$\star$-Liftings for Differential Privacy[*][†]

Gilles Barthe[1], Thomas Espitau[2], Justin Hsu[3], Tetsuya Sato[4], and Pierre-Yves Strub[5]

1 **IMDEA Software Institute, Madrid, Spain**
gjbarthe@gmail.com

2 **Sorbonne Universités, UPMC Paris 6, Paris, France**
t.espitau@gmail.com

3 **University of Pennsylvania, Philadelphia, PA, USA**
email@justinh.su

4 **Research Institute for Mathematical Sciences, Kyoto University, Kyoto, Japan**
satoutet@kurims.kyoto-u.ac.jp

5 **École Polytechnique, Palaiseau, France**
pierre-yves@strub.nu

Abstract

Recent developments in formal verification have identified *approximate liftings* (also known as *approximate couplings*) as a clean, compositional abstraction for proving differential privacy. There are two styles of definitions for this construction. Earlier definitions require the existence of one or more witness distributions, while a recent definition by Sato uses universal quantification over all sets of samples. These notions have different strengths and weaknesses: the universal version is more general than the existential ones, but the existential versions enjoy more precise composition principles.

We propose a novel, existential version of approximate lifting, called $\star$-*lifting*, and show that it is equivalent to Sato's construction for discrete probability measures. Our work unifies all known notions of approximate lifting, giving cleaner properties, more general constructions, and more precise composition theorems for both styles of lifting, enabling richer proofs of differential privacy. We also clarify the relation between existing definitions of approximate lifting, and generalize our constructions to approximate liftings based on f-divergences.

1998 ACM Subject Classification D.2.4 Software/Program Verification

Keywords and phrases Differential Privacy, Probabilistic Couplings, Formal Verification

Digital Object Identifier 10.4230/LIPIcs.ICALP.2017.102

1 Introduction

Differential privacy [7] is a rigorous notion of statistical privacy that delivers strong individual guarantees for privacy-preserving computations. Informally, differential privacy guarantees to every individual that their (non)-participation in a database will have a small (in a rigorous, quantitative sense) effect on the results obtained by third parties when querying the database. The formal definition of differential privacy is parametrized by two non-negative real numbers, (ϵ, δ). These parameters quantify the effect of individuals on the output of the private query;

[*] A full version of the paper is available at https://arxiv.org/abs/1705.00133.

[†] This work is partially supported by a grant from the NSF (TWC-1513694) and a grant from the Simons Foundation (#360368 to Justin Hsu).

44th International Colloquium on Automata, Languages, and Programming (ICALP 2017).
Editors: Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl;
Article No. 102; pp. 102:1–102:12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

smaller values give stronger privacy guarantees. The main strengths of differential privacy lie in its theoretical elegance, minimal assumptions, and flexibility for many applications.

Motivated by the importance of differential privacy, programming language researchers have developed approaches based on dynamic analysis, type systems, and program logics for formally proving differential privacy for programs. (We refer the interested reader to a recent survey [4] for an overview of this growing field.) In this paper, we consider approaches based on relational program logics [5, 6, 10, 2, 3, 11]. To capture the quantitative nature of differential privacy, these systems rely on a quantitative generalization of probabilistic couplings (see, e.g., [9, 13, 14]), called *approximate liftings* or (ϵ, δ)-liftings. Existing works have considered several potential definitions. While all definitions support compositional reasoning and enable program logics that can verify complex examples from the privacy literature, the various notions of approximate liftings have different strengths and weaknesses.

Broadly speaking, one class of definitions require the existence of one or two *witness distributions* that "couple' the two executions of programs. The earliest definition [5] supports accuracy-based reasoning for the Laplace mechanism, while subsequent definitions [6, 10] support more precise composition principles from differential privacy and can be generalized to other notions of distance on distributions. These definitions, and their associated program logics, were designed for discrete distributions.

In the course of extending these ideas to continuous distributions, Sato [11] proposes a radically different notion of approximate lifting, which does not rely on witness distributions. Instead, it uses a universal quantification over all sets of samples. Sato shows that this definition is strictly more general than the existential versions, but it is unclear (a) whether the gap can be closed and (b) whether his construction satisfies the same composition principles enjoyed by some existential definitions.

As a consequence, there is currently no single approximate lifting with the properties needed to support all existing formalized proofs of differential privacy. Furthermore, some of the most involved privacy proofs cannot be formalized at all, as their proofs require a combination of tools from several kinds of approximate liftings.

Outline of the paper

After reviewing the necessary mathematical preliminaries in Section 2, we introduce our main technical contribution: a new, existential definition of approximate lifting. This construction, which we call ⋆*-lifting*, is a generalization of an existing definition by Barthe and Olmedo [6, 10]. The key idea is to allow the witness distributions to have a larger domain, broadening the class of approximate liftings. By a maximum flow/minimum cut argument, we show that ⋆-liftings are equivalent to Sato's lifting over discrete distributions. This equivalence can be viewed as an approximate version of Strassen's theorem [12], a classical result in probability theory describing the existence of probabilistic couplings. We present the definition of ⋆-lifting and the proof of equivalence in Section 3.

Then, we show that ⋆-liftings satisfy desirable theoretical properties. We are able to leverage the equivalence of liftings in two ways. In one direction, Sato's definition gives simpler proofs of more general properties of ⋆-liftings. In the other direction, ⋆-liftings – like other existential definitions – can smoothly incorporate composition principles from the theory of differential privacy. Our connection shows that Sato's definition can use these principles in the discrete case. We describe the key theoretical properties of ⋆-liftings in Section 4.

Finally, we provide a thorough comparison of ⋆-lifting with existing definitions of approximate lifting in Section 5, and describe how to construct ⋆-liftings for more general version of approximate liftings based on f-divergences in Section 6.

Overall, the equivalence of $\star$-liftings and Sato's lifting, along with the natural theoretical properties satisfied by the common notion, suggest that these definitions are two views on the same concept: an approximate version of probabilistic coupling.

2 Background

To model probabilistic behavior, we work with *discrete sub-distributions.*

▶ **Definition 1.** A *sub-distribution* over a set A is defined by its mass function $\mu : A \to \mathbb{R}^+$, which gives the probability of the singleton events $a \in A$. This mass function must be s.t. $|\mu| \triangleq \sum_{a\in A} \mu(a)$ is well-defined and at most 1. In particular, the *support* $\operatorname{supp}(\mu) \triangleq \{a \in A \mid \mu(a) \neq 0\}$ must be discrete (i.e. finite or countably infinite). When the *weight* $|\mu|$ is equal to 1, we call μ a *(proper) distribution.* We let $\mathbb{D}(A)$ denote the set of sub-distributions over A. The probability of an event $E(x)$ w.r.t. μ, written $\mathbb{P}_{x\sim\mu}[E(x)]$ or $\mathbb{P}_\mu[E]$, is defined as $\sum_{x\in A|E(x)} \mu(x)$.

Simple examples of sub-distributions include the *null sub-distribution* $\mathbb{0}^A \in \mathbb{D}(A)$, which maps each element of A to 0, and the *Dirac distribution centered on* x, written $\mathbb{1}_x$, which maps x to 1 and all other elements to 0. One can equip distributions with a monadic structure using the Dirac distributions $\mathbb{1}_x$ for the unit and *distribution expectation* $\mathbb{E}_{x\sim\mu}[f(x)]$ for the bind; if μ is a distribution over A and f has type $A \to \mathbb{D}(B)$, then the bind defines a sub-distribution over B: $\mathbb{E}_{a\sim\mu}[f(a)] : b \mapsto \sum_a \mu(a) \cdot f(a)(b)$.

If $f : A \to B$, we can lift f to a function $f^\sharp : \mathbb{D}(A) \to \mathbb{D}(B)$ as follows: $f^\sharp(\mu) \triangleq \mathbb{E}_{a\sim\mu}[\mathbb{1}_{f(a)}]$ – or, equivalently, $f^\sharp(\mu) : b \mapsto \mathbb{P}_{a\sim\mu}[a \in f^{-1}(b)]$. For instance, when working with sub-distributions over pairs, this allows to obtain the probabilistic versions $\pi_1^\sharp$ and $\pi_2^\sharp$ (called *marginals*) of the usual projections π_1 and π_2. One can check that the *first* and *second marginals* $\pi_1^\sharp(\mu)$ and $\pi_2^\sharp(\mu)$ of a distribution μ over $A \times B$ are also given by the following equations: $\pi_1^\sharp(\mu)(a) = \sum_{b\in B} \mu(a, b)$ and $\pi_2^\sharp(\mu)(b) = \sum_{a\in A} \mu(a, b)$. When $f : A \to \mathbb{D}(B)$, we will abuse notation and write the lifting $f^\sharp : \mathbb{D}(A) \to \mathbb{D}(B)$ to mean $f^\sharp(\mu) \triangleq \mathbb{E}_{x\sim\mu}[f(x)]$.

Finally, if $\alpha : A \to \mathbb{R}^+$, we write $\alpha[X] \in \mathbb{R}^+ \cup \{\infty\}$ for $\sum_{x\in X} \alpha(x)$. Moreover, if $\alpha : A \times B \to \mathbb{R}^+$, we write $\alpha[X, Y]$ (resp. $\alpha[x, Y]$, $\alpha[X, y]$) for $\alpha[X \times Y]$ (resp. $\alpha[\{x\} \times Y$, $\alpha[X \times \{y\}]$). Note that for a sub-distribution $\mu \in \mathbb{D}(A)$ and an event $E \subseteq A$, $\mathbb{P}_\mu[E] = \mu[E]$.

We now review the definition of differential privacy.

▶ **Definition 2** (Dwork et al. [7])**.** A probabilistic computation $M : A \to \mathbb{D}(B)$ satisfies (ϵ, δ)-*differential privacy* w.r.t. an adjacency relation $\phi \subseteq A \times A$ iff for every pair of inputs $a, a' \in A$ such that $a \mathrel{\phi} a'$ and every subset of outputs $E \subseteq B$,

$$\mathbb{P}_{M(a)}[E] \leq e^\epsilon \cdot \mathbb{P}_{M(a')}[E] + \delta.$$

It is useful to define a notion of distance on distributions, reflecting differential privacy.

▶ **Definition 3** (Barthe and Olmedo [5], Barthe et al. [6], Olmedo [10])**.** Let $\epsilon \geq 0$. The ϵ-*DP divergence* $\Delta_\epsilon(\mu_1, \mu_2)$ between two sub-distributions $\mu_1, \mu_2 \in \mathbb{D}(B)$ is defined as

$$\sup_{E\subseteq B} \left(\mathbb{P}_{\mu_1}[E] - e^\epsilon \cdot \mathbb{P}_{\mu_2}[E]\right).$$

Then, differential privacy admits an alternative characterization based on DP divergence.

▶ **Lemma 4.** *A probabilistic computation* $M : A \to \mathbb{D}(B)$ *satisfies* (ϵ, δ)-differential privacy *w.r.t. an adjacency relation* $\phi \subseteq A \times A$ *iff* $\Delta_\epsilon(M(a), M(a')) \leq \delta$ *for every pair of inputs* $a, a' \in A$ *such that* $a \mathrel{\phi} a'$*.*

Our new definition of approximate lifting is inspired by a version of approximate liftings involving two witness distributions, proposed by Barthe and Olmedo [6], Olmedo [10].

▶ **Definition 5** (Barthe and Olmedo [6], Olmedo [10]). Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\epsilon, \delta \in \mathbb{R}^+$ and $\mathcal{R}$ be a binary relation over A & B. An (ϵ, δ)-*approximate* 2-*lifting* of μ_1 & μ_2 for $\mathcal{R}$ is a pair $(\mu_\triangleleft, \mu_\triangleright)$ of sub-distributions over $A \times B$ s.t.

1. $\pi_1^\sharp(\mu_\triangleleft) = \mu_1$ and $\pi_2^\sharp(\mu_\triangleright) = \mu_2$;
2. $\Delta_\epsilon(\mu_\triangleleft, \mu_\triangleright) \leq \delta$; and
3. $\text{supp}(\mu) \subseteq \mathcal{R}$.

We write $\mu_1 \mathcal{R}^{(2)}_{\epsilon,\delta} \mu_2$ if there exists an (ϵ, δ)-approximate (2-)lifting of μ_1 & μ_2 for $\mathcal{R}$; the (2) indicates that there are two witnesses in this definition of lifting.

Combined with Lemma 4, a probabilistic computation $M : A \to \mathbb{D}(B)$ is (ϵ, δ)-differentially private if and only if for every two adjacent inputs $a \,\phi\, a'$, there is an approximate lifting of the equality relation: $M(a) =^{(2)}_{\epsilon,\delta} M(a')$.

2-liftings can be generalized by varying the notion of distance given by Δ_ϵ; we will return to this point in Section 6. These liftings also satisfy useful theoretical properties, but some of the properties are not as general as we would like. For example, it is known that 2-liftings satisfy the following mapping property.

▶ **Theorem 6** (Barthe et al. [2]). *Let $\mu_1 \in \mathbb{D}(A_1)$, $\mu_2 \in \mathbb{D}(A_2)$, $f_1 : A_1 \to B_1$, $f_2 : A_2 \to B_2$* surjective *maps and $\mathcal{R}$ a binary relation on B_1 & B_2. Then*

$$f_1^\sharp(\mu_1) \,\mathcal{R}^{(2)}_{\epsilon,\delta}\, f_2^\sharp(\mu_2) \iff \mu_1 \,\mathcal{S}^{(2)}_{\epsilon,\delta}\, \mu_2$$

where $a_1 \,\mathcal{S}\, a_2 \overset{\triangle}{\iff} f_1(a_1) \,\mathcal{R}\, f_2(a_2)$.

This property can be used to pull back an approximate lifting on two distributions over B_1, B_2 to an approximate lifting on two distributions over A_1, A_2. For applications in program logics, B_1, B_2 could be the domain of a program variable, A_1, A_2 could be the set of memories, and f_1, f_2 could project a memory to a program variable. While the mapping theorem is quite useful, it is puzzling why it only applies to surjective maps. For instance, this theorem cannot be used when the maps f_1, f_2 embed a smaller space into a larger space.

For another example, there exist 2-liftings of the following form, sometimes called the *optimal subset coupling*.

▶ **Theorem 7** (Barthe et al. [2]). *Let $\mu \in \mathbb{D}(A)$ and consider two subsets $P_1 \subseteq P_2 \subseteq A$. Suppose that P_2 is a* strict subset *of A. Then, we have the following equivalence:*

$$\mathbb{P}_\mu[P_2] \leq e^\epsilon \cdot \mathbb{P}_\mu[P_1] \iff \mu \,\mathcal{R}^{(2)}_{\epsilon,0}\, \mu,$$

where $a_1 \,\mathcal{R}\, a_2 \overset{\triangle}{\iff} a_1 \in P_1 \iff a_2 \in P_2$.

In this construction, it is puzzling why the larger subset P_2 must be a *strict* subset of the domain A. For example, this theorem does not apply for $P_2 = A$, but we may be able to construct the approximate lifting if we simply embed A into a larger space B – even though μ has support over A! Furthermore, it is not clear why the subsets must be nested, nor is it clear why we can only relate μ to itself.

These shortcomings suggest that the definition of 2-liftings may be problematic. While the distance condition appears to be the most constraining requirement, the marginal and support conditions are responsible for the main issues.

Witnesses can only use pairs in the relation

For some relations $\mathcal{R}$, there may be elements a such that $a \mathrel{\mathcal{R}} b$ does not hold for any b, or vice versa. It can be impossible find witnesses with the correct marginals on these elements, even if the distance condition can be easily satisfied. For instance, we can sometimes construct a pair $\mu_\triangleleft$ and $\mu_\triangleright$ satisfying the distance requirement, but where $\mu_\triangleright$ needs additional mass to achieve the marginal requirement for an element b. Adding this mass anywhere preserves the distance bound, but there may not be an element a such that $a \mathrel{\mathcal{R}} b$.

No canonical choice of witnesses

A related problem is that the marginal requirement only constrains one marginal of each witness distribution. Along the other component, the witnesses may place the mass anywhere on any pair in the relation. As a result, witnesses to an approximate lifting $\mu_1 \mathrel{\mathcal{R}^{(2)}_{\epsilon,\delta}} \mu_2$ may have mass outside of $\operatorname{supp}(\mu_1) \times \operatorname{supp}(\mu_2)$, even though it seems that only elements in the support should be relevant to the lifting.

3 $\star$-Liftings and Strassen's Theorem

To improve the theoretical properties of 2-liftings, we propose a simple extension: allow witnesses to be distributions over a larger set.

▶ **Notation 8.** Let A be a set. We write $A^\star$ for $A \uplus \{\star\}$.

▶ **Definition 9** ($\star$-lifting)**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\epsilon, \delta \in \mathbb{R}^+$ and $\mathcal{R}$ be a binary relation over A & B. An (ϵ, δ)-*approximate* $\star$-*lifting* of μ_1 & μ_2 for $\mathcal{R}$ is a pair of sub-distributions $\eta_\triangleleft \in \mathbb{D}(A \times B^\star)$ and $\eta_\triangleright \in \mathbb{D}(A^\star \times B)$ s.t.

1. $\pi_1^\sharp(\eta_\triangleleft) = \mu_1$ and $\pi_2^\sharp(\eta_\triangleright) = \mu_2$;
2. $\operatorname{supp}(\eta_{\triangleleft|A\times B}), \operatorname{supp}(\eta_{\triangleright|A\times B}) \subseteq \mathcal{R}$; and
3. $\Delta_\epsilon(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta$, where $\overline{\eta_\bullet}$ is the canonical lifting of $\eta_\bullet$ to $A^\star \times B^\star$.

We write $\mu_1 \mathrel{R^{(\star)}_{\epsilon,\delta}} \mu_2$ if there exists an (ϵ, δ)-approximate lifting of μ_1 & μ_2 for $\mathcal{R}$.

By adding an element $\star$, we address both problems discussed at the end of the previous section. First, for every $a \in A$, witnesses may place mass at $(a, \star)$; for every $b \in B$, witnesses may place mass at $(\star, b)$. Second, $\star$ can serve as a generic element where all mass that lies outside the supports $\operatorname{supp}(\mu_1) \times \operatorname{supp}(\mu_2)$ may be placed, while preserving the marginal and distance requirements, giving more control over the form of the witnesses.

▶ **Lemma 10.** *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be distributions such that $\mu_1 \mathrel{\mathcal{R}^{(\star)}_{\epsilon,\delta}} \mu_2$. Then, there are witnesses with support contained in* $\operatorname{supp}(\mu_1)^\star \times \operatorname{supp}(\mu_2)^\star$.

3.1 Basic Properties

$\star$-liftings satisfy all basic properties satisfied by other notions of lifting. We start by proving that this new definition of lifting still characterizes differential privacy.

▶ **Lemma 11.** *A randomized algorithm $P : A \to \mathbb{D}(B)$ is (ϵ, δ)-differentially private for ϕ iff for all $a_1, a_2 \in A$, $a_1 \mathrel{\phi} a_2$ implies $P(a_1) =^{(\star)}_{\epsilon,\delta} P(a_2)$.*

The next lemma establishes several other basic properties of $\star$-liftings: monotonicity, and closure under relational and sequential composition.

▶ **Lemma 12.**

- *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, and $\mathcal{R}$ be a binary relation over A & B. If $\mu_1 \; \mathcal{R}^{(\star)}_{\epsilon,\delta} \; \mu_2$, then for any $\epsilon' \geq \epsilon$, $\delta' \geq \delta$ and $\mathcal{S} \supseteq \mathcal{R}$, we have $\mu_1 \; \mathcal{S}^{(\star)}_{\epsilon',\delta'} \; \mu_2$.*
- *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, $\mu_2 \in \mathbb{D}(C)$ and $\mathcal{R}$ (resp. $\mathcal{S}$) be a binary relation over A & B (resp. over B & C). If $\mu_1 \; \mathcal{R}^{(\star)}_{\epsilon,\delta} \; \mu_2$ and $\mu_2 \; \mathcal{S}^{(\star)}_{\epsilon',\delta'} \; \mu_3$, then $\mu_1 \; (\mathcal{S} \circ \mathcal{R})^{(\star)}_{\epsilon+\epsilon',\delta+e^{\epsilon}\cdot\delta'} \; \mu_3$.*
- *For $i \in \{1,2\}$, let $\mu_i \in \mathbb{D}(A_i)$ and $\eta_i : A_i \to \mathbb{D}(B_i)$. Let $\mathcal{R}$ (resp. $\mathcal{S}$) be a binary relation over A_1 & A_2 (resp. over B_1 & B_2). If $\mu_1 \; \mathcal{R}^{(\star)}_{\epsilon,\delta} \; \mu_2$ for some $\epsilon, \delta \geq 0$ and for any $(a_1, a_2) \in \mathcal{R}$, $\eta_1(a_1) \; \mathcal{S}^{(\star)}_{\epsilon',\delta'} \; \eta_2(a_2)$ for some $\epsilon', \delta' \geq 0$, then*

$$\mathbb{E}_{\mu_1}[\eta_1] \; \mathcal{S}^{(\star)}_{\epsilon+\epsilon',\delta+\delta'} \; \mathbb{E}_{\mu_2}[\eta_2].$$

3.2 Equivalence with Sato's Definition

In recent work on verifying differential privacy over general, continuous distributions, Sato [11] proposes an alternative definition of approximate lifting. In the special case of discrete distributions, where measurability of events can be forgotten, his definition can be stated as follows.

▶ **Definition 13** (Sato [11])**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$, $\mathcal{R}$ be a binary relation over A & B and $\epsilon, \delta \geq 0$. Then, there is an (ϵ, δ)-approximate lifting of μ_1 & μ_2 for $\mathcal{R}$ if

$$\forall X \subseteq A.\, \mu_1[X] \leq e^{\epsilon} \cdot \mu_2[\mathcal{R}(X)] + \delta.$$

Notice that this definition has no witness distributions at all; instead, it uses a universal quantifier over all subsets. We can show that ⋆-liftings are equivalent to Sato's definition in the case of discrete distributions. This equivalence is reminiscent of Strassen's theorem from probability theory, which characterizes the existence of probabilistic couplings.

▶ **Theorem 14** (Strassen [12])**.** *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$ be two proper distributions, and $\mathcal{R}$ let be a binary relation over A & B. Then there exists a joint distribution $\mu \in \mathbb{D}(A \times B)$ with support in $\mathcal{R}$ such that $\pi_1^{\sharp}(\mu) = \mu_1$ and $\pi_2^{\sharp}(\mu) = \mu_2$ if and only if*

$$\forall X \subseteq A.\, \mu_1[X] \leq \mu_2[\mathcal{R}(X)].$$

Our result (Theorem 19) can be viewed as a generalization of Strassen's theorem to approximate couplings. The key ingredient in our proof is the *max-flow min-cut* theorem for *countable* networks; we begin by reviewing the basic setting.

▶ **Definition 15** (Flow network)**.** A *flow network* is a structure $((V, E), \top, \bot, c)$ s.t. $\mathcal{N} = (V, E)$ is a loop-free directed graph without infinite simple path (or rays), $\top$ and $\bot$ are two distinct distinguished vertices of $\mathcal{N}$ s.t. no edge starts from $\bot$ and ends at $\top$, and $c : E \to \mathbb{R}^{+} \cup \{+\infty\}$ is a function assigning to each edge of $\mathcal{N}$ a capacity. The capacity c is extended to V^2 by assigning capacity 0 to any pair (u, v) s.t. $(u, v) \notin E$.

▶ **Definition 16** (Flow)**.** Given a flow network $\mathcal{N} \triangleq ((V, E), \top, \bot, c)$, a function $f : V^2 \to \mathbb{R}$ is a *flow* for $\mathcal{N}$ iff

1. $\forall u, v \in V.\, f(u, v) \leq c(u, v)$,
2. $\forall u, v \in V.\, f(u, v) = -f(v, u)$, and
3. $\forall u \in V.\, u \notin \{\top, \bot\} \implies \sum_{v \in V} f(u, v) = 0$ (Kirchhoff's Law).

The *mass* $|f|$ of a flow f is defined as $|f| \triangleq \sum_{v \in V} f(\top, v) \in \mathbb{R}\{\cup + \infty\}$.

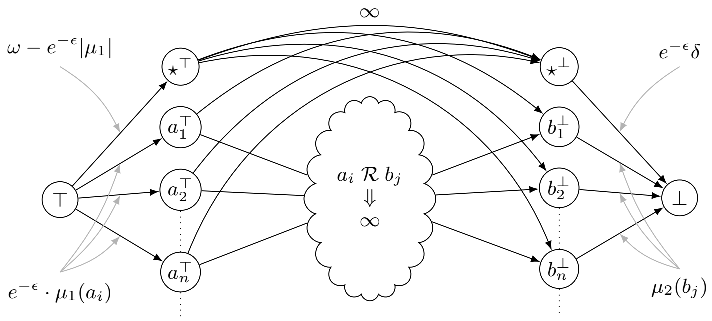

Figure 1 Flow Network in Theorem 19.

▶ **Definition 17** (Cut). Given a flow network $\mathcal{N} \triangleq ((V, E), \top, \perp, c)$, a *cut* for $\mathcal{N}$ is any set $C \subseteq V$ that partition V s.t. $\top \in V$ but $\perp \notin V$. The *cut-set* $\mathcal{E}(C)$ of a cut C is defined as: $\{(u, v) \in E \mid u \in S, v \notin S\}$. The *capacity* $|C| \in \mathbb{R}^+ \cup \{\infty\}$ of a cut is defined as $|C| \triangleq \sum_{(u,v)\in\mathcal{E}(C)} c(u, v)$.

For flow networks with finitely many vertices an edges, the maximum flow is equal to the minimum cut. Aharoni et al. [1] consider when this is the case for a countable network. For the flow networks that we consider in this paper – where there are no infinite directed paths – equality holds.

▶ **Theorem 18** (Weak Countable Max-Flow Min-Cut). *Let $\mathcal{N}$ be a network flow. Then,*

$$\sup\{|f| \mid f \textit{ is a flow for } \mathcal{N}\} = \inf\{|C| \mid C \textit{ is a cut for } \mathcal{N}\}$$

and both the supremum and infimum are reached.

We are now ready to prove an approximate version of Strassen's theorem, thereby showing equivalence between $\star$-liftings and Sato's liftings.

▶ **Theorem 19.** *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$, $\mathcal{R}$ be a binary relation over A & B and $\epsilon, \delta \in \mathbb{R}^+$. Then, $\mu_1 \, R^{(\star)}_{\epsilon,\delta} \, \mu_2$ iff $\forall X \subseteq A.\, \mu_1(X) \leq e^\epsilon \cdot \mu_2(\mathcal{R}(X)) + \delta$.*

Proof. We only detail the reverse direction. We can assume that A and B are countable; in the case where A and B are not both countable, we first consider the restriction of μ_1 and μ_2 to their respective supports – which are countable sets – and construct witnesses to the $\star$-lifting. The witnesses can then be extended to a coupling of μ_1 and μ_2 by adding a null mass to the extra points.

Let $\omega \triangleq |\mu_2| + e^{-\epsilon} \cdot \delta$ and let $\top$ and $\perp$ be fresh symbols. For any set X, define $X^\top$ and $X^\perp$ resp. as $\{x^\top \mid x \in X\}$ and $\{x^\perp \mid x \in X\}$. Let $\mathcal{N}$ be the flow network of Figure 1 whose resp. source and sink are $\top$ and $\perp$, whose set of vertices V is $\{\top, \perp\} \uplus (A^\star)^\top \uplus (B^\star)^\perp$, and whose set of edges E is $E_\top \uplus E_\perp \uplus E_\mathcal{R} \uplus E_\star$ with

$$E_\top \triangleq \{\top \mapsto_{\mu_1(a)} a^\top \mid a \in A\} \qquad E_\perp \triangleq \{b^\perp \mapsto_{e^{-\epsilon}\mu_2(b)} \perp \mid b \in B\}$$

$$E_\mathcal{R} \triangleq \{a^\top \mapsto_\infty b^\perp \mid a \,\mathcal{R}\, b \vee a = \star \vee b = \star\} \qquad E_\star \triangleq \{\top \mapsto_{(\omega - e^{-\epsilon}|\mu_1|)} \star^\top,\ \star^\perp \mapsto_{e^{-\epsilon}\delta} \perp\}.$$

Let C be a cut of $\mathcal{N}$ – in the following, we use C independently for the cut C and its cut-set $\mathcal{E}(C)$. We check $|C| \geq \omega$. If $C \cap E_\mathcal{R} \neq \emptyset$ then $|C| = \infty$. Note that $C \cap E_\star = \emptyset$ implies $C \cap E_\mathcal{R} \neq \emptyset$. If $(\top, \star^\top) \in C$ and $(\perp, \star^\perp) \notin C$ then we must have $E_\top \subseteq C$. This implies that $|C| \geq \omega$ since $E_\top \uplus \{(\top, \star^\top)\}$ is a cut with capacity ω. If $(\top, \star^\top) \notin C$ and $(\perp, \star^\perp) \in C$ then we have $|C| \geq \omega$ in the similar way as above. Otherwise (i.e. $C \cap E_{\mathcal{R}} = \emptyset$ and $E_\star \subseteq C$), for C to be a cut, we must have $\mathcal{R}(A - A^\dagger) \subseteq B^\dagger$ where $A^\dagger \triangleq \{x \in A \mid (\top, x^\top) \in C\}$ and $B^\dagger \triangleq \{y \in B \mid (y^\perp, \perp) \in C\}$. Thus,

$$\begin{aligned}
|C| &= e^{-\epsilon} \cdot \mu_1[A^\dagger] + \mu_2[B^\dagger] + |E_\star| \\
&\geq e^{-\epsilon} \cdot \mu_1[A^\dagger] + \mu_2[\mathcal{R}(A - A^\dagger)] + e^{-\epsilon} \cdot \delta + (\omega - e^{-\epsilon} \cdot |\mu_1|) \\
&\geq e^{-\epsilon} \cdot (\mu_1[A^\dagger] + \mu_1[A - A^\dagger]) + \omega - e^{-\epsilon} \cdot |\mu_1| = \omega.
\end{aligned}$$

Hence, $E_\top \uplus \{(\star^\perp, \perp)\}$ is a minimum cut with capacity ω. By Theorem 18, we obtain a maximum flow f with mass ω. Note that the flow f saturates the capacity of all edges in $E_\top$, $E_\perp$, and $E_\star$. Let $\hat{f} : (a, b) \in A^\star \times B^\star \mapsto f(a^\top, b^\perp)$. We now define the following distributions:

$$\begin{aligned}
\eta_\triangleleft &: A \times B^\star \to \mathbb{R}^+ & \eta_\triangleright &: A^\star \times B \to \mathbb{R}^+ \\
&(a, b) \mapsto e^\epsilon \cdot \hat{f}(a, b) & &(a, b) \mapsto \hat{f}(a, b).
\end{aligned}$$

We clearly have $\pi_1^\sharp(\eta_\triangleleft) = \mu_1$ and $\pi_2^\sharp(\eta_\triangleright) = \mu_2$. Moreover, by construction of the flow network $\mathcal{N}$, $\mathrm{supp}(\hat{f}_{|A \times B}) \subseteq \mathcal{R}$. Hence, $\mathrm{supp}(\eta_{\triangleleft|A \times B}), \mathrm{supp}(\eta_{\triangleright|A \times B}) \subseteq \mathcal{R}$. It remains to show that $\Delta_\epsilon(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta$. Let X be a subset of $A^\star \times B^\star$. Let $\overline{X_a} \triangleq \{a \in A \mid (a, \star) \in X\}$, $\overline{X_b} \triangleq \{b \in B \mid (\star, b) \in X\}$ and $\overline{X} \triangleq X \cap (A \times B)$. Then,

$$\begin{aligned}
\overline{\eta_\triangleleft}[X] - e^\epsilon \cdot \overline{\eta_\triangleright}[X] &= e^\epsilon \left(\hat{f}[\overline{X}] + \hat{f}[\overline{X_a} \times \{\star\}]\right) - e^\epsilon \left(\hat{f}[\overline{X}] + \hat{f}[\{\star\} \times \overline{X_b}]\right) \\
&\leq e^\epsilon \cdot \hat{f}[\overline{X_a} \times \{\star\}] \leq e^\epsilon \cdot \hat{f}[A \times \{\star\}] = \delta.
\end{aligned}$$

The last equality holds by Kirchhoff's law: $\hat{f}[A \times \{\star\}] = \sum_{a \in A} f(a^\top, \star^\perp) = f(\star^\perp, \perp) = e^{-\epsilon} \cdot \delta$. ◀

4 Properties of ⋆-Liftings

Our main theorem can be used to show a variety of natural properties of ⋆-liftings. To begin, we can generalize the mapping property from Theorem 6, lifting the requirement that the maps must be surjective.

▶ **Lemma 20.** *Let $\mu_1 \in \mathbb{D}(A_1)$, $\mu_2 \in \mathbb{D}(A_2)$, $f_1 : A_1 \to B_1$, $f_2 : A_2 \to B_2$ and $\mathcal{R}$ a binary relation on B_1 & B_2. Let $\mathcal{S}$ such that $a_1 \mathrel{\mathcal{S}} a_2 \overset{\triangle}{\iff} f_1(a_1) \mathrel{\mathcal{R}} f_2(a_2)$. Then*

$$f_1^\sharp(\mu_1) \mathrel{\mathcal{R}_{\epsilon,\delta}^{(\star)}} f_2^\sharp(\mu_2) \iff \mu_1 \mathrel{\mathcal{S}_{\epsilon,\delta}^{(\star)}} \mu_2.$$

Similarly, we can generalize the existing rules for up-to-bad reasoning (cf. Barthe et al. [2, Theorem 13]), which restrict the post-condition to be equality. There are two versions: the conditional event is either on the left side, or the right side. Note that the resulting index $\overline{\delta}$ are different in the two cases.

▶ **Lemma 21.** *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, $\theta \subseteq A$ and $\mathcal{R} \subseteq A \times B$. Assume that $\mu_1 \ (\theta_\triangleleft \implies \mathcal{R})_{\epsilon,\delta}^{(\star)} \ \mu_2$ for some parameters $\epsilon, \delta \geq 0$. Then, $\mu_1 \mathrel{\mathcal{R}_{\epsilon,\overline{\delta}}^{(\star)}} \mu_2$, where $\overline{\delta} \triangleq \delta + \mu_1[\overline{\theta}]$.*

▶ **Lemma 22.** *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, $\theta \subseteq B$ and $\mathcal{R} \subseteq A \times B$. Assume that $\mu_1 \ (\theta_\triangleright \implies \mathcal{R})_{\epsilon,\delta}^{(\star)} \ \mu_2$ for some parameters $\epsilon, \delta \geq 0$. Then, $\mu_1 \mathrel{\mathcal{R}_{\epsilon,\overline{\delta}}^{(\star)}} \mu_2$, where $\overline{\delta} \triangleq \delta + e^\epsilon \cdot \mu_2[\overline{\theta}]$.*

As a consequence, an approximately lifted relation can be conjuncted with a one-sided predicate if the δ parameter is increased. This principle is useful for constructing approximate liftings that express *accuracy* bounds: when $\theta_{a,\lhd}$ is an event that happens with high probability, we can assume that $\theta_{a,\lhd}$ holds if we increase the δ parameter of the approximate lifting.

▶ **Lemma 23.** *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, $\theta_a \subseteq A$, $\theta_b \subseteq B$ and $\mathcal{R} \subseteq A \times B$. Assume that $\mu_1 \ \mathcal{R}^{(\star)}_{\epsilon,\delta} \ \mu_2$. Then, $\mu_1 \ (\theta_{a,\lhd} \wedge \mathcal{R})^{(\star)}_{\epsilon,\delta_a} \ \mu_2$ and $\mu_1 \ (\theta_{b,\rhd} \wedge \mathcal{R})^{(\star)}_{\epsilon,\delta_b} \ \mu_2$ where $\delta_a \triangleq \delta + \mu_1[\overline{\theta_a}]$ and $\delta_b \triangleq \delta + e^{\epsilon} \cdot \mu_2[\overline{\theta_b}]$.*

$\star$-liftings also support a significant generalization of optimal subset coupling. Unlike the known construction for 2-liftings (Theorem 7), the two subsets need not be nested, and either subset may be the entire domain. Furthermore, the distributions μ_1, μ_2 need not be the same, or even have the same domain. Finally, the equivalence is valid for any parameters (ϵ, δ), not just $\delta = 0$.

▶ **Theorem 24** (Barthe et al. [2])**.** *Let $\mu_1 \in \mathbb{D}(A_1)$, $\mu_2 \in \mathbb{D}(A_2)$ and consider two subsets $P_1 \subseteq A_1, P_2 \subseteq A_2$. Then, we have the following equivalence:*

$$\mathbb{P}_{\mu_1}[P_1] \leq e^{\epsilon} \cdot \mathbb{P}_{\mu_2}[P_2] + \delta \wedge \mathbb{P}_{\mu_1}[A_1 - P_1] \leq e^{\epsilon} \cdot \mathbb{P}_{\mu_2}[A_2 - P_2] + \delta \iff \mu_1 \ \mathcal{R}^{(\star)}_{\epsilon,\delta} \ \mu_2,$$

where $a_1 \ \mathcal{R} \ a_2 \overset{\triangle}{\iff} a_1 \in P_1 \iff a_2 \in P_2$.

Proof. Immediate by Theorem 19. ◀

Finally, we can directly extend known composition theorems from differential privacy to $\star$-liftings. This connection is quite useful for lifting existing results from the privacy literature–which can be quite sophisticated – to approximate liftings.

▶ **Lemma 25.** *Pose $\mathbb{R}^+_2 \triangleq \mathbb{R}^+ \times \mathbb{R}^+$ and let $(\mathbb{R}^+_2)^*$ be the set of finite sequences over $\mathbb{R}^+_2$. Let $r : (\mathbb{R}^+_2)^* \to \mathbb{R}^+_2$ be a* DP-composition operator*, i.e. r is an operator such that for any sets A, D and family $\{f_i : D \times A \to \mathbb{D}(A)\}_{i<n}$ of functions, if for every $a \in A$ and $i < n$, $f_i(-, a) : D \to \mathbb{D}(A)$ is (ϵ_i, δ_i)-differentially private for some parameters $\epsilon_i, \delta_i \geq 0$ and fixed adjacency relation ϕ, then, for any $a \in A$, $F(-, a)$ is (ϵ^*, δ^*)-differentially private for ϕ, where $F : (d, a) \mapsto (\bigcirc_{i<n} \ (f_i(d, -))^{\sharp})(\mathbb{1}_a)$ is the the n-fold composition of the $[f_i]_{i<n}$ and $(\epsilon^*, \delta^*) \triangleq r([(\epsilon_i, \delta_i)]_{i<n})$.*

Let $n \in \mathbb{N}$ and assume given two families of sets $\{A_i\}_{i \leq n}$ and $\{B_i\}_{i \leq n}$, together with a family of binary relations $\{\mathcal{R}(i) \subseteq A_i \times B_i\}_{i \leq n}$. Fix two families of functions $\{g_i : A_i \to \mathbb{D}(A_{i+1})\}_{i<n}$ and $\{h_i : B_i \to \mathbb{D}(B_{i+1})\}_{i<n}$ s.t. for any $i < n$ and $(a, b) \in \mathcal{R}(i)$ we have:
1. *$g_i(a) \ \mathcal{R}(i+1)^{(\star)}_{\epsilon_i,\delta_i} \ h_i(b)$ for some parameters $\epsilon_i, \delta_i \geq 0$, and*
2. *$g_i(a)$ and $h_i(b)$ are proper distributions.*
Then, for $(a_0, b_0) \in \mathcal{R}_0$, there exists a $\star$-lifting

$$G(a_0) \ \mathcal{R}(n)^{(\star)}_{\epsilon^*,\delta^*} \ H(b_0)$$

where $(\epsilon^*, \delta^*) \triangleq r([(\epsilon_i, \delta_i)]_{i<n})$, and $G : A_0 \to \mathbb{D}(A_n)$ and $H : B_0 \to \mathbb{D}(B_n)$ are the n-fold compositions of $[g_i]_{i \leq n}$ and $[h_i]_{i \leq n}$ respectively – i.e. $G(a) \triangleq (\bigcirc_{i<n} \ g_i^{\sharp})(\mathbb{1}_a)$ and $H(b) \triangleq (\bigcirc_{i<n} \ h_i^{\sharp})(\mathbb{1}_b)$.

For some of the more sophisticated composition results (notably, the advanced composition theorem by Dwork et al. [8]), Lemma 25 is not quite strong enough and requires a slight adaptation of the notion of $\star$-lifting. We refer to the full version of the paper for more details.

5 Comparison with Existing Approximate Liftings

Now that we have seen $\star$-liftings, we briefly consider other definitions of approximate liftings. We have already seen 2-liftings, which involve two witnesses (Definition 5). Evidently, $\star$-liftings strictly generalize 2-liftings.

▶ **Theorem 26.** *For all binary relations $\mathcal{R}$ over A & B and parameters $\epsilon, \delta \geq 0$, we have $\mathcal{R}^{(2)}_{\epsilon,\delta} \subseteq \mathcal{R}^{(\star)}_{\epsilon,\delta}$. There exist relations and parameters where the inclusion is strict.*

Proof. The inclusion $\mathcal{R}^{(2)}_{\epsilon,\delta} \subseteq \mathcal{R}^{(\star)}_{\epsilon,\delta}$ is immediate. We have a strict inclusion $\mathcal{R}^{(2)}_{\epsilon,\delta} \subsetneq \mathcal{R}^{(\star)}_{\epsilon,\delta}$ even for $\delta = 0$ by considering the optimal subset coupling from Theorem 7. Consider a distribution μ over set A, and let $P_1 \subseteq P_2 = A$. There is an $(\epsilon, 0)$-approximate $\star$-lifting (by Theorem 24), but a $(\epsilon, 0)$-approximate 2-lifting does not exist if μ has non-zero mass outside of P_1: the first witness $\mu_\triangleleft$ must place non-zero mass at (a_1, a_2) with $a_1 \notin P_1$ in order to have $\pi_1^\sharp(\mu_\triangleleft) = \mu$, but we must have $a_2 \notin P_2$ for the support requirement, and there is no such a_2. ◀

It is more interesting to compare $\star$-liftings with the original definitions of (ϵ, δ)-approximate lifting, by Barthe et al. [5]. They introduce two notions, a symmetric lifting and an asymmetric lifting, each using a single witness distribution. We will focus on the asymmetric version.

▶ **Definition 27** (Barthe et al. [5])**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\epsilon, \delta \in \mathbb{R}^+$ and $\mathcal{R}$ be a binary relation over A & B. An (ϵ, δ)-*approximate* 1-*lifting* of μ_1 & μ_2 for $\mathcal{R}$ is a sub-distribution $\mu \in \mathbb{D}(A \times B)$ s.t.

1. $\pi_1^\sharp(\mu) \leq \mu_1$ and $\pi_2^\sharp(\mu) \leq \mu_2$;
2. $\Delta_\epsilon(\mu_1, \pi_1^\sharp(\mu)) \leq \delta$; and
3. $\text{supp}(\mu) \subseteq \mathcal{R}$.

In the first point we take the point-wise order on sub-distributions: if μ and μ' are sub-distributions over X, then $\mu \leq \mu'$ when $\mu(x) \leq \mu'(x)$ for all $x \in X$. We will write $\mu_1 \; \mathcal{R}^{(1)}_{\epsilon,\delta} \; \mu_2$ if there exists an (ϵ, δ)-approximate 1-lifting of μ_1 & μ_2 for $\mathcal{R}$; the (1) indicates that there is one witness for this lifting.

1-liftings bear a close resemblance to *probabilistic couplings* from probability theory, which also have a single witness. However, 1-liftings are less well-understood theoretically than 2-liftings – basic properties such as mapping (Theorem 20) are not known to hold; the subset coupling (Theorem 7) is not known to exist.

Somewhat surprisingly, 1-liftings are equivalent to $\star$-liftings (and hence by Theorem 19, also to Sato's approximate lifting).

▶ **Theorem 28.** *For all binary relations $\mathcal{R}$ over A & B and parameters $\epsilon, \delta \geq 0$, we have $\mathcal{R}^{(1)}_{\epsilon,\delta} = \mathcal{R}^{(\star)}_{\epsilon,\delta}$.*

6 $\star$-Lifting for f-Divergences

The definition of $\star$-lifting can be extended to lifting constructions based on general f-divergences, as previously proposed by Barthe and Olmedo [6], Olmedo [10]. Roughly, a f-divergence a function $\Delta_f(\mu_1, \mu_2)$ that measures the difference between two probability distributions μ_1 and μ_2. Much like we generalized their definition for (ϵ, δ)-liftings, we can define $\star$-lifting with f-divergences. Before going any further, let us first define formally f-divergences. We denote by $\mathcal{F}$ the set of non-negative convex functions vanishing at 1: $\mathcal{F} = \{f : \mathbb{R}^+ \to \mathbb{R}^+ \mid f(1) = 0\}$. We also adopt the following notational conventions: $0 \cdot f(0/0) \triangleq 0$, and $0 \cdot f(x/0) \triangleq x \cdot \lim_{t \to 0^+} t \cdot f(1/t)$; we write L_f for the limit.

▶ **Definition 29.** Given $f \in \mathcal{F}$, the f-*divergence* $\Delta_f(\mu_1, \mu_2)$ between two distributions μ_1 and μ_2 in $\mathbb{D}(A)$ is defined as:

$$\Delta_f(\mu_1, \mu_1) = \sum_{a \in A} \nu(a) f\left(\frac{\mu_1(a)}{\mu_2(a)}\right).$$

Examples of f-divergences include statistical distance ($f(t) = \frac{1}{2}|t-1|$), Kullback-Leibler divergence ($f(t) = \ln(t) - t + 1$), and Hellinger distance ($f(t) = \frac{1}{2}(\sqrt{t}-1)^2$).

▶ **Definition 30** ($\star$-lifting for f-divergences)**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be distributions, $\mathcal{R}$ be a binary relation over A & B, and $f \in \mathcal{F}$. An $(f;\delta)$-*approximate lifting* of μ_1 & μ_2 for $\mathcal{R}$ is a pair of distributions $\eta_\triangleleft \in \mathbb{D}(A \times B^\star)$ and $\eta_\triangleright \in \mathbb{D}(A^\star \times B)$ s.t.

- $\pi_1^\sharp(\eta_\triangleleft) = \mu_1$ and $\pi_2^\sharp(\eta_\triangleright) = \mu_2$;
- $\mathrm{supp}(\eta_{\triangleleft|A\times B}), \mathrm{supp}(\eta_{\triangleright|A\times B}) \subseteq \mathcal{R}$; and
- $\Delta_f(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta$,

where $\overline{\eta_\bullet}$ is the canonical lifting of $\eta_\bullet$ to $A^\star \times B^\star$. We will write: $\mu_1 \; R^{(\star)}_{f;\delta} \; \mu_2$ if there exists an $(f;\delta)$-approximate lifting of μ_1 & μ_2 for $\mathcal{R}$.

$\star$-liftings for f-divergences compose sequentially.

▶ **Lemma 31.** *Suppose f has divergence statistical distance, Kullback-Leibler, or Hellinger distance. For $i \in \{1,2\}$, let $\mu_i \in \mathbb{D}(A_i)$ and $\eta_i : A_i \to \mathbb{D}(B_i)$. Let $\mathcal{R}$ (resp. $\mathcal{S}$) be a binary relation over A_1 & A_2 (resp. over B_1 & B_2). If $\mu_1 \; \mathcal{R}^{(\star)}_{f;\delta} \; \mu_2$ for some $\delta \geq 0$ and for any $(a_1, a_2) \in \mathcal{R}$ we have $\eta_1(a_1) \; \mathcal{S}^{(\star)}_{f;\delta'} \; \eta_2(a_2)$ for some $\delta' \geq 0$, then*

$$\mathbb{E}_{\mu_1}[\eta_1] \; \mathcal{S}^{(\star)}_{f;\delta+\delta'} \; \mathbb{E}_{\mu_2}[\eta_2].$$

Much like the $\star$-liftings we saw before, $\star$-liftings for f-divergences have witness distributions with support determined by the support of μ_1 and μ_2 (cf. Lemma 10).

▶ **Lemma 32.** *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be distributions such that $\mu_1 \; R^{(\star)}_{f;\delta} \; \mu_2$. Then, there are witnesses with support contained in* $\mathrm{supp}(\mu_1)^\star \times \mathrm{supp}(\mu_2)^\star$.

Finally, the mapping property from Lemma 20 holds also for these $\star$-liftings. While the proof of Lemma 20 relies on the equivalence for Sato's definition, there is no such equivalence (or definition) for general f-divergences. Therefore, we must work directly with the witnesses of the approximate lifting.

▶ **Lemma 33.** *Let $\mu_1 \in \mathbb{D}(A_1)$, $\mu_2 \in \mathbb{D}(A_2)$, $g_1 : A_1 \to B_1$, $g_2 : A_2 \to B_2$ and $\mathcal{R}$ a binary relation on B_1 & B_2. Let $\mathcal{S}$ such that $a_1 \; \mathcal{S} \; a_2 \overset{\triangle}{\iff} g_1(a_1) \; \mathcal{R} \; g_2(a_2)$. Then*

$$g_1^\sharp(\mu_1) \; \mathcal{R}^{(\star)}_{f;\delta} \; g_2^\sharp(\mu_2) \iff \mu_1 \; \mathcal{S}^{(\star)}_{f;\delta} \; \mu_2.$$

7 Conclusion

We have proposed a new definition of approximate lifting that unifies existing constructions and satisfies an approximate variant of Strassen's theorem. Our notion is useful both to simplify the soundness proof of existing program logics and to strengthen some of their proof rules. We see at least two important directions for future work. First, adapting existing program logics (for instance, apRHL [5]) to use $\star$-liftings, and formalizing examples that were out of reach of previous systems. Second, our notion of $\star$-liftings only applies when distributions have discrete support. It would be interesting to see if $\star$-liftings – and the approximate Strassen's theorem – can be generalized to the continuous setting.

Acknowledgments. We thank the anonymous reviewers for their helpful suggestions.

References

1 Ron Aharoni, Eli Berger, Agelos Georgakopoulos, Amitai Perlstein, and Philipp Sprüssel. The max-flow min-cut theorem for countable networks. *J. Comb. Theory, Ser. B*, 101(1):1–17, 2011. doi:10.1016/j.jctb.2010.08.002.

2 Gilles Barthe, Noémie Fong, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. Advanced probabilistic couplings for differential privacy. In *ACM SIGSAC Conference on Computer and Communications Security (CCS), Vienna, Austria*, 2016. URL: https://arxiv.org/abs/1606.07143.

3 Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. Proving differential privacy via probabilistic couplings. In *IEEE Symposium on Logic in Computer Science (LICS), New York, New York*, 2016. URL: http://arxiv.org/abs/1601.05047.

4 Gilles Barthe, Marco Gaboardi, Justin Hsu, and Benjamin C. Pierce. Programming language techniques for differential privacy. *SIGLOG News*, 3(1):34–53, 2016. doi:10.1145/2893582.2893591.

5 Gilles Barthe, Boris Köpf, Federico Olmedo, and Santiago Zanella-Béguelin. Probabilistic relational reasoning for differential privacy. *ACM Transactions on Programming Languages and Systems*, 35(3):9, 2013. URL: http://software.imdea.org/~bkoepf/papers/toplas13.pdf.

6 Gilles Barthe and Federico Olmedo. Beyond differential privacy: Composition theorems and relational logic for f-divergences between probabilistic programs. In *International Colloquium on Automata, Languages and Programming (ICALP), Riga, Latvia*, volume 7966 of *Lecture Notes in Computer Science*, pages 49–60. Springer-Verlag, 2013. URL: http://certicrypt.gforge.inria.fr/2013.ICALP.pdf.

7 Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *IACR Theory of Cryptography Conference (TCC), New York, New York*, pages 265–284, 2006. doi:10.1007/11681878_14.

8 Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. Boosting and differential privacy. In *IEEE Symposium on Foundations of Computer Science (FOCS), Las Vegas, Nevada*, pages 51–60, 2010. URL: http://research.microsoft.com/pubs/155170/dworkrv10.pdf.

9 Torgny Lindvall. *Lectures on the coupling method.* Courier Corporation, 2002.

10 Federico Olmedo. *Approximate Relational Reasoning for Probabilistic Programs.* PhD thesis, Universidad Politécnica de Madrid, 2014. URL: http://software.imdea.org/~federico/thesis.pdf.

11 Tetsuya Sato. Approximate relational Hoare logic for continuous random samplings. In *Conference on the Mathematical Foundations of Programming Semantics (MFPS), Pittsburgh, Pennsylvania*, volume 325 of *Electronic Notes in Theoretical Computer Science*, pages 277–298. Elsevier, 2016. URL: https://arxiv.org/abs/1603.01445.

12 Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, pages 423–439, 1965. URL: http://projecteuclid.org/euclid.aoms/1177700153.

13 Hermann Thorisson. *Coupling, Stationarity, and Regeneration.* Springer-Verlag, 2000.

14 Cédric Villani. *Optimal transport: old and new.* Springer-Verlag, 2008.