Logical Methods in Computer Science
Volume 15, Issue 4, 2019, pp. 18:1–18:32
https://lmcs.episciences.org/

Submitted Mar. 16, 2018
Published Dec. 19, 2019

RELATIONAL $\star$-LIFTINGS FOR DIFFERENTIAL PRIVACY

GILLES BARTHE, THOMAS ESPITAU, JUSTIN HSU, TETSUYA SATO, AND PIERRE-YVES STRUB

IMDEA Software Institute, Spain and MPI for Security and Privacy, Germany
e-mail address: gilles.barthe@imdea.org

Sorbonne Universités, UPMC Paris 6, France
e-mail address: t.espitau@gmail.com

University of Wisconsin–Madison, USA
e-mail address: email@justinh.su

Seikei University, Japan
e-mail address: t_sato@st.seikei.ac.jp

École Polytechnique, France
e-mail address: 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. This construction can be defined in two styles. 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 each have their own strengths: the universal version is more general than the existential ones, while existential liftings are known to satisfy 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, yielding 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 consider more general approximate liftings based on f-divergences.

1. Introduction

Differential privacy (Dwork et al., 2006) is a strong, rigorous notion of statistical privacy. Informally, differential privacy guarantees to every individual that their participation in a database query will have a quantitatively small effect on the query results, limiting the amount that the query answer depends on their private data. The definition of differential privacy is parametrized by two non-negative real numbers, (ε, δ), which quantify the effect of individuals on the output of the private query: smaller values give stronger privacy guarantees.

Key words and phrases: probabilistic, relational lifting, differential privacy, max flow/min cut.

LOGICAL METHODS IN COMPUTER SCIENCE
DOI:10.23638/LMCS-15(4:18)2019

The main strengths of differential privacy lie in its theoretical elegance, minimal assumptions, and flexibility.

Recently, 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 (Barthe et al., 2016c) for an overview of this growing field.) In this paper, we consider approaches based on relational program logics (Barthe and Olmedo, 2013; Barthe et al., 2013, 2016a,b; Olmedo, 2014; Sato, 2016). To capture the quantitative nature of differential privacy, these systems rely on a quantitative generalization of probabilistic couplings (see, e.g., (Lindvall, 2002; Thorisson, 2000; Villani, 2008)), called *approximate liftings* or (ε, δ)-liftings. Prior 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, the first class of definitions require the existence of one or two *witness distributions* that "couple" the output distributions from program executions on two related inputs (intuitively, the true database and the true database omitting one individual's record). The earliest definition (Barthe et al., 2013) supports accuracy-based reasoning for the Laplace mechanism, while subsequent definitions (Barthe and Olmedo, 2013; Olmedo, 2014) 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 (2016) proposes a radically different notion of approximate lifting that 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, no single definition is known to satisfy 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 constructions satisfied by existential or universal liftings, but not both.

Outline of the paper. After introducing mathematical preliminaries in Section 2, we introduce our main technical contribution: a new, existential definition of approximate lifting. This construction, which we call $\star$-*lifting*, is a generalization of an existing definition by Barthe and Olmedo (2013); Olmedo (2014). The key idea is to slightly enlarge the domain of witness distributions with a single generic point, broadening the class of approximate liftings. By a maximum flow/minimum cut argument, we show that $\star$-liftings are equivalent to Sato's lifting over discrete distributions. This equivalence can be viewed as an approximate version of Strassen's theorem (Strassen, 1965), a classical result in probability theory characterizing the existence of probabilistic couplings. We present our definition and the proof of equivalence in Section 3.

Then, we show that $\star$-liftings satisfy desirable theoretical properties by leveraging the equivalence of liftings in two ways. In one direction, Sato's definition gives simpler proofs of more general properties of $\star$-liftings. In the other direction, $\star$-liftings—like previously proposed existential liftings—can smoothly incorporate composition principles from the theory of differential privacy. In particular, our connection shows that Sato's definition can use these principles in the discrete case. We describe the key theoretical properties of $\star$-liftings in Section 4.

Finally, we provide a thorough comparison of $\star$-lifting with other existing definitions of approximate lifting in Section 5, introduce a symmetric version of $\star$-lifting that satisfies the so-called advanced composition theorem from differential privacy Dwork et al. (2010) in Section 6, and generalize $\star$-liftings to approximate liftings based on f-divergences in Section 7.

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 data, we work with *discrete sub-distributions.*

Definition 2.1. A *sub-distribution* over a set A is defined by its mass function $\mu : A \to [0,1]$, 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* $\mathrm{supp}(\mu) \triangleq \{a \in A \mid \mu(a) \neq 0\}$ 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. *Events* E are predicates on 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 the usual 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 via $\mathbb{E}_{a\sim\mu}[f(a)] : b \mapsto \sum_a \mu(a) \cdot f(a)(b)$.

If $f : A \to B$, we can lift f to a function $f^\sharp : \mathbb{D}(A) \to \mathbb{D}(B)$ as $f^\sharp(\mu) \triangleq \mathbb{E}_{a\sim\mu}[\mathbb{1}_{f(a)}]$; more explicitly, $f^\sharp(\mu) : b \mapsto \mathbb{P}_{a\sim\mu}[a \in f^{-1}(b)]$. For instance, when working with sub-distributions over pairs, we have the probabilistic versions $\pi_1^\sharp : \mathbb{D}(A\times B) \to \mathbb{D}(A)$ and $\pi_2^\sharp : \mathbb{D}(A\times B) \to \mathbb{D}(B)$ (called the *marginals*) of the usual projections π_1 and π_2 by lifting. 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 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 $f^\sharp(\mu) \triangleq \mathbb{E}_{x\sim\mu}[f(x)]$; this is sometimes called the *Kleisli extension* of f.

Finally, we will often consider sums of weight functions over sets. If $\alpha : A \to \mathbb{R}^{\geq 0}$ maps A to the non-negative real numbers, we write $\alpha[X] \in \mathbb{R}^{\geq 0} \cup \{\infty\}$ for $\sum_{x\in X} \alpha(x)$. Moreover, if $\alpha : A \times B \to \mathbb{R}^{\geq 0}$, 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 2.2 (Dwork et al. (2006))**.** Let $\varepsilon, \delta \geq 0$ be real parameters. A probabilistic computation $M : A \to \mathbb{D}(B)$ satisfies (ε, δ)-*differential privacy* w.r.t. an adjacency relation $\phi \subseteq A \times A$ if for every pair of inputs $(a, a') \in \phi$ and every subset of outputs $E \subseteq B$, we have

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

Differential privacy is closely related to a relaxed version of distance—technically, an f-divergence—on distributions.

Definition 2.3 (Barthe and Olmedo (2013); Barthe et al. (2013); Olmedo (2014))**.** Let $\varepsilon \geq 0$. The *ε-DP divergence* $\Delta_\varepsilon(\mu_1, \mu_2)$ between two sub-distributions $\mu_1, \mu_2 \in \mathbb{D}(B)$ is defined as

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

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

Lemma 2.4. *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 $\Delta_\varepsilon(M(a), M(a')) \leq \delta$ for every pair of inputs $(a, a') \in \phi$.*

Our new definition of approximate lifting is inspired by a version of approximate liftings involving two witness distributions, proposed by Barthe and Olmedo (2013); Olmedo (2014).

Definition 2.5 (Barthe and Olmedo (2013); Olmedo (2014))**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$ and $\mathcal{R}$ be a binary relation over A and B. An (ε, δ)-*approximate* 2-*lifting* of μ_1 and 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) $\Delta_\varepsilon(\mu_\triangleleft, \mu_\triangleright) \leq \delta$; and
(3) $\mathrm{supp}(\mu_\triangleleft) \subseteq \mathcal{R}$ and $\mathrm{supp}(\mu_\triangleright) \subseteq \mathcal{R}$.

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

Combined with Lemma 2.4, 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: $M(a) =^{(2)}_{\varepsilon,\delta} M(a')$.

2-liftings can be generalized by varying the notion of distance given by Δ_ε; we will return to this point in Section 7. 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 2.6 (Barthe et al. (2016a))**.** *Let $\mu_1 \in \mathbb{D}(A_1)$, $\mu_2 \in \mathbb{D}(A_2)$ be distributions $f_1 : A_1 \to B_1$, $f_2 : A_2 \to B_2$ be* surjective *maps and $\mathcal{R}$ be a binary relation on B_1 and B_2. Then*

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

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 inject 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 2.7 (Barthe et al. (2016a))**.** *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:*

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

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 A'—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 must have support in the relation $\mathcal{R}$. For some relations $\mathcal{R}$, there may be elements a such that $a\, \mathcal{R}\, b$ does not hold for any b, or vice versa. It can be impossible to find witnesses with the correct marginals on these elements while satisfying the support condition, even if the distance condition is satisfied. At a high level, there are situations where it is possible to 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 some element b. Adding this mass anywhere preserves the distance bound between $\mu_\triangleleft$ and $\mu_\triangleright$—since it only increases the mass of $\mu_\triangleright$ while preserving the mass of $\mu_\triangleleft$, and the distance bound is asymmetric—but if there is no element a such that $a\, \mathcal{R}\, b$ then we cannot place this mass within the required support, and $\mu_\triangleleft$ and $\mu_\triangleright$ cannot be related by an approximate lifting.

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 $\mu_1\ \mathcal{R}^{(2)}_{\varepsilon,\delta}\ \mu_2$ are sometimes required to place mass outside of $\operatorname{supp}(\mu_1) \times \operatorname{supp}(\mu_2)$, even though intuitively only elements in the support of the related distributions should be relevant to the lifting. This theoretical flaw makes it difficult to establish basic mapping and support properties of approximate liftings.

Example 2.8. We illustrate these problems with a concrete example. Consider the geometric distribution with parameter $p = 1/2$, a distribution over the natural numbers which models the distribution of number of flips of a fair coin before the coin first comes up tails. Formally, the distribution $\gamma \in \mathbb{D}(\mathbb{N})$ is defined by $\gamma(k) = 1/2^{k+1}$. Consider the binary relation $\mathcal{R}= \{(x_1, x_2) \mid x_1 + 1 = x_2\}$ over $\mathbb{N}$. Now, γ cannot be related to itself via an approximate lifting $\gamma\ \mathcal{R}^{(2)}_{\varepsilon,\delta}\ \gamma$ for any parameters ε, δ. To see why, the second witness $\mu_\triangleright$ must satisfy the second marginal condition at $k = 0$, so it must put total weight $\gamma(0) = 1/2$ on pairs of the form $(-, 0)$. These pairs must belong to the relation $\mathcal{R}$, but there is no $x_1 \in \mathbb{N}$ such that $x_1 + 1 = 0$. However, there *is* an approximate lifting $\overline{\gamma}\ \overline{\mathcal{R}}^{(2)}_{\ln(2),0}\ \overline{\gamma}$, where γ and $\mathcal{R}_{+1}$ are extended to the integers $\mathbb{Z}$. For instance, the two joint distributions with support $\mu_\triangleleft(z, z + 1) = 1/2^{z+1}$ for $z \geq 0$, and $\mu_\triangleright(z, z + 1) = 1/2^{z+2}$ for $z \geq -1$ form witnesses.

This behavior is a sign that the notion of approximate lifting is not well-behaved: the support of $\overline{\gamma}$ remains the non-negative integers, but somehow embedding γ into a larger space enables additional approximate liftings.

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 3.1. For a set A, we write $A^\star$ for $A \uplus \{\star\}$. For a distribution $\eta \in \mathbb{D}(C)$ and a subset $C' \subseteq C$, we write $\eta_{|C'}$ for the *restriction* of η to C', i.e., the sub-distribution given by $\eta_{|C'}(c) = \eta(c)$ for $c \in C'$, and $\eta_{|C'}(c) = 0$ otherwise.

Definition 3.2 ($\star$-lifting)**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$ and $\mathcal{R}$ be a binary relation over A and B. An (ε, δ)-*approximate $\star$-lifting* of μ_1 and 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) $\mathrm{supp}(\eta_{\triangleleft|A\times B}), \mathrm{supp}(\eta_{\triangleright|A\times B}) \subseteq \mathcal{R}$; and
(3) $\Delta_\varepsilon(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta$, where $\overline{\eta_\bullet}$ is the extension of $\eta_\bullet$ to $\mathbb{D}(A^\star \times B^\star)$ given by the evident inclusions from $A \times B^\star$ and $A^\star \times B$ to $A^\star \times B^\star$.

We write $\mu_1 \; R^{(\star)}_{\varepsilon,\delta} \; \mu_2$ if there exists an (ε, δ)-approximate $\star$-lifting of μ_1 and 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 $(a, \star)$, and for every $b \in B$ witnesses may place mass at $(\star, b)$. Second, $\star$ serves as a generic element where all mass outside the supports $\mathrm{supp}(\mu_1) \times \mathrm{supp}(\mu_2)$ may located, giving more control over the form of the witnesses. Formally, $\star$-liftings satisfy the following natural support property.

Lemma 3.3. *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be distributions such that $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Then, there are witnesses with support contained in* $\mathrm{supp}(\mu_1)^\star \times \mathrm{supp}(\mu_2)^\star$.

Proof. See Appendix, p. 22 □

3.1. **Basic Properties.**

$\star$-liftings satisfy key properties enjoyed by existing notions of approximate lifting. To start, $\star$-liftings characterize differential privacy.[1]

Lemma 3.4. *A randomized algorithm $P : A \to \mathbb{D}(B)$ is (ε, δ)-differentially private w.r.t. ϕ if for all $(a_1, a_2) \in \phi$ we have $P(a_1) =^{(\star)}_{\varepsilon,\delta} P(a_2)$. (Here, we are turning the equality relation on B to an approximate lifting relating distributions over B.)*

Proof. See Appendix, p. 22 □

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

Lemma 3.5.
- *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, and $\mathcal{R}$ be a binary relation over A and B. If $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$, then for any $\varepsilon' \geq \varepsilon$, $\delta' \geq \delta$ and $\mathcal{S} \supseteq \mathcal{R}$, we have $\mu_1 \; \mathcal{S}^{(\star)}_{\varepsilon',\delta'} \; \mu_2$.*
- *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$, $\mu_3 \in \mathbb{D}(C)$ and $\mathcal{R}$ (resp. $\mathcal{S}$) be a binary relation over A and B (resp. over B and C). If $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$ and $\mu_2 \; \mathcal{S}^{(\star)}_{\varepsilon',\delta'} \; \mu_3$, then $\mu_1 \; (\mathcal{S} \circ \mathcal{R})^{(\star)}_{\varepsilon+\varepsilon',\delta+e^{\varepsilon}\cdot\delta'} \; \mu_3$.*

[1] The proofs of the next two lemmas use an equivalence that we will soon prove in Theorem 3.12. This is purely for convenience—these proofs could also be performed separately, and in any case Theorem 3.12 does not use these lemmas so there is no circularity.

- *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 and A_2 (resp. over B_1 and B_2). If $\mu_1\ \mathcal{R}^{(\star)}_{\varepsilon,\delta}\ \mu_2$ for some $\varepsilon, \delta \geq 0$ and for any $(a_1, a_2) \in\mathcal{R}$, we have $\eta_1(a_1)\ \mathcal{S}^{(\star)}_{\varepsilon',\delta'}\ \eta_2(a_2)$ for some $\varepsilon', \delta' \geq 0$, then*
$$\mathbb{E}_{\mu_1}[\eta_1]\ \mathcal{S}^{(\star)}_{\varepsilon+\varepsilon',\delta+\delta'}\ \mathbb{E}_{\mu_2}[\eta_2].$$

Proof. See Appendix, p. 22 □

3.2. **Equivalence with Sato's Definition.** In recent work on verifying differential privacy over continuous distributions, Sato (2016) proposes an alternative definition of approximate lifting. In the special case of discrete distributions—where all events are measurable—his definition can be stated as follows.

Definition 3.6 (Sato (2016))**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$, $\mathcal{R}$ be a binary relation over A and B and $\varepsilon, \delta \geq 0$. The distributions μ_1 and μ_2 are related by a *(witness-free) (ε, δ)-approximate lifting* for $\mathcal{R}$ if
$$\forall X \subseteq A.\, \mu_1[X] \leq e^{\varepsilon} \cdot \mu_2[\mathcal{R}(X)] + \delta.$$
We write $\mathcal{R}\ (X) = \{b \in B \mid \exists a \in X.\, (a, b) \in\mathcal{R}\} \subseteq B$.

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 3.7 (Strassen (1965))**.** *Let $\mu_1 \in \mathbb{D}(A)$, $\mu_2 \in \mathbb{D}(B)$ be two proper distributions, and let $\mathcal{R}$ be a binary relation over A and 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 3.12) 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; we begin by reviewing the basic setting.

Definition 3.8 (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 3.9 (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 $|f| \triangleq \sum_{v\in V} f(\top, v) \in \mathbb{R}\{\cup + \infty\}$.

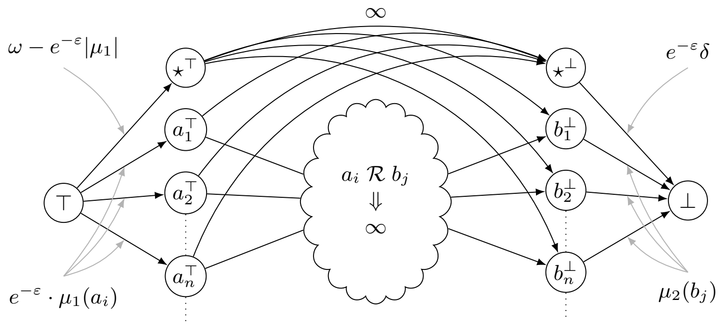

Figure 1: Flow Network in Theorem 3.12

Definition 3.10 (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* $\mathcal{E}(C) \subseteq E$ of a cut C is the set of edges crossing the cut: $\{(u, v) \in E \mid u \in C, v \notin C\}$. The *capacity* $|C| \in \mathbb{R}^{\geq 0} \cup \{\infty\}$ of a cut C is defined as $|C| \triangleq \sum_{(u,v)\in\mathcal{E}(C)} c(u, v)$.

For finite flow networks, the maximum flow is equal to the minimum cut (see, e.g., Kleinberg and Tardos (2005)). Aharoni et al. (2011) generalize this theorem to networks with countable vertices and edges under certain conditions. We will use a consequence of their result.

Theorem 3.11 (Weak Countable Max-Flow Min-Cut)**.** *Let $\mathcal{N}$ be flow network with (a) no infinite directed paths and (b) finite total capacity leaving $\top$ and entering $\perp$. Then,*

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

and both supremum and infimum are achieved by some flow and cut, respectively.

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

Theorem 3.12. *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$, $\mathcal{R}$ be a binary relation over A and B and $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$. Then, $\mu_1 \; R^{(\star)}_{\varepsilon,\delta} \; \mu_2$ iff $\forall X \subseteq A.\, \mu_1(X) \leq e^{\varepsilon} \cdot \mu_2(\mathcal{R}(X)) + \delta$.*

Proof. We detail the reverse direction; the forward direction is immediate. 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 an approximate coupling of μ_1 and μ_2 by adding a null mass to the extra points.

Let $\omega \triangleq |\mu_2| + e^{-\varepsilon} \cdot \delta$ 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

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

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

Let C be a cut of $\mathcal{N}$. In the following, we sometimes use C to denote both 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 $(\bot, \star^\bot) \in C$ 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 $B^\dagger \triangleq \{y \in B \mid (y^\bot, \bot) \in C\}$. Thus,

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

Hence, $E_\top \uplus \{(\star^\bot, \bot)\}$ is a minimum cut with capacity ω. By Theorem 3.11, we obtain a maximum flow f with mass ω. Note that the flow f saturates the capacity of all edges in $E_\top$, $E_\bot$, and $E_\star$. Let $\hat{f} : (a, b) \in A^\star \times B^\star \mapsto f(a^\top, b^\bot)$. We now define the following distributions:

$$\begin{aligned}\eta_\triangleleft &: A \times B^\star \to \mathbb{R}^{\geq 0} & \eta_\triangleright &: A^\star \times B \to \mathbb{R}^{\geq 0} \\ (a, b) &\mapsto e^\varepsilon \cdot \hat{f}(a, b) & (a, b) &\mapsto \hat{f}(a, b).\end{aligned}$$

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 $\Delta_\varepsilon(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta$. 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,

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

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

3.3. **Alternative Proof of Approximate Strassen's Theorem.** We can provide an alternative, arguably simpler proof of the reverse direction of the approximate Strassen's theorem (Theorem 3.12). Instead of relying on the max-flow min-cut theorem for countable networks by Aharoni et al. (2011), we apply the more standard result on finite networks and then pass from approximate liftings on finite restrictions of the two target distributions to an approximate lifting of the limit distributions, via a limiting argument. The results of this section have been formalized in the Coq proof assistant.[2]

We first start with a simple technical lemma.

Lemma 3.13. *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ (with respective support S_1 and S_2), $\mathcal{R}$ be a binary relation over A and B and $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$. Then, $(\mu_1)_{|S_1}\ R^{(\star)}_{\varepsilon,\delta}\ (\mu_2)_{|S_2}$ implies $\mu_1\ R^{(\star)}_{\varepsilon,\delta}\ \mu_2$.*

Proof. See Appendix, p. 23 □

We can now prove the theorem for distributions over finite domains:

Lemma 3.14 (Finite approximate Strassen's theorem)**.** *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions with finite supports, $\mathcal{R}$ be a binary relation over A and B and $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$. If for all $X \subseteq A$ we have $\mu_1[X] \leq e^\varepsilon \cdot \mu_2[\mathcal{R}(X)] + \delta$, then $\mu_1\ R^{(\star)}_{\varepsilon,\delta}\ \mu_2$.*

[2] `https://github.com/strub/xhl`

Proof. See Appendix, p. 23 □

To extend the finite case to distributions over countable sets, we need a lemma that will allow us to assume that the witnesses are within a multiplicative factor of each other, except possibly on pairs with $\star$.

Lemma 3.15. *Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ s.t. $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Then, there exists $\eta_\triangleleft$ and $\eta_\triangleright$ witnessing the lifting s.t. for $(a, b) \in A \times B$, we have:*

$$\eta_\triangleright(a, b) \leq \eta_\triangleleft(a, b) \leq e^\varepsilon \cdot \eta_\triangleright(a, b).$$

Proof. See Appendix, p. 23 □

We can now prove the reverse direction of Theorem 3.12.

Alternative proof of Theorem 3.12. By Lemma 3.13, without loss of generality, we can assume that A and B are countable. Hence, there exists a family $\{A_n\}_n$ (resp. $\{B_n\}_n$) of increasing finite subsets of A s.t. $\cup_i A_i = A$ (resp. of B s.t. $\cup_i B_i = B$). For $n \in \mathbb{N}$, we denote by μ_1^n and μ_2^n the domain restrictions of μ_1 to A_n and μ_2 to B_n, i.e., $\mu_1^n(a) = \mu_1(a)$ if $a \in A_n$, 0 otherwise and $\mu_2^n(b) = \mu_2(b)$ if $b \in B_n$, 0 otherwise.

Fix $n \in \mathbb{N}$ and let $X \subseteq A$. We have:

$$\begin{aligned}
\mu_1^n[X] \leq \mu_1[X] &\leq e^\varepsilon \cdot \mu_2[\mathcal{R}(x)] + \delta \\
&= e^\varepsilon \cdot (\mu_2[\mathcal{R}(X) \cap B_n] + \mu_2[\mathcal{R}(X) \cap \overline{B_n}]) + \delta \\
&= e^\varepsilon \cdot \mu_2^n[\mathcal{R}(X)] + \underbrace{(e^\varepsilon \cdot \mu_2[\mathcal{R}(X) \cap \overline{B_n}] + \delta)}_{\triangleq \; \delta_n}
\end{aligned}$$

Hence, by Lemma 3.14, we have $\mu_1^n \; \mathcal{R}^{(\star)}_{\varepsilon,\delta_n} \; \mu_2^n$. By Lemma 3.15, we can moreover assume that $\mu_1^n \; \mathcal{R}^{(\star)}_{\varepsilon,\delta_n} \; \mu_2^n$ is witnessed by sub-distributions $\eta^n_\triangleleft$ and $\eta^n_\triangleright$ such that

$$\forall a \in A, b \in B.\, \eta^n_\triangleright(a, b) \leq \eta_\triangleleft(a, b) \leq e^\varepsilon \cdot \eta_\triangleright(a, b). \tag{3.1}$$

Since $\mathcal{R}\ (X) \cap \overline{B_n} \xrightarrow[n\to\infty]{} \emptyset$, we also have $\delta_n \xrightarrow[n\to\infty]{} \delta$.

As a countable product of a sequentially compact sets, $[0, 1]^{A \times B^\star}$ and $[0, 1]^{A^\star \times B}$ are sequentially compact, and we can find a subsequence of indices $\{\omega_n\}_{n\in\mathbb{N}}$ s.t. both $\eta^{\omega_n}_\triangleleft$, $\eta^{\omega_n}_\triangleright$ resp. converge pointwise to sub-distributions $\eta_\triangleleft$ and $\eta_\triangleright$.

We now prove that these sub-distributions are witnesses for $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. It is clear that the supports of $\eta_{\triangleleft|A\times B}$ and $\eta_{\triangleright|A\times B}$ are included in $\mathcal{R}$. We now detail the marginals and distance conditions. First, note that

$$\eta^n_\triangleleft(a, b) \leq \sum_{b\in B^\star} \eta^n_\triangleleft(a, b) = \pi_1^\sharp(\eta^n_\triangleleft)(a) = \mu_1^n(a) \leq \mu_1(a) \quad [a \in A, b \in B^\star]$$

$$\eta^n_\triangleright(a, b) \leq \sum_{a\in A^\star} \eta^n_\triangleright(a, b) = \pi_2^\sharp(\eta^n_\triangleright)(b) = \mu_2^n(b) \leq \mu_2(b) \quad [a \in A^\star, b \in B].$$

Hence, Equation (3.1) yields

$$\eta^n_\triangleleft(a, b) \leq e^\varepsilon \cdot \eta^n_\triangleright(a, b) \leq e^\varepsilon \cdot \mu_2(b) \quad [a \in A, b \in B^\star] \tag{3.2a}$$

$$\eta^n_\triangleright(a, b) \leq \eta^n_\triangleleft(a, n) \leq \mu_1(a) \quad [a \in A^\star, b \in B]. \tag{3.2b}$$

Now, for the first marginal, let $a \in A$. We have:

$$\pi_1^\sharp(\eta_\triangleleft)(a) = \sum_{b \in B^\star} \eta_\triangleleft(a,b) = \sum_{b \in B^\star} \lim_{n\to\infty} \eta_\triangleleft^{\omega_n}(a,b)$$

From (3.2a), it is clear that $b \in B^\star \mapsto \eta_\triangleleft^{\omega_n}(a,b)$ is absolutely dominated by the summable function $[b \in B^\star \mapsto e^\varepsilon \cdot \mu_2(b)$ if $b \in B$, else $1]$. By the dominated convergence theorem, we can swap the limit and summation:

$$\begin{aligned}\pi_1^\sharp(\eta_\triangleleft)(a) &= \lim_{n\to\infty} \sum_{b \in B^\star} \eta_\triangleleft^{\omega_n}(a,b) = \lim_{n\to\infty} \pi_1^\sharp(\eta_\triangleleft)(a) \\ &= \lim_{n\to\infty} \mu_1^{\omega_n}(a) = \mu_1(a).\end{aligned}$$

For the second marginal, let $b \in B$. We have:

$$\pi_2^\sharp(\eta_\triangleright)(b) = \sum_{a \in A^\star} \eta_\triangleright(a,b) = \sum_{a \in A^\star} \lim_{n\to\infty} \eta_\triangleright^{\omega_n}(a,b)$$

From (3.2b), we have that $a \in A^\star \mapsto \eta_\triangleleft^{\omega_n}(a,b)$ is absolutely dominated by the summable function: $[a \in A^\star \mapsto \mu_1(a)$ if $a \in A$, else $1]$. Again by the dominated convergence theorem, we can swap the limit and summation:

$$\begin{aligned}\pi_2^\sharp(\eta_\triangleright)(b) &= \lim_{n\to\infty} \sum_{a \in A^\star} \eta_\triangleright^{\omega_n}(a,b) = \lim_{n\to\infty} \pi_2^\sharp(\eta_\triangleright)(b) \\ &= \lim_{n\to\infty} \mu_2^{\omega_n}(b) = \mu_1(b).\end{aligned}$$

We are left to prove the distance condition, i.e., that for any $X \subseteq A^* \times B^*$, we have $\overline{\eta_\triangleleft}[X] - e^\varepsilon \cdot \overline{\eta_\triangleright}[X] \leq \delta$. First, note that $\overline{\eta_\triangleleft}[X] = \lim_{n\to\infty} \overline{\eta_\triangleleft}^{\omega_n}[X]$ and $\overline{\eta_\triangleright}[X] = \lim_{n\to\infty} \overline{\eta_\triangleright}^{\omega_n}[X]$. Indeed, for $\overline{\eta_\triangleleft}[X]$, we have

$$\begin{aligned}\overline{\eta_\triangleleft}[X] &= \sum_{a,b \in X} \lim_{n\to\infty} \eta_\triangleleft^{\omega_n}(a,b) = \sum_{a \in A^\star} \sum_{b \in X_\triangleleft^a} \lim_{n\to\infty} \eta_\triangleleft^{\omega_n}(a,b) \\ &= \sum_{a \in A^\star} \lim_{n\to\infty} \sum_{b \in X_\triangleleft^a} \eta_\triangleleft^{\omega_n}(a,b) && \text{(dominated convergence)} \\ &= \lim_{n\to\infty} \sum_{a \in A^\star} \sum_{b \in X_\triangleleft^a} \eta_\triangleleft^{\omega_n}(a,b) && \text{(dominated convergence)} \\ &= \lim_{n\to\infty} \sum_{a,b \in X} \eta_\triangleleft^{\omega_n}(a,b) = \lim_{n\to\infty} \eta_\triangleleft^{\omega_n}(a,b)[X]\end{aligned}$$

where $X_\triangleleft^a \triangleq \{b \in B^\star \mid (a,b) \in X\}$. The first application of the dominated convergence theorem uses, like for the first marginal condition, the function $[b \in B^\star \mapsto e^\varepsilon \cdot \mu_2(b)$ if $b \in B$, else $1]$ as the dominating function (using Equation (3.2a)). The second application of the dominated convergence theorem uses $[a \in A^\star \mapsto \mu_1(a)$ if $a \in A$, else $0]$ as the dominating function. Indeed, if $a \in A$, then

$$\sum_{b \in X_\triangleleft^a} \overline{\eta_\triangleleft^{\omega_n}}(a,b) = \sum_{b \in X_\triangleleft^a} \eta_\triangleleft^{\omega_n}(a,b) \leq \sum_{b \in B^\star} \eta_\triangleleft^{\omega_n}(a,b) = \pi_1^\sharp(\eta_\triangleleft^{\omega_n})(a) = \mu_1^{\omega_n}(a) \leq \mu_1(a), \text{and}$$

$$\sum_{b \in X_\triangleleft^a} \overline{\eta_\triangleleft}(\star,b) = \sum_{b \in X_\triangleleft^a} 0 = 0.$$

Likewise, for $\overline{\eta_\triangleright}[X]$, we have

$$\begin{aligned}
\overline{\eta_\triangleright}[X] &= \sum_{a,b\in X} \lim_{n\to\infty} \eta_\triangleright^{\omega_n}(a,b) = \sum_{b\in B} \sum_{a\in X_\triangleright^b} \lim_{n\to\infty} \eta_\triangleright^{\omega_n}(a,b) \\
&= \sum_{b\in B} \lim_{n\to\infty} \sum_{a\in X_\triangleright^b} \eta_\triangleleft^{\omega_n}(a,b) && \text{(dominated convergence)} \\
&= \lim_{n\to\infty} \sum_{b\in B} \sum_{a\in X_\triangleright^b} \eta_\triangleleft^{\omega_n}(a,b) && \text{(dominated convergence)} \\
&= \lim_{n\to\infty} \sum_{a,b\in X} \eta_\triangleright^{\omega_n}(a,b) = \lim_{n\to\infty} \eta_\triangleright^{\omega_n}(a,b)[X]
\end{aligned}$$

where $X_\triangleright^b \triangleq \{a \in A^\star \mid (a,b) \in X\}$. Here, the first application of the dominated convergence theorem uses, as for the second marginal condition, the function $[a \in A^\star \mapsto \mu_1(a)$ if $a \in A$, else $1]$ as the dominating function (using equation (3.2b)). The second application of the dominated convergence theorem uses $[b \in B^\star \mapsto \mu_2(b)$ if $b \in B$, else $0]$ as the dominating function. Indeed, if $b \in B$, then

$$\sum_{a\in X_\triangleright^b} \overline{\eta_\triangleright^{\omega_n}}(a,b) = \sum_{a\in X_\triangleright^b} \eta_\triangleright^{\omega_n}(a,b) \leq \sum_{a\in A^\star} \eta_\triangleright^{\omega_n}(a,b) = \pi_2^\sharp(\eta_\triangleright^{\omega_n})(b) = \mu_2^{\omega_n}(b) \leq \mu_2(b), \text{and}$$

$$\sum_{a\in X_\triangleright^b} \overline{\eta_\triangleright^{\omega_n}}(a,\star) = \sum_{a\in X_\triangleright^b} 0 = 0.$$

Hence, we can conclude the distance condition by taking limits:

$$\begin{aligned}
\overline{\eta_\triangleleft}[X] - e^\varepsilon \cdot \overline{\eta_\triangleright}[X] &= \lim_{n\to\infty} \overline{\eta_\triangleleft^{\omega_n}}[X] - e^\varepsilon \cdot \lim_{n\to\infty} \overline{\eta_\triangleright^{\omega_n}}[X] \\
&= \lim_{n\to\infty} \left(\overline{\eta_\triangleleft^{\omega_n}} - e^\varepsilon \cdot \overline{\eta_\triangleright^{\omega_n}}[X] \right) \\
&\leq \lim_{n\to\infty} \delta_n = \delta.
\end{aligned}$$

□

This proof constructs witnesses to an approximate lifting relating two distributions (μ_1, μ_2) given a sequence of approximate liftings relating pairs of finite restrictions of μ_1 and μ_2. Using essentially the same argument, we can construct an approximate lifting relating (μ_1, μ_2) given a sequence of approximate liftings relating pairs of distributions converging to μ_1 and μ_2; the main difference is that a slightly more general form of the dominated convergence theorem is needed (see Hsu (2017, Lemma 5.1.7) for details).

4. Properties of ⋆-Liftings

Our main theorem can be used to show several natural properties of ⋆-liftings. To begin, we can improve the mapping property from Theorem 2.6, lifting the requirement that the maps must be surjective.

Lemma 4.1. *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 and 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*

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

Proof. See Appendix, p. 25 □

Similarly, we can generalize the existing rules for up-to-bad reasoning (cf. Barthe et al. (2016a, 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 indices $\overline{\delta}$ are different in the two cases. We write $\overline{\theta}$ for the complement of θ.

Lemma 4.2. *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 $\mu_1 \; (\theta_{\lhd} \implies \mathcal{R})^{(\star)}_{\varepsilon,\delta} \; \mu_2$ for some parameters $\varepsilon, \delta \geq 0$. Then $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\overline{\delta}} \; \mu_2$, where $\overline{\delta} \triangleq \delta + \mu_1[\overline{\theta}]$.*

Proof. See Appendix, p. 25 □

Lemma 4.3. *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 $\mu_1 \; (\theta_{\rhd} \implies \mathcal{R})^{(\star)}_{\varepsilon,\delta} \; \mu_2$ for some parameters $\varepsilon, \delta \geq 0$. Then, $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\overline{\delta}} \; \mu_2$, where $\overline{\delta} \triangleq \delta + e^{\varepsilon} \cdot \mu_2[\overline{\theta}]$.*

Proof. See Appendix, p. 26 □

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 based on *accuracy* bounds. For instance, suppose that we have an approximate lifting of $\mathcal{R}$ relating two distributions, $\mu_{\lhd}$ and $\mu_{\rhd}$. If θ_a is an event that happens with high probability in the first distribution $\mu_{\lhd}$—say, a certain noise variable is at most 100—we can incorporate $\theta_{a,\lhd}$ into the approximate lifting by increasing the δ parameter by the probability that θ_a *fails* to hold in $\mu_{\lhd}$. When reasoning in terms of approximate couplings, intuitively we can "assume" θ_a holds by "paying" with an increase in δ. A similar property holds for the second distribution $\mu_{\rhd}$.

We formalize these constructions with the following lemma.

Lemma 4.4. *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 $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Then, $\mu_1 \; (\theta_{a,\lhd} \cap \mathcal{R})^{(\star)}_{\varepsilon,\delta_a} \; \mu_2$ and $\mu_1 \; (\theta_{b,\rhd} \cap \mathcal{R})^{(\star)}_{\varepsilon,\delta_b} \; \mu_2$ where $\delta_a \triangleq \delta + \mu_1[\overline{\theta_a}]$ and $\delta_b \triangleq \delta + e^{\varepsilon} \cdot \mu_2[\overline{\theta_b}]$.*

Proof. See Appendix, p. 26 □

⋆-liftings also support a significant generalization of optimal subset coupling. Unlike the known construction for 2-liftings (Theorem 2.7), 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 4.5. *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:*

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

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 3.12. □

We can then recover the existing notion of optimal subset coupling (Barthe et al., 2016a) for ⋆-liftings as a special case.

Corollary 4.6 (Barthe et al. (2016a))**.** *Let $\mu \in \mathbb{D}(A)$ and consider two nested subsets $P_2 \subseteq P_1 \subseteq A$. Then, we have the following equivalence:*

$$\mathbb{P}_{\mu}[P_1] \leq e^{\varepsilon} \cdot \mathbb{P}_{\mu}[P_2] \iff \mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,0} \; \mu_2,$$

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 4.5, noting that

$$\mathbb{P}_\mu[A - P_1] \leq e^\varepsilon \cdot \mathbb{P}_\mu[A - P_2]$$

is automatic since $P_2 \subseteq P_1$ implies $\mathbb{P}_\mu[A - P_1] \leq \mathbb{P}_\mu[A - P_2]$. Note that there is no longer a need for P_1 to be a strict subset of A. □

Finally, we can directly extend known composition theorems from differential privacy to $\star$-liftings. This connection is quite useful for transferring existing composition results from the privacy literature to approximate liftings. We first define a general template describing how the privacy parameters ε, δ decay under sequential composition.

Definition 4.7. Let $\mathbb{R}^{\geq 0}_2 \triangleq \mathbb{R}^{\geq 0} \times \mathbb{R}^{\geq 0}$ and let $(\mathbb{R}^{\geq 0}_2)^*$ be the set of finite sequences over pairs of non-negative reals. A map $r : (\mathbb{R}^{\geq 0}_2)^* \to \mathbb{R}^{\geq 0}_2$ is a *DP-composition rule* if for all sets A, D, adjacency relations $\phi \subseteq D \times D$, and families of functions $\{f_i : D \times A \to \mathbb{D}(A)\}_{i<n}$, the following implication holds: if for every initial value $a \in A$ and $i < n$, $f_i(-, a) : D \to \mathbb{D}(A)$ is $(\varepsilon_i, \delta_i)$-differentially private w.r.t. ϕ, then $F(-, a)$ is $(\varepsilon^*, \delta^*)$-differentially private w.r.t. ϕ and any initial value $a \in A$ where $F : (d, a) \mapsto (\bigcirc_{i<n} (f_i(d, -))^\sharp)(\mathbb{1}_a)$ is the n-fold composition of the functions $[f_i]_{i<n}$ and $(\varepsilon^*, \delta^*) \triangleq r([(\varepsilon_i, \delta_i)]_{i<n})$.

Lemma 4.8. *Let* $r : (\mathbb{R}^{\geq 0}_2)^* \to \mathbb{R}^{\geq 0}_2$ *be a DP-composition rule. 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* $\{\mathcal{R}(i) \subseteq A_i \times B_i\}_{i\leq n}$*. Let* $\{g_i : A_i \to \mathbb{D}(A_{i+1})\}_{i<n}$ *and* $\{h_i : B_i \to \mathbb{D}(B_{i+1})\}_{i<n}$ *be two families of functions s.t. for all* $i < n$ *and* $(a, b) \in \mathcal{R}(i)$*, we have:*

(1) $g_i(a)\ \mathcal{R}(i+1)^{(\star)}_{\varepsilon_i,\delta_i}\ h_i(b)$ *for some parameters* $\varepsilon_i, \delta_i \geq 0$*, 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*

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

where $G : A_0 \to \mathbb{D}(A_n)$ *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* $H(b) \triangleq (\bigcirc_{i<n} h_i^\sharp)(\mathbb{1}_b)$*—and* $(\varepsilon^*, \delta^*) \triangleq r([(\varepsilon_i, \delta_i)]_{i<n})$.

Proof. We assume that $A_i = A$ is the same for all i, and $B_i = B$ is the same for all i. This is without loss of generality, since when A_i and B_i vary with i we may work with the disjoint unions $\sqcup_i A_i$ and $\sqcup_i B_i$ by restricting each $\mathcal{R}(i)$ to only relate pairs that are both in the i-th component. Define $D = \{0, 1\} = \{tt, \mathit{ff}\}$ and $\phi = \{(tt, \mathit{ff})\} \subseteq D \times D$.

For every $i < n$ and $(a_i, b_i) \in \mathcal{R}(i)$, the definition of $\star$-lifting gives two distributions $\mu_\triangleleft[a_i, b_i], \mu_\triangleright[a_i, b_i]$ witnessing $g_i(a)\ \mathcal{R}(i+1)^{(\star)}_{\varepsilon_i,\delta_i}$. We regard both witness distributions as elements of $\mathbb{D}(A^\star \times B^\star)$ via the evident embeddings. We define the maps $f_i : D \times (A^\star \times B^\star) \to \mathbb{D}(A^\star \times B^\star)$ by cases:

$$\begin{aligned}
f_i(tt, (a_i, b_i)) &= \mu_\triangleleft[a_i, b_i] \\
f_i(\mathit{ff}, (a_i, b_i)) &= \mu_\triangleright[a_i, b_i] \\
f_i(-, (a_i, \star)) &= g_i(a_i) \times \mathbb{1}_\star \\
f_i(-, (\star, b_i)) &= \mathbb{1}_\star \times h_i(b_i) \\
f_i(-, (\star, \star)) &= \mathbb{1}_{(\star,\star)}
\end{aligned}$$

where $\times$ denotes the product distribution and $(a_i, b_i) \in \mathcal{R}(i)$; otherwise, $f_i(-,(a,b)) = 0$.

Now for all $(a, b) \in A^\star \times B^\star$, the map $f_i(-,(a,b)) : D \to \mathbb{D}(A^\star \times B^\star)$ is $(\varepsilon_i, \delta_i)$-differentially private with respect to ϕ by the distance property on $\mu_\triangleleft[a_i, b_i]$ and $\mu_\triangleright[a_i, b_i]$ (and by definition when $(a_i, b_i) \notin \mathcal{R}(i)$). Hence, the DP-composition rule implies that $F : (d,(a,b)) \mapsto (\bigcirc_{i<n} (f_i(d,-))^\sharp)(\mathbb{1}_{(a,b)})$ is $(\varepsilon^*, \delta^*)$-differentially private with respect to ϕ for any $(a, b) \in A^\star \times B^\star$. For any $(a_0, b_0) \in \mathcal{R}(0)$, we claim that that $F(tt,(a_0, b_0))$ and $F(\mathit{ff},(a_0, b_0))$ witness the desired approximate lifting

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

The support and marginal conditions are not hard to show, and the distance condition follows from differential privacy of $F(-,(a_0, b_0)) : D \to \mathbb{D}(A^\star \times B^\star)$. □

Some of the more sophisticated composition results from differential privacy—for instance, the advanced composition theorem by Dwork et al. (2010)—do not apply to arbitrary adjacency relations ϕ, but only *symmetric* relations. Lemma 4.8 cannot lift such theorems to composition principles for approximate liftings. In Section 6 we will remedy this problem by working with a symmetric version of $\star$-lifting.

5. Comparison with Prior 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 2.5). Evidently, $\star$-liftings strictly generalize 2-liftings.

Theorem 5.1. *For all binary relations $\mathcal{R}$ over A and B and parameters $\varepsilon, \delta \geq 0$, we have $\mathcal{R}^{(2)}_{\varepsilon,\delta} \subseteq \mathcal{R}^{(\star)}_{\varepsilon,\delta}$. There exist relations and parameters where the inclusion is strict.*

Proof. The inclusion $\mathcal{R}^{(2)}_{\varepsilon,\delta} \subseteq \mathcal{R}^{(\star)}_{\varepsilon,\delta}$ is immediate. We have a strict inclusion $\mathcal{R}^{(2)}_{\varepsilon,\delta} \subsetneq \mathcal{R}^{(\star)}_{\varepsilon,\delta}$ even for $\delta = 0$ by considering the optimal subset coupling from Theorem 2.7. Consider a distribution μ over set A, and let $P_1 \subseteq P_2 = A$. There is an $(\varepsilon, 0)$-approximate $\star$-lifting (by Theorem 4.5), but a $(\varepsilon, 0)$-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. □

We can also compare $\star$-liftings with the original definitions of (ε, δ)-approximate lifting by Barthe et al. (2013). They introduce two notions, a symmetric lifting and an asymmetric lifting, each using a single witness distribution. We will focus on the asymmetric version here, and return to the symmetric version in Section 6.

Definition 5.2 (Barthe et al. (2013))**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$ and $\mathcal{R}$ be a binary relation over A and B. An (ε, δ)-*approximate* 1-*lifting* of μ_1 and 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) $\Delta_\varepsilon(\mu_1, \pi_1^\sharp(\mu)) \leq \delta$; and
(3) $\mathrm{supp}(\mu) \subseteq \mathcal{R}$.

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 $\mu_1\ \mathcal{R}^{(1)}_{\varepsilon,\delta}\ \mu_2$

if there exists an (ε, δ)-approximate 1-lifting of μ_1 and μ_2 for $\mathcal{R}$; the superscript $\cdot^{(1)}$ 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 more awkward to manipulate and less well-understood theoretically than their 2-lifting cousins—basic properties such as mapping (Lemma 4.1) are not known to hold; the subset coupling (Theorem 2.7) is not known to exist. Somewhat surprisingly, 1-liftings are equivalent to $\star$-liftings and hence by Theorem 3.12, also to Sato's approximate lifting.

Theorem 5.3. *For all binary relations $\mathcal{R}$ over A and B and parameters $\varepsilon, \delta \geq 0$, we have $\mathcal{R}^{(1)}_{\varepsilon,\delta} {=} \mathcal{R}^{(\star)}_{\varepsilon,\delta}$.*

Proof. See Appendix, p. 26 □

6. Symmetric $\star$-Lifting

The approximate liftings we have considered so far are all *asymmetric*. For instance, the approximate lifting $\mu_1 \ \mathcal{R}^{(\star)}_{\varepsilon,\delta} \ \mu_2$ may not imply the lifting $\mu_2 \ (\mathcal{R}^{-1})^{(\star)}_{\varepsilon,\delta} \ \mu_1$. Given witnesses (μ_L, μ_R) to the first lifting, we may consider the witnesses $(\nu_L, \nu_R) = (\mu_R^\top, \mu_L^\top)$ where the transpose map $(-)^\top : \mathbb{D}(A \times B) \to \mathbb{D}(B \times A)$ is defined in the obvious way. Then (ν_L, ν_R) almost witness the second lifting—the marginal and support conditions holds, but the distance bound is in the wrong direction:

$$\Delta_\varepsilon(\nu_R, \nu_L) = \Delta_\varepsilon(\mu_L^\top, \mu_R^\top) = \Delta_\varepsilon(\mu_L, \mu_R) \leq \delta.$$

In general, we cannot bound $\Delta_\varepsilon(\nu_L, \nu_R)$ and the symmetric lifting $\mu_2 \ (\mathcal{R}^{-1})^{(\star)}_{\varepsilon,\delta} \ \mu_1$ may not hold. To recover symmetry, we can define a symmetric version of $\star$-lifting.

Definition 6.1 (Symmetric $\star$-lifting)**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$ and $\mathcal{R}$ be a binary relation over A and B. An (ε, δ)-*approximate symmetric $\star$-lifting* of μ_1 and 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) $\mathrm{supp}(\eta_{\triangleleft | A \times B}), \mathrm{supp}(\eta_{\triangleright | A \times B}) \subseteq \mathcal{R}$; and
(3) $\Delta_\varepsilon(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta, \Delta_\varepsilon(\overline{\eta_\triangleright}, \overline{\eta_\triangleleft}) \leq \delta$, where $\overline{\eta_\bullet}$ is the canonical lifting of $\eta_\bullet$ to $A^\star \times B^\star$.

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

Symmetric $\star$-lifting is a special case of $\star$-lifting that can capture differential privacy under when the adjacency relation ϕ is *symmetric*: 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: $M(a) \ \overline{(=)}^{(\star)}_{\varepsilon,\delta} \ M(a')$. Unfortunately, the more advanced properties in Section 4 do not all hold when moving to symmetric liftings. However, we can show that symmetric $\star$-liftings are equivalent to the symmetric version of 1-witness lifting proposed by Barthe et al. (2013).

Definition 6.2 (Barthe et al. (2013))**.** Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\varepsilon, \delta \in \mathbb{R}^{\geq 0}$ and $\mathcal{R}$ be a binary relation over A and B. An (ε, δ)-*approximate symmetric* 1-*lifting* of μ_1 and 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) $\Delta_\varepsilon(\mu_1, \pi_1^\sharp(\mu)) \leq \delta$ and $\Delta_\varepsilon(\mu_2, \pi_2^\sharp(\mu)) \leq \delta$; and
(3) $\mathrm{supp}(\mu) \subseteq \mathcal{R}$.

We will write $\mu_1 \, \overline{\mathcal{R}}^{(1)}_{\varepsilon,\delta} \, \mu_2$ if there exists an (ε, δ)-approximate symmetric 1-lifting of μ_1 and μ_2 for $\mathcal{R}$; the superscript $\cdot^{(1)}$ indicates that there is one witness for this lifting.

Theorem 6.3 (cf. the asymmetric result Theorem 5.3)**.** *For all binary relations $\mathcal{R}$ over A and B and parameters $\varepsilon, \delta \geq 0$, we have $\overline{\mathcal{R}}^{(1)}_{\varepsilon,\delta} = \overline{\mathcal{R}}^{(\star)}_{\varepsilon,\delta}$.*

Proof. See Appendix, p. 27 □

The main use of symmetric approximate liftings is to support richer composition results that only apply to symmetric adjacency relations.

Definition 6.4. Let $\mathbb{R}_2^{\geq 0} \triangleq \mathbb{R}^{\geq 0} \times \mathbb{R}^{\geq 0}$ and let $(\mathbb{R}_2^{\geq 0})^*$ be the set of finite sequences over pairs of non-negative reals. A map $r : (\mathbb{R}_2^{\geq 0})^* \to \mathbb{R}_2^{\geq 0}$ is a *symmetric DP-composition rule* if for all sets A, D, symmetric adjacency relations $\phi \subseteq D \times D$, and families of functions $\{f_i : D \times A \to \mathbb{D}(A)\}_{i<n}$, the following implication holds: if for every initial value $a \in A$ and $i < n$, $f_i(-, a) : D \to \mathbb{D}(A)$ is $(\varepsilon_i, \delta_i)$-differentially private w.r.t. ϕ, then $F(-, a)$ is $(\varepsilon^*, \delta^*)$-differentially private w.r.t. ϕ and any initial value $a \in A$ where $F : (d, a) \mapsto (\bigcirc_{i<n} (f_i(d, -))^\sharp)(\mathbb{1}_a)$ is the n-fold composition of the functions $[f_i]_{i<n}$ and $(\varepsilon^*, \delta^*) \triangleq r([(\varepsilon_i, \delta_i)]_{i<n})$.

We have the following reduction, a symmetric version of Lemma 4.8.

Lemma 6.5. *Let $r : (\mathbb{R}_2^{\geq 0})^* \to \mathbb{R}_2^{\geq 0}$ be a symmetric DP-composition rule. 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) $g_i(a) \, \overline{\mathcal{R}(i+1)}^{(\star)}_{\varepsilon_i,\delta_i} \, h_i(b)$ for some parameters $\varepsilon_i, \delta_i \geq 0$, 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 symmetric $\star$-lifting

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

where $G : A_0 \to \mathbb{D}(A_n)$ 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 $H(b) \triangleq (\bigcirc_{i<n} h_i^\sharp)(\mathbb{1}_b)$ and $(\varepsilon^*, \delta^*) \triangleq r([(\varepsilon_i, \delta_i)]_{i<n})$.

Proof. Essentially the same as the proof of Lemma 4.8. Let $D = \{tt, f\!f\}$ as before, and take $\overline{\phi} = \{(tt, f\!f), (f\!f, tt)\}$ be a binary relation on D.

For every $i < n$ and $(a_i, b_i) \in \mathcal{R}(i)$, the definition of symmetric $\star$-lifting gives two distributions $\mu_\triangleleft[a_i, b_i], \mu_\triangleright[a_i, b_i]$ witnessing $g_i(a) \, \overline{\mathcal{R}(i+1)}^{(\star)}_{\varepsilon_i,\delta_i}$. We define the same maps $f_i : D \times (A^\star \times B^\star) \to \mathbb{D}(A^\star \times B^\star)$ as before:

$$\begin{aligned} f_i(tt, (a_i, b_i)) &= \mu_\triangleleft[a_i, b_i] \\ f_i(f\!f, (a_i, b_i)) &= \mu_\triangleright[a_i, b_i] \\ f_i(-, (a_i, \star)) &= g_i(a_i) \times \mathbb{1}_\star \\ f_i(-, (\star, b_i)) &= \mathbb{1}_\star \times h_i(b_i) \\ f_i(-, (\star, \star)) &= \mathbb{1}_{(\star,\star)} \end{aligned}$$

for $(a_i, b_i) \in \mathcal{R}(i)$; otherwise, $f_i(-,(a,b)) = 0$.

Compared to proof of Lemma 4.8, the crucial difference is that since we have witnesses to a *symmetric* approximate lifting, the resulting maps $f_i(-,(a,b)) : D \to \mathbb{D}(A^\star \times B^\star)$ are $(\varepsilon_i, \delta_i)$-differentially private with respect to the *symmetric* relation $\overline{\phi}$, not just the *asymmetric* relation ϕ. Hence, we may apply the symmetric DP-composition rule r and conclude as before. □

With this reduction we hand, we can generalize the advanced composition theorem from differential privacy to $\star$-liftings.

Theorem 6.6 (Advanced composition (Dwork et al., 2010))**.** *Consider a* symmetric *adjacency relation ϕ on databases D. Let $f_i : D \times A \to \mathbb{D}(A)$ be a sequence of n functions, such that for every $a \in A$ the functions $f_i(-,a) : D \to \mathbb{D}(A)$ are (ε, δ)-differentially private with respect to ϕ. Then, for every $a \in A$ and $\omega \in (0,1)$, running $f_1, \dots, f_n$ in sequence is $(\varepsilon^*, \delta^*)$-differentially private for*

$$\varepsilon^* = \left(\sqrt{2n\ln(1/\omega)}\right)\varepsilon + n\varepsilon(e^\varepsilon - 1) \quad \textit{and} \quad \delta^* = n\delta + \omega.$$

Corollary 6.7. *Let n be a natural number, $\varepsilon, \delta \geq 0$, and $\omega \in (0,1)$ be real parameters. Suppose we have:*

(1) *sets $\{A_i\}_i, \{B_i\}_i$ with i ranging from $0, \dots, n$;*
(2) *relations $\{\mathcal{R}(i)\}_i$ on A_i and B_i with i ranging from $0, \dots, n$; and*
(3) *functions $\{f_i : A_i \to \mathbb{D}(A_{i+1})\}_i, \{g_i : B_i \to \mathbb{D}(B_{i+1})\}_i$ with i ranging from $0, \dots, n-1$*

such that for all $(a,b) \in \mathcal{R}(i)$, we have

$$f_i(a)\ \overline{\mathcal{R}(i+1)}^{(\star)}_{\varepsilon,\delta}\ g_i(b)$$

and $f_i(a), g_i(b)$ proper distributions. Then, there is an approximate lifting of the compositions:

$$F(a_0)\ \overline{\mathcal{R}(n)}^{(\star)}_{\varepsilon',\delta'}\ G(b_0)$$

for every $(a_0, b_0) \in \mathcal{R}(0)$, where $F : A_0 \to \mathbb{D}(A_n)$ and $G : B_0 \to \mathbb{D}(B_n)$ are the n-fold (Kleisli) compositions of $\{f_i\}$ and $\{g_i\}$ respectively, and the lifting parameters are:

$$\varepsilon' \triangleq \varepsilon\sqrt{2n\ln(1/\omega)} + n\varepsilon(e^\varepsilon - 1) \qquad \delta' \triangleq n\delta + \omega.$$

Proof. By the advanced composition theorem for differential privacy (Theorem 6.6), the map $r([(\varepsilon,\delta)]_{i<n}) \triangleq (\varepsilon', \delta')$ is a symmetric DP composition rule. So, we can conclude by Lemma 6.5. □

7. $\star$-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 (2013); Olmedo (2014). Roughly, a f-divergence is 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 the definition for (ε, δ)-liftings, we can define $\star$-lifting with f-divergences. Let us first formally define f-divergences. We denote by $\mathcal{F}$ the set of non-negative convex functions vanishing at 1: $\mathcal{F} = \{f : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0} \mid f(1) = 0\}$. 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 7.1. 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

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

Examples of f-divergences include statistical distance ($f(t) = \frac{1}{2}\,|t-1|$), Kullback-Leibler divergence ($f(t) = t\ln(t) - t + 1$),[3] and Hellinger distance ($f(t) = \frac{1}{2}(\sqrt{t}-1)^2$).

Definition 7.2 ($\star$-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 and B, and $f \in \mathcal{F}$. An $(f;\delta)$-*approximate lifting* of μ_1 and 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 and μ_2 for $\mathcal{R}$.

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

Lemma 7.3. *Suppose f corresponds to 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 and A_2 (resp. over B_1 and 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*

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

Proof. Essentially the same as the proof of Lemma 4.8, lifting known composition results for these f-divergences (namely, Barthe and Olmedo (2013, Proposition 5)). □

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 3.3).

Lemma 7.4. *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* $\mathrm{supp}(\mu_1)^\star \times \mathrm{supp}(\mu_2)^\star$.

Proof. See Appendix, p. 29 □

Finally, the mapping property from Lemma 4.1 holds also for these $\star$-liftings. While the proof of Lemma 4.1 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 7.5. *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 and 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*

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

Proof. See Appendix, p. 30 □

[3] The additional term $-t + 1$ extends the classical definition of KL-divergence to sub-distributions (Barthe and Olmedo, 2013).

8. Conclusion

We have proposed a new definition of approximate lifting that unifies all known 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.

Subsequent to the original publication of this work, researchers have explored two extensions of $\star$-liftings. First, Albarghouthi and Hsu (2018) develop *variable approximate liftings*, a refinement of the $(\varepsilon, 0)$-approximate liftings where the ε parameter is real-valued function $A \times B \to \mathbb{R}^{\geq 0}$ and the distance condition on witnesses is generalized to

$$\mu_{\triangleleft}(a, b) \leq e^{\varepsilon(a,b)} \mu_{\triangleright}(a, b).$$

In effect, the approximation parameters may vary over pairs of samples instead of being a uniform upper bound. This refinement allows capturing more precise approximation bounds, in some cases simplifying proofs of differential privacy. An (ε, δ) version of variable approximate lifting is currently not known.

Second, Sato et al. (2018) explore 2-liftings in the continuous case modeling differential privacy and f-divergences, but also relaxations of differential privacy based on Rényi divergences (Bun and Steinke, 2016; Mironov, 2017). These 2-liftings subsume $\star$-liftings; a continuous analogue of the approximate Strassen theorem is strongly believed to hold but remains to be shown.

We see at least two important directions for future work. First, adapting existing program logics (in particular, apRHL (Barthe et al., 2013)) to use $\star$-liftings, and formalizing examples that were out of reach of previous systems. Second, symmetric $\star$-liftings seem to be an important notion—for instance, the advanced composition theorem of differential privacy (Dwork et al., 2010) applies to these liftings—but only existential versions of the definition are currently known. A universal definition, similar to Sato's definition for asymmetric liftings, would give more evidence that symmetric liftings are indeed a mathematically interesting abstraction, and also give a more convenient route to constructing such liftings.

Acknowledgments

We thank the anonymous reviewers for their helpful suggestions. This work was partially supported by a grant from the NSF (TWC-1513694), a grant from the Simons Foundation (#360368 to Justin Hsu), and the ERC Starting Grant ProFoundNet (#679127).

References

R. Aharoni, E. Berger, A. Georgakopoulos, A. Perlstein, and P. Sprüssel. The max-flow min-cut theorem for countable networks. *J. Comb. Theory, Ser. B*, 101(1):1–17, 2011.

A. Albarghouthi and J. Hsu. Synthesizing coupling proofs of differential privacy. *Proceedings of the ACM on Programming Languages*, 2(POPL), Jan. 2018. Appeared at ACM SIGPLAN–SIGACT Symposium on Principles of Programming Languages (POPL), Los Angeles, California.

G. Barthe and F. 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.

G. Barthe, B. Köpf, F. Olmedo, and S. Zanella-Béguelin. Probabilistic relational reasoning for differential privacy. *ACM Transactions on Programming Languages and Systems*, 35 (3):9, 2013.

G. Barthe, N. Fong, M. Gaboardi, B. Grégoire, J. Hsu, and P.-Y. Strub. Advanced probabilistic couplings for differential privacy. In *ACM SIGSAC Conference on Computer and Communications Security (CCS), Vienna, Austria*, 2016a.

G. Barthe, M. Gaboardi, B. Grégoire, J. Hsu, and P.-Y. Strub. Proving differential privacy via probabilistic couplings. In *IEEE Symposium on Logic in Computer Science (LICS), New York, New York*, 2016b.

G. Barthe, M. Gaboardi, J. Hsu, and B. C. Pierce. Programming language techniques for differential privacy. *SIGLOG News*, 3(1):34–53, 2016c.

M. Bun and T. Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *IACR Theory of Cryptography Conference (TCC), Beijing, China*, volume 9985 of *Lecture Notes in Computer Science*, pages 635–658. Springer-Verlag, 2016.

I. Csiszár and P. C. Shields. Information theory and statistics: A tutorial. *Foundations and Trends® in Communications and Information Theory*, 1(4):417–528, 2004.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *IACR Theory of Cryptography Conference (TCC), New York, New York*, pages 265–284, 2006.

C. Dwork, G. N. Rothblum, and S. Vadhan. Boosting and differential privacy. In *IEEE Symposium on Foundations of Computer Science (FOCS), Las Vegas, Nevada*, pages 51–60, 2010.

J. Hsu. *Probabilistic Couplings for Probabilistic Reasoning.* PhD thesis, University of Pennsylvania, 2017.

J. Kleinberg and E. Tardos. *Algorithm Design.* Addison-Wesley, 2005. ISBN 0321295358.

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

I. Mironov. Rényi differential privacy. In *IEEE Computer Security Foundations Symposium (CSF), Santa Barbara, California*, pages 263–275, 2017.

F. Olmedo. *Approximate Relational Reasoning for Probabilistic Programs.* PhD thesis, Universidad Politécnica de Madrid, 2014.

T. 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.

T. Sato, G. Barthe, M. Gaboradi, J. Hsu, and S. Katsumata. Reasoning about divergences for relaxations of differential privacy. 2018.

V. Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, pages 423–439, 1965.

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

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

Appendix A. Detailed Proofs

In the proofs, we will sometimes refer to the witnesses of a $\star$-lifting.

Notation A.1. Let $\mu_1 \in \mathbb{D}(A)$ and $\mu_2 \in \mathbb{D}(B)$ be sub-distributions, $\varepsilon, \delta \in \mathbb{R}^+$ and $\mathcal{R}$ be a binary relation over A and B. If two distributions $\eta_\triangleleft \in \mathbb{D}(A \times B^\star)$ and $\eta_\triangleright \in \mathbb{D}(A^\star \times B)$ are witnesses to the $\star$-lifting $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$, then we write:

$$\langle \eta_\triangleleft, \eta_\triangleright \rangle \blacktriangleleft^{\mathcal{R}}_{\varepsilon,\delta} \langle \mu_1 \;\&\; \mu_2 \rangle.$$

Proof of Lemma 3.3. Let $\mu_\triangleleft$ and $\mu_\triangleright$ be any pair of witnesses to the approximate lifting. We will construct witnesses $\eta_\triangleleft, \eta_\triangleright$ with the desired support. For ease of notation, let $S_i \triangleq \mathrm{supp}(\mu_i)$ for $i \in \{1, 2\}$. Define:

$$\eta_\triangleleft(a,b) = \begin{cases} \mu_\triangleleft(a,b) & : (a,b) \in S_1 \times S_2 \\ \mu_\triangleleft[a, B^\star - S_2] & : b = \star \end{cases}$$

$$\eta_\triangleright(a,b) = \begin{cases} \mu_\triangleright(a,b) & : (a,b) \in S_1 \times S_2 \\ \mu_\triangleright[A^\star - S_1, b] & : a = \star \end{cases}$$

Evidently, $\eta_\triangleleft$ and $\eta_\triangleright$ have support in $S_1^\star \times S_2^\star$. Additionally, it is straightforward to check that $\pi_1^\sharp(\eta_\triangleleft) = \pi_1^\sharp(\mu_\triangleleft) = \mu_1$ and $\pi_2^\sharp(\eta_\triangleright) = \pi_2^\sharp(\mu_\triangleright) = \mu_2$ so $\eta_\triangleleft$ and $\eta_\triangleright$ have the desired marginals.

It only remains to check the distance condition. By the definition of the distance Δ_ε, we know that there are non-negative values $\delta(a,b)$ such that (i) $\overline{\mu_\triangleleft}(a,b) \leq e^\varepsilon \overline{\mu_\triangleright}(a,b) + \delta(a,b)$ and (ii) $\sum_{a,b} \delta(a,b) \leq \delta$. We can define new constants:

$$\zeta(a,b) = \begin{cases} \delta(a,b) & : (a,b) \in S_1 \times S_2 \cup \{\star\} \times B \\ \delta[a, B^\star - S_2] & : b = \star. \end{cases}$$

Since $\overline{\mu_\triangleleft}(\star, b) = \overline{\eta_\triangleleft}(\star, b) = 0$ for all $b \in B^\star$, and $\overline{\mu_\triangleright}(a,b) = \overline{\eta_\triangleright}(a,b) = 0$ for all $b \notin S_2$, point (i) holds for the witnesses $\eta_\triangleleft, \eta_\triangleright$ and constants $\zeta(a,b)$. Since $\sum_{a,b} \zeta(a,b) = \sum_{a,b} \delta(a,b) \leq \delta$, point (ii) holds as well. Hence, $\Delta_\varepsilon(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) \leq \delta$ and we have witnesses for the desired approximate lifting. □

Proof of Lemma 3.4. **($\Longrightarrow$):** Let P be (ε, δ)-differentially private w.r.t. ϕ and let $(a_1, a_2) \in \phi$. Let X be a subset of B. By definition of differential privacy, we have $P(a_1)[X] \leq e^\varepsilon \cdot P(a_2)[X] + \delta = e^\varepsilon \cdot P(a_2)[(=)(X)] + \delta$. Recall that the image of a set X under a binary relation is simply the set of all elements related to some element in X. In particular, $(=)(X)$ is just X. Hence, by application of Theorem 3.12, we have $P(a_1) =^{(\star)}_{\varepsilon,\delta} P(a_2)$.
($\Longleftarrow$): By application of Theorem 3.12, we have that

$$\forall a_1, a_2 \in A, \forall X \subseteq B.\, (a_1, a_2) \in \phi \implies P(a_1)[X] \leq e^\varepsilon \cdot P(a_2)[X] + \delta.$$

This is the definition of P being (ε, δ)-differentially private w.r.t. ϕ. □

Proof of Lemma 3.5.

- Immediate.

- Let $\overline{\varepsilon} \triangleq \varepsilon + \varepsilon'$ and $\overline{\delta} \triangleq \delta + e^{\varepsilon} \cdot \delta'$. By Theorem 3.12, it is sufficient to show that $\mu_1(X) \le e^{\overline{\varepsilon}} \cdot \mu_1(\mathcal{S}(\mathcal{R}(X))) + \overline{\delta}$ for any set X. We have:
$$\begin{aligned} \mu_1[X] &\le e^{\varepsilon} \cdot \mu_2[\mathcal{R}(X)] + \delta && \text{(Theorem 3.12)} \\ &\le e^{\varepsilon} \cdot (e^{\varepsilon'} \cdot \mu_3[\mathcal{S}(\mathcal{R}(X))] + \delta') + \delta && \text{(Theorem 3.12)} \\ &= e^{\varepsilon+\varepsilon'} \cdot \mu_3[\mathcal{S}(\mathcal{R}(X))] + e^{\varepsilon} \cdot \delta' + \delta. \end{aligned}$$
- We know that $\exists \, \langle \mu_{\triangleleft}, \mu_{\triangleright} \rangle \blacktriangleleft^{\mathcal{R}}_{\varepsilon,\delta} \langle \mu_1 \,\&\, \mu_2 \rangle$. Likewise, for $a \triangleq (a_1, a_2) \in \mathcal{R}$, $\exists \, \langle \eta_{\triangleleft,a}, \eta_{\triangleright,a} \rangle \blacktriangleleft^{\mathcal{S}}_{\varepsilon',\delta'} \langle \eta_1(a_1) \,\&\, \eta_2(a_2) \rangle$. Let $\eta_{\triangleleft}$ and $\eta_{\triangleright}$ be the following distribution constructors:
$$\eta_{\triangleleft} : a \mapsto \begin{cases} \eta_{\triangleleft,a} & \text{if } a \in \mathcal{R} \\ \mathbb{0} & \text{otherwise} \end{cases} \qquad \eta_{\triangleright} : a \mapsto \begin{cases} \eta_{\triangleright,a} & \text{if } a \in \mathcal{R} \\ \mathbb{0} & \text{otherwise} \end{cases}$$
and let $\xi_{\triangleleft} \triangleq \mathbb{E}_{\mu_{\triangleleft}}[\eta_{\triangleleft}]$ (resp. $\xi_{\triangleright} \triangleq \mathbb{E}_{\mu_{\triangleright}}[\eta_{\triangleright}]$). We now prove that:
$$\langle \xi_{\triangleleft}, \xi_{\triangleright} \rangle \blacktriangleleft^{\mathcal{S}}_{\varepsilon+\varepsilon',\delta+\delta'} \langle \mathbb{E}_{\mu_1}[\eta_1] \,\&\, \mathbb{E}_{\mu_2}[\eta_2] \rangle.$$
The marginal and support conditions are immediate. The distance condition is obtained by an immediate application of the previous point. □

Proof of Theorem 3.12. To show the forward direction of Theorem 3.12, let $X \subseteq A$ and $\overline{\mathcal{R}(X)} = B - \mathcal{R}(X)$. Then, we have
$$\begin{aligned} \mu_1[X] &= \pi_1^{\sharp}(\eta_{\triangleleft})[X] = \eta_{\triangleleft}[X, B^{\star}] = \overline{\eta_{\triangleleft}}[X, B^{\star}] = \overline{\eta_{\triangleleft}}[X, \mathcal{R}(X) \uplus \overline{\mathcal{R}(X)} \uplus \{\star\}] \\ &= \overline{\eta_{\triangleleft}}[X, \mathcal{R}(X) \uplus \{\star\}] + \underbrace{\overline{\eta_{\triangleleft}}[X, \overline{\mathcal{R}(X)}]}_{=\,0} \le e^{\varepsilon} \cdot \overline{\eta_{\triangleright}}[X, \mathcal{R}(X) \uplus \{\star\}] + \delta \\ &\le e^{\varepsilon} \cdot \underbrace{\overline{\eta_{\triangleright}}[A^{\star}, \mathcal{R}(X)]}_{=\,\eta_{\triangleright}[A^{\star},\mathcal{R}(X)]} + e^{\varepsilon} \cdot \underbrace{\overline{\eta_{\triangleright}}[A^{\star}, \{\star\}]}_{=\,0} + \delta \\ &= e^{\varepsilon} \cdot \pi_2^{\sharp}(\eta_{\triangleright})[\mathcal{R}(X)] + \delta = e^{\varepsilon} \cdot \mu_2[\mathcal{R}(X)] + \delta, \end{aligned}$$
as desired. □

Proof of Lemma 3.13. If $\eta_L \in \mathbb{D}(S_1 \times S_2^{\star})$, $\eta_R \in \mathbb{D}(S_1^{\star} \times S_2)$ are the witnesses of $(\mu_1)_{|S_1} \; R^{(\star)}_{\varepsilon,\delta} \; (\mu_2)_{|S_2}$, their extensions to $\mathbb{D}(A \times B^{\star})$ and $\mathbb{D}(A^{\star} \times B)$, padding with 0, form witnesses for $\mu_1 \; R^{(\star)}_{\varepsilon,\delta} \; \mu_2$. □

Proof of Lemma 3.14. The case for distributions with finite domains can be proven using the same proof of Theorem 3.12, using the standard (finite) max-flow min-cut theorem. The result extends to distributions with finite supports via Lemma 3.13. □

Proof of Lemma 3.15. Let $\mu_{\triangleleft}$, $\mu_{\triangleright}$ be witnesses of $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. For $(a, b) \in A \times B$, let
$$\begin{aligned} \nu_{\triangleleft}(a, b) &\triangleq \min(\mu_{\triangleleft}(a, b), e^{\varepsilon} \cdot \mu_{\triangleright}(a, b)) \\ \nu_{\triangleright}(a, b) &\triangleq \min(\mu_{\triangleleft}(a, b), \mu_{\triangleright}(a, b)) \end{aligned}$$

and define $\eta_{\triangleleft}$, $\eta_{\triangleright}$ as

$$\eta_{\triangleleft} : (a,b) \in A \times B^{\star} \mapsto \begin{cases} \nu_{\triangleleft}(a,b) & \text{if } b \neq \star \\ \mu_1(a) - \sum_{x \in B} \nu_{\triangleleft}(a,x) & \text{otherwise,} \end{cases}$$

$$\eta_{\triangleright} : (a,b) \in A^{\star} \times B \mapsto \begin{cases} \nu_{\triangleright}(a,b) & \text{if } a \neq \star \\ \mu_2(b) - \sum_{x \in A} \nu_{\triangleright}(x,b) & \text{otherwise.} \end{cases}$$

Note that $\eta_{\triangleleft}$ and $\eta_{\triangleright}$ are well-defined as sub-distributions by the marginal conditions for $\mu_{\triangleleft}$ and $\mu_{\triangleright}$. To show that the witnesses are non-negative, let $a \in A$ and $b \in B$. We have:

$$\begin{aligned} \eta_{\triangleleft}(a,\star) &= \mu_1(a) - \sum_{b \in B} \nu_{\triangleleft}(a,b) \geq \mu_1(a) - \sum_{b \in B} \mu_{\triangleleft}(a,b) \\ &\geq \mu_1(a) - \sum_{b \in B^{\star}} \mu_{\triangleleft}(a,b) = \mu_1(a) - \mu_1(a) = 0, \\ \eta_{\triangleright}(\star,b) &= \mu_2(b) - \sum_{a \in A} \nu_{\triangleright}(a,b) \geq \mu_2(b) - \sum_{a \in A} \nu_{\triangleright}(a,b) \\ &\geq \mu_2(b) - \sum_{a \in A^{\star}} \mu_{\triangleright}(a,b) = \mu_2(b) - \mu_2(b) = 0. \end{aligned}$$

To show that the witnesses sum to at most 1, we have

$$\begin{aligned} \sum_{(a,b) \in A \times B^{\star}} \eta_{\triangleleft}(a,b) &= \sum_{a \in A} \mu_1(a) = |\mu_1| \leq 1, \\ \sum_{(a,b) \in A^{\star} \times B} \eta_{\triangleright}(a,b) &= \sum_{b \in B} \mu_2(b) = |\mu_2| \leq 1. \end{aligned}$$

Moreover, these witness distributions satisfy the claimed distance condition. Indeed for $(a,b) \in A \times B$, we have:

$$\eta_{\triangleright}(a,b) = \min(\mu_1(a), \mu_2(b)) \leq \min(\mu_1(a), e^{\varepsilon} \cdot \mu_2(b)) = \eta_{\triangleleft}(a,b) \tag{A.1a}$$

$$\eta_{\triangleleft}(a,b) = \min(\mu_1(a), e^{\varepsilon} \cdot \mu_2(b)) \leq \min(e^{\varepsilon} \cdot \mu_1(a), e^{\varepsilon} \cdot \mu_2(b)) = e^{\varepsilon} \cdot \eta_{\triangleright}(a,b). \tag{A.1b}$$

It remains to prove that $\eta_{\triangleleft}$, $\eta_{\triangleright}$ are witnesses for $\mu_1 \mathrel{\mathcal{R}^{(\star)}_{\varepsilon,\delta}} \mu_2$. The marginals conditions are obvious. For the support condition, let $a \in A$ and $b \in B$. Then, $\eta_{\triangleleft}(a,b) > 0$ (resp. $\eta_{\triangleright}(a,b) > 0$) implies $\mu_{\triangleleft}(a,b) > 0$ (resp. $\eta_{\triangleright}(a,b) > 0$) and hence that $a \mathrel{\mathcal{R}} b$.

The distance condition follows by a calculation. For $X \subseteq A^{\star} \times B^{\star}$, we have:

$$\begin{aligned} \overline{\eta_{\triangleleft}}[X] &= \eta_{\triangleleft}[X \cap (A \times B)] + \eta_{\triangleleft}[X \cap (A \times \{\star\})] \\ &\leq e^{\varepsilon} \cdot \eta_{\triangleright}[X \cap (A \times B)] + \eta_{\triangleleft}[X \cap (A \times \{\star\})] && \text{(By Eq. (A.1))} \\ &\leq e^{\varepsilon} \cdot \overline{\eta_{\triangleright}[X]} + \eta_{\triangleleft}[X \cap (A \times \{\star\})]. \end{aligned}$$

To conclude, it suffices to show that $\eta_{\triangleleft}[X \cap (A \times \{\star\})] \leq \delta$. For that, let

$$\zeta(a,b) \triangleq \max(\overline{\mu_{\triangleleft}}(a,b) - e^{\varepsilon} \cdot \overline{\mu_{\triangleright}}(a,b), 0).$$

Then, we can bound

$$\begin{aligned}
\eta_{\triangleleft}[X \cap (A \times \{\star\})] &= \sum_{a \in A} \eta_{\triangleleft}(a, \star) = \sum_{a \in A} \left(\mu_1(a) - \sum_{b \in B} \nu_{\triangleleft}(a, b) \right) \\
&= \sum_{a \in A} \left(\mu_1(a) - \sum_{b \in B} \mu_{\triangleleft}(a, b) - \zeta(a, b) \right) \\
&= \sum_{a \in A} \left(\mu_1(a) - \sum_{b \in B} \mu_{\triangleleft}(a, b) \right) + \sum_{(a,b) \in A \times B} \zeta(a, b) \\
&= \sum_{a \in A} \left(\sum_{b \in B^\star} \mu_{\triangleleft}(a, b) - \sum_{b \in B} \mu_{\triangleleft}(a, b) \right) + \sum_{(a,b) \in A \times B} \zeta(a, b) \\
&= \sum_{a \in A} \underbrace{\mu_{\triangleleft}(a, \star)}_{= \zeta(a, \star)} + \sum_{(a,b) \in A \times B} \zeta(a, b) = \sum_{(a,b) \in A \times B^\star} \zeta(a, b) \\
&\leq \sum_{(a,b) \in A^\star \times B^\star} \zeta(a, b).
\end{aligned}$$

Now, let $S \triangleq \{(a, b) \in A^\star \times B^\star \mid e^\varepsilon \cdot \mu_{\triangleright}(a, b) < \mu_{\triangleleft}(a, b)\}$. By the distance condition on $\mu_{\triangleleft}$ and $\mu_{\triangleright}$, we have $\overline{\mu_{\triangleleft}}[S] - e^\varepsilon \cdot \overline{\mu_{\triangleright}}[S] \leq \delta$. Hence we conclude the desired distance condition:

$$\begin{aligned}
\sum_{(a,b) \in A^\star \times B^\star} \zeta(a, b) &= \sum_{(a,b) \in S} \zeta(a, b) + \sum_{(a,b) \notin S} \underbrace{\zeta(ab)}_{= 0} \\
&= \sum_{(a,b) \in S} \overline{\mu_{\triangleleft}}(a, b) - e^\varepsilon \cdot \overline{\mu_{\triangleright}}(a, b) \\
&= \overline{\mu_{\triangleleft}}[S] - e^\varepsilon \cdot \overline{\mu_{\triangleright}}[S] \leq \delta.
\end{aligned}$$

□

Proof of Lemma 4.1. **($\Longrightarrow$):** Assume that $f_1^\sharp(\mu_1) \; \mathcal{R}_{\varepsilon,\delta}^{(\star)} \; f_2^\sharp(\mu_2)$ and let $X \subseteq A_1$. Then,

$$\begin{aligned}
\mu_1[X] &\leq \mu_1[f_1^{-1}(f_1(X))] = f_1^\sharp(\mu_1)[f_1(X)] \\
&\leq e^\varepsilon \cdot f_2^\sharp(\mu_2)[\mathcal{R}(f_1(X))] + \delta && \text{(Theorem 3.12)} \\
&= e^\varepsilon \cdot \mu_2[\underbrace{f_2^{-1}(\mathcal{R}(f_1(X)))}_{\subseteq \mathcal{S}(X)}] + \delta \leq e^\varepsilon \cdot \mu_2[\mathcal{S}\ (X)] + \delta.
\end{aligned}$$

Hence, by Theorem 3.12, $\mu_1 \; \mathcal{S}_{\varepsilon,\delta}^{(\star)} \; \mu_2$.

($\Longleftarrow$): Assume that $\mu_1 \; \mathcal{S}_{\varepsilon,\delta}^{(\star)} \; \mu_2$ and let $X \subseteq A_2$. Then,

$$\begin{aligned}
f_1^\sharp(\mu_1)[X] &= \mu_1[f_1^{-1}(X)] \\
&\leq e^\varepsilon \cdot \mu_2[\underbrace{\mathcal{S}\ (f_1^{-1}(X))}_{\subseteq f_2^{-1}(\mathcal{R}(X))}] + \delta \leq e^\varepsilon \cdot f_2^\sharp(\mu_2)[\mathcal{R}(X)] + \delta. && \text{(Theorem 3.12)}
\end{aligned}$$

Hence, by Theorem 3.12, $f_1^\sharp(\mu_1) \; \mathcal{R}_{\varepsilon,\delta}^{(\star)} \; f_2^\sharp(\mu_2)$. □

Proof of Lemma 4.2. By Theorem 3.12, it is sufficient to prove that

$$\mu_1[X] \leq e^\varepsilon \cdot \mu_2[\mathcal{R}(X)] + \mu_1[\overline{\theta}] + \delta$$

for any $X \subseteq A$. By direct computation:

$$\begin{aligned}\mu_1[X] &= \mu_1[X \cap \theta] + \mu_1[X \cap \overline{\theta}] \leq \mu_1[X \cap \theta] + \mu_1[\overline{\theta}] \\ &\leq e^{\varepsilon} \cdot \mu_2[\underbrace{(\theta_{\triangleleft} \implies \mathcal{R})(X \cap \theta)}_{= \, \mathcal{R}(X \cap \theta) \, \subseteq \, \mathcal{R}(X)}] + \delta + \mu_1[\overline{\theta}] \\ &\leq e^{\varepsilon} \cdot \mu_2[\mathcal{R}(X)] + \mu_1[\overline{\theta}] + \delta. \qquad \square\end{aligned}$$

Proof of Lemma 4.3. By Theorem 3.12, it is sufficient to prove that

$$\mu_1[X] \leq e^{\varepsilon} \cdot \mu_2[\mathcal{R}(X)] + e^{\varepsilon} \cdot \mu_2[\overline{\theta}] + \delta$$

for any $X \subseteq A$. Let X be such a set, then:

$$\begin{aligned}\mu_1[X] &\leq e^{\varepsilon} \cdot \mu_2[(\theta_{\triangleright} \implies \mathcal{R})(X)] + \delta \\ &\leq e^{\varepsilon} \cdot (\mu_2[(\theta_{\triangleright} \implies \mathcal{R})(X) \cap \theta] + \mu_2[\overline{\theta}]) + \delta \\ &\leq e^{\varepsilon} \mu_2[\underbrace{(\theta_{\triangleright} \implies \mathcal{R})(X) \cap \theta}_{\subseteq \mathcal{R}(X) \cap \theta}] + e^{\varepsilon} \cdot \mu_2[\overline{\theta}] + \delta \\ &\leq e^{\varepsilon} \mu_2[\mathcal{R}(X)] + e^{\varepsilon} \cdot \mu_2[\overline{\theta}] + \delta. \qquad \square\end{aligned}$$

Proof of Lemma 4.4. From $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$ and Lemma 3.5, we have $\mu_1 \; \mathcal{S}^{(\star)}_{\varepsilon,\delta} \; \mu_2$ where $\mathcal{S} \triangleq \theta_{a,\triangleleft} \implies \theta_{a,\triangleleft} \cap \mathcal{R}$. Hence, by Lemma 4.2, we obtain $\mu_1 \; (\theta_{a,\triangleleft} \cap \mathcal{R})^{(\star)}_{\varepsilon,\delta_a} \; \mu_2$. Using similar reasoning with $\theta_{b,\triangleright} \implies \theta_{b,\triangleright} \cap \mathcal{R}$ and Lemma 4.3, we have $\mu_1 \; (\theta_{b,\triangleright} \cap \mathcal{R})^{(\star)}_{\varepsilon,\delta_b} \; \mu_2$. □

Proof of Theorem 5.3. Suppose that (μ_L, μ_R) are witnesses to $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Define the witness $\eta \in \mathbb{D}(A \times B)$ as the point-wise minimum: $\eta(a,b) = \min(\mu_L(a,b), \mu_R(a,b))$. We will show that η is a witness to $\mu_1 \; \mathcal{R}^{(1)}_{\varepsilon,\delta} \; \mu_2$.

The support condition follows from the support condition for (μ_L, μ_R). The marginal conditions $\pi_1^{\sharp}(\eta) \leq \mu_1$ and $\pi_2^{\sharp}(\eta) \leq \mu_2$ also follow by the marginal conditions for (μ_L, μ_R). The only thing to check is the distance condition. By the distance condition on (μ_L, μ_R), there exist non-negative values $\delta(a,b)$ such that

$$\mu_L(a,b) \leq \exp(\varepsilon)\mu_R(a,b) + \delta(a,b)$$

and $\sum_{a,b} \delta(a,b) \leq \delta$. So, $\mu_R(a,b) \geq \exp(-\varepsilon)(\mu_L(a,b) - \delta(a,b))$. Now let $S \subseteq A$ be any subset. Then:

$$\begin{aligned}\mu_1(S) - \exp(\varepsilon)\pi_1^{\sharp}(\eta)(S) &= \sum_{a \in S} \mu_1(a) - \exp(\varepsilon) \sum_{b \in B} \min(\mu_L(a,b), \mu_R(a,b)) \\ &\leq \sum_{a \in S} \mu_1(a) - \exp(\varepsilon) \sum_{b \in B} \exp(-\varepsilon)(\mu_L(a,b) - \delta(a,b)) \\ &= \sum_{a \in S, b \in B} \delta(a,b) \leq \delta.\end{aligned}$$

Thus, η witnesses $\mu_1 \; \mathcal{R}^{(1)}_{\varepsilon,\delta} \; \mu_2$, so $\mathcal{R}^{(\star)}_{\varepsilon,\delta} \subseteq \mathcal{R}^{(1)}_{\varepsilon,\delta}$.

The other direction is more interesting. Let $\eta \in \mathbb{D}(A \times B)$ be the witness for $\mathcal{R}^{(1)}_{\varepsilon,\delta}$. By the distance condition $\Delta_{\varepsilon}(\mu_1, \pi_1^{\sharp}\eta) \leq \delta$, there exist non-negative values $\delta(a)$ such that

$$\mu_1(a) \leq \exp(\varepsilon)\pi_1^{\sharp}\eta(a) + \delta(a)$$

with equality when $\delta(a)$ is strictly positive, and $\sum_{a\in A}\delta(a) \leq \delta$. Define two witnesses $\mu_L \in \mathbb{D}(A \times B^\star), \mu_R \in \mathbb{D}(A^\star \times B)$ as follows:

$$\mu_L(a,b) = \begin{cases} \eta(a,b) \cdot \frac{\mu_1(a)-\delta(a)}{\pi_1^\sharp \eta(a)} & : b \neq \star \\ \mu_1(a) - \sum_{b\in B} \mu_L(a,b) & : b = \star \end{cases}$$

$$\mu_R(a,b) = \begin{cases} \eta(a,b) & : a \neq \star \\ \mu_2(b) - \sum_{a\in A} \mu_R(a,b) & : a = \star. \end{cases}$$

(As usual, if any denominator is zero, we take the probability to be zero as well.)

The support condition follows from the support condition of η. The marginal conditions hold by definition. Note that all probabilities are non-negative. For μ_L, note that if $\delta(a) > 0$ then $\mu_1(a) - \delta(a) = \exp(\varepsilon)\pi_1^\sharp\eta(a) \geq 0$ and hence

$$\mu_L(a,\star) = \mu_1(a) - \delta(a) \geq 0.$$

assuming $\pi_1^\sharp\eta(a) > 0$; if $\pi_1^\sharp\eta(a) = 0$ then $\mu_L(a,\star) = 0$. For μ_R, non-negativity holds because $\pi_2^\sharp\eta \leq \mu_2$.

We show the distance bound. Note that when $a, b \neq \star$, by definition $\mu_L(a,b)$ and $\mu_R(a,b)$ are both strictly positive or both equal to zero, and $\eta(a,b)$ is strictly positive or equal to zero accordingly. If $\mu_L(a,b), \mu_R(a,b), \eta(a,b)$ are all strictly positive, then we know

$$\frac{\mu_L(a,b)}{\eta(a,b)} = \frac{\mu_1(a) - \delta(a)}{\pi_1^\sharp\eta(a)} \leq \exp(\varepsilon).$$

Thus we always have

$$\mu_L(a,b) \leq \exp(\varepsilon)\eta(a,b) = \exp(\varepsilon)\mu_R(a,b).$$

We can also bound the mass on points $(a,\star)$. Let $S \subseteq A$ be any subset. Then:

$$\begin{aligned} \overline{\mu_L}(S \times \{\star\}) &= \sum_{a\in S} \mu_1(a) - \mu_1(a) \sum_{b\in B} \frac{\eta(a,b)}{\pi_1^\sharp\eta(a)} + \delta(a) \sum_{b\in B} \frac{\eta(a,b)}{\pi_1^\sharp\eta(a)} \\ &= \mu_1(S) - \mu_1(S) + \delta(S) \leq \exp(\varepsilon)\overline{\mu_R}(S \times \{\star\}) + \delta. \end{aligned}$$

So $\Delta_\varepsilon(\overline{\mu_L}, \overline{\mu_R}) \leq \delta$ as desired, and we have witnesses to $\mu_1 \; \mathcal{R}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Hence, $\mathcal{R}^{(1)}_{\varepsilon,\delta} \subseteq \mathcal{R}^{(\star)}_{\varepsilon,\delta}$. □

Proof of Theorem 6.3. Suppose that (μ_L, μ_R) are witnesses to $\mu_1 \; \overline{\mathcal{R}}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Define the witness $\eta \in \mathbb{D}(A \times B)$ as the point-wise minimum: $\eta(a,b) = \min(\mu_L(a,b), \mu_R(a,b))$. We will show that η is a witness to $\mu_1 \; \overline{\mathcal{R}}^{(1)}_{\varepsilon,\delta} \; \mu_2$.

The support condition follows from the support condition for (μ_L, μ_R). The marginal conditions $\pi_1^\sharp(\eta) \leq \mu_1$ and $\pi_2^\sharp(\eta) \leq \mu_2$ also follow by the marginal conditions for (μ_L, μ_R). The only thing to check is the distance condition. By the distance condition on (μ_L, μ_R), there exist non-negative values $\delta(a,b)$ such that

$$\mu_L(a,b) \leq \exp(\varepsilon)\mu_R(a,b) + \delta(a,b)$$

and $\sum_{a,b}\delta(a,b) \leq \delta$. So, $\mu_R(a,b) \geq \exp(-\varepsilon)(\mu_L(a,b) - \delta(a,b))$. Similarly, there are non-negative values $\delta'(a,b)$ such that

$$\mu_R(a,b) \leq \exp(\varepsilon)\mu_L(a,b) + \delta'(a,b)$$

and $\sum_{a,b}\delta'(a,b) \leq \delta$. So, $\mu_L(a,b) \geq \exp(-\varepsilon)(\mu_R(a,b) - \delta'(a,b))$.

Now let $S \subseteq A$ be any subset. Then:

$$\begin{aligned}
\mu_1(S) - \exp(\varepsilon)\pi_1^\sharp(\eta)(S) &= \sum_{a \in S} \mu_1(a) - \exp(\varepsilon) \sum_{b \in B} \min(\mu_L(a,b), \mu_R(a,b)) \\
&\leq \sum_{a \in S} \mu_1(a) - \exp(\varepsilon) \sum_{b \in B} \exp(-\varepsilon)(\mu_L(a,b) - \delta(a,b)) \\
&= \sum_{a \in S, b \in B} \delta(a,b) \leq \delta.
\end{aligned}$$

The other marginal is similar. For any subset $T \subseteq B$ we have

$$\begin{aligned}
\mu_2(T) - \exp(\varepsilon)\pi_2^\sharp(\eta)(T) &= \sum_{b \in T} \mu_2(b) - \exp(\varepsilon) \sum_{a \in A} \min(\mu_L(a,b), \mu_R(a,b)) \\
&\leq \sum_{b \in T} \mu_2(b) - \exp(\varepsilon) \sum_{a \in A} \exp(-\varepsilon)(\mu_R(a,b) - \delta'(a,b)) \\
&= \sum_{b \in T, a \in A} \delta'(a,b) \leq \delta.
\end{aligned}$$

Thus, η witnesses $\mu_1 \; \overline{\mathcal{R}}^{(1)}_{\varepsilon,\delta} \; \mu_2$.

The other direction is more interesting. Let $\eta \in \mathbb{D}(A \times B)$ be the single witness to $\overline{\mathcal{R}}^{(1)}_{\varepsilon,\delta}$. By the distance conditions $\Delta_\varepsilon(\mu_1, \pi_1^\sharp \eta) \leq \delta$ and $\Delta_\varepsilon(\mu_2, \pi_2^\sharp \eta) \leq \delta$, there exist non-negative values $\delta(a)$ and $\delta'(b)$ such that

$$\begin{aligned}
\mu_1(a) &\leq \exp(\varepsilon)\pi_1^\sharp \eta(a) + \delta(a) \\
\mu_2(b) &\leq \exp(\varepsilon)\pi_2^\sharp \eta(b) + \delta'(b),
\end{aligned}$$

there is equality when $\delta(a)$ or $\delta'(b)$ are strictly positive, and both $\sum_{a \in A} \delta(a)$ and $\sum_{b \in B} \delta'(b)$ are at most δ. Define two witnesses $\mu_L \in \mathbb{D}(A \times B^\star), \mu_R \in \mathbb{D}(A^\star \times B)$ as follows:

$$\begin{aligned}
\mu_L(a,b) &= \begin{cases} \eta(a,b) \cdot \frac{\mu_1(a) - \delta(a)}{\pi_1^\sharp \eta(a)} & : b \neq \star \\ \mu_1(a) - \sum_{b \in B} \mu_L(a,b) & : b = \star \end{cases} \\
\mu_R(a,b) &= \begin{cases} \eta(a,b) \cdot \frac{\mu_2(b) - \delta'(b)}{\pi_2^\sharp \eta(b)} & : a \neq \star \\ \mu_2(b) - \sum_{a \in A} \mu_R(a,b) & : a = \star. \end{cases}
\end{aligned}$$

(As usual, if any denominator is zero, we take the probability to be zero as well.)

The support condition follows from the support condition of η. The marginal conditions hold by definition. Note that all probabilities are non-negative. For instance in μ_L, note that if $\delta(a) > 0$ then $\mu_1(a) - \delta(a) = \exp(\varepsilon)\pi_1^\sharp \eta(a) \geq 0$ and hence

$$\mu_L(a, \star) = \mu_1(a) - \delta(a) \geq 0.$$

assuming $\pi_1^\sharp \eta(a) > 0$; if $\pi_1^\sharp \eta(a) = 0$ then $\mu_L(a, \star) = 0$. A similar argument shows that μ_R is non-negative.

So, it remains to check the distance bounds. Note that when $a, b \neq \star$, by definition $\mu_L(a,b)$ and $\mu_R(a,b)$ are both strictly positive or both equal to zero, and $\eta(a,b)$ is strictly positive or equal to zero accordingly. If $\mu_L(a,b), \mu_R(a,b), \eta(a,b)$ are all strictly positive, then we know

$$\begin{aligned}\frac{\mu_L(a,b)}{\eta(a,b)} &= \frac{\mu_1(a)-\delta(a)}{\pi_1^\sharp\eta(a)} \leq \exp(\varepsilon)\\ \frac{\mu_R(a,b)}{\eta(a,b)} &= \frac{\mu_2(b)-\delta'(b)}{\pi_2^\sharp\eta(b)} \leq \exp(\varepsilon).\end{aligned}$$

We can also lower bound the ratios:

$$\begin{aligned}\frac{\mu_L(a,b)}{\eta(a,b)} &= \frac{\mu_1(a)-\delta(a)}{\pi_1^\sharp\eta(a)} \geq 1\\ \frac{\mu_R(a,b)}{\eta(a,b)} &= \frac{\mu_2(b)-\delta'(b)}{\pi_2^\sharp\eta(b)} \geq 1;\end{aligned}$$

for instance when $\delta(a) > 0$ then the ratio is exactly equal to $\exp(\varepsilon) \geq 1$, and when $\delta(a) = 0$ then the ratio is at least 1 by the marginal property $\pi_1^\sharp\eta \leq \mu_1$. So we have $\mu_L(a,b)/\eta(a,b)$ and $\mu_R(a,b)/\eta(a,b)$ in $[1, \exp(\varepsilon)]$ when all distributions are strictly positive. Thus we always have

$$\begin{aligned}\mu_L(a,b) &\leq \exp(\varepsilon)\mu_R(a,b)\\ \mu_R(a,b) &\leq \exp(\varepsilon)\mu_L(a,b).\end{aligned}$$

We can also bound the mass on points $(a,\star)$. Let $S \subseteq A$ be any subset. $\overline{\mu_R}(S \times \{\star\}) \leq \exp(\varepsilon)\overline{\mu_L}(S \times \{\star\}) + \delta$ is clear. For the other direction:

$$\begin{aligned}\overline{\mu_L}(S \times \{\star\}) &= \sum_{a \in S} \mu_1(a) - \mu_1(a) \sum_{b \in B} \frac{\eta(a,b)}{\pi_1^\sharp\eta(a)} + \delta(a) \sum_{b \in B} \frac{\eta(a,b)}{\pi_1^\sharp\eta(a)}\\ &= \mu_1(S) - \mu_1(S) + \delta(S) \leq \exp(\varepsilon)\overline{\mu_R}(S \times \{\star\}) + \delta.\end{aligned}$$

The mass at points $(\star, b)$ can be bounded in a similar way. Let $T \subseteq B$ be any subset. Then, $\overline{\mu_L}(\{\star\} \times T) \leq \exp(\varepsilon)\overline{\mu_R}(\{\star\} \times T) + \delta$ is clear. For the other direction:

$$\begin{aligned}\overline{\mu_R}(\{\star\} \times T) &= \sum_{b \in T} \mu_2(b) - \mu_2(b) \sum_{a \in A} \frac{\eta(a,b)}{\pi_2^\sharp\eta(b)} + \delta'(b) \sum_{a \in A} \frac{\eta(a,b)}{\pi_2^\sharp\eta(b)}\\ &= \mu_2(T) - \mu_2(T) + \delta'(T) \leq \exp(\varepsilon)\overline{\mu_L}(\{\star\} \times T) + \delta.\end{aligned}$$

So $\Delta_\varepsilon(\overline{\mu_L}, \overline{\mu_R}) \leq \delta$ and $\Delta_\varepsilon(\overline{\mu_R}, \overline{\mu_L}) \leq \delta$ so we have witnesses to $\mu_1 \; \overline{\mathcal{R}}^{(\star)}_{\varepsilon,\delta} \; \mu_2$. Hence, $\overline{\mathcal{R}}^{(1)}_{\varepsilon,\delta} = \overline{\mathcal{R}}^{(\star)}_{\varepsilon,\delta}$. □

Proof of Lemma 7.4. Let $\mu_\triangleleft$ and $\mu_\triangleright$ be any pair of witnesses to the approximate lifting. We will construct witnesses $\eta_\triangleleft, \eta_\triangleright$ with the desired support. For ease of notation, let $S_i \triangleq \mathrm{supp}(\mu_i)$ for $i \in \{1, 2\}$. Define:

$$\eta_\triangleleft(a,b) = \begin{cases} \mu_\triangleleft(a,b) & : (a,b) \in S_1 \times S_2\\ \mu_\triangleleft[a, B^\star - S_2] & : b = \star \end{cases}$$

$$\eta_\triangleright(a,b) = \begin{cases} \mu_\triangleright(a,b) & : (a,b) \in S_1 \times S_2\\ \mu_\triangleright[A^\star - S_1, b] & : a = \star \end{cases}$$

Evidently, $\eta_\triangleleft$ and $\eta_\triangleright$ have support in $S_1^\star \times S_2^\star$. Additionally, it is straightforward to check that $\pi_1^\sharp(\eta_\triangleleft) = \pi_1^\sharp(\mu_\triangleleft) = \mu_1$ and $\pi_2^\sharp(\eta_\triangleright) = \pi_2^\sharp(\mu_\triangleright) = \mu_2$ so $\eta_\triangleleft$ and $\eta_\triangleright$ have the desired marginals.

It only remains to check the distance condition. We can compute:

$$\begin{aligned}
\Delta_f(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) &= \sum_{(a,b)\in S_1\times S_2} \eta_\triangleright(a,b)\cdot f\left(\frac{\eta_\triangleleft(a,b)}{\eta_\triangleright(a,b)}\right) \\
&\quad + \sum_{a\in S_1} \eta_\triangleright(a,\star)\cdot f\left(\frac{\eta_\triangleleft(a,\star)}{\eta_\triangleright(a,\star)}\right) + \sum_{b\in S_2} \eta_\triangleright(\star,b)\cdot f\left(\frac{\eta_\triangleleft(\star,b)}{\eta_\triangleright(\star,b)}\right) \\
&= \sum_{(a,b)\in S_1\times S_2} \mu_\triangleright(a,b)\cdot f\left(\frac{\mu_\triangleleft(a,b)}{\mu_\triangleright(a,b)}\right) + \sum_{a\in S_1} \eta_\triangleleft(a,\star)\cdot L_f + \sum_{b\in S_2} \eta_\triangleright(\star,b)\cdot f(0) \\
&= \sum_{(a,b)\in S_1\times S_2} \mu_\triangleright(a,b)\cdot f\left(\frac{\mu_\triangleleft(a,b)}{\mu_\triangleright(a,b)}\right) \\
&\quad + \sum_{a\in S_1}\sum_{b'\in B^\star - S_2} \mu_\triangleleft(a,b')\cdot L_f + \sum_{b\in S_2}\sum_{a'\in A^\star - S_1} \mu_\triangleright(a',b)\cdot f(0)
\end{aligned}$$

Now, note that for all $b' \in B^\star - S_2$, we know $\mu_\triangleright(a,b') = 0$. Similarly, for all $a' \in A^\star - S_1$, we know $\mu_\triangleleft(a',b) = 0$. Hence, the last line is equal to

$$\begin{aligned}
\Delta_f(\overline{\eta_\triangleleft}, \overline{\eta_\triangleright}) &= \sum_{(a,b)\in S_1\times S_2} \mu_\triangleright(a,b)\cdot f\left(\frac{\mu_\triangleleft(a,b)}{\mu_\triangleright(a,b)}\right) \\
&\quad + \sum_{a\in S_1}\sum_{b'\in B^\star - S_2} \mu_\triangleright(a,b')\cdot f\left(\frac{\mu_\triangleleft(a,b')}{\mu_\triangleright(a,b')}\right) \\
&\qquad + \sum_{b\in S_2}\sum_{a'\in A^\star - S_1} \mu_\triangleright(a',b)\cdot f\left(\frac{\mu_\triangleleft(a',b)}{\mu_\triangleright(a',b)}\right) \\
&= \Delta_f(\overline{\mu_\triangleleft}, \overline{\mu_\triangleright}) \leq \delta.
\end{aligned}$$

Thus, $\eta_\triangleleft$ and $\eta_\triangleright$ witness the desired $\star$-lifting. □

Proof of Lemma 7.5. For the reverse direction, take the witnesses $\mu_\triangleleft, \mu_\triangleright \in \mathbb{D}(A^\star \times A^\star)$ and define witnesses $\nu_\triangleleft \triangleq (g_1^\star \times g_2^\star)^\sharp\ (\mu_\triangleleft)$ and $\nu_\triangleright \triangleq (g_1^\star \times g_2^\star)^\sharp\ (\mu_\triangleright)$, where $g_1^\star \times g_2^\star$ takes a pair (a_1, a_2) to the pair $(g_1(a_1), g_2(a_2))$ and maps $\star$ to $\star$. The support and marginal requirements are clear. The only thing to check is the distance condition, but this follows from monotonicity of f-divergences—under the mapping $g_1^\star \times g_2^\star : A^\star \times A^\star \to B^\star \times B^\star$, the f-divergence can only decrease (see, e.g., Csiszár and Shields (2004)).

For the forward direction, let $\nu_\triangleleft, \nu_\triangleright \in \mathbb{D}(B^\star \times B^\star)$ be the witnesses to the second lifting. By Lemma 7.4, we may assume without loss of generality that $\mathrm{supp}(\nu_\triangleleft)$ and $\mathrm{supp}(\nu_\triangleright)$ are contained in

$$\mathrm{supp}(g_1^\sharp\ (\mu_1))^\star \times \mathrm{supp}(g_2^\sharp\ (\mu_2))^\star \subseteq g_1(A)^\star \times g_2(A)^\star.$$

We aim to construct a pair of witnesses $\mu_\triangleleft, \mu_\triangleright \in \mathbb{D}(A^\star \times A^\star)$ to the first lifting. The basic idea is to define $\mu_\triangleleft$ and $\mu_\triangleright$ based on equivalence classes of elements in A mapping to a particular $b \in B$, and then smooth out the probabilities within each equivalence class. To begin, for $a \in A$, define $[a]_g \triangleq g^{-1}(g(a))$ and $\alpha_i(a) \triangleq \Pr_{\mu_i}[\{a\} \mid [a]_{g_i}]$. We take $\alpha_i(a) = 0$ when $\mu_i([a]_{f_i}) = 0$, and we let $\alpha_i(\star) = 0$. We define $\mu_\triangleleft$ and $\mu_\triangleright$ as

$$\begin{aligned}
\mu_\triangleleft &: (a_1, a_2) \mapsto \alpha_\triangleleft(a_1, a_2)\cdot \nu_\triangleleft(g_1(a_1), g_2(a_2)) \\
\mu_\triangleright &: (a_1, a_2) \mapsto \alpha_\triangleright(a_1, a_2)\cdot \nu_\triangleright(g_1(a_1), g_2(a_2))
\end{aligned}$$

where

$$\alpha_{\triangleleft}(a_1, a_2) = \begin{cases} \alpha_1(a_1) \cdot \alpha_2(a_2) & : a_2 \neq \star \\ \alpha_1(a_1) & : a_2 = \star, \end{cases} \qquad \alpha_{\triangleright}(a_1, a_2) = \begin{cases} \alpha_1(a_1) \cdot \alpha_2(a_2) & : a_1 \neq \star \\ \alpha_2(a_2) & : a_1 = \star. \end{cases}$$

The support and marginal conditions follow from the support and marginal conditions of $\nu_{\triangleleft}$, $\nu_{\triangleright}$. For instance:

$$\begin{aligned} \sum_{a_2 \in A^\star} \mu_{\triangleleft}(a_1, a_2) &= \sum_{a_2 \in A^\star} \alpha_{\triangleleft}(a_1, a_2) \nu_{\triangleleft}(g_1(a_1), g_2(a_2)) \\ &= \alpha_1(a_1) \nu_{\triangleleft}(g_1(a_1), \star) + \sum_{a_2 \in A} \alpha_1(a_1) \alpha_2(a_2) \nu_{\triangleleft}(g_1(a_1), g_2(a_2)) \\ &= \alpha_1(a_1) \left(\nu_{\triangleleft}(g_1(a_1), \star) + \sum_{b_2 \in g_2(A)} \nu_{\triangleleft}(g_1(a_1), b_2) \sum_{a_2 \in g_2^{-1}(b_2)} \alpha_2(a_2) \right) \\ &= \alpha_1(a_1) \sum_{b_2 \in B^\star} \nu_{\triangleleft}(g_1(a_1), b_2) = \alpha_1(a_1) \mu_1([a_1]_{g_1}) = \mu_1(a_1). \end{aligned}$$

In the last line, we replace the sum over $b_2 \in g_2(A^\star)$ with a sum over $b_2 \in B^\star$; this holds since the support of $g_2^\sharp\ (\mu_2)$ is contained in $g_2(A)$, so we can assume that $\nu_{\triangleleft}(a, b_2) = 0$ for all b_2 outside of $g_2(A^\star)$. Then, we can conclude by the marginal condition $\pi_1^\sharp(\nu_{\triangleleft}) = g_1^\sharp\ (\mu_1)$. The second marginal is similar.

We now check the distance condition $\Delta_f(\overline{\mu_{\triangleleft}}, \overline{\mu_{\triangleright}}) \leq \delta$. We can split the f-divergence into $\Delta_f(\overline{\mu_{\triangleleft}}, \overline{\mu_{\triangleright}}) = P_0 + P_1 + P_2 + P_3$, where

$$P_0 \triangleq \mu_{\triangleright}(\star, \star) \cdot f\left(\frac{\mu_{\triangleleft}(\star, \star)}{\mu_{\triangleright}(\star, \star)}\right) \qquad P_1 \triangleq \sum_{(a_1, a_2) \in A \times A} \mu_{\triangleright}(a_1, a_2) \cdot f\left(\frac{\mu_{\triangleleft}(a_1, a_2)}{\mu_{\triangleright}(a_1, a_2)}\right)$$

$$P_2 \triangleq \sum_{a_1 \in A} \mu_{\triangleright}(a_1, \star) \cdot f\left(\frac{\mu_{\triangleleft}(a_1, \star)}{\mu_{\triangleright}(a_1, \star)}\right) \qquad P_3 \triangleq \sum_{a_2 \in A} \mu_{\triangleright}(\star, a_2) \cdot f\left(\frac{\mu_{\triangleleft}(\star, a_2)}{\mu_{\triangleright}(\star, a_2)}\right)$$

We will handle each term separately. Evidently $P_0 = 0$. For P_1, we have

$$\begin{aligned} P_1 &= \sum_{(a_1, a_2) \in A \times A} \alpha_{\triangleright}(a_1, a_2) \nu_{\triangleright}(g_1(a_1), g_2(a_2)) \cdot f\left(\frac{\alpha_{\triangleleft}(a_1, a_2) \nu_{\triangleleft}(g_1(a_1), g_2(a_2))}{\alpha_{\triangleright}(a_1, a_2) \nu_{\triangleright}(g_1(a_1), g_2(a_2))}\right) \\ &= \sum_{(a_1, a_2) | S^{=0}} \alpha_{\triangleleft}(a_1, a_2) \nu_{\triangleleft}(g_1(a_1), g_2(a_2)) \cdot L_f \\ &\quad + \sum_{(a_1, a_2) | S^{\neq 0}} \alpha_{\triangleright}(a_1, a_2) \nu_{\triangleright}(g_1(a_1), g_2(a_2)) \cdot f\left(\frac{\nu_{\triangleleft}(g_1(a_1), g_2(a_2))}{\nu_{\triangleright}(g_1(a_1), g_2(a_2))}\right) \end{aligned}$$

where the sets $S^{=0}$ and $S^{\neq 0}$ are

$$\begin{aligned} S^{=0} &\triangleq \{(a_1, a_2) \mid \nu_{\triangleright}(g_1(a_1), g_2(a_2)) = 0\} \\ S^{\neq 0} &\triangleq \{(a_1, a_2) \mid \nu_{\triangleright}(g_1(a_1), g_2(a_2)) \neq 0\}. \end{aligned}$$

By further rearranging,

$$
\begin{aligned}
P_1 &= \sum_{(b_1,b_2)\in(g_1\times g_2)(S^{=0})} \nu_{\triangleleft}(b_1,b_2)\cdot L_f \left(\sum_{a_1\in g_1^{-1}(b_1)} \alpha_1(a_1)\right)\left(\sum_{a_2\in g_1^{-1}(b_2)} \alpha_2(a_2)\right) \\
&\qquad + \sum_{(b_1,b_2)\in(g_1\times g_2)(S^{\neq 0})} \nu_{\triangleright}(b_1,b_2)\cdot f\left(\frac{\nu_{\triangleleft}(b_1,b_2)}{\nu_{\triangleright}(b_1,b_2)}\right)\left(\sum_{a_1\in g_1^{-1}(b_1)} \alpha_1(a_1)\right)\left(\sum_{a_2\in g_1^{-1}(b_2)} \alpha_2(a_2)\right) \\
&= \sum_{(b_1,b_2)\in(g_1\times g_2)(S^{=0})} \nu_{\triangleleft}(b_1,b_2)\cdot L_f + \sum_{(b_1,b_2)\in(g_1\times g_2)(S^{\neq 0})} \nu_{\triangleright}(b_1,b_2)\cdot f\left(\frac{\nu_{\triangleleft}(b_1,b_2)}{\nu_{\triangleright}(b_1,b_2)}\right) \\
&= \sum_{(b_1,b_2)\in(g_1\times g_2)(A\times A)} \nu_{\triangleright}(b_1,b_2)\cdot f\left(\frac{\nu_{\triangleleft}(b_1,b_2)}{\nu_{\triangleright}(b_1,b_2)}\right) \\
&= \sum_{(b_1,b_2)\in B\times B} \nu_{\triangleright}(b_1,b_2)\cdot f\left(\frac{\nu_{\triangleleft}(b_1,b_2)}{\nu_{\triangleright}(b_1,b_2)}\right).
\end{aligned}
$$

The final equality is because without loss of generality, we can assume (by Lemma 7.4) that $\nu_{\triangleleft}, \nu_{\triangleright}$ are zero outside of the support of $g_1^{\sharp}(\mu_1)$ and $g_2^{\sharp}(\mu_2)$, which have support contained in $(g_1\times g_2)(A\times A)$.

The remaining two terms P_2 and P_3 are simpler to bound. For P_2, note that $\overline{\mu_{\triangleright}}(a,\star)=0$ for all $a\in A$. Thus:

$$
\begin{aligned}
P_2 &= \sum_{a_1\in A} \alpha_{\triangleleft}(a_1,\star)\nu_{\triangleleft}(g_1(a_1),\star)\cdot L_f = \sum_{b_1\in g_1(A)}\sum_{a_1\in g_1^{-1}(b_1)} \alpha_1(a_1)\nu_{\triangleleft}(b_1,\star)\cdot L_f \\
&= \sum_{b_1\in g_1(A)} \nu_{\triangleleft}(b_1,\star)\cdot L_f = \sum_{b_1\in B} \nu_{\triangleleft}(b_1,\star)\cdot L_f = \sum_{b_1\in B} \overline{\nu_{\triangleright}}(b_1,\star)\cdot f\left(\frac{\nu_{\triangleleft}(b_1,\star)}{\overline{\nu_{\triangleright}}(b_1,\star)}\right)
\end{aligned}
$$

where the last equality is because $\overline{\nu_{\triangleright}}(b,\star)=0$ for all $b\in B$.

Similarly for P_3, using $\overline{\mu_{\triangleleft}}(\star,a)=\overline{\nu_{\triangleleft}}(\star,b)=0$ for all $a\in A$ and $b\in B$, we have:

$$
\begin{aligned}
P_3 &= \sum_{a_2\in A} \alpha_{\triangleright}(\star,a_2)\nu_{\triangleright}(\star,g_2(a_2))\cdot f(0) = \sum_{b_2\in g_2(A)}\sum_{a_2\in g_2^{-1}(b_2)} \alpha_2(a_2)\nu_{\triangleright}(\star,b_2)\cdot f(0) \\
&= \sum_{b_2\in g_2(A)} \nu_{\triangleright}(\star,b_2)\cdot f(0) = \sum_{b_2\in B} \nu_{\triangleright}(\star,b_2)\cdot f(0) = \sum_{b_2\in B} \nu_{\triangleright}(\star,b_2)\cdot f\left(\frac{\overline{\nu_{\triangleleft}}(\star,b_2)}{\nu_{\triangleright}(\star,b_2)}\right).
\end{aligned}
$$

Putting everything together, we conclude

$$\Delta_f(\overline{\mu_{\triangleleft}},\overline{\mu_{\triangleright}}) = \Delta_f(\overline{\nu_{\triangleleft}},\overline{\nu_{\triangleright}}) \leq \delta$$

by assumption, so $\mu_{\triangleleft},\mu_{\triangleright}$ witness the desired approximate lifting. □