 120503.
[8] T.S. Cubitt, J. Eisert, M.M. Wolf, The complexity of relating quantum channels to master equations, Comm. Math. Phys. 310 (2) (jan 2012) 383–418.
[9] P.D. Egleston, T.D. Lenker, S.K. Narayan, The nonnegative inverse eigenvalue problem, Linear Algebra Appl. 379 (mar 2004) 475–490.
[10] G. Elfving, Zur Theorie der Markoffschen Ketten, Acta Soc. Sci. Fennicae N. Ser. A 2 (8) (1937) 1–17.
[11] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W.H. Freeman & Co., jan 1979.
[12] V. Gorini, Completely positive dynamical semigroups of N-level systems, J. Math. Phys. 17 (5) (aug 1976) 821.
[13] W. Hart, M. van Hoeij, A. Novocin, Practical polynomial factoring in polynomial time, in: Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation – ISSAC '11, ACM Press New York, New York, USA, jun 2011, p. 163.
[14] Q.-M. He, E. Gunn, A note on the stochastic roots of stochastic matrices, J. Syst. Sci. Syst. Eng. 12 (2) (jun 2003) 210–223.
[15] N.J. Higham, Computing real square roots of a real matrix, Linear Algebra Appl. 88–89 (apr 1987) 405–430.
[16] N.J. Higham, L. Lin, On pth roots of stochastic matrices, Linear Algebra Appl. 435 (3) (aug 2011) 448–463.
[17] R.A. Jarrow, A Markov model for the term structure of credit risk spreads, Rev. Financ. Stud. 10 (2) (apr 1997) 481–523.
[18] S.K. Katti, Infinite divisibility of discrete distributions. III, 1977.
[19] J.F.C. Kingman, The imbedding problem for finite Markov chains, Probab. Theory Related Fields 1 (1) (1962) 14–24.
[20] L. Lin, Roots of stochastic matrices and fractional matrix powers, PhD thesis, University of Manchester, 2011.
[21] G. Lindblad, On the generators of quantum dynamical semigroups, Comm. Math. Phys. 48 (2) (1976) 119–130.
[22] L. Ljung, System Identification: Theory for the User, Prentice Hall, 1987.
[23] A.E. Loring, A Hand-Book of the Electromagnetic Telegraph, D. Van Nostrand, 1878.
[24] H. Minc, Nonnegative Matrices, Wiley, 1988.
[25] N.T. Müller, Uniform computational complexity of Taylor series, in: T. Ottmann (Ed.), Proceedings of the 13th International Colloquium on Automata, Languages and Programming, in: Lecture Notes in Computer Science, vol. 267, Springer Berlin Heidelberg, Berlin, Heidelberg, jan 1987, pp. 435–444.
[26] A. Müller-Hermes, D. Reeb, M.M. Wolf, Quantum subdivision capacities and continuous-time quantum coding, IEEE Trans. Inform. Theory 61 (1) (2015) 565–581.

[27] M.A. Nielsen, E. Knill, R. Laflamme, Complete quantum teleportation using nuclear magnetic resonance, Nature 396 (6706) (nov 1998) 15.
[28] F. Steutel, J. Kent, Infinite divisibility in theory and practice, Scand. J. Stat. 6 (2) (1979) 57–64.
[29] O. Thorin, On the infinite divisibility of the Pareto distribution, Scand. Actuar. J. 1977 (1) (jan 1977) 31–40.
[30] O. Thorin, On the infinite divisibility of the lognormal distribution, Scand. Actuar. J. 1977 (3) (mar 1977) 121–148.
[31] F.V. Waugh, M.E. Abel, On fractional powers of a matrix, J. Amer. Statist. Assoc. 62 (319) (sep 1967) 1018–1021.
[32] H. Whitney, Complex Analytic Varieties, Addison-Wesley Series in Mathematics, Addison–Wesley, 1972.
[33] M.M. Wolf, J.I. Cirac, Dividing quantum channels, Comm. Math. Phys. 279 (1) (feb 2008) 147–168.