Nawrotzki's Algorithm for the Countable Splitting Lemma, Constructively

Ana Sokolova ✉
Department of Computer Sciences, University of Salzburg, Austria

Harald Woracek ✉
Institute of Analysis and Scientific Computing, TU Wien, Austria

Abstract

We reprove the countable splitting lemma by adapting Nawrotzki's algorithm which produces a sequence that converges to a solution. Our algorithm combines Nawrotzki's approach with taking finite cuts. It is constructive in the sense that each term of the iteratively built approximating sequence as well as the error between the approximants and the solution is computable with finitely many algebraic operations.

2012 ACM Subject Classification Mathematics of computing → Mathematical analysis; Mathematics of computing → Probability and statistics; Theory of computation → Design and analysis of algorithms; Theory of computation → Semantics and reasoning

Keywords and phrases countable splitting lemma, distributions with given marginals, couplings

Digital Object Identifier 10.4230/LIPIcs.CALCO.2021.23

Category (Co)algebraic pearls

Acknowledgements We are indebted to Paul Levy for bringing this topic to us by asking us several years ago to figure out some details of Nawrotzki's algorithm, in particular its constructivity.

1 Explanation of what is going on ...

Given a measure μ on a product space $\prod_{i\in I} X_i$, the *j-th marginal* μ_j of μ is the push-forward of μ under the j-th canonical projection $\pi_j\colon \prod_{i\in I} X_i \to X_j$. Explicitly, this is

$$\mu_j(A) := \mu\big(\pi_j^{-1}(A)\big)$$

for all $A \subseteq X_j$ with $\pi_j^{-1}(A)$ being measureable.

In his fundamental paper [21] Strassen investigated the existence of measures on a product $X \times Y$ which have prescribed marginals and satisfy additional constraints of a certain form. The result stated in Theorem 1 below is a corollary of [21, Theorem 11] and known as *Strassen's theorem on stochastic domination.* Curiously, it is not even explicitly stated in Strassen's paper, but only mentioned in one sentence. We state a slightly more general variant taken from [20, Corollary 7][1]. To formulate it, we need some notation.

- Let X be a Hausdorff space, and let $\preccurlyeq$ be a partial order on X which is closed as a subset of $X \times X$. A subset $A \subseteq X$ is *upward closed w.r.t.* $\preccurlyeq$, if

  $$\forall x \in X, y \in A.\ y \preccurlyeq x \Rightarrow x \in A.$$

- For two positive Borel measures μ, ν on X we write $\mu \preceq \nu$, if for all upward closed Borel sets $A \subseteq X$ it holds that $\mu(A) \le \nu(A)$.

[1] A different proof can be found in [16].

9th Conference on Algebra and Coalgebra in Computer Science (CALCO 2021).
Editors: Fabio Gadducci and Alexandra Silva; Article No. 23; pp. 23:1–23:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

▶ **Theorem 1.** *Let X be a Hausdorff space, let $\preccurlyeq$ be a closed partial order on X, and let μ and ν be two probability (Borel-) measures on X. If $\mu \preceq \nu$, then there exists a probability (Borel-) measure Λ on $X \times X$ which has the marginals μ and ν, and whose support is contained in $\preccurlyeq$, i.e. there exists a subset of $\preccurlyeq$ which has Λ-measure 1.*

An important particular case of Theorem 1 is when the base space X is finite or countable with the discrete topology. In the finite case this result is known as the splitting lemma [11, Theorem 4.10], and the latter is what the term "countable splitting lemma" refers to.

Over the years such results were established in different variants and on different levels of generality. For example: Strassen's original theorem [21] is proven for Polish spaces, [14] for completely regular spaces, [20] for Hausdorff spaces, [17] for probability contents instead of measures, [15] for normal measure spaces under finiteness assumptions on $\preccurlyeq$, [6] for measures with values in vector lattices under restrictions on $\preccurlyeq$, [5] for measure spaces where solutions are only required to have the given marginals up to equivalence of measures, [8] for operator valued measures, [9] for products of finitely many Polish spaces and a different proof than Strassen, [13] for Polish spaces adding some further equivalences. Some predecessors of Strassen's work are [15, 19]. A recent line of research where solutions are only required to have the given marginal up to some error is followed in [10] and related papers.

Theorem 1 plays an important role in probability theory and has applications in various areas. For example, it prominently occurs in finance mathematics, e.g. [4, 7], or in computer science, e.g. [3, 1, 2, 10, 11, 12].

The proof of Theorem 1 relies in general on a rather heavy analytic machinery, in particular, on theorems exploiting compactness properties. If X is finite, a required solution Λ can – naturally – be found by an algorithm which terminates after finitely many steps. This fact can be based on various reasoning. For example on elementary manipulations with inequalities, as e.g. in [15, §3], or combinatorial results like the max-flow min-cut theorem or the subforest lemma, as e.g. in [18] or [11, Theorem 4.10].

In the present exposition we deal with the countable discrete case. Our aim is to give a recursive algorithm which produces a sequence $(\Delta_N)_{N\in\mathbb{N}}$ of (discrete) probability measures on $X \times X$ such that

1. each term of the sequence is computable from the inital data μ, ν with a finite number of algebraic operations;
2. the sequence $(\Delta_N)_{N\in\mathbb{N}}$ converges to a solution Λ in the ℓ^1-norm on $X \times X$, in particular it converges pointwise;
3. the speed of pointwise convergence can be controlled in a computable way.

To explain our contribution, it is worthwhile to revisit the presently availabe proofs for the countable discrete case. First, specialising the general proof(s) of Theorem 1 obviously does not lead to an algorithm, since tools like e.g. the Banach-Alaoglu Theorem are used. More interesting are the arguments given in the papers of Kellerer [15, §4] and Nawrotzki [19]. Both are non-constructive, but for different reasons.

- Kellerer's approach is to reduce to the finite cases. Given μ, ν on a countable set, he produces appropriately cut-off data μ_N, ν_N, $N \in \mathbb{N}$, and solves the problem for those. This gives a measure Λ_N on X, which solves the problem up to the index N. Each measure Λ_N can be computed in finitely many steps. Sending the cut-off point N to infinity leads to existence of a solution for the full data μ, ν. The masses of the measures Λ_N may oscillate, and therefore the sequence $(\Lambda_N)_{N\in\mathbb{N}}$ need not be convergent. However, each accumulation point of the sequence $(\Lambda_N)_{N\in\mathbb{N}}$ will be a solution.

  What makes the method non-constructive is that accumulation points *exist by compactness* (in this case applied in the form of the Heine-Borel Theorem).

- Nawrotzki's approach is to produce a sequence $(\Lambda_N)_{N\in\mathbb{N}}$, which does not necessarily solve the problem on any finite section, but still converges to a solution. His construction ensures that the masses of the measures Λ_N are nonincreasing on points of the diagonal and nondecreasing off the diagonal. This ensures that passing to subsequences is not necessary.

  What makes the method non-constructive is that defining the measures Λ_N requires to evaluate *sums of infinite series* and *infima of infinite sets* of real numbers.

Our idea to produce $(\Delta_N)_{N\in\mathbb{N}}$ with **1.–3.** above, is to combine the approaches: we apply Nawrotzki's algorithm to appropriately truncated sequences to ensure computability, and control the error which is made by passing to cut-off's to ensure convergence.

2 Nawrotzki's algorithm

In [19], which preceeds the work of Strassen, Nawrotzki proved a discrete version of Strassen's theorem. In our present language his result reads as follows.

▶ **Theorem 2.** *Let $\mu = (\mu_n)_{n\in\mathbb{N}}$ and $\nu = (\nu_n)_{n\in\mathbb{N}}$ be sequences of real numbers, such that*

$$\forall n \in \mathbb{N}.\ \mu_n \geq 0 \wedge \nu_n \geq 0 \quad \textit{and} \quad \sum_{n\in\mathbb{N}} \mu_n = \sum_{n\in\mathbb{N}} \nu_n = 1, \tag{1}$$

Moreover, let $\preccurlyeq$ be a partial order on $\mathbb{N}$.

If it holds that

$$\forall R \subseteq \mathbb{N} \text{ upwards closed w.r.t. } \preccurlyeq.\ \sum_{n\in R} \mu_n \leq \sum_{n\in R} \nu_n, \tag{2}$$

then there exists an infinite matrix $\Lambda = (\lambda_{n,m})_{n,m\in\mathbb{N}}$ of real numbers, such that

$$\forall n, m \in \mathbb{N}.\ \lambda_{n,m} \geq 0 \quad \textit{and} \quad \sum_{n,m\in\mathbb{N}} \lambda_{n,m} = 1, \tag{3}$$

$$\forall n, m \in \mathbb{N}.\ \lambda_{n,m} \neq 0 \Rightarrow n \preccurlyeq m, \tag{4}$$

$$\forall n \in \mathbb{N}.\ \sum_{m\in\mathbb{N}} \lambda_{n,m} = \mu_n, \tag{5}$$

$$\forall m \in \mathbb{N}.\ \sum_{n\in\mathbb{N}} \lambda_{n,m} = \nu_m. \tag{6}$$

In this section we present Nawrotzki's argument in a structured way including all details. This provides an in-depth understanding of his work, and this is necessary to make appropriate adaption to the algorithm later on (in Section 3).

▶ Remark 3. Before we dive into the formulas and proofs, which are a bit technical and lengthy, let us give an intuition for what is going to happen.

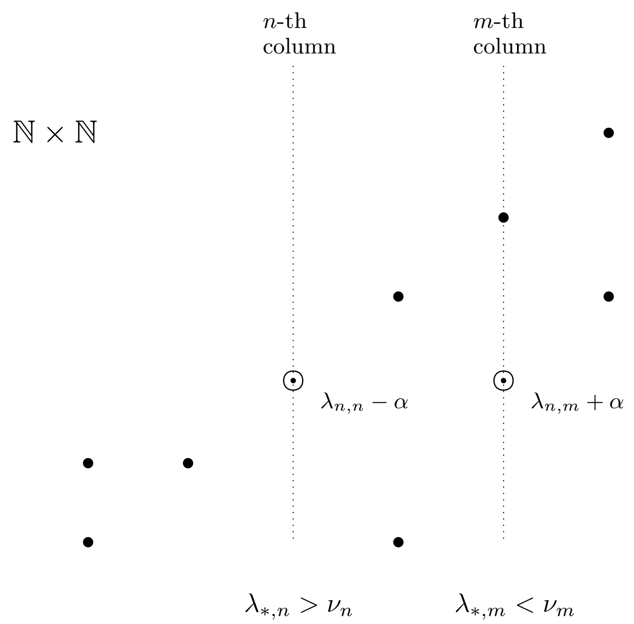
Assume we are given data μ_n, ν_m satisfying Equations (1) and (2) and a (probably bad) approximation of a solution $\lambda_{n,m}$ that satisfies Equations (3) and (4), as well as Equation (5). Note that achieving correctness of one marginal, i.e. satisfying Equation (5), is very easy; for example already the diagonal matrix with μ_n's on the diagonal will satisfy this.

If the column sums do not give the correct results as required by Equation (6), it must be that some of them are larger than the target value and some of them are smaller since the total sum is always 1. Now we want to modify the values $\lambda_{n,m}$ to improve the approximation, i.e., make the error in Equation (6) smaller while retaining all other properties. Most importantly, we have to ensure that Equation (2), also known as *stochastic dominance*, is inherited. In addition, we want to make the modification in such a way that:

1. At each place (n, m) entries change monotonically when repeating the step in the algorithm. This is achieved by having diagonal entries nonincreasing and off-diagonal entries nondecreasing. This will guarantee existence of a limit.
2. Make sure that the pattern of which column sums are too large and which are too small is inherited with exception that some column sums may become correct. This will guarantee that the algorithm can proceed appropriately.

The algorithm proceeds in steps. In each step exactly two values of the matrix change: one at the diagonal at position (n, n) and another in the same row at position (n, m) such that Equation (6) fails for n and m, as pictured below. The new values are $\lambda'_{n,n} = \lambda_{n,n} - \alpha$ and $\lambda'_{n,m} = \lambda_{n,m} + \alpha$, where α is chosen such that still $\lambda'_{*,n} \geq \nu_n$, $\lambda'_{*,m} \leq \nu_m$.

In the picture, filled circles indicate those points where our approximation has nonzero entries, circled dots mark the changes made by one step of the algorithm, and $\alpha > 0$ is the correction term whose exact definition (see Definition 7) is taylor made so that the above explained requirements are met.



The next result, Proposition 5, is the first crucial ingredient to Nawrotzki's algorithm (out of two; the second is Proposition 10 further below). It will ensure that in the limit a solution is obtained. To formulate it, we need additional notation.

▶ **Definition 4.** *Let $\preccurlyeq$ be a partial order on $\mathbb{N}$. For each $(n, m) \in \mathbb{N} \times \mathbb{N}$ with $n \prec m$, we denote*

$$\mathcal{R}_{n,m} := \big\{ R \subseteq \mathbb{N} \mid n \notin R, m \in R, R \textit{ upward closed w.r.t. } \preccurlyeq \big\}.$$

Note that $\mathcal{R}_{n,m}$ is always nonempty. For example, we have

$$\{ l \in \mathbb{N} \mid m \preccurlyeq l \} \in \mathcal{R}_{n,m}.$$

▶ **Proposition 5.** *Assume that μ, ν, and $\preccurlyeq$, satisfy Equation* (1) *and Equation* (2)*. If for each pair $(n,m) \in \mathbb{N} \times \mathbb{N}$ with $n \prec m$ at least one of*

$$\mu_n \leq \nu_n, \tag{7}$$

$$\mu_m \geq \nu_m, \tag{8}$$

$$\inf_{R \in \mathcal{R}_{n,m}} \sum_{l \in R} (\nu_l - \mu_l) = 0, \tag{9}$$

holds, then $\mu = \nu$.

Note here that all series in Equation (9) converge absolutely and that by Equation (2) the infimum in Equation (9) is nonnegative. Moreover, in an algorithm acting as explained in Remark 3 above (and defined in precise mathematical terms in Definition 7 below), using $\mathcal{R}_{n,m}$ instead of all upwards closed sets is sufficient to retain Equation (2). This is because for upwards closed sets which are not in $\mathcal{R}_{n,m}$, Equation (2) is trivially inherited.

In the proof of Proposition 5, we use the following simple fact.

▶ **Lemma 6.** *Assume that μ, ν, and $\preccurlyeq$, satisfy Equation* (1) *and Equation* (2)*. Further, let $R_1, R_2, \ldots$ be a (finite or infinite) sequence of upward closed (w.r.t. $\preccurlyeq$) subsets of $\mathbb{N}$, and set*

$$R := \bigcup_k R_k.$$

Then R is upward closed, and

$$\sum_{l \in R} (\nu_l - \mu_l) \leq \sum_k \sum_{l \in R_k} (\nu_l - \mu_l).$$

Proof. Since $|\nu_l - \mu_l| \leq \nu_l + \mu_l$, the series on the left side converges absolutely. Hence, we may rearrange summands without changing its value. Now write R as the disjoint union

$$R = \dot{\bigcup}_k R'_k$$

where

$$R'_k := R_k \setminus \bigcup_{j<k} R_j.$$

Then

$$\sum_{l \in R} (\nu_l - \mu_l) = \sum_k \sum_{l \in R'_k} (\nu_l - \mu_l).$$

For each k we have

$$\sum_{l \in R_k} (\nu_l - \mu_l) = \sum_{l \in R'_k} (\nu_l - \mu_l) + \sum_{R_k \cap \bigcup_{j<k} R_j} (\nu_l - \mu_l).$$

The set $R_k \cap \bigcup_{j<k} R_j$ is upward closed, and hence the second summand on the right side is nonnegative. This shows that

$$\sum_{l \in R'_k} (\nu_l - \mu_l) \leq \sum_{l \in R_k} (\nu_l - \mu_l)$$

for all k. ◀

Proof of Proposition 5. It is enough to show that $\mu_n \leq \nu_n$ for all $n \in \mathbb{N}$. Assume towards a contradiction that there exists $n \in \mathbb{N}$ with $\mu_n > \nu_n$, and fix one with this property. Moreover, choose $\epsilon > 0$ small enough, say,

$$\epsilon := \frac{1}{3}(\mu_n - \nu_n).$$

By the assumption of the proposition we know that for each $m \in \mathbb{N}$ with $m \succ n$ at least one of

- $\mu_m \geq \nu_m$,
- $\inf_{R \in \mathcal{R}_{n,m}} \sum_{l \in R}(\nu_l - \mu_l) = 0$,

must hold.

Consider the set where the second case takes place

$$H := \Big\{ m \in \mathbb{N} \mid n \prec m, \inf_{R \in \mathcal{R}_{n,m}} \sum_{l \in R}(\nu_l - \mu_l) = 0 \Big\}.$$

If $H = \emptyset$, it is easy to reach a contradiction. Namely, if $\mu_m \geq \nu_m$ for all $m \succ n$, then

$$\sum_{m \succcurlyeq n} \mu_m > \sum_{m \succcurlyeq n} \nu_m,$$

and this contradicts Equation (2).

If $H \neq \emptyset$, we argue as follows. For each $m \in H$ choose $R_m \in \mathcal{R}_{n,m}$, such that

$$\sum_{l \in R_m} (\nu_l - \mu_l) \leq \frac{\epsilon}{2^m},$$

and set $R := \bigcup_{m \in H} R_m$. Then $H \subseteq R$, $n \notin R$, and

$$\sum_{l \in R}(\nu_l - \mu_l) \leq \sum_{m \in H} \sum_{l \in R_m} (\nu_l - \mu_l) \leq \sum_{m \in H} \frac{\epsilon}{2^m} \leq 2\epsilon.$$

Consider the upward closed set

$$R' := R \cup \{ l \in \mathbb{N} \mid n \prec l \}.$$

If $l \in R' \setminus R$, then $n \prec l$ and $l \notin H$. Thus we must have $\mu_l \geq \nu_l$. From this we see that

$$0 \leq \sum_{l \in R'}(\nu_l - \mu_l) = \sum_{l \in R}(\nu_l - \mu_l) + \sum_{l \in R' \setminus R} (\nu_l - \mu_l) \leq \sum_{l \in R}(\nu_l - \mu_l) \leq 2\epsilon.$$

The set $R' \cup \{n\}$ is also upward closed. Using the above estimate, and recalling that $n \notin R'$, we reach the contradiction

$$0 \leq \sum_{l \in R' \cup \{n\}} (\nu_l - \mu_l) = \sum_{l \in R'}(\nu_l - \mu_l) + (\nu_n - \mu_n) \leq 2\epsilon + (\nu_n - \mu_n) = \frac{1}{3}(\nu_n - \mu_n) < 0. \blacktriangleleft$$

Nawrotzki's algorithm for the proof of Theorem 2 proceed in three steps:

1. Start with the diagonal matrix built from μ.
2. Iteratively modify this matrix in such a way, that the set of all points (n, m) where all of Equation (7)–Equation (9) fail (for certain modified sequences), gets smaller in each step.
3. Pass to the limit, so to reach a situation where Proposition 5 applies.

The single steps of the recursive process **2.** are realised by maps which act on $\ell^1(\mathbb{N}\times\mathbb{N})$. To define those maps, we first introduce an abbreviation for row- and column sums of a matrix. Given $\Lambda=(\lambda_{n,m})_{n,m\in\mathbb{N}}\in\ell^1(\mathbb{N}\times\mathbb{N})$, we denote

$$\lambda_{*,m}:=\sum_{n\in\mathbb{N}}\lambda_{n,m},\quad \lambda_{n,*}:=\sum_{m\in\mathbb{N}}\lambda_{n,m}.$$

Note that these series converge absolutely since $\Lambda\in\ell^1(\mathbb{N}\times\mathbb{N})$.

▶ **Definition 7.** *Let $\nu=(\nu_n)_{n\in\mathbb{N}}\in\ell^1(\mathbb{N})$ and $(n,m)\in\mathbb{N}\times\mathbb{N}$. We define maps*

$$\alpha^\nu_{n,m}\colon \ell^1(\mathbb{N}\times\mathbb{N})\to[0,\infty),\quad \Phi^\nu_{n,m}\colon \ell^1(\mathbb{N}\times\mathbb{N})\to\ell^1(\mathbb{N}\times\mathbb{N}).$$

- *For $\Lambda\in\ell^1(\mathbb{N}\times\mathbb{N})$ set*

$$\alpha^\nu_{n,m}(\Lambda):=\min\Big\{\lambda_{*,n}-\nu_n,\nu_m-\lambda_{*,m},\inf_{R\in\mathcal{R}_{n,m}}\sum_{l\in R}(\nu_l-\lambda_{*,l})\Big\},$$

if $n\preccurlyeq m$ and this minimum is positive, and set $\alpha^\nu_{n,m}:=0$ otherwise.
- *For $\Lambda\in\ell^1(\mathbb{N}\times\mathbb{N})$ let $\Phi^\nu_{n,m}(\Lambda)$ be the matrix with the entries*

$$\Big[\Phi^\nu_{n,m}\Big]_{l,k}(\Lambda):=\begin{cases}\lambda_{l,k}-\alpha^\nu_{n,m}(\Lambda) & \textit{if } (l,k)=(n,n),\\ \lambda_{l,k}+\alpha^\nu_{n,m}(\Lambda) & \textit{if } (l,k)=(n,m),\\ \lambda_{l,k} & \textit{otherwise.}\end{cases}$$

Note that $\Phi^\nu_{n,m}$ is well-defined, since $\alpha^\nu_{n,m}\neq 0$ implies that $n\neq m$, and since it is obvious that $\Phi^\nu_{n,m}(\Lambda)$ is again summable.

Let us collect some more obvious properties of the transformations $\Phi^\nu_{n,m}$.

▶ Remark 8. For each $\nu\in\ell^1(\mathbb{N})$ and $(n,m)\in\mathbb{N}\times\mathbb{N}$, the following statements hold.

1. $\operatorname{supp}\Phi^\nu_{n,m}(\Lambda)\subseteq\big(\operatorname{supp}\Lambda\big)\cup\{(n,n),(n,m)\}$,
2. $\forall l\in\mathbb{N}.\ \Big[\Phi^\nu_{n,m}(\Lambda)\Big]_{l,*}=\lambda_{l,*}$,
3. $\forall l\in\mathbb{N}.\ \Big[\Phi^\nu_{n,m}(\Lambda)\Big]_{*,l}=\begin{cases}\lambda_{*,l}-\alpha^\nu_{n,m}(\Lambda) & \text{if } l=n,\\ \lambda_{*,l}+\alpha^\nu_{n,m}(\Lambda) & \text{if } l=m,\\ \lambda_{*,l} & \text{otherwise.}\end{cases}$

Having $\alpha^\nu_{n,m}(\Lambda)=0$ just means that at the point (n,m) one of Equation (7)–Equation (9) holds for the sequences $(\lambda_{*,n})_{n\in\mathbb{N}}$ and $(\nu_n)_{n\in\mathbb{N}}$. Moreover, in this case, $\Phi^\nu_{n,m}$ does not change Λ. We are interested to see what happens if $\alpha^\nu_{n,m}(\Lambda)>0$.

▶ **Definition 9.** *Let $\nu\in\ell^1(\mathbb{N})$ and $\Lambda\in\ell^1(\mathbb{N}\times\mathbb{N})$. Then we set*

$$S(\Lambda):=\Big\{(n,m)\in\mathbb{N}\times\mathbb{N}\mid \alpha^\nu_{n,m}(\Lambda)>0\Big\}.$$

Moreover, we denote by $\pi_1(S(\Lambda))$ and $\pi_2(S(\Lambda))$ the projections of $S(\Lambda)$ onto the first and second, respectively, component.

To avoid bulky notation, we do not explicitly notate the dependency on ν. Moreover, observe that $S(\Lambda)$ is contained in $\preccurlyeq$ and does not intersect the diagonal, in fact,

$$\pi_1(S(\Lambda))\cap\pi_2(S(\Lambda))=\emptyset.$$

In the next proposition we show that $\Phi^\nu_{n,m}$ preserves several relevant properties and indeed shrinks the set $S(\Lambda)$.

▶ **Proposition 10.** *Let* $\nu = (\nu_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{N})$, $\Lambda \in \ell^1(\mathbb{N}\times\mathbb{N})$, *and assume that*

$$\forall n, m \in \mathbb{N}.\ \lambda_{n,m} \geq 0 \quad \text{and} \quad \sum_{n,m\in\mathbb{N}} \lambda_{n,m} = 1, \tag{10}$$

$$\forall n \in \pi_1(S(\Lambda)).\ \lambda_{*,n} = \lambda_{n,n}. \tag{11}$$

$$\forall R \subseteq \mathbb{N} \text{ upward closed w.r.t. } \preccurlyeq.\ \sum_{l\in R} \lambda_{*,l} \leq \sum_{l\in R} \nu_l, \tag{12}$$

Further, let $(n', m') \in \mathbb{N}\times\mathbb{N}$, *and assume that* $\alpha^\nu_{n',m'}(\Lambda) > 0$. *Then*

1. $\Phi^\nu_{n',m'}(\Lambda)$ *satisfies Equation* (10), *Equation* (11), *and Equation* (12),
2. $S\big(\Phi^\nu_{n',m'}(\Lambda)\big) \subseteq S(\Lambda) \setminus \{(n', m')\}$.

Proof. To shorten notation, we write

$$\Lambda' = (\lambda'_{n,m})_{n,m\in\mathbb{N}} := \Phi^\nu_{n',m'}(\Lambda).$$

We start with showing that Λ' satisfies Equation (10) and Equation (12). Let $(n, m) \neq (n', n')$. Then $\lambda'_{n,m} \geq \lambda_{n,m}$ and hence is nonnegative. For $(n, m) = (n', n')$ we use (11) to obtain

$$\lambda'_{n',n'} = \lambda_{n',n'} - \alpha^\nu_{n',m'}(\Lambda) = \lambda_{*,n'} - \alpha^\nu_{n',m'}(\Lambda) \geq \nu_{n'} \geq 0.$$

Obviously, applying $\Phi^\nu_{n',m'}$ does not change the total sums of the entries of a matrix. Thus

$$\sum_{n,m\in\mathbb{N}} \lambda'_{n,m} = \sum_{n,m\in\mathbb{N}} \lambda_{n,m} = 1.$$

We see that Equation (10) holds.

Let $R \subseteq \mathbb{N}$ be upward closed. If $R \notin \mathcal{R}_{n',m'}$, then

$$\sum_{l\in R} \lambda'_{*,l} \leq \sum_{l\in R} \lambda_{*,l} \leq \sum_{l\in R} \nu_l.$$

Next, for $R \in \mathcal{R}_{n',m'}$

$$\sum_{l\in R} \lambda'_{*,l} = \sum_{l\in R} \lambda_{*,l} + \alpha^\nu_{n',m'}(\Lambda), \tag{13}$$

and from this we find

$$\sum_{l\in R} \lambda'_{*,l} = \sum_{l\in R} \lambda_{*,l} + \alpha^\nu_{n',m'}(\Lambda) \leq \sum_{l\in R} \lambda_{*,l} + \sum_{l\in R} (\nu_n - \lambda_{*,l}) = \sum_{l\in R} \nu_l.$$

Thus Equation (12) holds.

Now we come to the proof of **2.**. This is the major part of the argument.

In the first step we show that $(n', m') \notin S(\Lambda')$. We make a case distinction according to which term is the minimum in the definition of $\alpha^\nu_{n',m'}(\Lambda)$.

- Case $\alpha^\nu_{n',m'}(\Lambda) = \lambda_{*,n'} - \nu_{n'}$:
  Then $\lambda'_{*,n'} = \nu_{n'}$, and hence $n' \notin \pi_1(S(\Lambda'))$. In particular, $(n', m') \notin S(\Lambda')$.
- Case $\alpha^\nu_{n',m'}(\Lambda) = \nu_{m'} - \lambda_{*,n'}$:
  Then $\lambda'_{*,m'} = \nu_{m'}$, and hence $m' \notin \pi_2(S(\Lambda'))$. In particular, $(n', m') \notin S(\Lambda')$.

- Case $\alpha^\nu_{n',m'}(\Lambda) = \inf_{\mathcal{R}_{n',m'}} \sum_{l\in R}(\nu_l - \lambda_{*,l})$:
  Recalling Equation (13), we find

$$\inf_{R\in\mathcal{R}_{n',m'}} \sum_{l\in R}(\nu_l - \lambda'_{*,l}) = \inf_{R\in\mathcal{R}_{n',m'}} \sum_{l\in R}\Big[(\nu_l - \lambda_{*,l}) - \alpha^\nu_{n',m'}(\Lambda)\Big] = 0.$$

  Thus also in this case $(n',m') \notin S(\Lambda')$.

In the second step, we show that $S(\Lambda') \subseteq S(\Lambda)$. Assume towards a contradiction that $(n,m) \in S(\Lambda') \setminus S(\Lambda)$. Explicitly this means that

$$n \prec m \ \wedge \ \lambda'_{*,n} > \nu_n \ \wedge \ \lambda'_{*,m} < \nu_m \ \wedge \ \inf_{R\in\mathcal{R}_{n,m}} \sum_{l\in R}(\nu_l - \lambda'_{*,l}) > 0$$
$$\wedge \ \Big[\ \lambda_{*,n} \le \nu_n \ \vee \ \lambda_{*,m} \ge \nu_m \ \vee \ \inf_{R\in\mathcal{R}_{n,m}} \sum_{l\in R}(\nu_l - \lambda_{*,l}) = 0\ \Big]$$

We distinguish cases according to the disjunction in the square bracket.

- Case $\lambda_{*,n} \le \nu_n$:
  The sum of the n-th column increases, and thus we must have $n = m'$. This implies

$$\lambda'_{*,n} = \lambda'_{*,m'} = \lambda_{*,m'} + \alpha^\nu_{n',m'}(\Lambda) \le \nu_{m'} = \nu_n,$$

  which contradicts the second term in the conjunction.
- Case $\lambda_{*,m} \ge \nu_m$:
  The sum of the m-th column decreases, and thus we must have $m = n'$. This implies

$$\lambda'_{*,m} = \lambda'_{*,n'} = \lambda_{*,n'} - \alpha^\nu_{n',m'}(\Lambda) \ge \nu_{n'} = \nu_m,$$

  which contradicts the third term in the conjunction.
- Case $\inf_{R\in\mathcal{R}_{n,m}} \sum_{l\in R}(\nu_l - \lambda_{*,l}) = 0$:
  Choose $R' \in \mathcal{R}_{n,m}$ such that

$$\sum_{l\in R'}(\nu_l - \lambda_{*,l}) < \inf_{R\in\mathcal{R}_{n,m}} \sum_{l\in R}(\nu_l - \lambda'_{*,l}).$$

  Then, in particular, the value of the sum over all $l \in R'$ decreases, and we must have $n' \in R'$ and $m' \notin R'$. Since R' is upward closed and $n' \prec m'$, this is a contradiction.

The proof of **2.** is complete.

It remains to deduce Equation (11). Let $n \in \pi_1(S(\Lambda'))$. Then also $n \in \pi_1(S(\Lambda))$, and therefore $n \ne m'$ and $\lambda_{*,n} = \lambda_{n,n}$. From the first property we obtain that the n-th column is modified at most at its diagonal entry, and now the second implies that $\lambda'_{*,n} = \lambda'_{n,n}$. ◀

Next, we investigate iterative application of maps $\Phi^\nu_{n,m}$. Start with $\nu \in \ell^1(\mathbb{N})$, $\Lambda^{(0)} \in \ell^1(\mathbb{N}\times\mathbb{N})$, and a sequence $((n_k,m_k))_{k\ge 1}$ of points in $\mathbb{N}\times\mathbb{N}$. From this data, we built the sequence $(\Lambda^{(k)})_{k\in\mathbb{N}}$ where

$$\Lambda^{(k)} := \Big[\Phi^\nu_{n_k,m_k} \circ \cdots \circ \Phi^\nu_{n_1,m_1}\Big](\Lambda^{(0)}). \tag{14}$$

It turns out that, in the situation of Theorem 2, sequences of this form converge. In fact, they do so because of a very simple reason, namely, monotonicity.

▶ **Lemma 11.** *Let $(\Lambda^{(k)})_{k\in\mathbb{N}}$ be a sequence in $\ell^1(\mathbb{N}\times\mathbb{N})$, such that*

$$\sup_{k\in\mathbb{N}}\|\Lambda^{(k)}\|_1<\infty,\quad \forall n,m,k\in\mathbb{N}.\ \lambda^{(k)}_{n,m}\geq 0,$$

and that there exists a partition $\mathbb{N}\times\mathbb{N}=A\dot{\cup}B$ such that $(\lambda^{(k)}_{n,m})_{k\in\mathbb{N}}$ is nondecreasing for all $(n,m)\in A$ and nonincreasing for all $(n,m)\in B$.

Then the limit $\Lambda:=\lim_{k\to\infty}\Lambda^{(k)}$ exists in the ℓ^1-norm.

Proof. Each of the sequences $(\lambda^{(k)}_{n,m})_{k\in\mathbb{N}}$ is monotone and bounded, hence convergent. Denote $\lambda_{n,m}:=\lim_{k\to\infty}\lambda^{(k)}_{n,m}$. We have to show that the pointwise limit $\Lambda=(\lambda_{n,m})_{n,m\in\mathbb{N}}$ is actually attained in the ℓ^1-norm. To this end we split the corresponding sum according to the given partition.

For each $(n,m)\in A$ the sequence $(\lambda^{(k)}_{n,m})_{k\in\mathbb{N}}$ is nondecreasing, and hence the monotone convergence theorem yields

$$\sum_{(n,m)\in A}\lambda_{n,m}=\lim_{k\to\infty}\sum_{(n,m)\in A}\lambda^{(k)}_{n,m}\leq\sup_{k\in\mathbb{N}}\|\Lambda^{(k)}\|_1<\infty.$$

Since $\lambda_{n,m}\geq\lambda_{n,m}-\lambda^{(k)}_{n,m}\geq 0$, we may now refer to the bounded convergence theorem to obtain that

$$\lim_{k\to\infty}\sum_{(n,m)\in A}\left|\lambda^{(k)}_{n,m}-\lambda_{n,m}\right|=0.$$

For each $(n,m)\in B$ and $k\in\mathbb{N}$ we have

$$\lambda^{(0)}_{n,m}\geq\lambda^{(k)}_{n,m}\geq\lambda^{(k)}_{n,m}-\lambda_{n,m}\geq 0.$$

Since $\sum_{(n,m)\in B}\lambda^{(0)}_{n,m}<\infty$, the bounded convergence theorem applies, and we find that

$$\lim_{k\to\infty}\sum_{(n,m)\in B}\left|\lambda^{(k)}_{n,m}-\lambda_{n,m}\right|=0.$$ ◀

▶ **Corollary 12.** *Assume that $\Lambda^{(0)}$ satisfies Equation (10) and Equation (11), let $((n_k,m_k))_{k\geq 1}$ be any sequence, and let $(\Lambda^{(k)})_{k\in\mathbb{N}}$ be defined by Equation (14). Then the limit*

$$\Lambda:=\lim_{k\to\infty}\Lambda^{(k)}$$

exists w.r.t. the ℓ^1-norm.

Proof. Since $\alpha^{\nu}_{n,m}(\Lambda)$ is always nonnegative, a partition of $\mathbb{N}\times\mathbb{N}$ required to apply Lemma 11 is obtained by taking the diagonal as the set A. ◀

Now we show that, when passing to a limit, the set $S(\Lambda)$ can be controlled.

▶ **Lemma 13.** *Let $(\Lambda^{(k)})_{k\in\mathbb{N}}$ be a sequence in $\ell^1(\mathbb{N}\times\mathbb{N})$ which converges in the ℓ^1-norm, and denote $\Lambda:=\lim_{k\to\infty}\Lambda^{(k)}$. Then*

$$S(\Lambda)\subseteq\bigcup_{N\in\mathbb{N}}\bigcap_{k\geq N}S(\Lambda^{(k)}).$$

Proof. Let $(n,m) \in S(\Lambda)$, and set $\epsilon := \frac{1}{2}\alpha^{\nu}_{n,m}(\Lambda)$. Choose $N \in \mathbb{N}$ such that

$$\forall k \geq N.\ \|\Lambda^{(k)} - \Lambda\|_1 \leq \epsilon.$$

Then for all $k \geq N$

$$\lambda^{(k)}_{*,n} \geq \lambda_{*,n} - \epsilon \geq \nu_n, \quad \lambda^{(k)}_{*,m} \leq \lambda_{*,m} + \epsilon \leq \nu_m,$$

and for all $R \in \mathcal{R}_{n,m}$

$$\sum_{l \in R} \left(\nu_l - \lambda^{(k)}_{*,l}\right) \geq \sum_{l \in R} (\nu_l - \lambda_{*,l}) - \epsilon \geq \epsilon > 0$$

Thus $(n,m) \in S(\Lambda^{(k)})$. ◀

We have collected all the neccessary tools needed for the proof of Theorem 2.

Proof of Theorem 2. Let μ, ν, and $\preccurlyeq$, be given, and assume that Equation (1) and Equation (2) hold.

Let $\Lambda^{(0)} = (\lambda^{(0)}_{n,m})_{n,m\in\mathbb{N}}$ be the diagonal matrix built from μ, i.e.,

$$\lambda^{(0)}_{n,m} := \begin{cases} \mu_n & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Choose a sequence of points $((n_k, m_k))_{k\geq 1}$ in $\mathbb{N} \times \mathbb{N}$ which covers $\prec$. For example, every enumeration of $\mathbb{N} \times \mathbb{N}$ certainly has this property. Now define $\Lambda^{(k)}$ by Equation (14) using this sequence.

By Proposition 10, each $\Lambda^{(k)}$ satisfies Equation (10), Equation (11), and Equation (12). Moreover,

$$S(\Lambda^{(k)}) \subseteq S(\Lambda^{(0)}) \setminus \big\{(n_1, m_1), \ldots, (n_k, m_k)\big\}.$$

The limit

$$\Lambda = (\lambda_{n,m})_{n,m\in\mathbb{N}} := \lim_{k\to\infty} \Lambda^{(k)}$$

exists in the ℓ^1-norm by Corollary 12, and $S(\Lambda) = \emptyset$ by Lemma 13.

Clearly, Equation (3)–Equation (5) hold for Λ. By virtue of Proposition 10, we may apply Proposition 5 with the sequences $(\lambda_{*,n})_{n\in\mathbb{N}}$ and $(\nu_n)_{n\in\mathbb{N}}$, and obtain that also Equation (6) holds. ◀

We refer to the procedure carried out in this proof as *Nawrotzki's algorithm* being *performed along* the sequence $((n_k, m_k))_{k\geq 1}$.

▶ Remark 14. For later use, we observe the following fact. Let $(\Lambda^{(k)})_{k\in\mathbb{N}}$ be a sequence produced by an application of Nawrotzki's algorithm. Then off-diagonal elements $\lambda^{(k)}_{n,m}$ change their value at most once when k runs through $\mathbb{N}$. Namely, only when $(n,m) = (n_k, m_k)$ and it happens that $\alpha^{\nu}_{n,m}(\Lambda^{(k-1)}) > 0$.

3 A constructive variant of the algorithm

Nawrotzki's proof of Theorem 2 is non-constructive for the following reason:

- The set $\mathcal{R}_{n,m}$ is in general infinite, and its elements themselves are in general infinite.

Because of this, computing the numbers $\alpha^{\nu}_{n,m}$ requires to evaluate the sum of infinite series and an infimum of an infinite set. Hence, it is not possible to compute any term of the sequence $(\Lambda^{(k)})_{k\in\mathbb{N}}$, which converges to a solution matrix Λ, with a finite number of algebraic operations.

Our aim is to give a proof of Theorem 2 which is more constructive in the following sense.

▶ **Theorem 15.** *Let $\mu, \nu, \preccurlyeq$ be given such that Equation (1) and Equation (2) hold. Then there exists a sequence $(\Delta^{(k)})_{k\in\mathbb{N}}$ of matrices in $\ell^1(\mathbb{N}\times\mathbb{N})$ with the following properties.*

1. *Each $\Delta^{(k)}$ can be computed from the given data μ and ν by a finite number of algebraic operations.*
2. *The limit $\Delta := \lim_{k\to\infty} \Delta^{(k)}$ exists in the ℓ^1-norm and satisfies Equation (3)–Equation (6).*

As usual we use the notation $\Delta^{(k)} = (\delta^{(k)}_{n,m})_{n,m\in\mathbb{N}}$ and $\Delta = (\delta_{n,m})_{n,m\in\mathbb{N}}$.

3. *For each fixed $(n,m) \in \mathbb{N}\times\mathbb{N}$ with $n \prec m$, and for each $\epsilon > 0$, a number k_0 with the property that*

$$\forall k \geq k_0.\ |\delta^{(k)}_{n,m} - \delta_{n,m}| \leq \epsilon$$

can be computed from the given data μ and ν by a finite number of algebraic operations

While the speed of pointwise convergence is controlled by the assertion in item **3.** (even in a constructive way), we have no control of the speed of ℓ^1-convergence.

The idea to prove this theorem is the simplest possible: we consider cut-off data μ_N, ν_N instead of μ, ν, apply Nawrotzki's algorithm to the truncated data, and then send the cut-off point to infinity. Realising this idea, however, requires some work.

We start with discussing convergence matters. The error when using cut-off's instead of the full data can be controlled using the following general perturbation lemma.

▶ **Lemma 16.** *Let $\nu, \tilde{\nu} \in \ell^1(\mathbb{N})$, $\Lambda, \tilde{\Lambda} \in \ell^1(\mathbb{N}\times\mathbb{N})$, and $(n,m) \in \mathbb{N}\times\mathbb{N}$. Then*

$$\left|\alpha^{\nu}_{n,m}(\Lambda) - \alpha^{\tilde{\nu}}_{n,m}(\tilde{\Lambda})\right| \leq \|\Lambda - \tilde{\Lambda}\|_1 + \|\nu - \tilde{\nu}\|_1. \tag{16}$$

Proof. We have

$$\begin{aligned}&\left|(\lambda_{*,n} - \nu_n) - (\tilde{\lambda}_{*,n} - \tilde{\nu}_n)\right| \\ &\qquad\qquad \leq \sum_{l\in\mathbb{N}} |\lambda_{l,n} - \tilde{\lambda}_{l,n}| + |\nu_n - \tilde{\nu}_n| \leq \|\Lambda - \tilde{\Lambda}\|_1 + \|\nu - \tilde{\nu}\|_1,\end{aligned}$$

and in the same way

$$\begin{aligned}&\left|(\lambda_{*,m} - \nu_m) - (\tilde{\lambda}_{*,m} - \tilde{\nu}_m)\right| \\ &\qquad\qquad \leq \sum_{l\in\mathbb{N}} |\lambda_{l,m} - \tilde{\lambda}_{l,m}| + |\nu_m - \tilde{\nu}_m| \leq \|\Lambda - \tilde{\Lambda}\|_1 + \|\nu - \tilde{\nu}\|_1.\end{aligned}$$

Next let $R \subseteq \mathbb{N}$. Then

$$\begin{aligned}&\left|\sum_{l\in R}(\nu_l - \lambda_{*,l}) - \sum_{l\in R}(\tilde{\nu}_l - \tilde{\lambda}_{*,l})\right| \leq \\ &\qquad\qquad \leq \sum_{l\in R}\sum_{k\in\mathbb{N}} |\lambda_{k,l} - \tilde{\lambda}_{k,l}| + \sum_{l\in R} |\nu_l - \tilde{\nu}_l| \leq \|\Lambda - \tilde{\Lambda}\|_1 + \|\nu - \tilde{\nu}\|_1.\end{aligned}$$

It follows that

$$\Big| \inf\Big(\{\lambda_{*,n} - \nu_n, \nu_m - \lambda_{*,m}\} \cup \Big\{\sum_{l\in R}(\nu_l - \lambda_{*,l}) \mid R \in \mathcal{R}_{n,m}\Big\}\Big)$$
$$- \inf\Big(\{\tilde\lambda_{*,n} - \tilde\nu_n, \tilde\nu_m - \tilde\lambda_{*,m}\} \cup \Big\{\sum_{l\in R}(\tilde\nu_l - \tilde\lambda_{*,l}) \mid R \in \mathcal{R}_{n,m}\Big\}\Big)\Big|$$
$$\le \|\Lambda - \tilde\Lambda\|_1 + \|\nu - \tilde\nu\|_1.$$

This is Equation (16) if $n \preccurlyeq m$. Otherwise $\alpha^{\nu}_{n,m} = \alpha^{\tilde\nu}_{n,m}(\tilde\Lambda) = 0$, and the required estimate holds trivially. ◀

▶ **Corollary 17.** *Let $\nu, \tilde\nu \in \ell^1(\mathbb{N})$, $\Lambda, \tilde\Lambda \in \ell^1(\mathbb{N}\times\mathbb{N})$, and $((n_k, m_k))_{k\ge1}$ be a sequence in $\mathbb{N}\times\mathbb{N}$. Let $(\Lambda^{(k)})_{k\in\mathbb{N}}$ and $(\tilde\Lambda^{(k)})_{k\in\mathbb{N}}$ be the sequences defined by Equation* (14) *starting from $\Lambda^{(0)} := \Lambda$ and $\tilde\Lambda^{(0)} := \tilde\Lambda$, respectively. Moreover, set*

$$\epsilon := \|\Lambda - \tilde\Lambda\|_1 + \|\nu - \tilde\nu\|_1.$$

Then

$$\forall k \in \mathbb{N}.\ \|\Lambda^{(k)} - \tilde\Lambda^{(k)}\|_1 + \|\nu - \tilde\nu\|_1 \le 3^k\epsilon.$$

Proof. For $k = 0$ this is the definition of ϵ. Then proceed inductively based on the estimate

$$\big\|\Phi^{\nu}_{n,m}(\Lambda) - \Phi^{\tilde\nu}_{n,m}(\tilde\Lambda)\big\|_1 \le \|\Lambda - \tilde\Lambda\|_1 + 2|\alpha^{\nu}_{n,m}(\Lambda) - \alpha^{\tilde\nu}_{n,m}(\tilde\Lambda)|,$$

which holds for all $\nu, \tilde\nu, \Lambda, \tilde\Lambda, n, m$. ◀

Now we turn to computability matters. To settle these, we need one more notation.

▶ **Definition 18.** *Let $L \subseteq \mathbb{N}$, and $n, m \in L$ with $n \prec m$. Then we set*

$$\mathcal{R}^L_{n,m} := \big\{R \subseteq L \mid n \notin R, m \in R, \forall k \in R, l \in L.\ k \preccurlyeq l \Rightarrow l \in R\big\}.$$

▶ **Lemma 19.** *Let $\nu \in \ell^1(\mathbb{N})$, $\Lambda \in \ell^1(\mathbb{N}\times\mathbb{N})$, let $L \subseteq \mathbb{N}$, and assume that*

$$\operatorname{supp}\nu \subseteq L, \quad \operatorname{supp}\Lambda \subseteq L \times L. \tag{17}$$

Then

$$\forall (n,m) \notin L\times L.\ \alpha^{\nu}_{n,m}(\Lambda) = 0, \tag{18}$$
$$\forall (n,m) \in \mathbb{N}\times\mathbb{N}.\ \operatorname{supp}\Phi^{\nu}_{n,m}(\Lambda) \subseteq L \times L, \tag{19}$$
$$\forall n, m \in L, n \prec m.\ \inf_{R\in\mathcal{R}_{n,m}} \sum_{l\in R}(\nu_l - \lambda_{*,l}) = \inf_{R\in\mathcal{R}^L_{n,m}} \sum_{l\in R}(\nu_l - \lambda_{*,l}). \tag{20}$$

Proof. The assumption on the supports of ν and Λ shows that

$$\forall n \notin L.\ \nu_n = \lambda_{*,n} = 0.$$

From this Equation (18), and in turn also Equation (19), follows. Moreover, for every subset $R \subseteq \mathbb{N}$

$$\sum_{l\in R}(\nu_l - \lambda_{*,l}) = \sum_{l\in R\cap L}(\nu_l - \lambda_{*,l}).$$

To establish Equation (20), we show that for all $n, m \in L$ with $n \prec m$

$$\mathcal{R}^L_{n,m} = \{R \cap L \mid \ R \in \mathcal{R}_{n,m}\}.$$

The inclusion "$\supseteq$" is clear. For the reverse inclusion observe that, for each $R \in \mathcal{R}^L_{n,m}$, the set

$$R' := \big\{l \in \mathbb{N} \mid \exists k \in R.\ \ k \preccurlyeq l\big\}$$

belongs to $\mathcal{R}_{n,m}$ and $R' \cap L = R$. ◀

▶ **Corollary 20.** *Let $\nu \in \ell^1(\mathbb{N})$ and $\Lambda \in \ell^1(\mathbb{N} \times \mathbb{N})$ be finitely supported. Then*
1. *for each $n \in \mathbb{N}$ the number $\lambda_{*,n}$ is a finite sum, and*
2. *for each $(n, m) \in \mathbb{N} \times \mathbb{N}$ the infimum in the definition of $\alpha^\nu_{n,m}(\Lambda)$ is the minimum of a finite number of finite sums.*

Proof. We can choose a finite set $L \subseteq \mathbb{N}$ such that Equation (17) holds. Then each set $\mathcal{R}^L_{n,m}$, and also each of its elements, is finite. ◀

Proof of Theorem 15. Consider truncated data: for $N \in \mathbb{N}$, let $\mu_N = (\mu_{N;n})_{n\in\mathbb{N}}$ and $\nu_N = (\nu_{N;n})_{n\in\mathbb{N}}$ be defined by

$$\mu_{N;n} := \begin{cases} \mu_n & \text{if } n < N, \\ 1 - \sum_{l<N} \mu_l & \text{if } n = N, \\ 0 & \text{if } n > N, \end{cases} \qquad \nu_{N;n} := \begin{cases} \nu_n & \text{if } n < N, \\ 1 - \sum_{l<N} \nu_l & \text{if } n = N, \\ 0 & \text{if } n > N. \end{cases}$$

We execute Nawrotzki's algorithm with the data μ_N, ν_N along the enumeration $((n_k, m_k))_{k\geq 1}$ of $\mathbb{N} \times \mathbb{N}$ which is defined by running through the scheme

$$\begin{tikzcd}
\bullet \arrow[r] & \bullet \arrow[r] & \bullet \arrow[d] & & \mathbb{N} \times \mathbb{N} \\
\bullet \arrow[u] & \bullet \arrow[l] & \bullet \arrow[d] & & \\
\bullet \arrow[r] & \bullet \arrow[u] & \bullet \arrow[r] & \bullet \arrow[u, dotted, no head] &
\end{tikzcd}$$

and dropping all points (n, m) which do not satisfy $n \prec m$.

This provides us with sequences $(\Lambda^{(k)}_N)_{k\in\mathbb{N}}$, $N \in \mathbb{N}$. According to Lemma 19 and Corollary 20, we have

$$\operatorname{supp} \Lambda^{(k)}_N \subseteq \{0, \dots, N\} \times \{0, \dots, N\},$$

and each $\Lambda^{(k)}_N$ can be computed by a finite number of algebraic operations.

Let $(\Lambda^{(k)})_{k\in\mathbb{N}}$ be the sequence obtained by running Nawrotzki's algorithm along the same sequence $((n_k, m_k))_{k\geq 1}$ but starting with the full data μ, ν. We have

$$\|\Lambda^{(0)} - \Lambda^{(0)}_N\|_1 = 2 \sum_{n>N} \mu_n, \quad \|\nu - \nu\|_1 = 2 \sum_{n>N} \nu_n,$$

and hence

$$\|\Lambda^{(0)} - \Lambda^{(0)}_N\|_1 + \|\nu - \nu\|_1 = 2 \sum_{n>N} (\mu_n + \nu_n) = 2\Big(2 - \sum_{n\leq N} (\mu_n + \nu_n)\Big) =: \epsilon_N.$$

Corollary 17 applies and leads to the basic estimate

$$\forall k \in \mathbb{N}, N \in \mathbb{N}.\ \ \|\Lambda^{(k)} - \Lambda^{(k)}_N\|_1 + \|\nu - \nu\|_1 \leq 3^k \epsilon_N. \tag{21}$$

The next step is to define a sequence $(\Delta_k)_{k\in\mathbb{N}}$. This is done as follows: given $k \in N$, choose $N_k \in \mathbb{N}$ with

$$\epsilon_{N_k} \leq \frac{1}{k \cdot 3^k},$$

and set $\Delta_k := \Lambda^{(k)}_{N_k}$.

The number N_k can be found in finitely many steps by summing up beginning sections of μ and ν. Together with what we already observed above, thus, each Δ_k can be computed in finitely many steps.

We know that the limit $\Lambda := \lim_{k\to\infty} \Lambda^{(k)}$ exists in the ℓ^1-norm and satisfies Equation (3) – Equation (6). The basic estimate Equation (21) yields

$$\|\Lambda^{(k)} - \Delta^{(k)}\|_1 \leq \frac{1}{k},$$

and we see that also $\lim_{k\to\infty} \Delta^{(k)} = \Lambda$ in the ℓ^1-norm.

Let $(n, m) \in \mathbb{N} \times \mathbb{N}$ with $n \prec m$ and $\epsilon > 0$ be given. Define $k_0 \in \mathbb{N}$ as the least integer larger or equal to

$$\max\left\{\frac{1}{\epsilon}, \big(\max\{n, m\}\big)^2\right\}.$$

Then $(n, m) \in \{(n_1, m_1), \ldots, (n_{k_0}, m_{k_0})\}$ and for all $k \geq k_0$

$$\|\Lambda^{(k)} - \Delta^{(k)}\|_1 \leq \epsilon.$$

Now recall Remark 14: the entry $\lambda^{(k)}_{n,m}$ is constant for $k \geq k_0$. This implies that, for all $k \geq k_0$,

$$|\lambda_{n,m} - \delta^{(k)}_{n,m}| = |\lambda^{(k)}_{n,m} - \delta^{(k)}_{n,m}| \leq \|\Lambda^{(k)} - \Delta^{(k)}\|_1 \leq \epsilon.$$

The proof of Theorem 15 is complete. ◀

References

1. G. Barthe, T. Espitau, J. Hsu, T. Sato, and P.-Y. Strub. *-liftings for differential privacy. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, volume 80 of *LIPIcs*, pages 102:1–102:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.ICALP.2017.102.
2. G. Barthe, T. Espitau, J. Hsu, T. Sato, and P.-Y. Strub. Relational ⋆⋆\star-liftings for differential privacy. *Log. Methods Comput. Sci.*, 15(4), 2019. doi:10.23638/LMCS-15(4:18)2019.
3. G. Barthe, M. Gaboardi, B. Grégoire, J. Hsu, and P.-Y. Strub. Proving differential privacy via probabilistic couplings. In Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors, *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 749–758. ACM, 2016. doi:10.1145/2933575.2934554.
4. M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016. doi:10.1214/14-AOP966.
5. P. Berti, L. Pratelli, P. Rigo, and F. Spizzichino. Equivalent or absolutely continuous probability measures with given marginals. *Depend. Model.*, 3(1):47–58, 2015. doi:10.1515/demo-2015-0004.

6 E. D'Aniello and J.D.M. Wright. Finding measures with given marginals. *Q. J. Math.*, 51(4):405–416, 2000. doi:10.1093/qjmath/51.4.405.

7 M.H.A. Davis and D.G. Hobson. The range of traded option prices. *Math. Finance*, 17(1):1–14, 2007. doi:10.1111/j.1467-9965.2007.00291.x.

8 S. Friedland, J. Ge, and L. Zhi. Quantum Strassen's theorem. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 23(3):2050020, 29, 2020. doi:10.1142/S0219025720500204.

9 N. Gaffke and L. Rüschendorf. On the existence of probability measures with given marginals. *Statist. Decisions*, 2(1-2):163–174, 1984.

10 J. Hsu. *Probabilistic couplings for probabilistic reasoning.* Phd thesis, University of Pennsylvania, 2017. URL: https://repository.upenn.edu/edissertations/3017/.

11 C. Jones. *Probabilistic Non-determinism.* Phd thesis, University of Edinburgh, 1989.

12 A. Jung and R. Tix. The troublesome probabilistic powerdomain. In *Third Workshop on Computation and Approximation (Comprox III) (Birmingham, 1997)*, volume 13 of *Electron. Notes Theor. Comput. Sci.*, page 22. Elsevier Sci. B. V., Amsterdam, 1998.

13 T. Kamae, U. Krengel, and G.L. O'Brien. Stochastic inequalities on partially ordered spaces. *Ann. Probability*, 5(6):899–912, 1977. doi:10.1214/aop/1176995659.

14 J. Kawabe. A type of Strassen's theorem for positive vector measures with values in dual spaces. *Proc. Amer. Math. Soc.*, 128(11):3291–3300, 2000. doi:10.1090/S0002-9939-00-05384-3.

15 H.G. Kellerer. Funktionen auf Produkträumen mit vorgegebenen Marginal-Funktionen. *Math. Ann.*, 144:323–344, 1961. doi:10.1007/BF01470505.

16 H.G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984. doi:10.1007/BF00532047.

17 H. König. On the marginals of probability contents on lattices. *Mathematika*, 58(2):319–323, 2012. doi:10.1112/S0025579311002427.

18 V.T. Koperberg. Couplings and matchings. Bachelor thesis, Universiteit Leiden, 2016. URL: https://www.universiteitleiden.nl/binaries/content/assets/science/mi/scripties/koperbergbach.pdf.

19 K. Nawrotzki. Eine Monotonieeigenschaft zufälliger Punktfolgen. *Math. Nachr.*, 24:193–200, 1962. doi:10.1002/mana.19620240305.

20 H.J. Skala. The existence of probability measures with given marginals. *Ann. Probab.*, 21(1):136–142, 1993. URL: http://links.jstor.org/sici?sici=0091-1798(199301)21:1<136:TEOPMW>2.0.CO;2-I&origin=MSN.

21 V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965. doi:10.1214/aoms/1177700153.