Non-commutative disintegrations: Existence and uniqueness in finite dimensions

Arthur J. Parzygnat and Benjamin P. Russo

Abstract. Motivated by advances in categorical probability, we introduce non-commutative almost everywhere (a.e.) equivalence and disintegrations in the setting of C^*-algebras. We show that C^*-algebras (resp. W^*-algebras) and a.e. equivalence classes of 2-positive (resp. positive) unital maps form a category. We prove that non-commutative disintegrations are a.e. unique whenever they exist. We provide an explicit characterization for when disintegrations exist in the setting of finite-dimensional C^*-algebras, and we give formulas for the associated disintegrations.

Contents

1. Introduction and outline

Regular conditional probabilities, optimal hypotheses, disintegrations of one measure over another consistent with a measure-preserving map, conditional expectations, perfect error-correcting codes, and sufficient statistics are all examples of a single mathematical notion. We call this notion a *disintegration*. Although we only make the connection between our definition of disintegration and the first three examples listed, relationships to the other notions are described in [38], and further connections to entropy were made in [16, 17, 40, 41]. In this paper, our primary focus is to provide necessary and sufficient conditions for the existence and uniqueness of disintegrations in the setting of finite-dimensional C^*-algebras.

Developing this and related ideas is part of a larger program in extending Bayesian statistics to the *non-commutative* setting [17, 20, 38, 39, 42–44] in such a way so that it is compatible with a recently developed *categorical* framework for classical statistics [6, 15, 16]. These recent advances in classical categorical functional analysis and measure theory provide a suitable notion of disintegration [2,6,8,15,21,24,27,52,55,56], whose diagrammatic formulation can be transferred from a category of probability spaces to a category of states on C^*-algebras. This is achieved by utilizing a fully faithful (contravariant) functor from the former to the latter [18,35]. This categorical perspective offers a candidate for generalizing disintegrations to non-commutative probability theory without relying on the specific measure-theoretic details of classical probability theory. Since a disintegration is a special kind of *Bayesian inverse* [6,20,38,44], this article serves as a step towards a theory of non-commutative Bayesian inversion.

Briefly, the definition of a disintegration of a state ω over another state ξ consistent with a unital *-homomorphism F preserving these states is a completely positive unital map R in the *reverse* direction that is both state preserving and a left inverse of F modulo the null space of ξ. If *-homomorphisms are written as straight arrows $\rightarrow$ and completely positive unital maps are written as squiggly arrows $\rightsquigarrow$, this definition of a disintegration can be summarized diagrammatically as
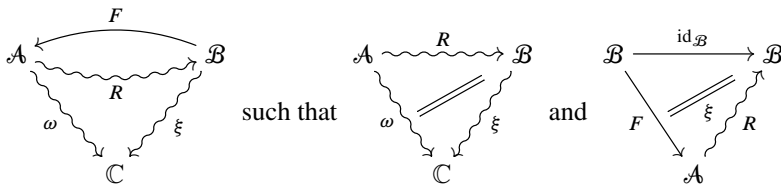
in the category of finite-dimensional C^*-algebras and completely positive unital maps. The right-most diagram commutes almost everywhere (a.e.), in a sense that we make precise in this article. We introduce and develop non-commutative a.e. equivalence in order to properly address the uniqueness properties of disintegrations.

The interpretation of completely positive unital maps as quantum conditional probabilities is not new [28], but we take this perspective further and include the relationships between states and partially reversible dynamics analogous to what *regular* conditional probabilities accomplish in classical statistics. Our core result is Theorem 4.3, which specializes to the case where $\mathcal{A}$ and $\mathcal{B}$ are matrix algebras and F sends $B \in \mathcal{B}$ to diag$(B, \ldots, B)$. If we express our states ω and ξ in terms of density matrices ρ and σ, respectively, Theorem 4.3 says that a unique disintegration exists if and only if there exists a density matrix τ such that $\rho = \tau \otimes \sigma$. This is closely related to a well-known result on the existence of state-preserving conditional expectations [47], but our notion generalizes it due to our weakened assumption of a.e. equivalence (see also [20]).

Our subsequent results are generalizations of this theorem and culminate in Theorem 5.61, which assumes $\mathcal{A}$ and $\mathcal{B}$ are arbitrary finite-dimensional C^*-algebras and F is an arbitrary unital *-homomorphism. We provide explicit formulas for disintegrations

and we analyze several examples, including one involving entanglement, which has its origins in the work of Einstein, Podolsky, and Rosen [11]. In Example 5.57, we show how the standard classical theorem on the existence and uniqueness of disintegrations (Theorem 5.1) is a direct corollary of our theorem. We conclude by exploring consequences of our characterization theorem in the context of measurement in quantum information theory. Finally, the appendix reviews stochastic maps (Markov kernels) and justifies our usage of the terminology "disintegration" by showing that the diagrammatic notion agrees with a general measure-theoretic one.

2. Non-commutative a.e. equivalence

For classical probability spaces, a.e. equivalence specifies the degree of uniqueness of disintegrations, Bayesian inverses, and conditional distributions. The same is true in the quantum/non-commutative setting. In this section, we first recall some relevant definitions involving states on C^*-algebras and completely positive maps from Paulsen [46] and Sakai [51] to establish notation and terminology. Afterwards, we define a.e. equivalence for linear maps between C^*-algebras in Definition 2.9. We provide a more computationally useful definition for finite-dimensional C^*-algebras in Lemma 2.20. In the rest of this section, we analyze several properties of a.e. equivalence.

Definition 2.1. A C^*-*algebra* is an algebra $\mathcal{A}$ equipped with a unit $1_{\mathcal{A}} \in \mathcal{A}$, an involution $^* : \mathcal{A} \to \mathcal{A}$, and a norm $\| \cdot \| : \mathcal{A} \to \mathbb{R}$ such that it is a unital *-algebra, it is closed with respect to the topology induced by its norm, and it satisfies the C^*-identity, which says $\|a^*a\| = \|a\|^2$ for all $a \in \mathcal{A}$. Given a C^*-algebra $\mathcal{A}$, a *positive element* of $\mathcal{A}$ is an element $a \in \mathcal{A}$ for which there exists an $x \in \mathcal{A}$ such that $a = x^*x$. The set of positive elements in $\mathcal{A}$ is denoted by $\mathcal{A}^+$. Given another C^*-algebra $\mathcal{B}$, a *positive map* $\varphi : \mathcal{B} \rightsquigarrow \mathcal{A}$ is a linear map such that $\varphi(\mathcal{B}^+) \subseteq \mathcal{A}^+$. A linear map $\varphi : \mathcal{B} \rightsquigarrow \mathcal{A}$ is *unital* if and only if $\varphi(1_{\mathcal{B}}) = 1_{\mathcal{A}}$. A *state* on a C^*-algebra $\mathcal{A}$ is a positive linear unital functional $\omega : \mathcal{A} \rightsquigarrow \mathbb{C}$. A *-*homomorphism* from $\mathcal{A}$ to $\mathcal{B}$ is a function $f : \mathcal{A} \to \mathcal{B}$ preserving the C^*-algebra structure, namely f is linear, f is multiplicative $f(aa') = f(a)f(a')$, f is unital $f(1_{\mathcal{A}}) = 1_{\mathcal{B}}$, and $f(a^*) = f(a)^*$ for all $a, a' \in \mathcal{A}$. If $\omega : \mathcal{A} \rightsquigarrow \mathbb{C}$ and $\xi : \mathcal{B} \rightsquigarrow \mathbb{C}$ are states, then a linear map $\varphi : \mathcal{B} \rightsquigarrow \mathcal{A}$ is said to be *state-preserving* whenever $\varphi \circ \omega = \xi$, and the notation $(\mathcal{B}, \xi) \overset{\varphi}{\rightsquigarrow} (\mathcal{A}, \omega)$ will be used to indicate this.

All C^*-algebras and *-homomorphisms will be unital unless specified otherwise. Note that positive (and linear) maps on C^*-algebras are denoted by squiggly arrows $\rightsquigarrow$, while *-homomorphisms are denoted by straight arrows $\to$.

Example 2.2. For each $n \in \mathbb{N}$, let $\mathcal{M}_n(\mathbb{C})$ denote the set of $n \times n$ complex matrices. The involution applied to $A \in \mathcal{M}_n(\mathbb{C})$ is given by the conjugate transpose and is written as $A^\dagger$ instead of A^* to be consistent with the standard notation used in quantum theory. A *matrix algebra* is a C^*-algebra of the form $\mathcal{M}_n(\mathbb{C})$ for some $n \in \mathbb{N}$. If $\mathcal{A}$ is another C^*-algebra,

then $\mathcal{M}_n(\mathbb{C}) \otimes \mathcal{A} \cong \mathcal{M}_n(\mathcal{A})$, the algebra of $n \times n$ matrices with entries in $\mathcal{A}$, admits a C^*-algebra structure by matrix operations and a norm that can be obtained in many ways (cf. Paulsen [46, Chapter 1]).

Our convention for the tensor product (also called the Kronecker product) of matrices will be

$$
\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}
$$

$$
= \begin{bmatrix} a_{11} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} & \cdots & a_{1m} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} \\ \vdots & & \vdots \\ a_{m1} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} & \cdots & a_{mm} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} \end{bmatrix}, \qquad (2.3)
$$

which is induced by the isomorphism $\mathbb{C}^m \otimes \mathbb{C}^n \to \mathbb{C}^{mn}$ determined by

$$\vec{e}_1 \otimes \vec{e}_1 \mapsto \vec{e}_1, \quad \ldots \quad \vec{e}_1 \otimes \vec{e}_n \mapsto \vec{e}_n, \quad \vec{e}_2 \otimes \vec{e}_1 \mapsto \vec{e}_{n+1}, \quad \ldots \quad \vec{e}_m \otimes \vec{e}_n \mapsto \vec{e}_{mn}. \quad (2.4)$$

Here, $\vec{e}_i$ denotes the standard i-th unit vector in $\mathbb{C}^n$ regardless of n.

Definition 2.5. Let $\mathcal{A}$ and $\mathcal{B}$ be C^*-algebras. Given $n \in \mathbb{N}$, a linear map $\varphi : \mathcal{B} \rightsquigarrow \mathcal{A}$ is *n-positive* if and only if $\mathrm{id}_{\mathcal{M}_n(\mathbb{C})} \otimes \varphi : \mathcal{M}_n(\mathbb{C}) \otimes \mathcal{B} \rightsquigarrow \mathcal{M}_n(\mathbb{C}) \otimes \mathcal{A}$ is positive. The map φ is *completely positive* if and only if φ is n-positive for all $n \in \mathbb{N}$. A completely positive (unital) map will be abbreviated as a CP (CPU) map.

The Choi–Kraus theorem gives a characterization of completely positive maps between *matrix* algebras. This will be used often, so we state it here to set notation [7, 26].

Theorem 2.6. *Fix* $n, m \in \mathbb{N}$. *A linear map* $R : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_m(\mathbb{C})$ *is completely positive if and only if there exists a finite collection* $\{R_i : \mathbb{C}^n \to \mathbb{C}^m\}$ *of linear maps such that*

$$R = \sum_i \mathrm{Ad}_{R_i}. \qquad (2.7)$$

Here, $\mathrm{Ad}_{R_i}(A) := R_i A R_i^\dagger$ *for all* $A \in \mathcal{M}_n(\mathbb{C})$. *The map* R *is CPU if and only if, in addition,* $\sum_i R_i R_i^\dagger = \mathbb{1}_m$.

A collection $\{R_i\}$ satisfying (2.7) is called a *Kraus decomposition* for R.

Remark 2.8. The standard assumption in quantum information theory is to work with completely positive *trace-preserving* maps instead of *unital* maps. The former class, typically called *quantum operations/channels* (cf. Nielsen and Chuang [32, Section 8.2])

is used in the Schrödinger representation when transforming physical states, while the latter is used in the Heisenberg representation when transforming physical observables. We briefly explain the relationship between the two. A completely positive map $R :$ $\mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_m(\mathbb{C})$ is unital if and only if the adjoint map $R^* : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ is trace-preserving. The adjoint map R^* is defined with respect to the *Hilbert–Schmidt* (a.k.a. *Frobenius*) inner product on a matrix algebra, which is given by $\langle A, B \rangle := \operatorname{tr}(A^\dagger B)$ for all square matrices (of the same dimension) A and B. Therefore, R^* is the unique map satisfying $\langle R^*(A), B \rangle = \langle A, R(B) \rangle$ for all $A \in \mathcal{M}_m(\mathbb{C})$ and all $B \in \mathcal{M}_n(\mathbb{C})$. If $R = \sum_i \operatorname{Ad}_{R_i}$ is a Kraus decomposition of R, then $R^* = \sum_i \operatorname{Ad}_{R_i^\dagger}$ is a Kraus decomposition of R^* because $\langle R^*(A), B \rangle = \langle A, R(B) \rangle = \sum_i \operatorname{tr}(A^\dagger R_i B R_i^\dagger) = \sum_i \operatorname{tr}(R_i^\dagger A^\dagger R_i B) = \sum_i \langle R_i^\dagger A R_i, B \rangle$ by the cyclicity of the trace. By a similar argument, $\langle R^*(\mathbb{1}_m), B \rangle = \langle \mathbb{1}_m, R(B) \rangle = \operatorname{tr}(R^*(B))$ for all $B \in \mathcal{M}_n(B)$. From this, it follows that R is unital if and only if R^* is trace-preserving.

In particular, if $\omega : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$ is a state, then its adjoint $\omega^* : \mathbb{C} \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ is determined by the image $\omega^*(1)$ of the unit 1 in $\mathbb{C}$, which is positive. Furthermore, since ω is unital, ω^* is trace-preserving. Hence, $\operatorname{tr}(\omega^*(1)) = 1$. In other words, $\omega^*(1)$ is a trace 1 positive matrix. This is called the *density matrix* associated to ω. Finally, $\omega = \operatorname{tr}(\omega^*(1) \cdot)$ as states on $\mathcal{M}_n(\mathbb{C})$.

We now proceed to defining a.e. equivalence of linear maps on C^*-algebras.

Definition 2.9. Let $\mathcal{A}$ and $\mathcal{B}$ be C^*-algebras, let $F, F' : \mathcal{B} \rightsquigarrow \mathcal{A}$ be two linear maps, and let $\omega : \mathcal{A} \rightsquigarrow \mathbb{C}$ be a state on $\mathcal{A}$ (or more generally a positive linear functional). Let

$$\mathcal{N}_\omega := \big\{ a \in \mathcal{A} : \omega(a^* a) = 0 \big\}$$

denote the *null space* of ω. Since $\mathcal{N}_\omega$ is a left ideal of $\mathcal{A}$ (see [36, Construction 3.1] for details), denote the quotient vector space by $\mathcal{A}/\mathcal{N}_\omega$. The maps F and F' are said to be *equal a.e. with respect to ω* or *equal ω-a.e.* if and only if the diagram (in the category of vector spaces and linear maps)

$$(2.10)$$

commutes, i.e., if and only if $F(b) - F'(b) \in \mathcal{N}_\omega$ for all $b \in \mathcal{B}$. The map $\mathcal{A} \twoheadrightarrow \mathcal{A}/\mathcal{N}_\omega$ in (2.10) is the quotient map of $\mathcal{A}$ onto $\mathcal{A}/\mathcal{N}_\omega$. When F and F' are equal ω-a.e., the notation $F \underset{\omega}{=} F'$ will be used.

The justification for the above terminology of a.e. equivalence is explained in the following illustrative example of finite probability spaces (cf. the appendix for terminology).

Example 2.11. Let $\mathcal{A} := \mathbb{C}^X$ and $\mathcal{B} := \mathbb{C}^Y$ be the commutative C^*-algebras of complex-valued functions on the finite sets X and Y, respectively, let $P : \mathbb{C}^X \rightsquigarrow \mathbb{C}$ be a state

on X, and let $F, G : \mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$ be two positive unital maps. Then there exists a unique probability measure p on X such that $\sum_{x \in X} \varphi(x) p_x = P(\varphi)$ for all $\varphi \in \mathbb{C}^X$ (see [35, Section 2.6] for details). Namely, $p_x := P(e_x)$, where e_x is the function on X defined by $X \ni x' \mapsto e_x(x') := \delta_{xx'}$. Similarly, there exist unique stochastic maps $f, g : X \rightsquigarrow Y$ such that

$$\big(F(\psi)\big)(x) = \sum_{y \in Y} \psi(y) f_{yx} \quad \forall \psi \in \mathbb{C}^Y, \ \forall x \in X,$$

namely

$$X \ni x \mapsto \big(Y \ni y \mapsto f_{yx} := F(e_y)(x)\big)$$

and similarly for G with g. One can show that the null space of P is given by

$$\mathcal{N}_P := \big\{\varphi \in \mathbb{C}^X : P(\varphi^*\varphi) = 0\big\} = \big\{\varphi \in \mathbb{C}^X : \varphi|_{X \setminus N_p} = 0\big\} = \mathrm{span}\bigg(\bigcup_{x \in N_p} \{e_x\}\bigg),$$

where $N_p \subseteq X$ is the measure-theoretic null space of p and $\varphi|_{X \setminus N_p}$ denotes the restriction of φ to $X \setminus N_p$. Hence, the quotient $\mathbb{C}^X / \mathcal{N}_P$ is isomorphic to functions on $X \setminus N_p$ by the isomorphism

$$\mathbb{C}^X / \mathcal{N}_P \to \mathbb{C}^{X \setminus N_p}$$
$$[\varphi] \mapsto \varphi\big|_{X \setminus N_p}.$$

As a result, the two positive unital maps $F, G : \mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$ are equal P-a.e. if and only if the associated stochastic maps $\tilde{f}, \tilde{g} : X \setminus N_p \rightsquigarrow Y$ defined by the restrictions of f and g to $X \setminus N_p$, respectively, are equal. This precisely means that $f \underset{p}{=} g$.

We now proceed to establishing several important facts regarding non-commutative a.e. equivalence. First, if two maps are a.e. equivalent in terms of some state, then they pull back that state to the same state.

Lemma 2.12. *Let $\mathcal{A}$ and $\mathcal{B}$ be C^*-algebras, let $\xi : \mathcal{B} \rightsquigarrow \mathbb{C}$ be a state (or more generally a positive functional), and let $\varphi, \psi : \mathcal{A} \rightsquigarrow \mathcal{B}$ be linear maps. If $\varphi \underset{\xi}{=} \psi$, then*

$$\xi \circ \varphi = \xi \circ \psi.$$

Proof. Let $a \in \mathcal{A}$. Then

$$\big|\xi\big(\varphi(a) - \psi(a)\big)\big|^2 \leq \xi\big(\big(\varphi(a) - \psi(a)\big)^*\big(\varphi(a) - \psi(a)\big)\big) = 0$$

by the Cauchy–Schwarz inequality for positive functionals (cf. Fillmore [12, Proposition 5.2.1]) and because $\varphi(a) - \psi(a) \in \mathcal{N}_\xi$. Hence, $\xi(\varphi(a)) = \xi(\psi(a))$. Since a was arbitrary, $\xi \circ \varphi = \xi \circ \psi$. ∎

The support of a state will also be useful in when formulating and proving our disintegration theorem.

Lemma 2.13. *Let* $\omega : \mathcal{A} \rightsquigarrow \mathbb{C}$ *be a state on a finite-dimensional C^*-algebra $\mathcal{A}$ (or more generally a W^*-algebra). Then there exists a unique projection $P_\omega \in \mathcal{A}$ (this means that $P_\omega^* = P_\omega$ and $P_\omega^2 = P_\omega$) such that $\mathcal{N}_\omega = \mathcal{A}(1_{\mathcal{A}} - P_\omega)$. Equivalently, P_ω is characterized by*

$$\omega(a) = \omega(a P_\omega) = \omega(P_\omega a) = (\omega \circ \mathrm{Ad}_{P_\omega})(a) \quad \forall a \in \mathcal{A}. \tag{2.14}$$

Proof. See Sakai [51, Section 1.14]. ∎

Remark 2.15. If $\mathcal{A}$ is not a finite-dimensional C^*-algebra in Lemma 2.13, then such a projection P_ω for a state ω need not exist. Indeed, if $\mathcal{A} = C(X)$, continuous complex-valued functions on a connected compact Hausdorff space X, then there are no non-trivial projections and yet there are many states generating non-trivial null spaces. Such a projection does exist, however, if $\mathcal{A}$ is a W^*-algebra. Hence, several (but not all) of the results that follow involving such supports also hold for W^*-algebras.

Definition 2.16. Using the same notation from Lemma 2.13, P_ω is called the *support of* ω. Its complement will be denoted by $P_\omega^\perp := 1_{\mathcal{A}} - P_\omega$.

Example 2.17. When $\mathcal{A} = \mathcal{M}_n(\mathbb{C})$ is a matrix algebra with a state $\omega : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$, then $\omega = \mathrm{tr}(\rho \cdot)$ for some unique density matrix $\rho \in \mathcal{M}_n(\mathbb{C})$ (cf. Remark 2.8). In this case, $P_\omega^\perp$ is the projection onto the zero eigenspace of ρ and P_ω satisfies $P_\omega \rho = \rho = \rho P_\omega$.

Lemma 2.18. *Let P be a projection in a C^*-algebra (or W^*-algebra) $\mathcal{B}$ and let $B \in \mathcal{B}$. Then $BP = 0$ implies that $B \in \mathcal{B} P^\perp$.*

Proof. This follows from the fact that every $B \in \mathcal{B}$ can be uniquely expressed as a sum of four terms

$$B = (P + P^\perp)B(P + P^\perp) = PBP + PBP^\perp + P^\perp BP + P^\perp BP^\perp, \tag{2.19}$$

the non-zero ones of which are linearly independent. ∎

The decomposition (2.19) will be used frequently in this work,[1] particularly in conjunction with the support of a state. For example, we have the following alternative and computationally useful characterization of a.e. equivalence.

Lemma 2.20. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras, let $\xi : \mathcal{B} \rightsquigarrow \mathbb{C}$ be a state, and let $\varphi, \psi : \mathcal{A} \rightsquigarrow \mathcal{B}$ be linear maps. Let $P_\xi \in \mathcal{B}$ denote the support of ξ.*

    (i)   *Then $\varphi \underset{\xi}{=} \psi$ if and only if[2] $\varphi(A) P_\xi = \psi(A) P_\xi$ for all $A \in \mathcal{A}$.*

    (ii)   *If $\varphi \underset{\xi}{=} \psi$, then $\mathrm{Ad}_{P_\xi} \circ \varphi = \mathrm{Ad}_{P_\xi} \circ \psi$.*

---

[1]The usage of such a decomposition is certainly not new. More recently, they have made an appearance in the study of Pierce and corner algebras, also in the context of conditional expectations [48]. We thank Chris Heunen for informing us of this reference.

[2]In the case where $\mathcal{A} = \mathcal{M}_m(\mathbb{C})$ and $\mathcal{B} = \mathcal{M}_n(\mathbb{C})$, this says that the two $n \times n$ matrices (viewed as operators on $\mathbb{C}$) $\varphi(A)$ and $\psi(A)$ agree when restricted to the subspace $P_\xi \mathbb{C}^n \subseteq \mathbb{C}^n$.

Proof. For the first claim, if $\varphi \underset{\xi}{=} \psi$, then $\varphi(A) - \psi(A) = BP_\xi^\perp$ for some $B \in \mathcal{B}$ since $\mathcal{N}_\xi = \mathcal{B}P_\xi^\perp$ by Lemma 2.13. Multiplying by P_ξ on the right gives item (i). Conversely, if $\varphi(A)P_\xi = \psi(A)P_\xi$ holds, then $(\varphi(A) - \psi(A))P_\xi = 0$. Hence, $\varphi(A) - \psi(A) \in \mathcal{B}P_\xi^\perp$ by Lemma 2.18. Item (ii) follows from item (i) by multiplying $\varphi(A)P_\xi = \psi(A)P_\xi$ on the left by P_ξ. ∎

Since CP maps between matrix algebras have particularly simple forms (cf. Theorem 2.6), it will also be useful to have a more quantitative version of Lemma 2.20. To state it, we first recall a general fact about the relationship between two Kraus decompositions of a CP map of matrix algebras.

Lemma 2.21. *Let $\varphi : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ be a CP map and suppose that*

$$\sum_{i=1}^p \mathrm{Ad}_{V_i} = \varphi = \sum_{j=1}^q \mathrm{Ad}_{W_j}$$

are two Kraus decompositions of φ with $p \geq q$. Then there exists a $q \times p$ matrix U that is a coisometry (meaning $UU^\dagger = \mathbb{1}_q$, i.e., the rows of U are orthonormal) such that $V_i = \sum_{j=1}^q u_{ji} W_j$ for all $i \in \{1, 2, \ldots, p\}$. Here u_{ji} denotes the ji-th entry of U.

Proof. The reader is referred to [37, Sections 6 and 7] for any unexplained details and terminology. First note that every such Kraus decomposition $\varphi = \sum_{i=1}^p \mathrm{Ad}_{V_i}$ can be expressed as a Stinespring representation $\varphi = \pi \circ \mathrm{Ad}_V$, where π and V are defined by

$$\mathcal{M}_m(\mathbb{C}) \ni A \overset{\pi}{\mapsto} \mathbb{1}_p \otimes A \in \mathcal{M}_p(\mathbb{C}) \otimes \mathcal{M}_m(\mathbb{C})$$

and

$$\mathbb{C}^p \otimes \mathbb{C}^m \cong \mathbb{C}^{pm} \xrightarrow{V := [\, V_1 \cdots V_p \,]} \mathbb{C}^n,$$

respectively (and similarly for $\varphi = \sum_{j=1}^q \mathrm{Ad}_{W_j}$ and $W := \begin{bmatrix} W_1 & \cdots & W_q \end{bmatrix}$). By the universal property of Stinespring representations (see [37, Theorem 6.29 and the end of Section 7]), there exists a coisometry[3] $\mathbb{C}^p \xrightarrow{U} \mathbb{C}^q$ such that the diagram

$$
\begin{array}{ccc}
 & \mathbb{C}^q \otimes \mathbb{C}^m & \\
{}^{W}\nearrow & & \uparrow {}^{U \otimes \mathbb{1}_m} \\
\mathbb{C}^n \xleftarrow{\quad} & & \\
{}_{V}\searrow & & \\
 & \mathbb{C}^p \otimes \mathbb{C}^m &
\end{array}
\qquad (2.22)
$$

commutes. Writing

$$U \otimes \mathbb{1}_m = \begin{bmatrix} u_{11}\mathbb{1}_m & \cdots & u_{1p}\mathbb{1}_m \\ \vdots & & \vdots \\ u_{q1}\mathbb{1}_m & \cdots & u_{qp}\mathbb{1}_m \end{bmatrix},$$

---

[3]Technically, the theorem referenced claims that there exists a partial isometry. However, this partial isometry can be extended to a coisometry by similar techniques to those employed in [37, Example 7.27 and Theorem 7.30].

we see that commutativity of (2.22) gives

$$\begin{bmatrix} V_1 & \cdots & V_p \end{bmatrix} = \begin{bmatrix} W_1 & \cdots & W_q \end{bmatrix} (U \otimes \mathbb{1}_n),$$

which is the result claimed. ∎

Lemma 2.23. *Fix a positive integer n, let $\xi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be a state, and let $\varphi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ be a CPU map such that $\varphi \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$. Let $P_\xi \in \mathcal{M}_n(\mathbb{C})$ denote the support of ξ and let φ have a Kraus decomposition $\varphi = \sum_{i=1}^p \mathrm{Ad}_{V_i}$. Then there exist complex numbers $\{\alpha_i\}_{i \in \{1,\ldots,p\}}$ such that*

$$P_\xi V_i = \alpha_i P_\xi \quad \forall i \in \{1, \ldots, p\} \quad and \quad \sum_{i=1}^p |\alpha_i|^2 = 1.$$

Proof. Since $\varphi \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$, Lemma 2.20 implies that

$$\mathrm{Ad}_{P_\xi} \circ \varphi = \sum_{i=1}^p \mathrm{Ad}_{P_\xi V_i} = \mathrm{Ad}_{P_\xi}.$$

Furthermore, Lemma 2.21 implies that there exists a $1 \times p$ coisometry U such that $P_\xi V_i = u_{1i} P_\xi$, where u_{1i} is the $1i$-th entry of U. Set $\alpha_i := u_{1i}$. Since U is a coisometry, the first (and only) column of U is a unit vector, i.e., $\sum_{i=1}^p |u_{1i}|^2 = 1$, which proves the claim. ∎

Remark 2.24. Using the notation of Lemma 2.23, $P_\xi V_i = \alpha_i P_\xi$ says that

$$V_i = \alpha_i P_\xi + P_\xi^\perp V_i P_\xi + P_\xi^\perp V_i P_\xi^\perp. \tag{2.25}$$

If we choose a basis in which the density matrix $\xi^*(1)$ is diagonal with its non-zero eigenvalues all appearing on the top left, then (2.25) reads

$$V_i = \begin{bmatrix} \alpha_i \mathbb{1}_r & 0 \\ V_i^{\mathrm{bl}} & V_i^{\mathrm{br}} \end{bmatrix},$$

where r is the rank of $\xi^*(1)$, and where V_i^{bl} is an $(n-r) \times r$ matrix while V_i^{br} is an $(n-r) \times (n-r)$ matrix.

It turns out that Lemma 2.23 holds even when the support P_ξ from the equations is removed (cf. Theorem 2.32). To prove this, it seems convenient to recall the notion of a pre-Hilbert C^*-algebra module due to Paschke [45].

Definition 2.26. Let $\mathcal{A}$ be a C^*-algebra. A *pre-Hilbert $\mathcal{A}$-module* is a left[4] $\mathcal{A}$-module $\mathcal{E}$ together with a linear-conjugate linear map[5] $\langle\!\langle \cdot, \cdot \rangle\!\rangle : \mathcal{E} \times \mathcal{E} \to \mathcal{A}$ satisfying the following properties:

---

[4]Paschke defines a *right* module structure instead of a left one. This does change some properties, but we have modified them appropriately. One such property is the Paschke–Cauchy–Schwarz inequality in (2.29).

[5]This means that $\langle\!\langle s + t, u \rangle\!\rangle = \langle\!\langle s, u \rangle\!\rangle + \langle\!\langle t, u \rangle\!\rangle$ for all $s, t, u \in \mathcal{E}$. The other properties usually associated with sesqui-linearity (with conjugate linearity in the second coordinate) follow from the other conditions in the definition since the algebra $\mathcal{A}$ is unital.

(i) $\langle\!\langle s, s \rangle\!\rangle \geq 0$ for all $s \in \mathcal{E}$,

(ii) $\langle\!\langle s, t \rangle\!\rangle = \langle\!\langle t, s \rangle\!\rangle^*$ for all $s, t \in \mathcal{E}$,

(iii) $\langle\!\langle as, t \rangle\!\rangle = a \langle\!\langle s, t \rangle\!\rangle$ for all $s, t \in \mathcal{E}$ and $a \in \mathcal{A}$, and

(iv) $\langle\!\langle s, s \rangle\!\rangle = 0$ if and only if $s = 0$ (this is called *non-degeneracy* of $\langle\!\langle \cdot, \cdot \rangle\!\rangle$).

$\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is called the $\mathcal{A}$-*valued inner product* on $\mathcal{E}$.

Remark 2.27. It follows from this definition that $\langle\!\langle s, at \rangle\!\rangle = \langle\!\langle s, t \rangle\!\rangle a^*$ for all $s, t \in \mathcal{E}$ and $a \in \mathcal{A}$.

Lemma 2.28. *Let $\mathcal{A}$ be a C*-algebra and $\mathcal{E}$ a pre-Hilbert module over $\mathcal{A}$. Then*

$$\mathcal{E} \ni s \mapsto \|s\|_{\mathcal{E}} := \sqrt{\|\langle\!\langle s, s \rangle\!\rangle\|}$$

defines a norm on $\mathcal{E}$. Furthermore,

$$\langle\!\langle t, s \rangle\!\rangle \langle\!\langle s, t \rangle\!\rangle \leq \|t\|_{\mathcal{E}}^2 \langle\!\langle s, s \rangle\!\rangle \tag{2.29}$$

for all $s, t \in \mathcal{E}$.

Proof. See Paschke [45, Proposition 2.3] or Fillmore [12, Section 3.14]. ∎

Example 2.30. Fix $n, m, p \in \mathbb{N}$. Let $_m\mathcal{M}_n(\mathbb{C})$ denote the vector space of $m \times n$ complex matrices. Set $\mathcal{A} := \mathcal{M}_m(\mathbb{C})$ and $\mathcal{E} := {}_m\mathcal{M}_n(\mathbb{C})^p$, the vector space direct sum of p copies of $_m\mathcal{M}_n(\mathbb{C})$. Denote elements of $\mathcal{E}$ by $\vec{A} := (A_1, \ldots, A_p)$ so that $A_i \in {}_m\mathcal{M}_n(\mathbb{C})$ for all $i \in \{1, \ldots, p\}$. Define the left $\mathcal{A}$-module structure on $\mathcal{E}$ to be $B\vec{A} := (BA_1, \ldots, BA_p)$ for all $\vec{A} \in \mathcal{E}$ and $B \in \mathcal{A}$. Define the $\mathcal{A}$-valued inner product by

$$\mathcal{E} \times \mathcal{E} \ni (\vec{A}, \vec{B}) \mapsto \langle\!\langle \vec{A}, \vec{B} \rangle\!\rangle := \sum_{i=1}^{p} A_i B_i^{\dagger}. \tag{2.31}$$

Straightforward matrix algebra shows that $\mathcal{E}$ is indeed a pre-Hilbert $\mathcal{A}$-module with these structures. In fact, $\mathcal{E}$ is also a *right* $\mathcal{M}_n(\mathbb{C})$-module satisfying

$$\langle\!\langle \vec{A}C, \vec{B} \rangle\!\rangle = \langle\!\langle \vec{A}, \vec{B}C^{\dagger} \rangle\!\rangle \quad \forall \vec{A}, \vec{B} \in \mathcal{E} \text{ and } C \in \mathcal{M}_n(\mathbb{C}).$$

However, $\mathcal{E}$ is *not* a (right) pre-Hilbert module with respect to this action.

Theorem 2.32. *Fix $n \in \mathbb{N}$, let $\xi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be a state, and let $\varphi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ be a CPU map such that $\varphi \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$. Let φ have a Kraus decomposition $\varphi = \sum_{i=1}^{p} \mathrm{Ad}_{V_i}$. Then there exist complex numbers $\{\alpha_i\}_{i \in \{1, \ldots, p\}}$ such that*

$$V_i = \alpha_i \mathbb{1}_n \quad \forall i \in \{1, \ldots, p\} \quad \text{and} \quad \sum_{i=1}^{p} |\alpha_i|^2 = 1.$$

In particular, $\varphi = \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$.

Proof. Let $P_\xi \in \mathcal{M}_n(\mathbb{C})$ denote the support of ξ. In order to proceed avoiding as many indices and sums as possible, we will first introduce a certain pre-Hilbert C^*-algebra module based on the number p of Kraus operators assumed for φ. Set $\mathcal{E} := \mathcal{M}_n(\mathbb{C})^p$ and equip this with the pre-Hilbert $\mathcal{M}_n(\mathbb{C})$-module structure from Example 2.30. By Remark 2.24,

$$\vec{V} = P_\xi \vec{\alpha} + \vec{V}^{\text{bl}} + \vec{V}^{\text{br}}, \tag{2.33}$$

where

$$\vec{\alpha} := (\alpha_1 \mathbb{1}_n, \ldots, \alpha_p \mathbb{1}_n), \quad \vec{V}^{\text{bl}} := P_\xi^\perp \vec{V} P_\xi, \quad \text{and} \quad \vec{V}^{\text{br}} := P_\xi^\perp \vec{V} P_\xi^\perp.$$

Note that $C\vec{\alpha} = \vec{\alpha}C$ for all $C \in \mathcal{M}_n(\mathbb{C})$. The identities

$$\langle\!\langle \vec{\alpha}, \vec{\alpha} \rangle\!\rangle = \mathbb{1}_n, \quad \langle\!\langle P_\xi \vec{\alpha}, \vec{V}^{\text{br}} \rangle\!\rangle = 0, \quad \text{and} \quad \langle\!\langle \vec{V}^{\text{bl}}, \vec{V}^{\text{br}} \rangle\!\rangle = 0 \tag{2.34}$$

follow directly from the definitions. The fact that φ is unital means that

$$\mathbb{1}_n = \sum_i V_i V_i^\dagger.$$

In terms of the $\mathcal{M}_n(\mathbb{C})$-valued inner product, this becomes

$$\mathbb{1}_n = \langle\!\langle \vec{V}, \vec{V} \rangle\!\rangle \overset{(2.33)}{=} \langle\!\langle P_\xi \vec{\alpha} + \vec{V}^{\text{bl}} + \vec{V}^{\text{br}}, P_\xi \vec{\alpha} + \vec{V}^{\text{bl}} + \vec{V}^{\text{br}} \rangle\!\rangle$$

$$\overset{(2.34)}{=} \underbrace{\langle\!\langle P_\xi \vec{\alpha}, P_\xi \vec{\alpha} \rangle\!\rangle}_{P_\xi} + \langle\!\langle P_\xi \vec{\alpha}, \vec{V}^{\text{bl}} \rangle\!\rangle + \langle\!\langle \vec{V}^{\text{bl}}, P_\xi \vec{\alpha} \rangle\!\rangle + \langle\!\langle \vec{V}^{\text{bl}}, \vec{V}^{\text{bl}} \rangle\!\rangle + \langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle. \tag{2.35}$$

Since

$$P_\xi \vec{V}^{\text{bl}} = \vec{0} \quad \text{and} \quad P_\xi \vec{V}^{\text{br}} = \vec{0},$$

it follows that

$$\langle\!\langle \vec{X}, \vec{V}^{\text{bl}} \rangle\!\rangle P_\xi = 0 \quad \text{and} \quad \langle\!\langle \vec{X}, \vec{V}^{\text{br}} \rangle\!\rangle P_\xi = 0 \quad \forall \vec{X} \in \mathcal{E} \tag{2.36}$$

by the properties of the pre-Hilbert module structure (see Remark 2.27). Hence, multiplying (2.35) by P_ξ on the right and simplifying gives $0 = \langle\!\langle \vec{V}^{\text{bl}}, P_\xi \vec{\alpha} \rangle\!\rangle$ and similarly $\langle\!\langle P_\xi \vec{\alpha}, \vec{V}^{\text{bl}} \rangle\!\rangle = 0$. Furthermore, $\langle\!\langle \vec{V}^{\text{bl}}, P_\xi^\perp \vec{\alpha} \rangle\!\rangle = 0$ follows immediately from the definition of $\vec{V}^{\text{bl}}$. Putting these two together gives

$$\langle\!\langle \vec{V}^{\text{bl}}, \vec{\alpha} \rangle\!\rangle = 0. \tag{2.37}$$

Hence, the unitality of R condition (2.35) simplifies to

$$P_\xi^\perp = \langle\!\langle \vec{V}^{\text{bl}}, \vec{V}^{\text{bl}} \rangle\!\rangle + \langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle. \tag{2.38}$$

Now, write $A \in \mathcal{M}_n(\mathbb{C})$ as (cf. equation (2.19))

$$A = P_\xi A P_\xi + P_\xi A P_\xi^\perp + P_\xi^\perp A P_\xi + P_\xi^\perp A P_\xi^\perp \tag{2.39}$$

in terms of the support P_ξ of ξ and its orthogonal complement $P_\xi^\perp$. Using this decomposition,

$$\varphi(A)P_\xi = \sum_{i=1}^p V_i A V_i^\dagger P_\xi = \langle\!\langle \vec{V} A, \vec{V} \rangle\!\rangle P_\xi \overset{\text{Rmk 2.27}}{=} \langle\!\langle \vec{V} A, P_\xi \vec{V} \rangle\!\rangle$$

$$\overset{(2.33)}{=} \langle\!\langle \vec{V} A, P_\xi \vec{\alpha} \rangle\!\rangle \overset{(2.31)}{=} \langle\!\langle \vec{V}, \vec{\alpha} \rangle\!\rangle A P_\xi \overset{(2.34)\,\&\,(2.37)}{=} \left(P_\xi + \langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle \right) A P_\xi$$

$$\overset{(2.39)}{=} \left(P_\xi + \langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle \right)\left(P_\xi A P_\xi + P_\xi^\perp A P_\xi \right) \overset{(2.36)}{=} P_\xi A P_\xi + \langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle P_\xi^\perp A P_\xi \quad (2.40)$$

for all $A \in \mathcal{M}_n(\mathbb{C})$. But since $\varphi \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$, this equals $P_\xi A P_\xi + P_\xi^\perp A P_\xi$ by Lemma 2.20 item (i). Identifying terms, $\langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle P_\xi^\perp A P_\xi = P_\xi^\perp A P_\xi$ for all $A \in \mathcal{M}_n(\mathbb{C})$, i.e., $\langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle$ acts as the identity on $n \times n$ matrices of the form $P_\xi^\perp A P_\xi$. This combined with the fact that $\langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle = P_\xi^\perp \langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle P_\xi^\perp$ implies that

$$\langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle = P_\xi^\perp. \tag{2.41}$$

This implies that $\langle\!\langle \vec{\alpha}, \vec{V}^{\text{br}} \rangle\!\rangle^* \langle\!\langle \vec{\alpha}, \vec{V}^{\text{br}} \rangle\!\rangle = P_\xi^\perp$. Hence, by the Paschke–Cauchy–Schwarz inequality (Lemma 2.28),

$$P_\xi^\perp \leq \|\vec{\alpha}\|_\xi^2 \langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle \overset{(2.34)}{=} \langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle.$$

On the other hand, (2.38) entails $\langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle \leq P_\xi^\perp$ by condition (i) in Definition 2.26. These two inequalities force

$$\langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle = P_\xi^\perp \quad \text{and} \quad \langle\!\langle \vec{V}^{\text{bl}}, \vec{V}^{\text{bl}} \rangle\!\rangle = 0. \tag{2.42}$$

By non-degeneracy of the $\mathcal{M}_n(\mathbb{C})$-valued inner product on $\mathcal{E}$, this forces $\vec{V}^{\text{bl}} = \vec{0}$, i.e.,

$$P_\xi V_i P_\xi^\perp = 0 \quad \forall i \in \{1, \ldots, p\}.$$

Finally, using these relations and the properties of the $\mathcal{M}_n(\mathbb{C})$-valued inner product,

$$\langle\!\langle P_\xi^\perp \vec{\alpha} - \vec{V}^{\text{br}}, P_\xi^\perp \vec{\alpha} - \vec{V}^{\text{br}} \rangle\!\rangle$$

$$\overset{(2.34)}{=} P_\xi^\perp - \langle\!\langle \vec{\alpha}, \vec{V}^{\text{br}} \rangle\!\rangle - \langle\!\langle \vec{V}^{\text{br}}, \vec{\alpha} \rangle\!\rangle + \langle\!\langle \vec{V}^{\text{br}}, \vec{V}^{\text{br}} \rangle\!\rangle \overset{(2.41)\,\&\,(2.42)}{=} 0,$$

which, by non-degeneracy of the $\mathcal{M}_n(\mathbb{C})$-valued inner product, proves that

$$\vec{V}^{\text{br}} = P_\xi^\perp \vec{\alpha}, \tag{2.43}$$

i.e.,

$$P_\xi^\perp V_i P_\xi^\perp = \alpha_i P_\xi^\perp \quad \forall i \in \{1, \ldots, p\}.$$

Putting this all together gives

$$\vec{V} = \vec{\alpha}, \quad \text{i.e., } V_i = \alpha_i \mathbb{1}_n \ \forall i \in \{1, \ldots, p\}.$$

The fact that $\varphi = \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$ follows immediately from this result. ∎

Remark 2.44. It should be stressed how surprising Theorem 2.32 is. Even if ξ is a pure state, so that its support is a rank one projection, a.e. equivalence of a CPU map to the identity is strong enough to enforce *equality* of that CPU map to the identity, regardless of how large the dimension of the Hilbert space is. We feel this gives us a precise sense of how "probability zero" objects (the projection onto the orthogonal complement of the support of ξ in this case) in quantum theory cannot be disregarded in the way that they can be in classical probability theory [23]. Theorem 2.32 has recently been reformulated as a CP completion problem and substantially generalized in [4].

Remark 2.45. Theorem 2.32 might seem to suggest that if $\xi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$ is a state and if $\varphi, \psi : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ are two CPU maps, then $\varphi \underset{\xi}{=} \psi$ implies that $\varphi = \psi$. The following example shows that this is false in general. Let $s \in \{1, \dots, m-1\}$ and let $\chi : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be the map that takes the trace of the bottom right part of a matrix, namely

$$\chi(A) := \sum_{i=s+1}^{m} a_{ii},$$

where a_{ii} is the ii-th entry of A. Note that this map is positive and therefore CP since the codomain of χ is $\mathbb{C}$ (cf. Stinespring [53, Theorem 3]). Note, however, that χ is not unital. Similarly, the trace map tr $: \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}$ is CP (but not unital). Now, consider the two maps

$$\mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$$
$$A \overset{\varphi}{\mapsto} \frac{1}{m} \operatorname{tr}(A)\mathbb{1}_n \quad \text{and}$$
$$A \overset{\psi}{\mapsto} \frac{1}{m} \operatorname{tr}(A)P_\xi + \frac{1}{m-s}\chi(A)P_\xi^\perp.$$

Note that φ and ψ are not equal. Nevertheless, φ and ψ are ξ-a.e. equivalent because $\varphi(A)P_\xi = \psi(A)P_\xi$ for all $A \in \mathcal{M}_m(\mathbb{C})$. A simple calculation shows that ψ and φ are unital. Furthermore, they are both completely positive as their p-ampliations are

$$\varphi_p(\,\cdot\,) = \frac{1}{m} \operatorname{tr}_p(\,\cdot\,) \otimes \mathbb{1}_n \quad \text{and} \quad \psi_p(\,\cdot\,) = \frac{1}{m} \operatorname{tr}_p(\,\cdot\,) \otimes P_\xi + \frac{1}{m-s}\chi_p(\,\cdot\,) \otimes P_\xi^\perp,$$

respectively. Here, tr_p and χ_p are the p-ampliations of tr and χ, which are positive.

Remark 2.46. The conclusion of Theorem 2.32 is false if φ is assumed to only be CP but not unital. A simple counter-example is the CP map

$$\mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$$
$$A \overset{\varphi}{\mapsto} A + \operatorname{tr}(A)P_\xi^\perp.$$

Here, φ is ξ-a.e. equivalent to $\operatorname{id}_{\mathcal{M}_n(C)}$ but is not equal to it.

Remark 2.47. Using the same notation and assumptions as in Theorem 2.32, if φ is ξ-a.e. equivalent to a *-isomorphism, then it equals that *-isomorphism. However, if

$\xi : \mathcal{M}_{mp}(\mathbb{C}) \rightsquigarrow \mathbb{C}$ is a state and $\varphi : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_{mp}(\mathbb{C})$ is a CPU map that is ξ-a.e. equivalent to a *-homomorphism, then it is not necessarily equal to that *-homomorphism (unless $p = 1$). A simple counter-example is $\varphi(B) := \mathrm{diag}(B, \mathrm{tr}(B)\mathbb{1}_m, \dots, \mathrm{tr}(B)\mathbb{1}_m)$ and ξ the state represented by the density matrix $\frac{1}{m} \mathrm{diag}(\mathbb{1}_m, 0, \dots, 0)$. Then φ is ξ-a.e. equivalent to the *-homomorphism $B \mapsto \mathrm{diag}(B, \dots, B)$, but it is not equal to it (unless $m = 1$ or $p = 1$).

The following corollary of Theorem 2.32 is similar to a fact used frequently in the area of reversible quantum operations (cf. Nayak and Sen [31], the proof of Theorem 2.1]).

Corollary 2.48. *Let* $F : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ *and* $R : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_m(\mathbb{C})$ *be CPU maps with Kraus decompositions*

$$R = \sum_{i=1}^{p} \mathrm{Ad}_{R_i} \quad and \quad F = \sum_{j=1}^{q} \mathrm{Ad}_{F_j}.$$

If $R \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_m(\mathbb{C})}$ *for some state* $\xi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$, *then there exist complex numbers* $\{\alpha_{ij}\}_{i \in \{1,\dots,p\}, \, j \in \{1,\dots,q\}}$ *such that*

$$R_i F_j = \alpha_{ij} \mathbb{1}_m \quad and \quad \sum_{i,j} |\alpha_{ij}|^2 = 1.$$

In particular, $R \circ F = \mathrm{id}_{\mathcal{M}_m(\mathbb{C})}$.

Proof. This follows immediately from Theorem 2.32. ∎

3. Categories of C^*-algebras, states, and morphisms

We prove that *non-commutative probability spaces*, C^*-algebras equipped with states, and a.e. equivalence classes of CPU maps (in fact, 2-positive unital maps) form a category. In fact, finite-dimensional C^*-algebras and a.e. equivalence classes of positive unital maps form a category. The following Cauchy–Schwarz-type inequality, due to Kadison [25], for positive unital and 2-positive unital maps is useful in proving many of these claims.

Lemma 3.1. *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be* C^*-algebras *and let* $\varphi : \mathcal{A} \rightsquigarrow \mathcal{B}$ *be a positive unital map.*

(i)   *If* $a \in \mathcal{A}$ *is self-adjoint, then* $\varphi(a)^2 \leq \varphi(a^2)$.

(ii)   *If* φ *is 2-positive, then* $\varphi(a)^* \varphi(a) \leq \varphi(a^*a)$ *for all* $a \in \mathcal{A}$.

Proof. See [54, Theorem 1.3.1 and Corollary 1.3.2] and [46, Proposition 3.3]. ∎

Proposition 3.2. *Let* $\mathcal{C}$, $\mathcal{B}$, *and* $\mathcal{A}$ *be* C^*-algebras, *let* $\omega : \mathcal{A} \rightsquigarrow \mathbb{C}$ *be a state on* $\mathcal{A}$, *and let* $G, G' : \mathcal{C} \rightsquigarrow \mathcal{B}$ *and* $F, F' : \mathcal{B} \rightsquigarrow \mathcal{A}$ *be 2-positive (or Schwarz-positive) unital maps. If* $F \underset{\omega}{=} F'$ *and* $G \underset{\xi}{=} G'$, *where* $\xi := \omega \circ F$, *then* $F \circ G \underset{\omega}{=} F' \circ G'$.

Proof. By assumption, the diagrams

$$
\begin{array}{c}
\xrightarrow{G} \mathcal{B} \searrow \\
\mathcal{C} \qquad\qquad \mathcal{B}/\mathcal{N}_\xi \quad \text{and} \quad \mathcal{B} \qquad\qquad \mathcal{A}/\mathcal{N}_\omega \\
\searrow_{G'} \mathcal{B} \nearrow
\end{array}
\qquad (3.3)
$$

both commute. For the composite, we have

$$
\mathcal{C} \quad\quad \mathcal{B}/\mathcal{N}_\xi \quad\quad \mathcal{A}/\mathcal{N}_\omega.
$$

The left part of this diagram commutes by commutativity of the left diagram in (3.3). It would be convenient to have a function $\mathcal{B}/\mathcal{N}_\xi \to \mathcal{A}/\mathcal{N}_\omega$ to fill in the diagram. In this regard, let $\widetilde{F}, \widetilde{F'} : \mathcal{B}/\mathcal{N}_\xi \to \mathcal{A}/\mathcal{N}_\omega$ be the functions defined by

$$
\mathcal{B}/\mathcal{N}_\xi \ni [b]_\xi \mapsto \widetilde{F}\big([b]_\xi\big) := \big[F(b)\big]_\omega
$$

and

$$
\mathcal{B}/\mathcal{N}_\xi \ni [b]_\xi \mapsto \widetilde{F'}\big([b]_\xi\big) := \big[F'(b)\big]_\omega.
$$

To see that $\widetilde{F}$ is well defined, let $b \in \mathcal{N}_\xi$, i.e., $\xi(b^*b) = 0$. Then

$$
\omega\big(F(b)^* F(b)\big) \le \omega\big(F(b^*b)\big) = \xi(b^*b) = 0
$$

by Lemma 3.1 applied to F and the fact that F is state-preserving so that $\omega \circ F = \xi$. A similar conclusion can be made for $\widetilde{F'}$. Since $F(b)^* F(b) \ge 0$ and ω is a positive functional, this shows that $F(b) \in \mathcal{N}_\omega$, which proves that $\widetilde{F}$ and $\widetilde{F'}$ are well defined. In fact, by commutativity of the right diagram in (3.3), $\widetilde{F} = \widetilde{F'}$. Hence, all the subdiagrams in the diagram

$$
\mathcal{C} \quad\quad \mathcal{B}/\mathcal{N}_\xi \xrightarrow[\widetilde{F'}]{\widetilde{F}} \mathcal{A}/\mathcal{N}_\omega
$$

commute so that $F \circ G \underset{\omega}{=} F' \circ G'$. ∎

Assuming finite-dimensionality, we can prove more. Although the previous proposition is enough for the sequel, the following theorem is an interesting result in its own right.

Theorem 3.4. *Let $\mathcal{C}$, $\mathcal{B}$, and $\mathcal{A}$ be finite-dimensional C^*-algebras, let $\omega : \mathcal{A} \rightsquigarrow \mathbb{C}$ be a state on $\mathcal{A}$, and let $G, G' : \mathcal{C} \rightsquigarrow \mathcal{B}$ and $F, F' : \mathcal{B} \rightsquigarrow \mathcal{A}$ be positive unital maps with $G \underset{\xi}{=} G'$ and $F \underset{\omega}{=} F'$, where $\xi := \omega \circ F = \omega \circ F'$. Then $F \circ G \underset{\omega}{=} F' \circ G'$.*

We will break up this proof into several lemmas, some of which are of independent interest. For the first lemma (the proof of which is immediate), recall that if $\varphi : \mathcal{A} \rightsquigarrow \mathcal{B}$ is a linear map between C^*-algebras, then φ is *-*preserving* if and only if $\varphi(a)^* = \varphi(a^*)$ for all $a \in \mathcal{A}$. Also, a vector subspace $V \subseteq \mathcal{B}$ is *self-adjoint* if and only if $v \in V$ implies that $v^* \in V$.

Lemma 3.5. *Let $\mathcal{A}$ and $\mathcal{B}$ be C^*-algebras and let $\varphi : \mathcal{A} \rightsquigarrow \mathcal{B}$ be a linear map. If φ is *-preserving, then the image is a self-adjoint subspace of $\mathcal{B}$.*

Lemma 3.6. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras, let $\xi : \mathcal{B} \rightsquigarrow \mathbb{C}$ be a state, and let $\varphi : \mathcal{A} \rightsquigarrow \mathcal{B}$ be a *-preserving linear map. If φ is ξ-a.e. equivalent to 0, then $\mathrm{Im}(\varphi) \subseteq P_\xi^\perp \mathcal{B} P_\xi^\perp$.*

Proof of Lemma 3.6. By assumption, $\mathrm{Im}(\varphi) \subseteq \mathcal{B} P_\xi^\perp$. If $\varphi(a) = P_\xi b P_\xi^\perp + P_\xi^\perp b P_\xi^\perp$ for some $b \in \mathcal{B}$, then $\varphi(a^*) = \varphi(a)^* \in \mathcal{N}_\xi$ as well by Lemma 3.5. But $\varphi(a)^* = P_\xi^\perp b^* P_\xi + P_\xi^\perp b^* P_\xi^\perp$. Hence $P_\xi^\perp b^* P_\xi = 0$. By taking the adjoint of this, we get $P_\xi b P_\xi^\perp = 0$. Thus, $\varphi(a) = P_\xi^\perp b P_\xi^\perp$ for some $b \in \mathcal{B}$. ∎

Lemma 3.7. *Let $F : \mathcal{B} \rightsquigarrow \mathcal{A}$ be a positive map between C^*-algebras and let P be a projection in $\mathcal{A}$. If $b \in \mathcal{B}$, then $F(b)$ can be uniquely decomposed as*

$$F(b) = F^{\mathrm{tl}}(b) + F^{\mathrm{tr}}(b) + F^{\mathrm{bl}}(b) + F^{\mathrm{br}}(b),$$

where

$$F^{\mathrm{tl}}(b) := PF(b)P, \qquad F^{\mathrm{tr}}(b) := PF(b)P^\perp,$$
$$F^{\mathrm{bl}}(b) := P^\perp F(b)P, \quad F^{\mathrm{br}}(b) := P^\perp F(b)P^\perp.$$

Furthermore, $F^{\mathrm{tr}}(b)^* = F^{\mathrm{bl}}(b^*)$ for all $b \in \mathcal{B}$ and the maps $F^{\mathrm{tl}}, F^{\mathrm{br}} : \mathcal{B} \rightsquigarrow \mathcal{A}$ are positive.

Proof. The decomposition itself is just (2.19). From this and *-preservation of F,

$$F^{\mathrm{tr}}(b)^* = \big(PF(b)P^\perp\big)^* = P^\perp F(b)^* P = P^\perp F(b^*)P = F^{\mathrm{bl}}(b^*).$$

Furthermore, F^{tl} and F^{br} are positive maps since $F^{\mathrm{tl}} = \mathrm{Ad}_P \circ F$ and $F^{\mathrm{br}} = \mathrm{Ad}_{P^\perp} \circ F$ are composites of positive maps. ∎

Lemma 3.8. *Let $(\mathcal{B}, \xi)$ and $(\mathcal{A}, \omega)$ be finite-dimensional C^*-algebras equipped with states and let $F : \mathcal{B} \rightsquigarrow \mathcal{A}$ be positive unital and state-preserving. Then*

$$F(P_\xi^\perp \mathcal{B} P_\xi^\perp) \subseteq P_\omega^\perp \mathcal{A} P_\omega^\perp.$$

In particular, $F(P_\xi^\perp \mathcal{B} P_\xi^\perp) \subseteq \mathcal{N}_\omega$.

Proof. Let $b \in P_\xi^\perp \mathcal{B} P_\xi^\perp$. First, assume that b is self-adjoint. Then

$$\omega\big(F(b)^* F(b)\big) = \omega\big(F(b^*)F(b)\big) = \omega\big(F(b)^2\big) \leq \omega\big(F(b^2)\big) = \xi(b^2) = \xi(b^*b) = 0,$$

where the inequality follows from part (i) of Lemma 3.1, the equality after it follows from the fact that F is state-preserving, and the final equality follows from $b \in \mathcal{N}_\xi$. This proves that $F(b) \in \mathcal{N}_\omega$ for self-adjoint $b \in P_\xi^\perp \mathcal{B} P_\xi^\perp$. Hence, $F^{\text{tl}}(b) = 0$ and $F^{\text{bl}}(b) = 0$ for b self-adjoint by Lemma 3.7. Second, assume that b is skew-adjoint. Then $b = ib'$ for some self-adjoint $b' \in \mathcal{B}$ (namely, $b' := -ib$). Then

$$F^{\text{tl}}(b) = F^{\text{tl}}(ib') = iF^{\text{tl}}(b') = 0 \quad \text{and similarly} \quad F^{\text{bl}}(b) = 0$$

for skew-adjoint b by the previous fact since F^{tl} and F^{bl} are linear. Since every b can be decomposed as the linear combination of a self-adjoint and skew-adjoint element, this proves that F^{tl} and F^{bl} are both equal to the zero map. Finally, for any $b \in P_\xi^\perp \mathcal{B} P_\xi^\perp$,

$$F^{\text{tr}}(b) = F^{\text{tr}}\big((b^*)^*\big) = F^{\text{bl}}(b^*)^* = 0$$

by Lemma 3.7 and the facts just proved.   ∎

Proof of Theorem 3.4. We are required to prove that $F(G(c))P_\omega = F'(G'(c))P_\omega$ for all $c \in \mathcal{C}$. First, note that

$$F'\big(G'(c)\big)P_\omega = F\big(G'(c)\big)P_\omega \quad \text{since } F \underset{\omega}{=\!=} F'$$

$$= F\big(G'(c)P_\xi + G'(c)P_\xi^\perp\big)P_\omega. \tag{3.9}$$

Therefore,

$$F\big(G(c)\big)P_\omega - F'\big(G'(c)\big)P_\omega$$
$$= F\big((G(c) - G'(c))P_\xi + (G(c) - G'(c))P_\xi^\perp\big)P_\omega \quad \text{by (3.9)}$$
$$= F\big((G(c) - G'(c))P_\xi^\perp\big)P_\omega \quad \text{since } G \underset{\xi}{=\!=} G'$$
$$= F\big(\text{Ad}_{P_\xi^\perp}(G(c) - G'(c))\big)P_\omega \quad \text{by Lemma 3.6}$$
$$= \text{Ad}_{P_\omega^\perp}\big(F\big(\text{Ad}_{P_\xi^\perp}(G(c) - G'(c))\big)\big)P_\omega \quad \text{by Lemma 3.8}$$
$$= 0.$$

This proves that composition of a.e.-equivalence classes of positive unital maps between finite-dimensional C^*-algebras is well defined.   ∎

Remark 3.10. By Lemma 2.13 and Remark 2.15, Theorem 3.4 holds if $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are W^*-algebras.

Definition 3.11. A *non-commutative probability space* is a pair $(\mathcal{A}, \omega)$, with $\mathcal{A}$ a C^*-algebra and ω a state on $\mathcal{A}$. A *state-preserving map* $(\mathcal{B}, \xi) \rightsquigarrow (\mathcal{A}, \omega)$ is a map (linear, positive, CP, *-homomorphism, etc.) $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ such that $\xi = \omega \circ F$.

Corollary 3.12. *The following facts hold.*

(i)   *The collection of non-commutative probability spaces and state-preserving maps forms a category.*

(ii)   *The collection of non-commutative probability spaces and a.e. equivalence classes of 2-positive unital maps forms a category.*

(iii)   *The collection of finite-dimensional non-commutative probability spaces (or non-commutative probability spaces on W^*-algebras) and a.e. equivalence classes of PU maps forms a category.*

(iv)   *The opposite of the category of finite probability spaces and probability-preserving stochastic maps embeds fully into the category of non-commutative probability spaces and state-preserving PU maps. It is an equivalence on the subcategory of finite-dimensional commutative C^*-algebras.*

(v)   *Two probability-preserving stochastic maps are a.e. equivalent if and only if their associated PU maps are a.e. equivalent.*

The functor in item (iv) is uniquely determined by sending a finite set X to the C^*-algebra $\mathbb{C}^X$ and sending a stochastic map $f : X \rightsquigarrow Y$ to the PU map $\mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$ uniquely determined by sending the basis vector e_y to the function $\sum_{x \in X} f_{yx} e_x \in \mathbb{C}^X$ (cf. Example 2.11).

Proof. This follows from Example 2.11, Proposition 3.2, and Theorem 3.4, as well as [35, Section 2]. ∎

Remark 3.13. Note that when $f : X \to Y$ is a function, the functor from Corollary 3.12 item (iv) produces the *-homomorphism $\mathbb{C}^Y \to \mathbb{C}^X$ sending $\varphi \in \mathbb{C}^Y$ to $\varphi \circ f$, the pullback of φ along f.

Remark 3.14. The diagrammatic definition of a.e. equivalence discussed in Remark A.10 cannot be transferred to our categories of C^*-algebras and positive maps. To see this, first note that the cartesian product of sets also goes to the tensor product of C^*-algebras (up to a natural isomorphism); i.e., the (contravariant) functor sending stochastic maps to PU maps extends to a monoidal functor (the product of stochastic maps is defined by the product of the associated probability measures). In particular, the diagonal map $\Delta_X : X \to X \times X$ becomes the multiplication map $\mathbb{C}^X \otimes \mathbb{C}^X \to \mathbb{C}^X$. Therefore, a natural candidate for $\mathcal{A} \otimes \mathcal{A} \rightsquigarrow \mathcal{A}$ would be the linear map that takes the product of the elements. However, this map is not positive in general. This is closely related to the no-cloning/no-broadcasting theorem in quantum mechanics [3, 9, 10, 34, 38, 57]. As a result, it is not a morphism in any of our categories. Our definition of a.e. equivalence in Definition 2.9,

though not explicitly categorical, gives a direct definition of a.e. equivalence in terms of null spaces and is suitable for our purposes of non-commutative probability. Nevertheless, it has recently been proven that this result does agree with the categorical definition of a.e. equivalence when instantiated in quantum Markov categories [38, Theorem 5.12].

4. Non-commutative disintegrations on matrix algebras

Here, we define optimal hypothesis, disintegration, and regular conditional probability in the non-commutative setting. In Theorem 4.3, we provide a necessary and sufficient condition for a disintegration to exist on matrix algebras. The state on the initial algebra must be separable with the induced state as a factor. This result holds for *-homomorphisms of a special kind. In this same theorem, it is shown that a disintegration is *unique* whenever one exists. In the proof, we construct an explicit formula for *any* disintegration on matrix algebras. Theorem 4.16 covers the more general case of arbitrary *-homomorphisms between matrix algebras. Briefly, the existence no longer requires the initial state to be separable. However, it is separable after a specific unitary operation that transforms the *-homomorphism to one of the kind discussed in Theorem 4.3.

Definition 4.1. Given a state-preserving *-homomorphism $(\mathcal{B}, \xi) \xrightarrow{F} (\mathcal{A}, \omega)$ on C^*-algebras, a *hypothesis* for $(\mathcal{B}, \xi) \xrightarrow{F} (\mathcal{A}, \omega)$ is a CPU map $R : \mathcal{A} \rightsquigarrow \mathcal{B}$ such that[6] $R \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$. A hypothesis for $(\mathcal{B}, \xi) \xrightarrow{F} (\mathcal{A}, \omega)$ is *optimal* if and only if $\xi \circ R = \omega$. A CPU map $R : \mathcal{A} \rightsquigarrow \mathcal{B}$ is a *disintegration of ω over ξ* if and only if $\xi \circ R = \omega$ holds. A CPU map $R : \mathcal{A} \rightsquigarrow \mathcal{B}$ is a *disintegration of ω over ξ consistent with F* if and only if R is a disintegration of ω over ξ such that $R \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$. More concisely, a *disintegration* refers to a disintegration of ω over ξ consistent with F.

The following example illustrates how Definition 4.1 extends the classical definition of a disintegration to the non-commutative setting.

Example 4.2. Let X and Y be finite sets (with the discrete σ-algebras) with probability measures $p : \{\bullet\} \rightsquigarrow X$ and $q : \{\bullet\} \rightsquigarrow Y$. Let $f : X \to Y$ be a function and let $r : Y \rightsquigarrow X$ be a stochastic map. Let $P : \mathbb{C}^X \rightsquigarrow \mathbb{C}$, $Q : \mathbb{C}^Y \rightsquigarrow \mathbb{C}$, $F : \mathbb{C}^Y \to \mathbb{C}^X$, and $R : \mathbb{C}^X \rightsquigarrow \mathbb{C}^Y$ denote the corresponding PU maps (cf. Example 2.11). Functoriality as discussed after Corollary 3.12 immediately implies the following.

(i) F is state-preserving if and only if f is measure-preserving.

(ii) R is a disintegration of P over Q consistent with F in the sense of Definition 4.1 if and only if r is a disintegration of p over q consistent with f in the sense described in the appendix (see also Theorem 5.1 later for a simpler description in terms of finite sets).

---

[6]This definition of hypothesis is a non-commutative generalization of the definition from [1]. In [1], the definition also requires equality rather than a.e. equality, so our notion is also a weakening in this sense.

Several natural questions arise when comparing our definition of disintegration to the one from finite probability spaces, measure-preserving maps, and stochastic maps. First of all, given a state-preserving *-homomorphism $(\mathcal{B}, \xi) \xrightarrow{F} (\mathcal{A}, \omega)$, does there *exist* a disintegration R over ω consistent with ξ? Second, if a disintegration exists, is it unique or at least unique up to a.e. equivalence? Third, is a disintegration of a *-isomorphism (a.e. equivalent to) the inverse? All of these are true in the commutative case. Before addressing the general case of finite-dimensional C^*-algebras, the present section focuses on the setting of matrix algebras.

Theorem 4.3. *Fix $n, p \in \mathbb{N}$. Let F be the *-homomorphism given by the block diagonal inclusion*

$$\mathcal{M}_n(\mathbb{C}) \ni B \xrightarrow{F} \begin{bmatrix} B & & 0 \\ & \ddots & \\ 0 & & B \end{bmatrix} \equiv \mathbb{1}_p \otimes B \in \mathcal{M}_{np}(\mathbb{C}) \cong \mathcal{M}_p(\mathbb{C}) \otimes \mathcal{M}_n(\mathbb{C}) \quad (4.4)$$

and let $(\mathcal{M}_n(\mathbb{C}), \mathrm{tr}(\rho \,\cdot\,) \equiv \omega) \xrightarrow{F} (\mathcal{M}_{np}(\mathbb{C}), \xi \equiv \mathrm{tr}(\sigma \,\cdot\,))$ be state-preserving. Then the following facts hold.

(i) *A disintegration R of ω over ξ consistent with F exists if and only if there exists a density matrix $\tau \in \mathcal{M}_p(\mathbb{C})$ such that $\rho = \tau \otimes \sigma$.*

(ii) *When such a τ exists, the disintegration is unique and is given by the formula*

$$\mathcal{M}_{np}(\mathbb{C}) \ni A \equiv \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pp} \end{bmatrix} \mapsto R(A) := \sum_{j,k=1}^{p} \tau_{kj} A_{jk}$$

$$\equiv \mathrm{tr}_{\mathcal{M}_p(\mathbb{C})}\big((\tau \otimes \mathbb{1}_n)A\big), \quad (4.5)$$

where A_{jk} is the jk-th $n \times n$ block of A using the isomorphisms $\mathcal{M}_{np}(\mathbb{C}) \cong \mathcal{M}_p(\mathcal{M}_n(\mathbb{C})) \cong \mathcal{M}_p(\mathbb{C}) \otimes \mathcal{M}_n(\mathbb{C})$ and $\mathrm{tr}_{\mathcal{M}_p(\mathbb{C})} : \mathcal{M}_p(\mathbb{C}) \otimes \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ is the partial trace, *uniquely determined by sending $C \otimes B \in \mathcal{M}_p(\mathbb{C}) \otimes \mathcal{M}_n(\mathbb{C})$ to $\mathrm{tr}(C)B$. Furthermore, $R \circ F = \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$.*

(iii) *When such a τ exists, a Kraus decomposition of R is given by*

$$R = \mathrm{Ad}_{[\mathbb{1}_n \; 0 \; \cdots \; 0](\sqrt{\tau} \otimes \mathbb{1}_n)} + \cdots + \mathrm{Ad}_{[0 \; \cdots \; 0 \; \mathbb{1}_n](\sqrt{\tau} \otimes \mathbb{1}_n)}.$$

As a consequence of uniqueness, we prove all disintegrations on matrix algebras are strict left inverses of their associated *-homomorphism. Note that uniqueness is meant in the literal sense, not in the a.e. sense. This is surprising due to Remark 2.45, which says two a.e. equivalent CPU maps on matrix algebras need not be equal. The additional conditions for a disintegration are strong enough to imply equality. The following proof also provides a construction of the density matrix τ from R.

Proof of Theorem 4.3. (i) ($\Rightarrow$) Suppose that a disintegration $R : \mathcal{M}_{np}(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ exists. Let

$$R = \sum_{i=1}^{n^2 p} \mathrm{Ad}_{\left[V_{i1} \;\; \cdots \;\; V_{ip}\right]} \tag{4.6}$$

be a Kraus decomposition of R with $V_{ij} \in \mathcal{M}_n(\mathbb{C})$ for all $i \in \{1, \ldots, n^2 p\}$ and $j \in \{1, \ldots, p\}$ ($n^2 p$ is the minimal number of Kraus operators needed in this case). For the moment, let $R_i := [\, V_{i1} \cdots V_{ip} \,]$. Also note that F has a Kraus decomposition $F = \sum_{j=1}^{p} \mathrm{Ad}_{F_j}$, where the (adjoint of the) Kraus operators are given by

$$F_j^\dagger := \begin{bmatrix} 0 & \cdots & \mathbb{1}_n & \cdots & 0 \end{bmatrix} \tag{4.7}$$

with $\mathbb{1}_n$ in the j-th $n \times n$ block. By Corollary 2.48, there exist numbers

$$\{\alpha_{ij} \in \mathbb{C}\}_{i \in \{1,\ldots,n^2 p\},\, j \in \{1,\ldots,p\}}$$

such that

$$V_{ij} = R_i F_j = \alpha_{ij} \mathbb{1}_n \quad \forall i, j \quad \text{and} \quad \sum_{i,j} |\alpha_{ij}|^2 = 1 \tag{4.8}$$

due to the form of our matrices in (4.6) and (4.7). We now impose the condition $\xi \circ R = \omega$, which is equivalent to $R^*(\sigma) = \rho$, where

$$R^* = \sum_{i=1}^{n^2 p} \mathrm{Ad}_{R_i^\dagger} \tag{4.9}$$

is the adjoint of R with respect to the Hilbert–Schmidt inner product (cf. Remark 2.8). Therefore,

$$\rho = R^*(\sigma) \overset{(4.9)}{=} \sum_{i=1}^{n^2 p} \mathrm{Ad}_{R_i^\dagger}(\sigma) \overset{(4.6)}{=} \sum_{i=1}^{n^2 p} \begin{bmatrix} V_{i1}^\dagger \\ \vdots \\ V_{ip}^\dagger \end{bmatrix} \sigma \begin{bmatrix} V_{i1} & \cdots & V_{ip} \end{bmatrix}$$

$$\overset{(4.8)}{=} \sum_{i=1}^{n^2 p} \begin{bmatrix} \overline{\alpha_{i1}} \mathbb{1}_n \\ \vdots \\ \overline{\alpha_{ip}} \mathbb{1}_n \end{bmatrix} \sigma \begin{bmatrix} \alpha_{i1} \mathbb{1}_n & \cdots & \alpha_{ip} \mathbb{1}_n \end{bmatrix}$$

$$= \sum_{i=1}^{n^2 p} \begin{bmatrix} |\alpha_{i1}|^2 \sigma & \cdots & \overline{\alpha_{i1}} \alpha_{ip} \sigma \\ \vdots & & \vdots \\ \overline{\alpha_{ip}} \alpha_{i1} \sigma & \cdots & |\alpha_{ip}|^2 \sigma \end{bmatrix}$$

$$\overset{(2.3)}{=} \underbrace{\left(\sum_{i=1}^{n^2 p} \begin{bmatrix} |\alpha_{i1}|^2 & \cdots & \overline{\alpha_{i1}} \alpha_{ip} \\ \vdots & & \vdots \\ \overline{\alpha_{ip}} \alpha_{i1} & \cdots & |\alpha_{ip}|^2 \end{bmatrix} \right)}_{=:\tau} \otimes \sigma, \tag{4.10}$$

showing that ρ is separable and has a tensor product factorization with σ as a factor. Now, $\tau \in \mathcal{M}_p(\mathbb{C})$ is a positive matrix because it is a positive sum of positive operators, namely

$$\tau = \sum_{i=1}^{n^2 p} \begin{bmatrix} \overline{\alpha_{i1}} \\ \vdots \\ \overline{\alpha_{ip}} \end{bmatrix} \begin{bmatrix} \alpha_{i1} & \cdots & \alpha_{ip} \end{bmatrix}.$$

Furthermore, (4.8) implies that

$$\text{tr}(\tau) = \sum_{j=1}^{p} \sum_{i=1}^{n^2 p} |\alpha_{ij}|^2 = 1,$$

which shows that τ is a density matrix.

($\Longleftarrow$) Conversely, suppose that there exists a density matrix τ such that $\rho = \tau \otimes \sigma$. Define $R : \mathcal{M}_{np}(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ as in (4.5). The map R is linear by construction and unital since

$$R(\mathbb{1}_{np}) = \sum_{j,k=1}^{p} \tau_{kj} \delta_{jk} \mathbb{1}_p = \sum_{j} \tau_{jj} \mathbb{1}_p = \mathbb{1}_p$$

because $\text{tr}(\tau) = 1$. A similar calculation shows that

$$R\big(F(B)\big) = \sum_{j,k=1}^{p} \tau_{kj} \delta_{jk} B = B$$

for all $B \in \mathcal{M}_n(\mathbb{C})$. Hence, R is actually a left inverse of F. In order for R to preserve the states, it must be that $\omega \circ R = \xi$; i.e., $\text{tr}(\rho A) = \text{tr}(\sigma R(A))$ for all $A \in \mathcal{M}_{np}(\mathbb{C})$. This follows from

$$\text{tr}(\rho A) = \text{tr}\big((\tau \otimes \sigma)A\big) = \text{tr}\left(\begin{bmatrix} \tau_{11}\sigma & \cdots & \tau_{1p}\sigma \\ \vdots & & \vdots \\ \tau_{p1}\sigma & \cdots & \tau_{pp}\sigma \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pp} \end{bmatrix} \right)$$

$$= \text{tr}\left(\begin{bmatrix} \sum_{j=1}^{p} \tau_{1j}\sigma A_{j1} & \cdots & \sum_{j=1}^{p} \tau_{1j}\sigma A_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^{p} \tau_{pj}\sigma A_{j1} & \cdots & \sum_{j=1}^{p} \tau_{pj}\sigma A_{jp} \end{bmatrix} \right)$$

$$= \sum_{j,k=1}^{p} \tau_{kj} \text{tr}(\sigma A_{jk}) = \text{tr}\big(\sigma R(A)\big).$$

The final step is to prove that R is CP. This follows from the fact that the partial trace satisfies a partially cyclic property, namely

$$\text{tr}_{\mathcal{M}_p(\mathbb{C})}\big((\tau \otimes \mathbb{1}_n)A\big) = \text{tr}_{\mathcal{M}_p(\mathbb{C})}\big(A(\tau \otimes \mathbb{1}_n)\big).$$

Thus,

$$\text{tr}_{\mathcal{M}_p(\mathbb{C})}\big((\tau \otimes \mathbb{1}_n)A\big) = \text{tr}_{\mathcal{M}_p(\mathbb{C})}\big((\sqrt{\tau} \otimes \mathbb{1}_n)A(\sqrt{\tau} \otimes \mathbb{1}_n)\big)$$
$$= \big(\text{tr}_{\mathcal{M}_p(\mathbb{C})} \circ \text{Ad}_{\sqrt{\tau}\otimes\mathbb{1}_n}\big)(A)$$

shows that R is the composite of two CP maps, and is therefore CP. The claim also follows from showing that the Choi matrix associated to R is $\Phi(R) = \tau^T \otimes \Phi(\text{id}_{\mathcal{M}_n(\mathbb{C})})$, which is positive (the proof is omitted).

(ii) Suppose that R' is another disintegration of ω over ξ consistent with F. Let $\{\alpha_{ij}\}$ and $\{\alpha'_{ij}\}$ be coefficients obtained from Choi's theorem as in (4.8). Construct the density matrices τ and τ' as in the proof of part (i) of this theorem. Then $\tau' \otimes \sigma = \rho = \tau \otimes \sigma$, i.e., $(\tau - \tau') \otimes \sigma = 0$. Since σ is non-zero, this means that $\tau - \tau' = 0$; i.e., $\tau' = \tau$. Hence, each of the entries of τ and τ' are equal; i.e., $\tau_{jk} = \tau'_{jk}$ for all j, k, or in terms of the α's and α'''s,

$$\sum_{i=1}^{n^2 p} \overline{\alpha_{ij}}\alpha_{ik} = \sum_{i=1}^{n^2 p} \overline{\alpha'_{ij}}\alpha'_{ik} \quad \forall j, k \in \{1, \ldots, p\}. \tag{4.11}$$

Now, let A be a matrix in $\mathcal{M}_{np}(\mathbb{C}) \cong \mathcal{M}_p(\mathcal{M}_n(\mathbb{C}))$ as in (4.5) so that each $A_{jk} \in \mathcal{M}_n(\mathbb{C})$. Then, after some algebra

$$R(A) = \sum_{i=1}^{n^2 p} \begin{bmatrix} \alpha_{i1}\mathbb{1}_n & \cdots & \alpha_{ip}\mathbb{1}_n \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pp} \end{bmatrix} \begin{bmatrix} \overline{\alpha_{i1}}\mathbb{1}_n \\ \vdots \\ \overline{\alpha_{ip}}\mathbb{1}_n \end{bmatrix}$$

$$= \sum_{j,k=1}^{p} \left(\sum_{i=1}^{n^2 p} \alpha_{ij}\overline{\alpha_{ik}}\right) A_{jk} \stackrel{(4.11)}{=} \sum_{j,k=1}^{p} \left(\sum_{i=1}^{n^2 p} \alpha'_{ij}\overline{\alpha'_{ik}}\right) A_{jk} = R'(A),$$

which shows that $R = R'$. Hence, disintegrations are unique when they exist. The fact that $R \circ F = \text{id}_{\mathcal{M}_n(\mathbb{C})}$ follows from uniqueness of disintegrations and Corollary 2.48.

(iii) The formula for the Kraus decomposition follows from the results just proven and a Kraus decomposition for the partial trace. ∎

Example 4.12. Let

$$\rho := \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

be the density matrix on $\mathbb{C}^4$ corresponding to the projection operator onto the one-dimensional subspace of $\mathbb{C}^2 \otimes \mathbb{C}^2$ spanned by the vector[7]

$$\vec{u}_{\text{EPR}} := \frac{1}{\sqrt{2}}(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1).$$

---

[7]This is the spin EPR state discussed by Nielsen and Chuang in [32, Section 1.3.6].

Let $F : \mathcal{M}_2(\mathbb{C}) \to \mathcal{M}_4(\mathbb{C})$ be the map defined by

$$\mathcal{M}_2(\mathbb{C}) \ni B \mapsto F(B) := \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}, \tag{4.13}$$

which corresponds to the assignment $\mathcal{M}_2(\mathbb{C}) \ni B \mapsto \mathbb{1}_2 \otimes B \in \mathcal{M}_2(\mathbb{C}) \otimes \mathcal{M}_2(\mathbb{C})$ under the isomorphism from (2.4). Let σ be the density matrix on $\mathbb{C}^2$ given by $\sigma := \frac{1}{2}\mathbb{1}_2$. Let $\omega := \mathrm{tr}(\rho \cdot)$ and $\xi := \mathrm{tr}(\sigma \cdot)$ be the corresponding states. Then, $\mathcal{N}_\xi = \{0\}$ and $(\mathcal{M}_2(\mathbb{C}), \mathrm{tr}(\rho \cdot) \equiv \omega) \xrightarrow{F} (\mathcal{M}_4(\mathbb{C}), \xi \equiv \mathrm{tr}(\sigma \cdot))$ is state-preserving, but there does not exist a disintegration of ω over ξ consistent with F.

Example 4.14. Fix $p_1, p_2, p_3, p_4 \geq 0$ with $p_1 + p_2 + p_3 + p_4 = 1$, $p_1 + p_3 > 0$, and $p_2 + p_4 > 0$. Let

$$\rho = \begin{bmatrix} p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \\ 0 & 0 & 0 & p_4 \end{bmatrix} \quad \text{and} \quad \sigma = \begin{bmatrix} p_1 + p_3 & 0 \\ 0 & p_2 + p_4 \end{bmatrix}$$

be density matrices with associated states given by $\omega := \mathrm{tr}(\rho \cdot)$ and $\xi := \mathrm{tr}(\sigma \cdot)$, respectively. Let $F : \mathcal{M}_2(\mathbb{C}) \to \mathcal{M}_4(\mathbb{C})$ be the diagonal inclusion from (4.13). Then $\xi = \omega \circ F$ and $\mathcal{N}_\xi = \{0\}$. Furthermore, a CPU disintegration $R : \mathcal{M}_4(\mathbb{C}) \rightsquigarrow \mathcal{M}_2(\mathbb{C})$ of ω over ξ consistent with F exists if and only if $p_1 p_4 = p_2 p_3$. When this holds, the map

$$R = \mathrm{Ad}_{\sqrt{p_1 + p_2}[\mathbb{1}_2 \ 0]} + \mathrm{Ad}_{\sqrt{p_3 + p_4}[0 \ \mathbb{1}_2]}$$

is the unique disintegration of ω over ξ consistent with F. Furthermore, the density matrix $\tau \in \mathcal{M}_2(\mathbb{C})$ given by

$$\tau = \begin{bmatrix} \frac{p_1}{p_1 + p_3} & 0 \\ 0 & \frac{p_4}{p_2 + p_4} \end{bmatrix}$$

satisfies $\tau \otimes \sigma = \rho$.

Remark 4.15. Theorem 4.3 reproduces a well-known result in quantum information theory in the special case when the density matrices ρ and σ are invertible (see Example 9.6 in Petz's text for example [47]).[8] The surprising result we have shown is the fact that this still holds regardless of the sizes of the null-spaces associated to the density matrices and, moreover, the disintegration is uniquely determined.

The following result is a generalization of Theorem 4.3 on the existence of disintegrations to allow for *-homomorphisms F that are not necessary of the block diagonal form.

---

[8]The techniques we have used to prove our results do not use Takesaki's theorem nor the modular group (see [47, Theorem 9.2]). Instead, we worked directly with Kraus operators, a familiar tool in the quantum information theory community. A deeper analysis relating conditional expectations to disintegrations has recently been presented in [20].

Theorem 4.16. *Fix n, $p \in \mathbb{N}$. Let $(\mathcal{M}_n(\mathbb{C}), \mathrm{tr}(\rho \cdot) \equiv \omega) \xrightarrow{F} (\mathcal{M}_{np}(\mathbb{C}), \xi \equiv \mathrm{tr}(\sigma \cdot))$ be a state-preserving *-homomorphism. A disintegration of ω over ξ consistent with F exists if and only if there exists a unitary $U \in \mathcal{M}_{np}(\mathbb{C})$ and a density matrix $\tau \in \mathcal{M}_p(\mathbb{C})$ such that $F = \mathrm{Ad}_U \circ i$ and $U^\dagger \rho U = \tau \otimes \sigma$. Here $i : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_{np}(\mathbb{C})$ is the block diagonal inclusion* (4.4). *Furthermore, if a disintegration exists, it is unique.*

Proof. For any unital *-homomorphism $F : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_{np}(\mathbb{C})$, there exists a unitary $U \in \mathcal{M}_{np}(\mathbb{C})$ such that $F = \mathrm{Ad}_U \circ i$ (cf. Fillmore [12, Section 1.1.2]). Hence, the diagram

$$
\begin{array}{ccc}
\mathcal{M}_{np}(\mathbb{C}) & \xleftarrow{\quad i \quad} & \mathcal{M}_n(\mathbb{C}) \\
& & \\
\omega \circ \mathrm{Ad}_U = \mathrm{tr}(U^\dagger \rho U \cdot) \searrow & & \swarrow \xi \equiv \mathrm{tr}(\sigma \cdot) \\
& \mathbb{C} &
\end{array}
$$

commutes. By Theorem 4.3, a disintegration $R : \mathcal{M}_{np}(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ of $\omega \circ \mathrm{Ad}_U$ over ξ consistent with i exists if and only if there exists a density matrix τ such that $U^\dagger \rho U = \tau \otimes \sigma$. Explicitly, this means that $\xi \circ R = \omega \circ \mathrm{Ad}_U$ and $R \circ i \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$. Setting $R_U := R \circ \mathrm{Ad}_{U^\dagger} : \mathcal{M}_{np}(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ and applying $\mathrm{Ad}_{U^\dagger}$ to the right of $\xi \circ R = \omega \circ \mathrm{Ad}_U$ gives $\xi \circ R_U = \omega$. Similarly, $R \circ i \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$ holds if and only if $R \circ \mathrm{Ad}_{U^\dagger} \circ \mathrm{Ad}_U \circ i \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$ holds; i.e., $R_U \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{M}_n(\mathbb{C})}$. The map R_U is CPU if and only if R is CPU. Thus, R_U defines a disintegration of ω over ξ consistent with F if and only if $U^\dagger \rho U = \tau \otimes \sigma$. Finally, the uniqueness of R_U follows from the uniqueness of R by part (ii) of Theorem 4.3. ∎

Note that an immediate consequence of this theorem is when F is a *-isomorphism, then F^{-1} is the unique disintegration. Thus, a disintegration can be viewed as a generalization of time reversal (see also the work [42] for an extension of time-reversal symmetry to a larger class of morphisms).

Remark 4.17. Theorem 4.16 says that there exists a tensor factorization $\tau \otimes \sigma = U^\dagger \rho U$ if and only if a disintegration exists. It is not necessary for ρ to be separable in this case (compare this to Theorem 4.3, where $\rho = \tau \otimes \sigma$ was separable). In general, $U^\dagger \rho U$ being separable does not imply that ρ is separable—the unitary evolution of a separable state can cause that state to become entangled due to interactions between subsystems.

Remark 4.18. Theorem 4.16 bears a striking resemblance to Theorem 2.1 in the work of Nayak and Sen [31]. However, there are three main differences. First, they work with completely positive trace-preserving (not necessarily unital) maps $F : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$, where $m \leq n$, while we focus on the class of unital *-homomorphisms. Second, they assume that $R : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathcal{M}_m(\mathbb{C})$ is a strict left inverse of F while we initially assume that R is a left inverse up to a.e. equivalence. We showed that this condition is actually equivalent for matrix algebras in Corollary 2.48 but we will see that a.e. equivalence is necessary for arbitrary finite-dimensional C^*-algebras. Third, and most importantly, Nayak

and Sen do not require R and F to preserve any specified states while we do. This forces an additional constraint that our map R must satisfy making it even less obvious whether such a CP map R exists. Therefore, it seems that neither of our results subsume each other but are complementary and cover different situations.

If a deterministic process (a *-homomorphism) evolves a pure state into a mixed state between matrix algebras, is it possible for there to exist a disintegration that evolves the mixed state back into the pure state? The following corollary is a "no-go theorem" for such disintegrations.

Corollary 4.19. *Given a state-preserving *-homomorphism*

$$\big(\mathcal{M}_n(\mathbb{C}), \operatorname{tr}(\sigma \cdot)\big) \xrightarrow{F} \big(\mathcal{M}_{np}(\mathbb{C}), \operatorname{tr}(\rho \cdot)\big),$$

with ρ pure, if a disintegration exists, then σ must necessarily be pure as well.

Proof. By Theorem 4.16, there exist a unitary $U \in \mathcal{M}_{np}(\mathbb{C})$ and a density matrix $\tau \in \mathcal{M}_p(\mathbb{C})$ such that $F(A) = U \operatorname{diag}(A, \ldots, A)U^\dagger$ and $\rho = U^\dagger(\tau \otimes \sigma)U$. Since ρ is pure, it is a rank 1 projection operator. Its rank also equals $\operatorname{rank}(\rho) = \operatorname{rank}(\tau)\operatorname{rank}(\sigma)$, which equals 1 if and only if both $\operatorname{rank}(\tau)$ and $\operatorname{rank}(\sigma)$ are equal to 1. Hence τ and σ are pure. ∎

Remark 4.20. One might object to the conclusion of Corollary 4.19 and ask a more elementary question without referring to disintegrations. Namely, does there exist a mixed state $\xi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$ and a CPU map $\varphi : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathcal{M}_n(\mathbb{C})$ such that $\xi \circ \varphi$ is a pure state? The reason to ask such a question is that if its answer is no, then one does not even need a disintegration for it to be impossible to evolve a mixed state into a pure state. The following example addresses this. Let $\omega : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be *any* pure state and let $\xi : \mathcal{M}_n(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be *any* mixed state. Set $\varphi(A) := \omega(A)\mathbb{1}_n$, which is a CPU map satisfying $\omega = \xi \circ \varphi$. Note that this situation is described by the diagram
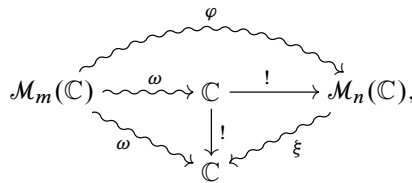
i.e., φ factors through $\mathbb{C}$. In this diagram, ! is the unique unital map from $\mathbb{C}$ into any C^*-algebra.

5. Disintegrations on finite-dimensional C^*-algebras

In the present section, we will extend Theorems 4.3 and 4.16 to the case of *-homomorphisms between arbitrary finite-dimensional C^*-algebras, which are all isomorphic to

finite direct sums of matrix algebras. We begin by analyzing CP maps between such direct sums in Lemma 5.11, their adjoints with respect to a generalized Hilbert–Schmidt inner product in Lemma 5.16, and the general form of states on direct sums in Lemma 5.17. After these preliminary results are established, we study the structure of Kraus decompositions of hypotheses in Lemmas 5.25 and 5.34. Proposition 5.39 provides a generalization of the "tracing out" operation for direct sums, i.e., the induced state via pullback from a *-homomorphism and a state on the target. After all this preparation, our main result, Theorem 5.45, is provided. Theorem 5.61 generalizes this disintegration theorem to arbitrary (unital) *-homomorphisms. But first, we recall the classical disintegration theorem.

Theorem 5.1. *Let* $(X, p) \xrightarrow{f} (Y, q)$ *be a probability-preserving function. Then the following facts hold.*

(i) *The assignment*

$$X \times Y \ni (x, y) \mapsto r_{xy} := \begin{cases} p_x \delta_{yf(x)}/q_y & \text{if } q_y > 0, \\ 1/|X| & \text{otherwise} \end{cases} \qquad (5.2)$$

defines a disintegration $r : Y \rightsquigarrow X$ *of* p *over* q *consistent with* f.

(ii) *The stochastic map* r *is the unique one up to a set of measure zero with respect to* q *satisfying* $f \circ r \underset{q}{=} \text{id}_Y$ *and* $r \circ q = p$; *i.e.,* $r \underset{q}{=} r'$ *for any other disintegration* $r' : Y \rightsquigarrow X$.

(iii) *Suppose that* f' *is another measure-preserving function satisfying* $f \underset{p}{=} f'$. *Let* r *be a disintegration of* f *and let* r' *be a disintegration of* f'. *Then* $r \underset{q}{=} r'$.

We will omit the details of this proof, which are neither difficult nor new. However, some of the lemmas used in proving it provide insight into the proof of our main theorem on non-commutative disintegrations in Theorem 5.45. These lemmas motivate the formula (5.2) and also assist in proving a.e. uniqueness. They show that a measure-preserving function $(X, p) \xrightarrow{f} (Y, q)$ is surjective onto a set of full measure, and they illustrate that a hypothesis $r : Y \rightsquigarrow X$ forces r_y to be supported on $f^{-1}(\{y\})$ for almost all $y \in Y$. This allows us to think of a disintegration more visually as follows (see Figure 1).[9] First, a probability space can be viewed as a finite number of water droplets, each of which has some volume (probability); the total volume is normalized to one. One can visualize a morphism $(X, p) \xrightarrow{f} (Y, q)$ as combining some of the water droplets, summing their volumes in the process. A hypothesis is a choice of physically splitting the water droplets back to the original set, but possibly with different volumes. A perfect splitting of the water droplets in which the volumes are reproduced exactly is an optimal hypothesis. From a topologist's point of view, a hypothesis is a stochastic section of f, which assigns a probability measure on the fiber (as opposed to a specific element) over each point that has non-zero q measure.

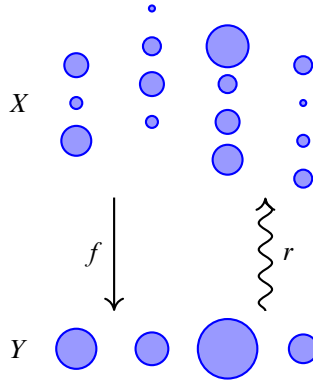[9]We learned this point of view from Gromov [22].

Figure 1. A probability-preserving function f and a disintegration r of f can be depicted in terms of combining water droplets via f in the vertical direction and splitting them back apart, also in the vertical direction, via r to reproduce the original configuration. Figure 1 is from [41].

Lemma 5.3. *Let $r : Y \rightsquigarrow X$ be a hypothesis for $(X, p) \xrightarrow{f} (Y, q)$. Then the probability measure r_y is supported on $f^{-1}(\{y\})$ for all $y \in Y \setminus N_q$.*

Lemma 5.4. *Let $(X, p) \xrightarrow{f} (Y, q)$ be a morphism in **FinProb**. Then, for each $y \in Y \setminus N_q$, there exists an $x \in X$ such that $f(x) = y$; i.e., f is surjective onto a set of full q-measure.*

Lemma 5.5. *Let $r : Y \rightsquigarrow X$ be a disintegration of $(X, p) \xrightarrow{f} (Y, q)$. Then*

$$r_{xf(x)} q_{f(x)} = p_x \quad \forall x \in X.$$

Notation 5.6. Throughout the rest of this section, let

$$\mathcal{A} := \mathcal{M}_{m_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{m_s}(\mathbb{C}) \quad \text{and} \quad \mathcal{B} := \mathcal{M}_{n_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{n_t}(\mathbb{C}) \tag{5.7}$$

denote direct sums of matrix algebras. An element $\vec{A} \in \mathcal{A}$ will be denoted by a column vector

$$\vec{A} \equiv \begin{pmatrix} A_1 \\ \vdots \\ A_s \end{pmatrix}$$

and similarly for elements of $\mathcal{B}$. The vector notation is often used for emphasis. An arbitrary linear map $\varphi : \mathcal{A} \rightsquigarrow \mathcal{B}$ will be written in matrix form as

$$\varphi \equiv \begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1s} \\ \vdots & & \vdots \\ \varphi_{t1} & \cdots & \varphi_{ts} \end{pmatrix}, \tag{5.8}$$

where $\varphi_{ji} : \mathcal{M}_{m_i}(\mathbb{C}) \rightsquigarrow \mathcal{M}_{n_j}(\mathbb{C})$ is a linear map for all i, j. The notation indicates the action of φ on $\vec{A}$ as

$$\varphi(\vec{A}) = \begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1s} \\ \vdots & & \vdots \\ \varphi_{t1} & \cdots & \varphi_{ts} \end{pmatrix} \begin{pmatrix} A_1 \\ \vdots \\ A_s \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{s} \varphi_{1i}(A_i) \\ \vdots \\ \sum_{i=1}^{s} \varphi_{ti}(A_i) \end{pmatrix}.$$

Let $(\mathcal{B}, \xi) \xrightarrow{F} (\mathcal{A}, \omega)$ be a state-preserving *-homomorphism defined by[10]

$$\mathcal{B} \ni \vec{B} \mapsto F\left(\begin{pmatrix} B_1 \\ \vdots \\ B_t \end{pmatrix}\right) := \begin{pmatrix} \mathrm{diag}(\overbrace{B_1, \ldots, B_1}^{c_{11} \text{ times}}, \ldots, \overbrace{B_t, \ldots, B_t}^{c_{1t} \text{ times}}) \\ \vdots \\ \mathrm{diag}(\underbrace{B_1, \ldots, B_1}_{c_{s1} \text{ times}}, \ldots, \underbrace{B_t, \ldots, B_t}_{c_{st} \text{ times}}) \end{pmatrix}, \qquad (5.9)$$

where the non-negative integer c_{ij} is called the *multiplicity* of F of the factor $\mathcal{M}_{n_j}(\mathbb{C})$ inside $\mathcal{M}_{m_i}(\mathbb{C})$ (cf. Fillmore [12, Sections 1.1.2 and 1.1.3]). In particular, the dimensions are related by the formula

$$m_i = \sum_{j=1}^{t} c_{ij} n_j \quad \forall i \in \{1, \ldots, s\}. \qquad (5.10)$$

Since F is linear, it also has a matrix representation

$$F \equiv \begin{pmatrix} F_{11} & \cdots & F_{1t} \\ \vdots & & \vdots \\ F_{s1} & \cdots & F_{st} \end{pmatrix}$$

with $F_{ij} : \mathcal{M}_{n_j}(\mathbb{C}) \to \mathcal{M}_{m_i}(\mathbb{C})$ a (not necessarily unital) *-homomorphism for all i, j.

Lemma 5.11. *Let* $\mathcal{A}$, $\mathcal{B}$, *and* φ *be as in* (5.7) *and* (5.8). *Then* φ *is CP if and only if* φ_{ji} *is CP for all* $i \in \{1, \ldots, s\}$ *and* $j \in \{1, \ldots, t\}$. *Furthermore,* φ *is unital if and only if*

$$\mathbb{1}_{n_j} = \sum_{i=1}^{s} \varphi_{ji}(\mathbb{1}_{m_i}) \quad \forall j \in \{1, \ldots, t\}. \qquad (5.12)$$

A Kraus decomposition of φ_{ji} in this case will be expressed as

$$\varphi_{ji} = \sum_{l_{ji}=1}^{m_i n_j} \mathrm{Ad}_{V_{ji;l_{ji}}}, \qquad (5.13)$$

---

[10]We will work with more general *-homomorphisms later, but we will see that all (unital) *-homomorphisms are unitarily equivalent to ones of this form. Hence, we do not lose much generality by focusing on these.

where the $V_{ji;l_{ji}} : \mathbb{C}^{m_i} \to \mathbb{C}^{n_j}$ are linear maps. This allows the unitality condition (5.12) to be expressed as

$$\mathbb{1}_{n_j} = \sum_{i=1}^{s} \sum_{l_{ji}=1}^{m_i n_j} V_{ji;l_{ji}} V_{ji;l_{ji}}^{\dagger} \quad \forall j \in \{1, \ldots, t\}. \tag{5.14}$$

The following facts are easy to check and are analogous to what happens in the usual matrix algebra case. We include them here for completeness.

Lemma 5.15. *Let* $\mathcal{A} := \mathcal{M}_{m_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{m_s}(\mathbb{C})$. *Then the assignment*

$$\mathcal{A} \times \mathcal{A} \ni (\vec{A}, \vec{A}') \mapsto \langle \vec{A}, \vec{A}' \rangle := \sum_{i=1}^{s} \mathrm{tr}(A_i^{\dagger} A_i')$$

defines an inner product on $\mathcal{A}$. *This is called the* Hilbert–Schmidt *(a.k.a.* Frobenius*) inner product on* $\mathcal{A}$.

Lemma 5.16. *Let* $\mathcal{A}$, $\mathcal{B}$, *and* φ *be as in* (5.7) *and* (5.8). *Then there exists a unique linear map* $\varphi^* : \mathcal{B} \rightsquigarrow \mathcal{A}$ *satisfying*

$$\langle \vec{B}, \varphi(\vec{A}) \rangle = \langle \varphi^*(\vec{B}), \vec{A} \rangle \quad \forall \vec{A} \in \mathcal{A}, \ \vec{B} \in \mathcal{B}.$$

The linear map φ^* *is called the* adjoint *of* φ. *Furthermore,*

$$\varphi^* = \begin{pmatrix} \varphi_{11}^* & \cdots & \varphi_{t1}^* \\ \vdots & & \vdots \\ \varphi_{1s}^* & \cdots & \varphi_{ts}^* \end{pmatrix},$$

where $\varphi_{ji}^* : \mathcal{M}_{n_j}(\mathbb{C}) \rightsquigarrow \mathcal{M}_{m_i}(\mathbb{C})$ *is the usual (Hilbert–Schmidt) adjoint of* φ_{ji}.[11] *In particular, if* φ *is a CP map where* φ_{ji} *has Kraus decomposition as in* (5.13)*, then*

$$\varphi_{ji}^* = \sum_{l_{ji}=1}^{m_i n_j} \mathrm{Ad}_{V_{ji;l_{ji}}^{\dagger}}.$$

Finally, φ *is CPU if and only if* φ^* *is CP and trace-preserving in the sense that*

$$\langle \vec{B}, 1_{\mathcal{B}} \rangle = \langle \varphi^*(\vec{B}), 1_{\mathcal{A}} \rangle \quad \forall \vec{B} \in \mathcal{B};$$

i.e.,

$$\sum_{j=1}^{t} \mathrm{tr}(B_j) = \sum_{i=1}^{s} \sum_{j=1}^{t} \mathrm{tr}\left(\varphi_{ji}^*(B_j)\right)$$

in terms of the components of φ^* *and* $\vec{B}$.

---

[11]φ_{ji}^* will be the notation used for the adjoint of φ_{ji} as opposed to the more precise $(\varphi_{ji})^*$. It is the adjoint of the ij-th entry of φ^*, which itself could be denoted by $(\varphi^*)_{ij}$. Hence, $(\varphi^*)_{ij} = \varphi_{ji}^* = (\varphi_{ji})^*$.

Lemma 5.17. *Let* $\xi : \mathcal{B} \rightsquigarrow \mathbb{C}$ *be a state with* $\mathcal{B}$ *as in* (5.7). *Then there exists unique non-negative real numbers* $q_1, \ldots, q_t$ *and (not necessarily unique) density matrices* $\sigma_1 \in \mathcal{M}_{n_1}(\mathbb{C}), \ldots, \sigma_t \in \mathcal{M}_{n_t}(\mathbb{C})$ *such that*

$$\sum_{j=1}^{t} q_j = 1 \quad and \quad \xi(\vec{B}) = \sum_{j=1}^{t} q_j \operatorname{tr}(\sigma_j B_j) \quad \forall \vec{B} \in \mathcal{B}. \tag{5.18}$$

Furthermore, for every j *such that* $q_j > 0$, *the density matrix* σ_j *is the unique one satisfying these conditions.*

Proof. Since ξ is a state, it is CPU. The adjoint $\xi^* : \mathbb{C} \rightsquigarrow \mathcal{M}_{n_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{n_t}(\mathbb{C})$ of ξ is CP and trace-preserving by Lemma 5.16. Let $\vec{\varsigma} := \xi^*(1)$, let $\varsigma_j \in \mathcal{M}_{n_j}(\mathbb{C})$ denote the j-th component of $\vec{\varsigma}$, and set $q_j := \operatorname{tr}(\varsigma_j)$. By the trace-preserving condition of ξ^*, the first equation in (5.18) holds. By the positivity of ξ^*, each ς_j is a non-negative matrix. If $q_j > 0$, set

$$\sigma_j := \frac{\varsigma_j}{q_j}.$$

Otherwise, if $q_j = 0$, then ς_j is the zero matrix. In this case, let σ_j be *any* density matrix. The conclusions of this lemma follow from these assignments. ∎

Notation 5.19. Due to Lemma 5.17, a state ξ as above might occasionally be denoted by $\xi \equiv \sum_{j=1}^{t} q_j \operatorname{tr}(\sigma_j \cdot)$, where $\xi(\vec{B})$ is understood to be given as in (5.18). Furthermore, the subset $N_q := \{j \in \{1, \ldots, t\} : q_j = 0\}$ will occasionally be used.

Lemma 5.20. *Using the same notation from Lemma 5.17, the support* P_ξ *of* ξ *is given by the vector of matrices whose* j-*th component is given by*

$$(P_\xi)_j = \begin{cases} P_{\xi_j} & if\ q_j > 0, \\ 0 & if\ q_j = 0, \end{cases}$$

where P_{ξ_j} *is the support of* σ_j *on* $\mathcal{M}_{n_j}(\mathbb{C})$.

Notation 5.21. Let $\mathcal{B}$ be as in (5.7). For each $j, k \in \{1, \ldots, t\}$, let

$$\iota_j : \mathcal{M}_{n_j}(\mathbb{C}) \hookrightarrow \mathcal{B} \quad and \quad \pi_k : \mathcal{B} \twoheadrightarrow \mathcal{M}_{n_k}(\mathbb{C}) \tag{5.22}$$

be the inclusion of the j-th factor and projection of the k-th factor, respectively.

Lemma 5.23. *Given* $\mathcal{A}, \mathcal{B}, F, \xi, P_\xi, (P_\xi)_k, \iota_j,$ *and* π_k *as in* (5.7), (5.9), (5.22), *and Lemma 5.20, a CPU map* $R : \mathcal{A} \rightsquigarrow \mathcal{B}$ *satisfies* $R \circ F \underset{\xi}{=} \operatorname{id}_{\mathcal{B}}$ *if and only if*

$$\mathcal{R}_{(P_\xi)_k} \circ \pi_k \circ R \circ F \circ \iota_j = \begin{cases} \mathcal{R}_{(P_\xi)_k} & if\ k = j, \\ 0 & if\ k \neq j, \end{cases} \tag{5.24}$$

for all $j, k \in \{1, \ldots, t\}$. *Here,* $\mathcal{R}_{(P_\xi)_k}$ *is the right-multiplication map defined by*

$$\mathcal{R}_{(P_\xi)_k}(B_k) := B_k (P_\xi)_k \quad \forall B_k \in \mathcal{M}_{n_k}(\mathbb{C}).$$
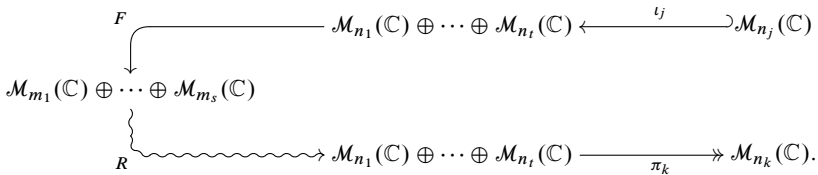
Proof. The condition $R \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$ holds if and only if (cf. Lemma 2.20) $\mathcal{R}_{P_\xi} \circ R \circ F = \mathcal{R}_{P_\xi}$, which holds if and only if $\mathcal{R}_{P_\xi} \circ R \circ F \circ \iota_j = \mathcal{R}_{P_\xi} \circ \iota_j$ for all $j \in \{1, \ldots, t\}$. Finally, this is equivalent to

$$
\begin{array}{ccc}
\overbrace{\pi_k \circ \mathcal{R}_{P_\xi} \circ R \circ F \circ \iota_j} & \pi_k \circ \mathcal{R}_{P_\xi} \circ \iota_j & \\
\Big\| & \Big\| & \forall j, k \in \{1, \ldots, t\}, \\
\mathcal{R}_{(P_\xi)_k} \circ \pi_k \circ R \circ F \circ \iota_j & \mathcal{R}_{(P_\xi)_k} \circ \pi_k \circ \iota_j &
\end{array}
$$

which is equivalent to the claim (5.24) because

$$
\pi_k \circ \iota_j = \begin{cases} \mathrm{id}_{\mathcal{M}_{n_k}(\mathbb{C})} & \text{if } k = j, \\ 0 & \text{if } k \neq j, \end{cases}
$$

$\blacksquare$

It may be helpful to visualize the map $\pi_k \circ R \circ F \circ \iota_j \equiv (R \circ F)_{kj}$ as the following composite of CP (not necessarily unital) maps:

$$
\begin{array}{ccc}
F & \mathcal{M}_{n_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{n_t}(\mathbb{C}) \xleftarrow{\quad \iota_j \quad} \mathcal{M}_{n_j}(\mathbb{C}) \\
\downarrow & \\
\mathcal{M}_{m_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{m_s}(\mathbb{C}) & \\
\Big\{ & \\
R & \mathcal{M}_{n_1}(\mathbb{C}) \oplus \cdots \oplus \mathcal{M}_{n_t}(\mathbb{C}) \xrightarrow{\quad \pi_k \quad} \mathcal{M}_{n_k}(\mathbb{C}).
\end{array}
$$

The following is an analogue of Lemma 2.23 to direct sums of matrix algebras.

Lemma 5.25. *Using the same notation as in Lemma 5.23 and assuming $R \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$, write*

$$
R = \begin{pmatrix} R_{11} & \cdots & R_{1s} \\ \vdots & & \vdots \\ R_{t1} & \cdots & R_{ts} \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} F_{11} & \cdots & F_{1t} \\ \vdots & & \vdots \\ F_{s1} & \cdots & F_{st} \end{pmatrix},
$$

where the completely positive (not necessarily unital) maps $R_{ki} : \mathcal{M}_{m_i}(\mathbb{C}) \rightsquigarrow \mathcal{M}_{n_k}(\mathbb{C})$ and $F_{ij} : \mathcal{M}_{n_j}(\mathbb{C}) \to \mathcal{M}_{m_i}(\mathbb{C})$ have Kraus decompositions,[12]

$$
R_{ki} = \sum_{\beta_{ki}=1}^{n_k m_i} \mathrm{Ad}_{R_{ki;\beta_{ki}}} \quad \text{and} \quad F_{ij} = \sum_{\gamma_{ij}=1}^{m_i n_j} \mathrm{Ad}_{F_{ij;\gamma_{ij}}} \tag{5.26}
$$

with $R_{ki;\beta_{ki}} : \mathbb{C}^{m_i} \to \mathbb{C}^{n_k}$ and $F_{ij;\gamma_{ij}} : \mathbb{C}^{n_j} \to \mathbb{C}^{m_i}$ linear maps.

---

[12]We will see in the text surrounding (5.32) that there exists a Kraus decomposition of F_{ij} such that the index γ_{ij} runs from 1 to c_{ij} instead of $m_i n_j = \sum_{k=1}^t c_{ik} n_k n_j$. This just means that $F_{ij;\gamma_{ij}}$ is zero when γ_{ij} exceeds c_{ij}.

(i)   *For each index $k \in \{1, \ldots, t\} \setminus N_q$, there exist a collection of complex numbers $\{\alpha_{k;i,\beta_{ki},\gamma_{ik}}\}$, indexed by $\gamma_{ik} \in \{1, \ldots, c_{ik}\}$, $\beta_{ki} \in \{1, \ldots, n_k m_i\}$, and $i \in \{1, \ldots, s\}$, such that*

$$P_{\xi_k} R_{ki;\beta_{ki}} F_{ik;\gamma_{ik}} = \alpha_{k;i,\beta_{ki},\gamma_{ik}} P_{\xi_k} \tag{5.27}$$

for all $\beta_{ki} \in \{1, \ldots, n_k m_i\}$, $\gamma_{ik} \in \{1, \ldots, n_k m_i\}$, $i \in \{1, \ldots, s\}$, and

$$\sum_{i=1}^{s} \sum_{\beta_{ki}=1}^{n_k m_i} \sum_{\gamma_{ik}=1}^{m_i n_k} |\alpha_{k;i,\beta_{ki},\gamma_{ik}}|^2 = 1. \tag{5.28}$$

(ii) *For every pair $j \in \{1, \ldots, t\}$ and $k \in \{1, \ldots, t\} \setminus N_q$ with $j \neq k$,*

$$P_{\xi_k} R_{ki,\beta_{ki}} F_{ij,\gamma_{ij}} = 0$$

for all $\beta_{ki} \in \{1, \ldots, n_k m_i\}$, $\gamma_{ij} \in \{1, \ldots, n_j m_i\}$, $i \in \{1, \ldots, s\}$.

Proof. Computing $\pi_k \circ R \circ F \circ \iota_j$ for $j, k \in \{1, \ldots, t\}$ gives

$$\pi_k \circ R \circ F \circ \iota_j = \sum_{i=1}^{s} R_{ki} \circ F_{ij} = \sum_{i=1}^{s} \sum_{\beta_{ki}=1}^{n_k m_i} \sum_{\gamma_{ij}=1}^{m_i n_j} \mathrm{Ad}_{R_{ki;\beta_{ki}} F_{ij;\gamma_{ij}}}. \tag{5.29}$$

Suppose that $q_k > 0$. Then, (5.24) becomes

$$\mathcal{R}_{P_{\xi_k}} \circ \pi_k \circ R \circ F \circ \iota_j = \begin{cases} \mathcal{R}_{P_{\xi_k}} & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases} \tag{5.30}$$

(i)   In the case $k = j$, (5.30) entails

$$\mathrm{Ad}_{P_{\xi_k}} \circ \pi_k \circ R \circ F \circ \iota_k = \mathrm{Ad}_{P_{\xi_k}}$$

upon multiplying by P_{ξ_k} on the left. Combining this with (5.29) and Lemma 2.21 (by following a similar proof to that of Lemma 2.23), there exist complex numbers $\alpha_{k;i,\beta_{ki},\gamma_{ik}}$ satisfying (5.27) and (5.28).

(ii) In the case $k \neq j$, (5.30) becomes

$$\mathrm{Ad}_{P_{\xi_k}} \circ \pi_k \circ R \circ F \circ \iota_j = 0$$

upon multiplying by P_{ξ_k} on the left. This implies that

$$P_{\xi_k} R_{ki;\beta_{ki}} F_{ij;\gamma_{ij}} = 0$$

by (5.29) and Lemma 2.21.   ∎

At this point, it is helpful to make the conclusions of Lemma 5.25 even more explicit by further explicating $P_{\xi_k} R_{ki;\beta_{ki}} F_{ij;\gamma_{ij}}$. The Kraus operator $R_{ki;\beta_{ki}} : \mathbb{C}^{m_i} \to \mathbb{C}^{n_k}$ can be

partitioned into block sums of matrices based on the multiplicity of F in the following way:

$$R_{ki;\beta_{ki}} = \left[\underbrace{\overbrace{\underbrace{V_{ki;\beta_{ki};11}}_{n_k \times n_1 \text{ matrix}} \cdots \underbrace{V_{ki;\beta_{ki};1c_{i1}}}_{n_k \times n_1 \text{ matrix}}}^{n_k \times (c_{i1}n_1) \text{ matrix}} \cdots \overbrace{\underbrace{V_{ki;\beta_{ki};t1}}_{n_k \times n_t \text{ matrix}} \cdots \underbrace{V_{ki;\beta_{ki};tc_{it}}}_{n_k \times n_t \text{ matrix}}}^{n_k \times (c_{it}n_t) \text{ matrix}} \right]$$

due to (5.10). Based on this partitioning, the unitality condition on R reads

$$\mathbb{1}_{n_k} = \sum_{i=1}^{s} \sum_{\beta_{ki}=1}^{m_i n_k} \sum_{j=1}^{t} \sum_{\gamma_{ij}=1}^{c_{ij}} V_{ki;\beta_{ki};j\gamma_{ij}} V_{ki;\beta_{ki};j\gamma_{ij}}^{\dagger} \quad \forall k \in \{1, \dots, t\} \tag{5.31}$$

due to (5.14). Furthermore, the definition of F from (5.9) says that

$$F_{ij}(B_j) = \text{diag}(0, \dots, 0, \dots, B_j, \dots, B_j, \dots, 0, \dots, 0) \quad \forall B_j \in \mathcal{M}_{n_j}(\mathbb{C}).$$

This implies that the (adjoint of the) Kraus operators $F_{ij;\gamma_{ij}} : \mathbb{C}^{n_j} \to \mathbb{C}^{m_i}$ of

$$F_{ij} : \mathcal{M}_{n_j}(\mathbb{C}) \to \mathcal{M}_{m_i}(\mathbb{C})$$

have the partitioned form

$$F_{ij;\gamma_{ij}}^{\dagger}$$
$$= \left[\overbrace{\underbrace{0}_{n_j \times n_1} \cdots \underbrace{0}_{n_j \times n_1}}^{n_j \times (c_{i1}n_1)} \cdots \overbrace{\underbrace{0}_{n_j \times n_j} \cdots \underbrace{\mathbb{1}_{n_j}}_{} \cdots \underbrace{0}_{n_j \times n_j}}^{n_j \times (c_{ij}n_j)} \cdots \overbrace{\underbrace{0}_{n_j \times n_t} \cdots \underbrace{0}_{n_j \times n_t}}^{n_j \times (c_{it}n_t)} \right], \tag{5.32}$$

where the identity matrix $\mathbb{1}_{n_j}$ is in the γ_{ij}-th $n_j \times n_j$ subblock inside the $n_j \times (c_{ij}n_j)$ block indicated (and all other entries are 0). In particular, the index γ_{ij} runs from 1 to c_{ij} (as opposed to $m_i n_j$). Therefore, the product $R_{ki;\beta_{ki}} F_{ij;\gamma_{ij}}$ is

$$R_{ki;\beta_{ki}} F_{ij;\gamma_{ij}} = V_{ki;\beta_{ki};j\gamma_{ij}}, \tag{5.33}$$

which is an $n_k \times n_j$ matrix. The following result is a generalization of equation (4.8) to the direct sum case.

Lemma 5.34. *Under the same assumptions as in Lemma 5.25, for every $k \in \{1, \dots, t\} \setminus N_q$, there exist a collection of complex numbers $\{\alpha_{k;i,\beta_{ki},\gamma_{ik}}\}$, indexed by $\gamma_{ik} \in \{1, \dots, c_{ik}\}$, $\beta_{ki} \in \{1, \dots, n_k m_i\}$, and $i \in \{1, \dots, s\}$, such that*

$$R_{ki;\beta_{ki}}$$
$$= \left[\overbrace{\underbrace{0}_{n_k \times n_1} \cdots \underbrace{0}_{n_k \times n_1}}^{n_k \times (c_{i1}n_1)} \cdots \overbrace{\underbrace{\alpha_{k;i,\beta_{ki},1}\mathbb{1}_{n_k}}_{n_k \times n_k} \cdots \underbrace{\alpha_{k;i,\beta_{ki},c_{ik}}\mathbb{1}_{n_k}}_{n_k \times n_k}}^{n_k \times (c_{ik}n_k)} \cdots \overbrace{\underbrace{0}_{n_k \times n_t} \cdots \underbrace{0}_{n_k \times n_t}}^{n_k \times (c_{it}n_t)} \right], \tag{5.35}$$

for all $\beta_{ki} \in \{1, \ldots, n_k m_i\}, i \in \{1, \ldots, s\}$, and

$$\sum_{i=1}^{s} \sum_{\beta_{ki}=1}^{n_k m_i} \sum_{\gamma_{ik}=1}^{c_{ik}} |\alpha_{k;i,\beta_{ki},\gamma_{ik}}|^2 = 1. \tag{5.36}$$

Proof. In analogy to the proof of Theorem 2.32, for every $k \in \{1, \ldots, t\} \setminus N_q$, let $\mathcal{E}_k$ be the pre-Hilbert $\mathcal{M}_{n_k}(\mathbb{C})$-module consisting of vectors of $n_k \times n_k$ matrices whose vector components are labeled by the triple of indices γ_{ik}, β_{ki}, i. Let $\vec{V}_k$ be the vector whose vector components are the $n_k \times n_k$ matrices $V_{ki;\beta_{ki};k\gamma_{ik}}$. The first case of Lemma 5.25 implies that there exists a vector $\vec{\alpha}_k \in \mathcal{E}_k$ whose vector components are constant multiples of the identity matrix satisfying

$$\vec{V}_k = P_{\xi_k} \vec{\alpha}_k + \vec{V}_k^{\text{bl}} + \vec{V}_k^{\text{br}},$$

where

$$\vec{V}_k^{\text{bl}} := P_{\xi_k}^{\perp} \vec{V}_k P_{\xi_k}, \quad \vec{V}_k^{\text{br}} := P_{\xi_k}^{\perp} \vec{V}_k P_{\xi_k}^{\perp}, \quad \text{and} \quad \langle\!\langle \vec{\alpha}_k, \vec{\alpha}_k \rangle\!\rangle = \mathbb{1}_{n_k}.$$

Similarly, for every pair $(k, j) \in \{1, \ldots, t\} \times \{1, \ldots, t\}$ such that $j \neq k$ and such that $q_k > 0$, let $\mathcal{E}_{kj}$ be the pre-Hilbert $\mathcal{M}_{n_k}(\mathbb{C})$-module consisting of vectors of $n_k \times n_j$ matrices whose vector components are labeled by the triple of indices γ_{ij}, β_{ki}, i. Let $\vec{V}_{kj}$ be the vector of the $n_k \times n_j$ matrices whose components are given by $V_{ki;\beta_{ki};j\gamma_{ij}}$. The second case of Lemma 5.25 implies that $\vec{V}_{kj} = \vec{V}_{kj}^{\text{bl}} + \vec{V}_{kj}^{\text{br}}$, where

$$\vec{V}_{kj}^{\text{bl}} := P_{\xi_k}^{\perp} \vec{V}_{kj} P_{\xi_j} \quad \text{and} \quad \vec{V}_{kj}^{\text{br}} := P_{\xi_k}^{\perp} \vec{V}_{kj} P_{\xi_k}^{\perp}.$$

The equalities

$$\langle\!\langle P_{\xi_k} \vec{\alpha}_k, \vec{V}_k^{\text{br}} \rangle\!\rangle = \langle\!\langle \vec{V}_k^{\text{bl}}, \vec{V}_k^{\text{br}} \rangle\!\rangle = \langle\!\langle \vec{V}_{kj}^{\text{bl}}, \vec{V}_{kj}^{\text{br}} \rangle\!\rangle = 0$$

all follow immediately from the definitions. Unitality of R takes on the form

$$\mathbb{1}_{n_k} = \langle\!\langle \vec{V}_k, \vec{V}_k \rangle\!\rangle + \sum_{j=1, j \neq k}^{t} \langle\!\langle \vec{V}_{kj}, \vec{V}_{kj} \rangle\!\rangle \tag{5.37}$$

by (5.31). By expanding out (5.37) and multiplying on the right by P_{ξ_k}, completely similar arguments to those in the proof of Theorem 2.32, specifically the discussion surrounding equations (2.35) through (2.37), prove that $\langle\!\langle \vec{V}_k^{\text{bl}}, \vec{\alpha}_k \rangle\!\rangle = 0$. Hence, the unitality condition (5.37) simplifies to

$$P_{\xi_k}^{\perp} = \langle\!\langle \vec{V}_k^{\text{bl}}, \vec{V}_k^{\text{bl}} \rangle\!\rangle + \langle\!\langle \vec{V}_k^{\text{br}}, \vec{V}_k^{\text{br}} \rangle\!\rangle + \sum_{j=1, j \neq k}^{t} \left(\langle\!\langle \vec{V}_{kj}^{\text{bl}}, \vec{V}_{kj}^{\text{bl}} \rangle\!\rangle + \langle\!\langle \vec{V}_{kj}^{\text{br}}, \vec{V}_{kj}^{\text{br}} \rangle\!\rangle \right) \tag{5.38}$$

analogously to (2.35). Now, computing $\pi_k \circ R \circ F \circ \iota_j$ in terms of the pre-Hilbert module inner product gives

$$(\pi_k \circ R \circ F \circ \iota_j)(A_j) = \langle\!\langle \vec{V}_{kj} A_j, \vec{V}_{kj} \rangle\!\rangle \quad \forall A_j \in \mathcal{M}_{n_j}(\mathbb{C})$$

for all k, j (when $j = k$, remove one of the indices from $\vec{V}_{kk}$ by (5.33)). When $j = k$, multiplying this equation on the right by P_{ξ_k} (which equals $(P_{\xi})_k$ since $q_k > 0$) and combining

this with Lemma 5.23 gives $\langle\langle \vec{V}_k^{\text{br}}, \vec{\alpha}_k \rangle\rangle = P_{\xi_k}^{\perp}$ by following an argument exactly analogous to (2.40) and the text surrounding this equation. Similarly, combining this result with the Paschke–Cauchy–Schwarz inequality gives $P_{\xi_k}^{\perp} \leq \langle\langle \vec{V}_k^{\text{br}}, \vec{V}_k^{\text{br}} \rangle\rangle$. On the other hand, (5.38) says that $P_{\xi_k}^{\perp} \geq \langle\langle \vec{V}_k^{\text{br}}, \vec{V}_k^{\text{br}} \rangle\rangle$. Therefore, following analogous lines of thought to those from (2.42) to (2.43) gives

$$\vec{V}_k^{\text{br}} = P_{\xi_k}^{\perp}\vec{\alpha}_k, \quad \vec{V}_k^{\text{bl}} = \vec{0}, \quad \vec{V}_{kj}^{\text{bl}} = \vec{0}, \quad \text{and} \quad \vec{V}_{kj}^{\text{br}} = \vec{0}.$$

Therefore, $\vec{V}_k = \vec{\alpha}_k$ and $\vec{V}_{kj} = \vec{0}$. Expanding out the vector entries coming from the definitions of $\mathcal{E}_k$ and $\mathcal{E}_{kj}$ completes the proof. ∎

Given a state-preserving *-homomorphism $(\mathcal{B}, \xi) \xrightarrow{F} (\mathcal{A}, \omega)$, it may be useful to know how the density matrices associated to ξ and ω are related. The following fact describes this relationship. It is a generalization of the "tracing out degrees of freedom" method in quantum theory.

Proposition 5.39. *Let $\mathcal{A}$, $\mathcal{B}$, F, ω, and ξ be as in Notation 5.6, and let*

$$\omega \equiv \sum_{i=1}^{s} p_i \operatorname{tr}(\rho_i \cdot) \quad \text{and} \quad \xi \equiv \sum_{j=1}^{t} q_j \operatorname{tr}(\sigma_j \cdot) \tag{5.40}$$

be decompositions of the states ω and ξ as described in Lemma 5.17. Then the following facts hold.

(i) *For each $i \in \{1, \ldots, s\}$, there exists a $j \in \{1, \ldots, t\}$ such that $c_{ij} > 0$.*

(ii) *If there exists a $j \in \{1, \ldots, s\}$ such that $c_{ij} = 0$ for all $i \in \{1, \ldots, s\}$, then $q_j = 0$.*

(iii) *Finally,*

$$q_j \sigma_j = \sum_{i=1}^{s} \sum_{\gamma_{ij}=1}^{c_{ij}} p_i \rho_{i;jj;\gamma_{ij}\gamma_{ij}} \quad \forall j \in \{1, \ldots, t\}, \tag{5.41}$$

where $\rho_{i;jj;\gamma_{ij}\gamma_{ij}}$ is the $n_j \times n_j$ matrix obtained from ρ_i in the following way. Since $m_i = \sum_{k=1}^{t} c_{ik} n_k$, each $m_i \times m_i$ matrix ρ_i has a block matrix decomposition

$$\rho_i = \begin{bmatrix} \rho_{i;11} & \cdots & \rho_{i;1t} \\ \vdots & & \vdots \\ \rho_{i;t1} & \cdots & \rho_{i;tt} \end{bmatrix}, \tag{5.42}$$

where $\rho_{i;jk}$ is a $(c_{ij}n_j) \times (c_{ik}n_k)$ matrix. This matrix further breaks up into subblocks

$$\rho_{i;jk} = \begin{bmatrix} \rho_{i;jk;11} & \cdots & \rho_{i;jk;1c_{ik}} \\ \vdots & & \vdots \\ \rho_{i;jk;c_{ij}1} & \cdots & \rho_{i;jk;c_{ij}c_{ik}} \end{bmatrix},$$

where $\rho_{i;jk;\gamma_{ij}\gamma_{ik}}$ is an $n_j \times n_k$ matrix.

Remark 5.43. The contrapositive of part (ii) of Proposition 5.39 will be used occasionally in certain technical points later. It states that if $q_j > 0$, there exists at least one $i \in \{1, \ldots, s\}$ such that $c_{ij} > 0$. In other words, F is injective a.e. This should be compared to Lemma 5.4. Furthermore, using partial traces, Equation (5.41) becomes

$$q_j \sigma_j = \sum_{i=1}^{s} p_i \, \mathrm{tr}_{\mathcal{M}_{c_{ij}}(\mathbb{C})}(\rho_{i;jj}),$$

where ρ_i is decomposed as in (5.42).

Proof of Proposition 5.39. (i) Since $m_i > 0$ and $m_i = \sum_{j=1}^{t} c_{ij} n_j$, there must exist a non-zero c_{ij} for some $j \in \{1, \ldots, t\}$.

(ii) Suppose that there exists a $j \in \{1, \ldots, s\}$ such that $c_{ij} = 0$ for all $i \in \{1, \ldots, s\}$. Then $F_{ij}(\mathbb{1}_{n_j}) = 0$ for all $i \in \{1, \ldots, s\}$. Since $\omega \circ F = \xi$, this shows that $\xi(\mathbb{1}_{n_j}) = 0$. But $\xi(\mathbb{1}_{n_j}) = q_j \, \mathrm{tr}(\sigma_j) = q_j$ so that $q_j = 0$.

(iii) This follows from taking the adjoint of the equation $\omega \circ F = \xi$, which gives

$$F^* \circ \omega^* = \xi^* \implies F^*(\omega^*(1)) = \xi^*(1).$$

Expanding out these expressions and extracting the j-th term gives

$$q_j \sigma_j = \sum_{i=1}^{s} p_i F_{ij}^*(\rho_i) \quad \forall j \in \{1, \ldots, t\}. \tag{5.44}$$

Applying (5.26) and (5.32) gives the desired result. ∎

A consequence of Lemma 5.34 is the following fact regarding the existence and uniqueness of disintegrations on finite-dimensional C^*-algebras. It is a generalization of Theorem 4.3 to direct sums of matrix algebras and is the main theorem of the present work.

Theorem 5.45. *Let $\mathcal{A}$, $\mathcal{B}$, F, ω, and ξ be as in Notation 5.6 and Proposition 5.39.*

(i) *A disintegration R of ω over ξ consistent with F exists if and only if for each $i \in \{1, \ldots, s\}$ and $j \in \{1, \ldots, t\}$ there exist non-negative matrices $\tau_{ji} \in \mathcal{M}_{c_{ij}}(\mathbb{C})$ such that[13]*

$$\mathrm{tr}\left(\sum_{i=1}^{s} \tau_{ji}\right) = 1 \quad \forall j \in \{1, \ldots, t\} \setminus N_q \tag{5.46}$$

and

$$p_i \rho_i = \mathrm{diag}(q_1 \tau_{1i} \otimes \sigma_1, \ldots, q_t \tau_{ti} \otimes \sigma_t) \quad \forall i \in \{1, \ldots, s\}. \tag{5.47}$$

(ii) *Furthermore, if R' is another disintegration of ω over ξ consistent with F, then $R' \underset{\xi}{=} R$ and*

$$R'_{ji} = R_{ji} \quad \forall i \in \{1, \ldots, s\} \quad \forall j \in \{1, \ldots, t\} \setminus N_q. \tag{5.48}$$

---

[13] N_q was introduced in Notation 5.19.

(iii)   *Finally, if such a disintegration R exists, a formula for the ji-th component of the disintegration is given by*

$$R_{ji}(A_i) = \operatorname{tr}_{\mathcal{M}_{c_{ij}}(\mathbb{C})}\left((\tau_{ji} \otimes \mathbb{1}_{n_j})A_{i;jj}\right) \tag{5.49}$$

for all $j \in \{1, \ldots, t\} \setminus N_q$ and for all $i \in \{1, \ldots, s\}$. Here, $A_{i;jj}$ is uniquely defined by the decomposition as a $t \times t$ matrix

$$A_i \equiv \begin{bmatrix} A_{i;11} & \cdots & A_{i;1t} \\ \vdots & & \vdots \\ A_{i;t1} & \cdots & A_{i;tt} \end{bmatrix}, \tag{5.50}$$

where the kl-th subblock, $A_{i;kl}$, is a $(c_{ik}n_k) \times (c_{il}n_l)$ matrix.

Proof. Proving the first item will provide proofs of the subsequent claims.

($\Rightarrow$) Suppose that a disintegration R exists. The condition $\xi \circ R = \omega$ is equivalent to $R^*(\xi^*(1)) = \omega^*(1)$ by Lemma 5.16. Hence, using the notation from (5.40), this equation gives

$$R^*\left(\xi^*(1)\right) \equiv \begin{pmatrix} R_{11}^* & \cdots & R_{t1}^* \\ \vdots & & \vdots \\ R_{1s}^* & \cdots & R_{ts}^* \end{pmatrix} \begin{pmatrix} q_1\sigma_1 \\ \vdots \\ q_t\sigma_t \end{pmatrix} = \begin{pmatrix} p_1\rho_1 \\ \vdots \\ p_s\rho_s \end{pmatrix} \equiv \omega^*(1),$$

which is equivalent to

$$p_i\rho_i = \sum_{j=1}^{t} q_j R_{ji}^*(\sigma_j) \quad \forall i \in \{1, \ldots, s\}. \tag{5.51}$$

To compute $R_{ji}^*(\sigma_j)$, we can follow an analogous computation to that from (4.10). First, when $q_j > 0$, we obtain

$$R_{ji}^*(\sigma_j) = \sum_{\beta_{ji}=1}^{n_j m_i} \operatorname{Ad}_{R_{ji;\beta_{ji}}^\dagger}(\sigma_j)$$

$$\overset{(5.35)}{=} \sum_{\beta_{ji}=1}^{n_j m_i} \begin{bmatrix} 0 \\ \overline{\alpha_{j;i,\beta_{ji},1}}\mathbb{1}_{n_j} \\ \vdots \\ \overline{\alpha_{j;i,\beta_{ji},c_{ij}}}\mathbb{1}_{n_j} \\ 0 \end{bmatrix} \sigma_j \begin{bmatrix} 0 & \alpha_{j;i,\beta_{ji},1}\mathbb{1}_{n_j} & \cdots & \alpha_{j;i,\beta_{ji},c_{ij}}\mathbb{1}_{n_j} & 0 \end{bmatrix},$$

where the top 0 block in the left matrix is a $(\sum_{k=1}^{j-1} c_{ik}n_k) \times n_j$ matrix and the bottom 0 block in the left matrix is a $(\sum_{k=j+1}^{t} c_{ik}n_k) \times n_j$ matrix. Keeping track of these sizes,

we obtain

$$
R_{ji}^*(\sigma_j) = \sum_{\beta_{ji}=1}^{n_j m_i} \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & |\alpha_{j;i,\beta_{ji},1}|^2 \sigma_j & \cdots & \overline{\alpha_{j;i,\beta_{ji},1}}\alpha_{j;i,\beta_{ji},c_{ij}}\sigma_j & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \overline{\alpha_{j;i,\beta_{ji},c_{ij}}}\alpha_{j;i,\beta_{ji},1}\sigma_j & \cdots & |\alpha_{j;i,\beta_{ji},c_{ij}}|^2 \sigma_j & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (5.52)
$$

where the top-left 0 matrix is a $(\sum_{k=1}^{j-1} c_{ik}n_k) \times (\sum_{k=1}^{j-1} c_{ik}n_k)$ matrix and the bottom-right 0 matrix is a $(\sum_{k=j+1}^{t} c_{ik}n_k) \times (\sum_{k=j+1}^{t} c_{ik}n_k)$ matrix. Define the $c_{ij} \times c_{ij}$ matrix τ_{ji} to be

$$
\tau_{ji} := \sum_{\beta_{ji}=1}^{n_j m_i} \begin{bmatrix} |\alpha_{j;i,\beta_{ji},1}|^2 & \cdots & \overline{\alpha_{j;i,\beta_{ji},1}}\alpha_{j;i,\beta_{ji},c_{ij}} \\ \vdots & & \vdots \\ \overline{\alpha_{j;i,\beta_{ji},c_{ij}}}\alpha_{j;i,\beta_{ji},1} & \cdots & |\alpha_{j;i,\beta_{ji},c_{ij}}|^2 \end{bmatrix}
$$

so that the $\gamma_{ij}\eta_{ij}$-th entry of τ_{ji} is given by

$$
\tau_{ji;\gamma_{ij}\eta_{ij}} = \sum_{\beta_{ji}=1}^{n_j m_i} \overline{\alpha_{j;i;\beta_{ji},\gamma_{ij}}}\alpha_{j;i,\beta_{ji},\eta_{ij}}. \tag{5.53}
$$

Notice that τ_{ji} is defined only when $c_{ij} > 0$ and when $q_j > 0$. Furthermore, when it is defined, τ_{ji} is a non-negative matrix and

$$
\sum_{i=1}^{s} \mathrm{tr}(\tau_{ji}) = \sum_{i=1}^{s} \sum_{\beta_{ji}=1}^{n_j m_i} \sum_{\gamma_{ij}=1}^{c_{ij}} |\alpha_{j;i,\beta_{ji},\gamma_{ij}}|^2 \overset{(5.36)}{=} 1, \tag{5.54}
$$

which shows that $\sum_{i=1}^{s}\tau_{ji}$ is a density matrix (again, when $q_j > 0$). The sum in (5.54) is guaranteed to have at least one term due to Remark 5.43. Second, when $q_j = 0$, then $q_j R_{ji}^*(\sigma_j) = 0$ so that this term does not contribute to the sum in (5.51). Therefore, in this case, τ_{ji} can be chosen to be an arbitrary non-negative matrix provided that $c_{ij} > 0$. If $c_{ij} = 0$, then τ_{ji} does not exist and any expression involving such a τ_{ji} should be excluded. Then,

$$
p_i \rho_i = \sum_{j=1}^{t} q_j R_{ji}^*(\sigma_j) \overset{(5.52)}{=} \begin{bmatrix} q_1 \tau_{1i} \otimes \sigma_1 & & 0 \\ & \ddots & \\ 0 & & q_t \tau_{ti} \otimes \sigma_t \end{bmatrix} \quad \forall i \in \{1,\dots,s\}.
$$

Note that this sum after the first equality is not empty by part (i) of Proposition 5.39.

($\Longleftarrow$) For the converse, suppose the non-negative matrices $\tau_{ji} \in \mathcal{M}_{c_{ij}}(\mathbb{C})$ satisfying (5.46) and (5.47) exist. Denote the $\gamma_{ij}\eta_{ij}$-th entry of τ_{ji} by $\tau_{ji;\gamma_{ij}\eta_{ij}}$. For each pair of indices $i \in \{1,\dots,s\}$ and $j \in \{1,\dots,t\}$, define $R'_{ji} : \mathcal{M}_{m_i}(\mathbb{C}) \rightsquigarrow \mathcal{M}_{n_j}(\mathbb{C})$ in the following way. Write an element $A_i \in \mathcal{M}_{m_i}(\mathbb{C})$ as in (5.50). Then, write $A_{i;kl}$ as a $c_{ik} \times c_{il}$

matrix consisting of $n_k \times n_l$ matrices indexed as in $A_{i;kl;\gamma_{ik}\eta_{il}}$. Set[14]

$$R'_{ji}(A_i) := \begin{cases} \sum_{\gamma_{ij},\eta_{ij}=1}^{c_{ij}} \tau_{ji;\eta_{ij}\gamma_{ij}} A_{i;jj;\gamma_{ij}\eta_{ij}} & \text{if } q_j > 0, \\ \frac{1}{sm_i} \operatorname{tr}(A_i) \mathbb{1}_{n_j} & \text{if } q_j = 0. \end{cases} \tag{5.55}$$

A direct calculation shows that this formula equals (5.49) when $j \in \{1, \ldots, t\} \setminus N_q$. Set $R' : \mathcal{A} \rightsquigarrow \mathcal{B}$ to be the $t \times s$ matrix of linear maps whose ji-th entry is R'_{ji} from (5.55). Then R' is a disintegration of ω over ξ consistent with F. The proof of this is similar to the proof of Theorem 4.3, though one must keep track of indices more carefully. Unitality of R' follows from

$$\sum_{i=1}^{s} R'_{ji}(\mathbb{1}_{m_i}) = \sum_{i=1}^{s} \sum_{\gamma_{ij},\eta_{ij}=1}^{c_{ij}} \tau_{ji;\eta_{ji}\gamma_{ij}} \mathbb{1}_{m_i;jj;\gamma_{ij}\eta_{ij}}$$

$$= \sum_{i=1}^{s} \sum_{\gamma_{ij}=1}^{c_{ij}} \tau_{ji;\gamma_{ji}\gamma_{ij}} \mathbb{1}_{n_j}$$

$$= \sum_{i=1}^{s} \operatorname{tr}(\tau_{ji}) \mathbb{1}_{n_j} = \mathbb{1}_{n_j}$$

whenever $q_j > 0$ because $\sum_{i=1}^{s} \operatorname{tr}(\tau_{ji}) = 1$. When $q_j = 0$, one obtains

$$\sum_{i=1}^{s} R'_{ji}(\mathbb{1}_{m_i}) = \sum_{i=1}^{s} \frac{1}{sm_i} \operatorname{tr}(\mathbb{1}_{m_i}) \mathbb{1}_{n_j} = \mathbb{1}_{n_j}.$$

We will now show that $\pi_j \circ R' \circ F = \pi_j$ for all j satisfying $q_j > 0$. First, note that

$$(R' \circ F)(\vec{B}) = \begin{pmatrix} R'_{11} & \cdots & R'_{1s} \\ \vdots & & \vdots \\ R'_{t1} & \cdots & R'_{ts} \end{pmatrix} \begin{pmatrix} \operatorname{diag}(\overbrace{B_1, \ldots, B_1}^{c_{11} \text{ times}}, \ldots, \overbrace{B_t, \ldots, B_t}^{c_{1t} \text{ times}}) \\ \vdots \\ \operatorname{diag}(\underbrace{B_1, \ldots, B_1}_{c_{s1} \text{ times}}, \ldots, \underbrace{B_t, \ldots, B_t}_{c_{st} \text{ times}}) \end{pmatrix}.$$

Focusing on the j-th term when $q_j > 0$, one obtains

$$\sum_{i=1}^{s} R'_{ji}\left(\operatorname{diag}(\underbrace{B_1, \ldots, B_1}_{c_{i1} \text{ times}}, \ldots, \underbrace{B_t, \ldots, B_t}_{c_{it} \text{ times}}) \right) = \sum_{i=1}^{s} \sum_{\gamma_{ij}=1}^{c_{ij}} \tau_{ji;\gamma_{ji}\gamma_{ij}} B_j$$

$$= \sum_{i=1}^{s} \operatorname{tr}(\tau_{ji}) B_j = B_j.$$

---

[14]Note that the swapping of the γ_{ij} and η_{ij} indices in equation (5.55) is not a typo. In addition, note that if $q_j > 0$ and $c_{ij} = 0$, then the sum in the top case is empty and gives, by definition of an empty sum, 0.

Although the equality $\pi_j \circ R' \circ F = \pi_j$ fails when $q_j = 0$, the equation $R' \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$ still holds. Furthermore, R' is state-preserving because

$$\omega(\vec{A}) = \sum_{i=1}^{s} \mathrm{tr}(p_i \rho_i A_i) \overset{(5.47)}{=} \sum_{i=1}^{s} \sum_{j=1}^{t} q_j \, \mathrm{tr}\left((\tau_{ji} \otimes \sigma_j) A_{i;jj}\right)$$

$$= \sum_{\substack{i=1 \\ j=1 \\ q_j>0}}^{s} \sum_{j=1}^{t} q_j \, \mathrm{tr}\left(\sum_{\gamma_{ij}=1, \eta_{ij}=1}^{c_{ij}} \tau_{ji;\eta_{ij}\gamma_{ij}} \sigma_j A_{i;jj;\gamma_{ij}\eta_{ij}}\right)$$

$$\overset{(5.55)}{=} \sum_{\substack{i=1 \\ q_j>0}}^{s} \sum_{j=1}^{t} q_j \, \mathrm{tr}\left(\sigma_j R'_{ji}(A_i)\right) = (\xi \circ R')(\vec{A})$$

for all $\vec{A} \in \mathcal{A}$. To show that R' is CP, it suffices to show that each R'_{ji} is CP by Lemma 5.11. This follows from the equality between the formulas (5.49) and (5.55) whenever $j \in \{1, \dots, t\} \setminus N_q$. The case when $j \in N_q$ gives a CP map as well since the trace in (5.55) is a CP map.

Finally, we prove the uniqueness condition (5.48) for disintegrations. The condition $R' \underset{\xi}{=} R$ is equivalent to

$$\sum_{i=1}^{s} R'_{ji}(A_i)(P_\xi)_j = \sum_{i=1}^{s} R_{ji}(A_i)(P_\xi)_j \quad \forall \vec{A} \in \mathcal{A}, \; j \in \{1, \dots, t\}.$$

When $q_j = 0$, this equality holds trivially because $(P_\xi)_j = 0$. When $q_j \neq 0$, Lemma 5.34 guarantees the existence of complex numbers $\{\alpha_{j;i,\beta_{ji},\gamma_{ij}}\}$ and Kraus operators $\{R_{ji;\beta_{ji}}\}$ for R_{ji} satisfying the conditions in the statement of that theorem. Therefore, by carefully working out the matrix operations, one obtains

$$R_{ji}(A_i) \overset{\text{Lemma 5.34}}{=} \sum_{\beta_{ji}=1}^{n_j m_i} \sum_{\gamma_{ij}=1}^{c_{ij}} \sum_{\eta_{ij}=1}^{c_{ij}} \alpha_{j;i;\beta_{ji},\gamma_{ij}} \overline{\alpha_{j;i;\beta_{ji},\eta_{ij}}} A_{i;jj;\gamma_{ij}\eta_{ij}}$$

$$\overset{(5.53)}{=} \sum_{\gamma_{ij}=1}^{c_{ij}} \sum_{\eta_{ij}=1}^{c_{ij}} \tau_{ji;\eta_{ij}\gamma_{ij}} A_{i;jj;\gamma_{ij}\eta_{ij}} \overset{(5.55)}{=} R'_{ji}(A_i) \quad \forall A_i \in \mathcal{M}_{m_i}(\mathbb{C}).$$

This concludes the proof of the theorem. ∎

Remark 5.56. By applying the trace to both sides of (5.47), one obtains

$$p_i = \sum_{j=1}^{t} q_j \, \mathrm{tr}(\tau_{ji}) \quad \forall i \in \{1, \dots, s\}.$$

An immediate corollary of Theorem 5.45 is the standard existence and uniqueness theorem of regular conditional probabilities from classical finite probability. We work this out in full detail as an example.

Example 5.57. Using the notation from Theorem 5.45, suppose that $m_i = 1$ and $n_j = 1$ for all i, j. Then, $\rho_i = 1 = \sigma_j$ for all i, j. Furthermore, since each $m_i = 1$, the multiplicity is drastically restricted since $m_i = \sum_{j=1}^{t} c_{ij} n_j$. By this equality, for each i, there exists a unique j such that $c_{ij} = 1$ and all other $c_{ik} = 0$. In other words, there exists a unique function $f : \{1, \ldots, s\} \to \{1, \ldots, t\}$ such that

$$c_{ij} = \delta_{f(i)j} \equiv \begin{cases} 1 & \text{if } f(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

This implies that

$$F = \begin{bmatrix} \delta_{f(1)1} & \cdots & \delta_{f(1)t} \\ \vdots & & \vdots \\ \delta_{f(s)1} & \cdots & \delta_{f(s)t} \end{bmatrix} \quad \text{and} \quad F^* = \begin{bmatrix} \delta_{f(1)1} & \cdots & \delta_{f(s)1} \\ \vdots & & \vdots \\ \delta_{f(1)t} & \cdots & \delta_{f(s)t} \end{bmatrix}. \tag{5.58}$$

Hence,

$$q_j \overset{(5.44)}{=} \sum_{i=1}^{s} p_i F_{ij}^*(1) \overset{(5.58)}{=} \sum_{i=1}^{s} p_i \delta_{f(i)j} = \sum_{i \in f^{-1}(j)} p_i, \tag{5.59}$$

which reproduces the probability-preserving condition $(X, p) \xrightarrow{f} (Y, q)$, assuming $X = \{1, \ldots, s\}$ and $Y = \{1, \ldots, t\}$. In what follows, we will construct, without any additional assumptions, non-negative matrices $\tau_{ji} \in \mathcal{M}_{c_{ij}}(\mathbb{C})$ satisfying (5.46) and (5.47) as well as a disintegration

$$R = \begin{bmatrix} r_{11} & \cdots & r_{1s} \\ \vdots & & \vdots \\ r_{t1} & \cdots & r_{ts} \end{bmatrix}.$$

This will prove that a disintegration automatically exists in this case. First note that if $j \neq f(i)$, the set $\mathcal{M}_{c_{ij}}(\mathbb{C})$ is just a singleton so that we only have a chance of constructing τ_{ji} when $j = f(i)$. In this case, $c_{if(i)} = 1$ and such a matrix will be a 1×1 matrix, i.e., a non-negative number. We set

$$\tau_{ji} := \begin{cases} p_i/q_j & \text{if } q_j > 0 \text{ and } j = f(i), \\ \# & \text{if } q_j = 0 \text{ and } j = f(i), \\ \text{DNE} & \text{if } c_{ij} = 0, \end{cases} \tag{5.60}$$

where $\#$ can be chosen to be any non-negative number. Note that if there exists a $j \in \{1, \ldots, t\}$ for which $c_{ij} = 0$ for all $i \in \{1, \ldots, s\}$, then $q_j = 0$ by part (ii) of Proposition 5.39. For such j, τ_{ij} cannot be defined for any $i \in \{1, \ldots, s\}$. Nevertheless,

$$\text{tr}\left(\sum_{i=1}^{s} \tau_{ji}\right) = \sum_{i=1}^{s} \tau_{ji} = \sum_{i \in f^{-1}(j)} \frac{p_i}{q_j} \overset{(5.59)}{=} 1 \quad \forall j \in \{1, \ldots, s\} \setminus N_q$$

proves (5.46). Secondly, because there exists a unique j for each i such that $c_{ij} = 1$,

$$\text{diag}(q_1 \tau_{1i} \otimes \sigma_1, \ldots, q_t \tau_{ti} \otimes \sigma_t)$$

$$= q_{f(i)} \tau_{f(i)i} = q_{f(i)} \begin{cases} p_i/q_{f(i)} & \text{if } q_{f(i)} > 0, \\ \# & \text{if } q_{f(i)} = 0, \end{cases}$$

$$= \begin{cases} p_i & \text{if } q_{f(i)} > 0, \\ 0 & \text{if } q_{f(i)} = 0, \end{cases} \quad \forall i \in \{1, \ldots, s\}.$$

Note that if $q_{f(i)} = 0$, then $p_i = 0$ by (5.59). Hence, this proves (5.47). Although this already proves a disintegration exists via Theorem 5.45, it is fruitful to construct it based on the proof of Theorem 5.45 and compare it to the classical disintegration from Theorem 5.1. Using the construction of a disintegration from (5.55), we get

$$r_{ji} \overset{(5.55)}{=} \begin{cases} \tau_{ji} & \text{if } q_j > 0 \text{ and } c_{ij} = 1, \\ 0 & \text{if } q_j > 0 \text{ and } c_{ij} = 0, \\ 1/s & \text{if } q_j = 0, \end{cases}$$

$$\overset{(5.60)}{=} \begin{cases} p_i/q_j & \text{if } q_j > 0 \text{ and } j = f(i), \\ 0 & \text{if } q_j > 0 \text{ and } j \neq f(i), \\ 1/s & \text{if } q_j = 0, \end{cases}$$

$$= \begin{cases} p_i \delta_{f(i)j}/q_j & \text{if } q_i > 0, \\ 1/s & \text{if } q_j = 0. \end{cases}$$

This reproduces formula (5.2) for an ordinary disintegration.

Finally, we end this section with a generalization of Theorems 5.45 and 4.16 by allowing for arbitrary (unital) *-homomorphisms $F : \mathcal{B} \to \mathcal{A}$.

Theorem 5.61. *Let $\mathcal{A}$, $\mathcal{B}$, F, ω, and ξ be as in Notation 5.6 and Proposition 5.39 except that F is now an arbitrary (unital) *-homomorphism, but not necessarily of the form (5.9). Then, a disintegration R exists if and only if there exist unitary matrices $U_i \in \mathcal{M}_{m_i}(\mathbb{C})$ and non-negative matrices $\tau_{ji} \in \mathcal{M}_{c_{ij}}(\mathbb{C})$ such that $\text{Ad}_{\tilde{U}} \circ F$ is of the form (5.9),*

$$\text{tr}\left(\sum_{i=1}^{s} \tau_{ji} \right) = 1 \quad \forall j \in \{1, \ldots, t\} \setminus N_q,$$

and

$$p_i U_i^{\dagger} \rho_i U_i = \text{diag}(q_1 \tau_{1i} \otimes \sigma_1, \ldots, q_t \tau_{ti} \otimes \sigma_t) \quad \forall i \in \{1, \ldots, s\}. \tag{5.62}$$

Furthermore, any two such disintegrations are unique ξ-a.e.

Proof. This follows from an argument analogous to the proof of Theorem 4.16. ∎

6. Example: Measurement in quantum mechanics

It is instructive to work out the following example due to its connection with measurement in quantum mechanics (it may be helpful at this point to review Example 2.11 for notation). We will also avoid using the results of Theorem 5.45 and will instead provide a self-contained analysis since this is simple enough in this special case. Fix $m \in \mathbb{N}$, let $A \in \mathcal{M}_m(\mathbb{C})$ be a self-adjoint matrix with spectrum $\sigma(A) \subseteq \mathbb{R}$, and let $\omega = \mathrm{tr}(\rho \cdot) : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be a state. The matrix A induces the *-homomorphism uniquely determined by

$$
\begin{array}{c}
\mathbb{C}^{\sigma(A)} \xrightarrow{F} \mathcal{M}_m(\mathbb{C}) \\
e_\lambda \mapsto P_\lambda,
\end{array}
\tag{6.1}
$$

where P_λ is the orthogonal projection onto the λ eigenspace associated to A. This pulls back the state ω to a probability measure q on $\sigma(A)$ whose evaluation on $\lambda \in \sigma(A)$ will be denoted by q_λ. The pullback state will be denoted by $\langle q, \cdot \rangle$, where $\langle \cdot, \cdot \rangle$ is the natural inner product on $\mathbb{C}^{\sigma(A)}$ induced by the basis $\{e_\lambda\}_{\lambda \in \sigma(A)}$. Physically, the number q_λ is interpreted as the probability that the state ω takes the value λ when the observable A is measured. If a disintegration $R : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}^{\sigma(A)}$ exists, it is uniquely determined by the collection of PU maps R_λ defined by

$$
\begin{array}{c}
\mathcal{M}_m(\mathbb{C}) \xrightarrow{R} \mathbb{C}^{\sigma(A)} \xrightarrow{\mathrm{ev}_\lambda} \mathbb{C} \\
A \xmapsto{\qquad R_\lambda \qquad} \langle e_\lambda, R(A) \rangle
\end{array}
$$

and indexed by $\lambda \in \sigma(A)$. Because these are states on $\mathcal{M}_m(\mathbb{C})$, they uniquely determine a density matrix $\rho_\lambda \in \mathcal{M}_m(\mathbb{C})$; i.e.,

$$
R_\lambda = \mathrm{tr}(\rho_\lambda \cdot), \quad \rho_\lambda \geq 0, \quad \mathrm{tr}(\rho_\lambda) = 1, \quad \text{and} \quad R = \sum_{\lambda \in \sigma(A)} \mathrm{ev}_\lambda^* \circ R_\lambda.
\tag{6.2}
$$

Because R must be state-preserving to be a disintegration, this entails

$$
\mathrm{tr}(\rho \cdot) = \langle q, R(\cdot) \rangle \overset{(6.2)}{=} \sum_{\lambda \in \sigma(A)} q_\lambda \mathrm{tr}(\rho_\lambda \cdot) \implies \rho = \sum_{\lambda \in \sigma(A)} q_\lambda \rho_\lambda.
\tag{6.3}
$$

The other condition for R to be a disintegration is $R \circ F \underset{\langle q, \cdot \rangle}{=} \mathrm{id}_{\mathbb{C}^{\sigma(A)}}$, which says that

$$
\sum_{\lambda \in \sigma(A)} b_\lambda e_\lambda - R\left(F\left(\sum_{\lambda \in \sigma(A)} b_\lambda e_\lambda \right) \right) \in \mathcal{N}_{\langle q, \cdot \rangle} \quad \forall \sum_{\lambda \in \sigma(A)} b_\lambda e_\lambda \in \mathbb{C}^{\sigma(A)}.
$$

Expanding this out, relabeling indices, and using part (i) of Lemma 2.20 gives

$$
\sum_{\lambda \in \sigma(A) \setminus N_q} b_\lambda e_\lambda = \sum_{\lambda \in \sigma(A) \setminus N_q} \left(\sum_{\mu \in \sigma(A)} b_\mu \mathrm{tr}(\rho_\lambda P_\mu) \right) e_\lambda.
$$

Linear independence of the e_λ then gives the constraints

$$b_\lambda = \sum_{\mu \in \sigma(A)} b_\mu \operatorname{tr}(\rho_\lambda P_\mu) \quad \forall \lambda \in \sigma(A) \setminus N_q.$$

Since the b's can be chosen arbitrarily and independently (indeed, set $b_\mu := \delta_{\mu\nu}$ for various ν), we conclude that

$$\operatorname{tr}(\rho_\lambda P_\mu) = \delta_{\lambda\mu} \quad \forall \mu, \lambda \in \sigma(A) \setminus N_q.$$

Since ρ_λ is a positive matrix, $\operatorname{tr}(\rho_\lambda P_\mu) = \operatorname{tr}(P_\mu \rho_\lambda P_\mu) = 0$ if and only if $P_\mu \rho_\lambda P_\mu = 0$ whenever $\mu \neq \lambda$. In what follows, we will prove that $\rho_\lambda = P_\lambda \rho_\lambda P_\lambda$. To see this, first let $\vec{u} \in \operatorname{Im}(P_\mu)$ and $\vec{v} \in \operatorname{Im}(P_\nu)$, where $\mu, \nu \in \sigma(A) \setminus \{\lambda\}$. Then $P_\mu, P_\nu \leq P_\lambda^\perp$ and

$$
\begin{aligned}
0 &\leq \langle \vec{u} + \vec{v}, P_\lambda^\perp \rho_\lambda P_\lambda^\perp (\vec{u} + \vec{v}) \rangle \\
&= \langle \vec{u}, P_\lambda^\perp \rho_\lambda P_\lambda^\perp \vec{u} \rangle + \langle \vec{u}, P_\lambda^\perp \rho_\lambda P_\lambda^\perp \vec{v} \rangle + \langle \vec{v}, P_\lambda^\perp \rho_\lambda P_\lambda^\perp \vec{u} \rangle + \langle \vec{v}, P_\lambda^\perp \rho_\lambda P_\lambda^\perp \vec{v} \rangle \\
&= \langle \vec{u}, P_\mu \rho_\lambda P_\mu \vec{u} \rangle + \langle \vec{u}, P_\mu \rho_\lambda P_\nu \vec{v} \rangle + \langle \vec{v}, P_\nu \rho_\lambda P_\mu \vec{u} \rangle + \langle \vec{v}, P_\nu \rho_\lambda P_\nu \vec{v} \rangle \\
&= 2 \Re e \langle \vec{u}, P_\mu \rho_\lambda P_\nu \vec{v} \rangle,
\end{aligned}
$$

where we have freely used the facts $P_\lambda^\perp \vec{u} = P_\mu \vec{u} = \vec{u}$ and $P_\lambda^\perp \vec{v} = P_\nu \vec{v} = \vec{v}$ together with the self-adjointness and orthogonality of these projections. Since $\vec{u}$ and $\vec{v}$ can be arbitrary, positivity of $P_\lambda^\perp \rho_\lambda P_\lambda^\perp$ guarantees that $P_\mu \rho_\lambda P_\nu = 0$ for all $\mu, \nu \in \sigma(A) \setminus \{\lambda\}$. So far, we have shown that $\rho_\lambda = P_\lambda \rho_\lambda P_\lambda + P_\lambda^\perp \rho_\lambda P_\lambda + P_\lambda \rho_\lambda P_\lambda^\perp$. What is left to show is that $P_\mu \rho_\lambda P_\lambda = 0$ for all $\mu \in \sigma(A) \setminus \{\lambda\}$ (which would imply that $P_\lambda \rho_\lambda P_\mu = 0$ by taking the adjoint). Now, let $\vec{u} \in \operatorname{Im}(P_\mu)$ and $\vec{v} \in \operatorname{Im}(P_\lambda)$, where $\mu \in \sigma(A) \setminus \{\lambda\}$. Positivity of ρ_λ gives

$$0 \leq \langle \vec{u} + \vec{v}, \rho_\lambda (\vec{u} + \vec{v}) \rangle = 2 \Re e \langle \vec{u}, P_\mu \rho_\lambda P_\lambda \vec{v} \rangle + \langle \vec{v}, \rho_\lambda \vec{v} \rangle$$

by a similar calculation and using the previous result. Since $\vec{u}$ can be chosen freely, it can be chosen so that the left term becomes arbitrarily negative unless $P_\mu \rho_\lambda P_\lambda = 0$. This concludes the argument that $\rho_\lambda = P_\lambda \rho_\lambda P_\lambda$. Thus, ρ_λ and $\rho_{\lambda'}$ have mutually orthogonal supports for $\lambda \neq \lambda'$ provided that $\lambda, \lambda' \in \sigma(A) \setminus N_q$. Hence, although we have no restrictions on ρ_λ when $\lambda \in N_q$, we still obtain

$$\rho \overset{(6.3)}{=} \sum_{\lambda \in \sigma(A)} q_\lambda \rho_\lambda = \sum_{\lambda \in \sigma(A) \setminus N_q} q_\lambda \rho_\lambda = \sum_{\lambda \in \sigma(A) \setminus N_q} q_\lambda P_\lambda \rho_\lambda P_\lambda, \tag{6.4}$$

which agrees with the result (5.47) with respect to a spectral basis, or more accurately (5.62), in this special case since $s = 1$ so that there is only one i index and $\sigma_j = 1$ for all j because σ_j is a 1×1 matrix. Thus, the τ_{ji} matrices reduce to the ρ_λ matrices. To make a more explicit connection to quantum information theory, we recall the definition of a Lüders projection, which is a model for the ensemble of the induced states of a system after a measurement has taken place [29].

Definition 6.5. Let $\rho \in \mathcal{M}_m(\mathbb{C})$ be a density matrix and let $A \in \mathcal{M}_m(\mathbb{C})$ be self-adjoint with spectrum $\sigma(A)$. The *Lüders projection of ρ with respect to the measurement of A* is the density matrix

$$\rho' := \sum_{\lambda \in \sigma(A)} P_\lambda \rho P_\lambda.$$

In summary, we have obtained the following theorem based on our above analysis.

Theorem 6.6. *Let $A \in \mathcal{M}_m(\mathbb{C})$ be a self-adjoint matrix with spectrum $\sigma(A)$, let $F :$ $\mathbb{C}^{\sigma(A)} \to \mathcal{M}_m(\mathbb{C})$ be as in (6.1), and let $\omega = \mathrm{tr}(\rho \cdot) : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}$ be a state with $\langle q, \cdot \rangle := \omega \circ F$ the induced state on $\mathbb{C}^{\sigma(A)}$. Then F has a disintegration of ω over $\langle q, \cdot \rangle$ consistent with F if and only if ρ equals its Lüders projection with respect to the measurement of A.*

Proof. We will use the same notation as earlier in this section.

($\Rightarrow$) Assume a disintegration exists. By (6.4), $P_\lambda \rho P_\lambda = q_\lambda \rho_\lambda$ for all $\lambda \in \sigma(A) \setminus N_q$. Hence,

$$\rho_\lambda = \frac{P_\lambda \rho P_\lambda}{q_\lambda} = \frac{P_\lambda \rho P_\lambda}{\mathrm{tr}(\rho P_\lambda)} \quad \forall \lambda \in \sigma(A) \setminus N_q \tag{6.7}$$

because ρ_λ is a density matrix. Furthermore, since a disintegration exists,

$$\rho \overset{(6.3)}{=} \sum_{\lambda \in \sigma(A) \setminus N_q} q_\lambda \rho_\lambda \overset{(6.7)}{=} \sum_{\lambda \in \sigma(A) \setminus N_q} q_\lambda \frac{P_\lambda \rho P_\lambda}{q_\lambda} = \sum_{\lambda \in \sigma(A) \setminus N_q} P_\lambda \rho P_\lambda.$$

($\Leftarrow$) Suppose that ρ equals its Lüders projection; i.e., suppose that

$$\rho = \sum_{\lambda \in \sigma(A)} P_\lambda \rho P_\lambda.$$

For each $\lambda \in \sigma(A)$, set $R_\lambda : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}$ to be the linear map defined by

$$\mathcal{M}_m(\mathbb{C}) \ni B \mapsto R_\lambda(B) := \begin{cases} \mathrm{tr}\left(\frac{P_\lambda \rho P_\lambda}{q_\lambda} B\right) & \text{if } q_\lambda > 0, \\ \frac{1}{m} \mathrm{tr}(B) & \text{if } q_\lambda = 0. \end{cases} \tag{6.8}$$

Then the linear map $R : \mathcal{M}_m(\mathbb{C}) \rightsquigarrow \mathbb{C}^{\sigma(A)}$ defined by $R := \sum_{\lambda \in \sigma(A)} \mathrm{ev}_\lambda^* \circ R_\lambda$ is a disintegration of F. To see this, first notice that R is positive, which implies that it is CP since $\mathbb{C}^{\sigma(A)}$ is commutative (cf. Stinespring [53, Theorem 3]). Second, R is unital because

$$R(\mathbb{1}_m) = \sum_{\lambda \in \sigma(A) \setminus N_q} R_\lambda(\mathbb{1}_m) e_\lambda + \sum_{\lambda \in N_q} R_\lambda(\mathbb{1}_m) e_\lambda$$

$$\overset{(6.8)}{=} \sum_{\lambda \in \sigma(A) \setminus N_q} e_\lambda + \sum_{\lambda \in N_q} e_\lambda = \sum_{\lambda \in \sigma(A)} e_\lambda$$

since $\mathrm{tr}(P_\lambda \rho P_\lambda) = q_\lambda$ for all $\lambda \in \sigma(A)$. To show that R satisfies

$$R \circ F \underset{\langle q, \cdot \rangle}{=} \mathrm{id}_{\mathbb{C}^{\sigma(A)}},$$

we will show that $R(F(e_\mu)) - e_\mu \in \mathcal{N}_{\langle q, \cdot \rangle}$ for all $\mu \in \sigma(A)$. Setting $d_\mu := \mathrm{tr}(P_\mu)$, the degeneracy/mutliplicity of $\mu \in \sigma(A)$, we obtain

$$R\big(F(e_\mu)\big) \overset{(6.1)}{=} R(P_\mu) = \sum_{\lambda \in \sigma(A) \setminus N_q} R_\lambda(P_\mu)e_\lambda + \sum_{\lambda \in N_q} R_\lambda(P_\mu)e_\lambda$$

$$\overset{(6.8)}{=} \sum_{\lambda \in \sigma(A) \setminus N_q} \delta_{\mu\lambda}e_\lambda + \sum_{\lambda \in N_q} \frac{d_\mu}{m}e_\lambda = \begin{cases} e_\mu + \frac{d_\mu}{m}\sum_{\lambda \in N_q} e_\lambda & \text{if } q_\mu > 0, \\ \frac{d_\mu}{m}\sum_{\lambda \in N_q} e_\lambda & \text{if } q_\mu = 0. \end{cases}$$

Hence, $R(F(e_\mu)) - e_\mu \in \mathbb{C}^{N_q}$, which is the null space of $\langle q, \cdot \rangle$. Thus, R is a disintegration. ∎

Appendix: Equivalent definitions of disintegration

In this appendix, we review the definition of a disintegration from measure theory (cf. Fremlin [14, Definition 452E]). Tables 1 and 2 provide two, a-priori different, definitions of a disintegration with varying input data and consistency conditions. This appendix serves to explain how these definitions are related to each other. More precisely, Theorems A.6 and A.11 state that the definitions in the respective tables are equivalent. Theorem A.18 says that this diagrammatic definition of a disintegration is equivalent to the definition of a regular conditional probability (cf. Panagaden [33, Definition 2.1]).

In all that follows, (X, Σ, μ) and (Y, Ω, ν) are measure spaces with no additional assumptions other than μ and ν are non-negative measures. Furthermore, $f : X \to Y$ is taken to be measure-preserving so that the pushforward $f_*\mu$ of μ along f is ν; i.e., $\nu(F) = \mu(f^{-1}(F))$ for all $F \in \Omega$.

Definition A.1. Let (X, Σ) and (Y, Ω) be measurable spaces. A *transition kernel* r from (Y, Ω) to (X, Σ), written as $r : Y \rightsquigarrow X$, is a function $r : Y \times \Sigma \to [0, \infty]$ such that

(i) $r(y, \cdot) : \Sigma \to [0, \infty]$ is a measure for all $y \in Y$ and

(ii) $r(\cdot, E) : Y \to [0, \infty]$ is measurable for all $E \in \Sigma$.

The notation $r_y(E) := r(y, E)$ will be implemented. A transition kernel as above is called a *stochastic map* (also *Markov kernel*) when r_y is a probability measure for all $y \in Y$.

Transition kernels are generalizations of measurable functions in that they assign to each point in the source/domain a measure on the target/codomain (cf. Example A.4). If a function is to be thought of as a deterministic process, then a transition kernel whose associated measures are probability measures can be interpreted as a non-deterministic (i.e., stochastic) process, where one only knows the probabilities associated with the possible outcomes of that process.

Example A.2. Let (X, Σ) be a measurable space and let $\{\bullet\}$ denote a one element set with the unique σ-algebra. There is a bijection between the set of measures on (X, Σ) and the set of transition kernels $\{\bullet\} \rightsquigarrow X$ from $\{\bullet\}$ to X. This allows measures to be viewed as morphisms.

Definition A.3. Let (X, Σ), (Y, Ω), and (Z, Ξ) be measurable spaces. Let $\mu : X \rightsquigarrow Y$ and $\nu : Y \rightsquigarrow Z$ be two transition kernels. The *composite* of μ followed by ν, written as $\nu \circ \mu : X \rightsquigarrow Z$, is defined by

$$X \times \Xi \ni (x, E) \mapsto (\nu \circ \mu)_x(E) := \int_Y \nu_y(E) \, d\mu_x(y).$$

This equation is known as the *Chapman–Kolmogorov* equation.

The fact that the composite of transition kernels defines a transition kernel follows from the monotone convergence theorem [50, Theorem 1.26]. Rather than proving this here, we will recall techniques from analysis that can be used to prove this when we prove Theorem A.6.

Example A.4. Let (X, Σ) and (Y, Ω) be measurable spaces and let $\mu : \{\bullet\} \rightsquigarrow X$ be a measure on X and let $f : X \to Y$ be a measurable function. Then f can be viewed as the transition kernel $f : X \rightsquigarrow Y$ given by

$$X \times \Omega \ni (x, E) \mapsto f_x(E) := \chi_E\big(f(x)\big) := \begin{cases} 1 & \text{if } f(x) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, $f \circ \mu$ is the pushforward $f_*\mu$ of the measure μ along the map f because

$$\Omega \ni E \mapsto (f \circ \mu)(E) = \int_X f_x(E) \, d\mu(x) = \int_X (\chi_E \circ f) \, d\mu$$

$$= \int_Y \chi_E \, d(f_*\mu) = \mu\big(f^{-1}(E)\big).$$

A special case of this occurs for the diagonal map $\Delta_Y : Y \to Y \times Y$. This pushes forward a probability measure $q : \{\bullet\} \rightsquigarrow Y$ to the diagonal subset

$$\Delta_Y(Y) := \big\{(y, y) \in Y \times Y \; : \; y \in Y\big\}$$

of $Y \times Y$ so that $(\Delta_Y \circ q)(A \times B) = q(A \cap B)$ for all $A, B \in \Omega$. This map is used to instantiate a categorical formulation of a.e. equivalence (cf. Remark A.10).

Example A.5. Let X, Y, and Z be finite sets equipped with the discrete σ-algebra and suppose that all transition kernels are stochastic maps. Then Definitions A.1 and A.3 reproduce the notion of *stochastic matrices* including their compositions. Indeed, since the σ-algebra on Y is discrete,

$$f_x(E) = \sum_{y \in E} f_x\big(\{y\}\big)$$

so that the probability measure f_x is determined by its values on points of Y. We therefore write $f_{yx} := f_x(\{y\})$ to denote the yx entry of f in matrix form. Second, the composite $X \xrightarrow{f} Y \xrightarrow{g} Z$ sends $x \in X$ to the probability measure on Z determined by

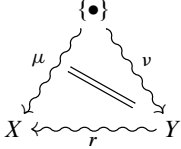$$Z \ni z \xmapsto{(g \circ f)(x)} (g \circ f)_{zx} := \sum_{y \in Y} g_{zy} f_{yx}.$$

	Functional	Measure-theoretic	Diagrammatic
Data	transition kernel $r : Y \rightsquigarrow X$	transition kernel $r : Y \rightsquigarrow X$	transition kernel $r : Y \rightsquigarrow X$
Conditions	$\int_X h \, d\mu = \int_Y (\int_X h \, dr_y) d\nu(y)$ $\forall$ measurable $h : X \to [0, \infty]$	$\mu(E) = \int_Y r_y(E) \, d\nu(y)$ $\forall E \in \Sigma$	$\{\bullet\}$ with μ, ν arrows to $X \xleftarrow{r} Y$ i.e., $r \circ \nu = \mu$
References	[19, 49]	[5, 14, 30, 56]	[8]

Table 1. Three definitions of a disintegration of (X, Σ, μ) over (Y, Ω, ν).

Following Example A.4, when Y is a finite set equipped with the discrete σ-algebra, the pushforward of $q : \{\bullet\} \rightsquigarrow Y$ along $\Delta_Y : Y \to Y \times Y$ simplifies to

$$(\Delta_Y \circ q)_{(y,y')} = \delta_{yy'} q_y.$$

With these definitions in place, we can compare several definitions of disintegrations. Table 1 below describes three equivalent definitions of a disintegration of one measure over another together with a list of references that use said definition.

Theorem A.6. *Given a transition kernel $Y \xrightarrow{r} X$ from a measure space (Y, Ω, ν) to a measure space (X, Σ, μ), the three conditions in Table 1 are equivalent.*

Proof. The equivalence between the measure-theoretic definition and the diagrammatic definition is immediate from the definition of the composition of transition kernels. Therefore, it suffices to prove the equivalence between the measure-theoretic and functional definitions. By setting $h := \chi_E$ with $E \in \Sigma$, the measure-theoretic condition follows from the functional definition. The only slightly non-trivial part of the proof of this equivalence is showing that the measure-theoretic definition implies the functional one. First, a straightforward computation, using $r \circ \nu = \mu$, shows that

$$\int_X s \, d\mu = \int_Y \left(\int_X s \, dr_y \right) d\nu(y) \tag{A.7}$$

for all simple functions $s : X \to [0, \infty)$. The general case for arbitrary measurable $h : X \to [0, \infty]$ follows from the monotone convergence theorem, though one needs to be careful about how to choose a monotone sequence of simple functions (s_n) converging pointwise to h. Such a sequence can be obtained as in the proof of Theorem 2.10 in the work of Folland [13] (cf. [35, Lemma 4.10]). From such a choice, it follows that

$$\mathbb{N} \ni n \mapsto \left(Y \ni y \mapsto \int_X s_n \, dr_y \right)$$

is a monotone increasing sequence of measurable functions on Y (see [35, equation (4.94) in the proof of part (iii) of Proposition 4.79] for details). Using all of these facts gives

$$\int_X h \, d\mu$$
$$= \lim_{n\to\infty} \int_X s_n \, d\mu \quad \text{by definition of } \int \text{ w.r.t. } \mu$$
$$= \lim_{n\to\infty} \int_Y \left(\int_X s_n \, dr_y \right) d\nu(y) \quad \text{by (A.7) for simple } s_n$$
$$= \int_Y \lim_{n\to\infty} \left(\int_X s_n \, dr_y \right) d\nu(y) \quad \text{by the monotone convergence theorem}$$
$$= \int_Y \left(\int_X h \, dr_y \right) d\nu(y) \quad \text{by definition of } \int \text{ w.r.t. } r_y$$

for arbitrary measurable $h : X \to [0, \infty]$. ∎

When one is equipped with the additional datum of a measure-preserving map $f : X \to Y$, there is another coherence condition that can be enforced on disintegrations. This assumption is to demand that a disintegration $r : Y \rightsquigarrow X$ be *consistent* with the map f. From our diagrammatic perspective, this means that r is a (stochastic) section of f a.e. This is described in Table 2.

Definition A.8. Let (X, Σ, μ) and (Y, Ω, ν) be two measure spaces. Two transition kernels $f, g : X \rightsquigarrow Y$ are said to be *μ-a.e. equivalent*, written as $f \underset{\mu}{=} g$, if and only if for each $F \in \Omega$, there exists a measurable set $N_F \in \Sigma$ such that

$$f_x(F) = g_x(F) \quad \forall x \in X \setminus N_F \quad \text{and} \quad \mu(N_F) = 0.$$

Example A.9. The definition of a.e. equivalence takes a particularly simple form for finite sets. Let X and Y be finite sets and let μ be a measure on X. Let

$$N_\mu := \{x \in X : \mu_x = 0\}$$

denote the *null-set* of (X, μ). Two transition kernels $f, g : X \rightsquigarrow Y$ are *μ-a.e. equivalent* if and only if

$$\{x \in X : f_x \neq g_x\} \subseteq N_\mu.$$

The notation $f \underset{\mu}{=} g$ is used whenever f and g are μ-a.e. equivalent. Here, $f_x \neq g_x$ means that f_x and g_x are different measures on Y; i.e., there exists a $y \in Y$ such that $f_{yx} \neq g_{yx}$.

Remark A.10. The definition of a.e. equivalence in Definition A.8 is a bit subtle in the general measure-theoretic case. Another reasonable option would be to say that f and g are μ-a.e. equivalent if and only if there exists an $N \in \Sigma$ such that $f_x = g_x$ (equality of measures) for all $x \in X \setminus N$ and $\mu(N) = 0$. However, this definition is too strong for the conditions in Table 2 to be equivalent for arbitrary measure spaces. The definition we have

	Measure-theoretic	Diagrammatic
Data	transition kernel $r : Y \rightsquigarrow X$	transition kernel $r : Y \rightsquigarrow X$
Conditions besides r is a disintegration of μ over ν	for each $F \in \Omega$ $\exists$ ν-null set $N_F \in \Omega$ s.t. $r_y(f^{-1}(F)) = 1$ $\forall y \in (Y \setminus N_F) \cap F$	i.e., $f \circ r \underset{\nu}{=} \mathrm{id}_Y$
References	[14, 30]	[8]

Table 2. Two definitions of a disintegration of (X, Σ, μ) over (Y, Ω, ν) consistent with a measure-preserving measurable map $f : X \to Y$.

chosen agrees with the diagrammatic definition of Cho and Jacobs [6, Section 5], which says that the diagram
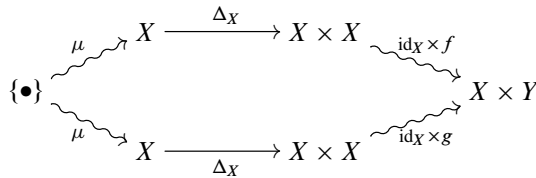
commutes (the product of stochastic maps can be defined using joint probability measures as is done in [6, Section 2]).

More explicitly, the condition $f \circ r \underset{\nu}{=} \mathrm{id}_Y$ says that for each $F \in \Omega$, there exists a ν-null set $M_F \in \Omega$ such that $(f \circ r)_y(F) = \chi_F(y)$ for all $y \in Y \setminus M_F$. Expanding out $(f \circ r)_y(F)$ using Example A.4 and the definition of transition kernels, this is equivalent to $r_y(f^{-1}(F)) = \chi_F(y)$ for all $y \in Y \setminus M_F$. Therefore, it is immediate that the diagrammatic definition implies that the measure-theoretic one.

Theorem A.11. *Let (Y, Ω, ν) and (X, Σ, μ) be measure spaces. Given a measure-preserving measurable map $X \xrightarrow{f} Y$, together with a disintegration $Y \overset{r}{\rightsquigarrow} X$ of μ over ν, the conditions in Table 2 are equivalent.*

Proof. By the comment preceding the statement of this theorem, the equivalence will follow from proving that the measure-theoretic definition implies the diagrammatic one; i.e., for each $F \in \Omega$, there exists a ν-null set $M_F \in \Omega$ such that $r_y(f^{-1}(F)) = \chi_F(y)$ for all $y \in Y \setminus M_F$ (cf. [14, Proposition 452G]). In more detail, by assumption, there exist ν-null sets $N_F, N_{Y \setminus F}, N_Y \in \Omega$ such that

$$r_y(X) = r_y\big(f^{-1}(Y)\big) = 1 \quad \forall y \in (Y \setminus N_Y) \cap Y \equiv Y \setminus N_Y, \tag{A.12}$$
$$r_y\big(f^{-1}(F)\big) = 1 \quad \forall y \in (Y \setminus N_F) \cap F, \tag{A.13}$$

and

$$r_y\big(f^{-1}(Y \setminus F)\big) = 1 \quad \forall y \in (Y \setminus N_{Y \setminus F}) \cap (Y \setminus F) \equiv Y \setminus (N_{Y \setminus F} \cup F).$$

Therefore,

$$1 - r_y\big(f^{-1}(F)\big) = r_y(X) - r_y\big(f^{-1}(F)\big) = r_y\big(f^{-1}(Y \setminus F)\big) = 1$$
$$\forall y \in (Y \setminus N_Y) \cap \big(Y \setminus (N_{Y \setminus F} \cup F)\big);$$

i.e.,

$$r_y\big(f^{-1}(F)\big) = 0 \quad \forall y \in (Y \setminus N_Y) \cap \big(Y \setminus (N_{Y \setminus F} \cup F)\big) \equiv Y \setminus (N_Y \cup N_{Y \setminus F} \cup F). \quad \text{(A.14)}$$

Set

$$M_F := N_F \cup N_{Y \setminus F} \cup N_Y,$$

which, being the finite union of ν-null sets, is ν-null. If $y \in (Y \setminus M_F) \cap F$, then, in particular, $y \in (Y \setminus N_F) \cap F$ so that $r_y(f^{-1}(F)) = 1$. If $y \in (Y \setminus M_F) \cap (Y \setminus F) \equiv Y \setminus (N_Y \cup N_{Y \setminus F} \cup N_Y \cup F)$, then, in particular, $y \in Y \setminus (N_Y \cup N_{Y \setminus F} \cup F)$ so that $r_y(f^{-1}(F)) = 0$. Putting these together,

$$r_y\big(f^{-1}(F)\big) = \chi_F(y) \quad \forall y \in Y \setminus M_F.$$

Therefore, the measure-theoretic definition implies the diagrammatic one. ∎

Remark A.15. The equality (A.12) says that r_y is a probability measure for all $y \in Y \setminus N_Y$.

A consistent disintegration is also related to the notion of a regular conditional probability.

Definition A.16. Let (X, Σ, μ) and (Y, Ω, ν) be measure spaces and let $f : X \to Y$ be a measure-preserving map. A *regular conditional probability* is a transition kernel $r : Y \rightsquigarrow X$ for which there exists a ν-null set $N \in \Omega$ such that r_y is a probability measure for all $y \in Y \setminus N$ and

$$\mu\big(E \cap f^{-1}(F)\big) = \int_F r_y(E) \, d\nu(y) \quad \forall E \in \Sigma \text{ and } \forall F \in \Omega. \quad \text{(A.17)}$$

Theorem A.18. *Let (X, Σ, μ) and (Y, Ω, ν) be measure spaces and let $f : X \to Y$ be a measure-preserving map. Then $r : Y \rightsquigarrow X$ is a regular conditional probability if and only if it is a disintegration of μ over ν consistent with f.*

Proof. ($\Rightarrow$) Suppose that r is a regular conditional probability. Then

$$\mu(E) = \mu(E \cap X) = \mu\big(E \cap f^{-1}(Y)\big) \overset{\text{(A.17)}}{=} \int_Y r_y(E) \, d\nu(y) = (r \circ \nu)(E)$$

for all $E \in \Sigma$. Now, fix $F \in \Omega$ and let $N \in \Omega$ be a ν-null set such that r_y is a probability

measure for all $y \in Y \setminus N$. Then

$$
\begin{array}{c}
\mu\big(f^{-1}(F) \cap f^{-1}(F)\big) \\[8pt]
\qquad\qquad\qquad \text{\tiny (A.17) with } E = f^{-1}(F) \\[6pt]
\mu\big(f^{-1}(F)\big) \qquad\qquad \int_F r_y\big(f^{-1}(F)\big)\, d\nu(y) \\[6pt]
\text{\tiny since } \nu = f \circ \mu \\[6pt]
\nu(F) \;=\!=\; \int_F d\nu
\end{array}
\qquad . \qquad \text{(A.19)}
$$

Since r_y is a probability measure ν-a.e., $r_y(f^{-1}(F)) \le 1$ for all $y \in (Y \setminus N) \cap F$ so that the quantity in (A.19) is finite. This allows us to meaningfully take the difference of these terms. Therefore, (A.19) implies that $\int_F (1 - r_y(f^{-1}(F)))\, d\nu(y) = 0$. Furthermore, since the integrand is non-negative, there exists a ν-null set $M_F \in \Omega$ such that

$$
r_y\big(f^{-1}(F)\big) = 1 \quad \forall y \in \big(Y \setminus (N \cup M_F)\big) \cap F.
$$

Hence, $f \circ r \underset{\nu}{=} \mathrm{id}_Y$ so that r is a consistent disintegration.

($\Longleftarrow$) Conversely, suppose that r is a consistent disintegration. By Remark A.15, r_y is a probability measure ν-a.e. Hence,

$$
\begin{array}{c}
\mu\big(E \cap f^{-1}(F)\big) \qquad\qquad\qquad\qquad \int_F r_y(E)\, d\nu(y) \\[8pt]
\text{\tiny }\mu = r \circ \nu \qquad\qquad\qquad\qquad\qquad\qquad \text{\tiny set theory} \\[6pt]
\int_Y r_y\big(E \cap f^{-1}(F)\big)\, d\nu(y) \qquad \int_F r_y\big((E \cap f^{-1}(F)) \cup (E \cap f^{-1}(Y \setminus F))\big)\, d\nu(y) \\[6pt]
\text{\tiny (A.14) for } F \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{\tiny } r_y \text{ is countably additive} \\[6pt]
\int_F r_y\big(E \cap f^{-1}(F)\big)\, d\nu(y) \;=\!=\; \begin{array}{l} \int_F r_y\big(E \cap f^{-1}(F)\big)\, d\nu(y) \\ + \int_F r_y\big(E \cap f^{-1}(Y \setminus F)\big)\, d\nu(y) \end{array} \\[4pt]
\text{\tiny (A.14) for } Y \setminus F
\end{array}
$$

for arbitrary $E \in \Sigma$ and $F \in \Omega$. This proves that r is a regular conditional probability. ∎

From this perspective, the results in this paper can be viewed as an approach to non-commutative regular conditional probabilities.

mechanics. Finally, we thank John Baez, Tobias Fritz, and Tom Leinster for the inspiration for this project, which came from a thorough investigation of the calculations in [2, Section 3].

References

[1] J. C. Baez and T. Fritz, A Bayesian characterization of relative entropy. *Theory Appl. Categ.* **29** (2014), no. 16, 422–457 Zbl 1321.94023 MR 3246591

[2] J. C. Baez, T. Fritz, and T. Leinster, A characterization of entropy in terms of information loss. *Entropy* **13** (2011), no. 11, 1945–1957 Zbl 1301.94043 MR 2868846

[3] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.* **76** (1996), 2818–2821

[4] B. V. R. Bhat and A. Chongdar, A minimal completion theorem and almost everywhere equivalence for completely positive maps. 2023, arXiv:2306.15952

[5] S. D. Chatterji, Disintegration of measures and lifting. In *Vector and operator valued measures and applications (Proc. Sympos., Alta, Utah, 1972)*, pp. 69–83, Academic Press, New York, 1973 Zbl 0288.28018 MR 0341087

[6] K. Cho and B. Jacobs, Disintegration and Bayesian inversion via string diagrams. *Math. Structures Comput. Sci.* **29** (2019), no. 7, 938–971 Zbl 1452.18008 MR 3987300

[7] M. D. Choi, Completely positive linear maps on complex matrices. *Linear Algebra Appl.* **10** (1975), 285–290 Zbl 0327.15018 MR 376726

[8] F. Clerc, V. Danos, F. Dahlqvist, and I. Garnier, Pointless learning. In *Foundations of software science and computation structures*, pp. 355–369, Lecture Notes in Comput. Sci. 10203, Springer, Berlin, 2017 Zbl 1486.68146 MR 3655552

[9] B. Coecke and R. W. Spekkens, Picturing classical and quantum Bayesian inference. *Synthese* **186** (2012), no. 3, 651–696 Zbl 1275.60006 MR 2945398

[10] D. Dieks, Communication by EPR devices. *Phys. Lett. A* **92** (1982), no. 6, 271–272

[11] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** (1935), 777–780 Zbl 0012.04201

[12] P. A. Fillmore, *A user's guide to operator algebras*. Canad. Math. Soc. Ser. Monogr. Adv. Texts, John Wiley & Sons, New York, 1996 Zbl 0853.46053 MR 1385461

[13] G. B. Folland, *Real analysis: Modern techniques and their applications*. 2nd edn., Pure Appl. Math. (N.Y.), John Wiley & Sons, New York, 1999 Zbl 0924.28001 MR 1681462

[14] D. H. Fremlin, *Measure theory. Vol. 4*. Torres Fremlin, Colchester, 2006 Zbl 1166.28001 MR 2462372

[15] T. Fritz, A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Adv. Math.* **370** (2020), article no. 107239 Zbl 1505.60004 MR 4104820

[16] J. Fullwood and A. J. Parzygnat, The information loss of a stochastic map. *Entropy* **23** (2021), no. 8, article no. 1021 MR 4303485

[17] J. Fullwood and A. J. Parzygnat, On dynamical measures of quantum information. 2023, arXiv:2306.01831

[18] R. Furber and B. Jacobs, From Kleisli categories to commutative C^*-algebras: probabilistic Gelfand duality. *Log. Methods Comput. Sci.* **11** (2015), no. 2, 1:5, 28 Zbl 1391.68034 MR 3357133

[19] N. Gagné and P. Panangaden, A categorical characterization of relative entropy on standard Borel spaces. In *The Thirty-third Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIII)*, pp. 135–153, Electron. Notes Theor. Comput. Sci. 336, Elsevier, Amsterdam, 2018 Zbl 07513459 MR 3796551

[20] L. Giorgetti, A. J. Parzygnat, A. Ranallo, and B. P. Russo, Bayesian inversion and the Tomita–Takesaki modular group. *Q. J. Math.* (2023), DOI 10.1093/qmath/haad014

[21] M. Giry, A categorical approach to probability theory. In *Categorical aspects of topology and analysis (Ottawa, Ont., 1980)*, pp. 68–85, Lecture Notes in Math. 915, Springer, Berlin, 1982 Zbl 0486.60034 MR 659884

[22] M. Gromov, Probability, symmetry, linearity lecture 1/6. 2014, Institut des Hautes Études Scientifiques, Lecture recording available at https://www.youtube.com/watch?v=aJAQVletzdY

[23] M. Gromov, Probability, symmetry, linearity lecture 2/6. 2014, Institut des Hautes Études Scientifiques, Lecture recording available at https://www.youtube.com/watch?v=Vci3C6yAzRE

[24] B. Jacobs, From probability monads to commutative effectuses. *J. Log. Algebr. Methods Program.* **94** (2018), 200–237 Zbl 1382.68073 MR 3726099

[25] R. V. Kadison, A generalized Schwarz inequality and algebraic invariants for operator algebras. *Ann. of Math. (2)* **56** (1952), 494–503 Zbl 0047.35703 MR 51442

[26] K. Kraus, *States, effects, and operations*. Lecture Notes in Phys. 190, Springer, Berlin, 1983 Zbl 0545.46049 MR 725167

[27] F. W. Lawvere, The category of probabilistic mappings. 1962, preprint

[28] M. S. Leifer, Quantum dynamics as an analog of conditional probability. *Phys. Rev. A* **74** (2006), article no. 042310

[29] G. Lüders, Concerning the state-change due to the measurement process. *Ann. Phys.* **15** (2006), no. 9, 663–670 Zbl 1108.81005

[30] J. E. Miller, *Nonstandard techniques in lifting theory*. Ph.D. thesis, The University of Texas at Austin, 2011

[31] A. Nayak and P. Sen, Invertible quantum operators and perfect encryption of quantum states. *Quantum Inf. Comput.* **7** (2007), no. 1-2, 103–110 Zbl 1152.81785 MR 2302678

[32] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge University Press, New York, 2010 Zbl 1288.81001

[33] P. Panangaden, The category of Markov kernels. In *PROBMIV'98: First International Workshop on Probabilistic Methods in Verification (Indianapolis, IN)*, pp. 171–187, Electron. Notes Theor. Comput. Sci. 22, Elsevier, Amsterdam, 1999 Zbl 0920.68067 MR 1722605

[34] J. L. Park, The concept of transition in quantum mechanics. *Found. Phys.* **1** (1970), no. 1, 23–33

[35] A. J. Parzygnat, Discrete probabilistic and algebraic dynamics: A stochastic commutative Gelfand–Naimark theorem. 2017, arXiv:1708.00091

[36] A. J. Parzygnat, From observables and states to Hilbert space and back: a 2-categorical adjunction. *Appl. Categ. Structures* **26** (2018), no. 6, 1123–1157 Zbl 1408.81017 MR 3875848

[37] A. J. Parzygnat, Stinespring's construction as an adjunction. *Compositionality* **1** (2019), no. 2 Zbl 1490.46053 MR 4275262

[38] A. J. Parzygnat, Inverses, disintegrations, and Bayesian inversion in quantum Markov categories. 2020, arXiv:2001.08375

[39] A. J. Parzygnat, Conditional distributions for quantum systems. In *Proceedings 18th International Conference on Quantum Physics and Logic*, p. 13, Electron. Proc. Theor. Comput. Sci. 343, 2021

[40] A. J. Parzygnat, Towards a functorial description of quantum relative entropy. In *Geometric science of information*, pp. 557–564, Lecture Notes in Comput. Sci. 12829, Springer, Cham, 2021 Zbl 1485.94038 MR 4424362

[41] A. J. Parzygnat, A functorial characterization of von Neumann entropy. *Cah. Topol. Géom. Différ. Catég.* **63** (2022), no. 1, 89–128 Zbl 1481.18007 MR 4551804

[42] A. J. Parzygnat and F. Buscemi, Axioms for retrodiction: achieving time-reversal symmetry with a prior. *Quantum* **7** (2023), article no. 1013

[43] A. J. Parzygnat and J. Fullwood, From time-reversal symmetry to quantum Bayes' rules. *PRX Quantum* **4** (2023), no. 2, article no. 020334

[44] A. J. Parzygnat and B. P. Russo, A non-commutative Bayes' theorem. *Linear Algebra Appl.* **644** (2022), 28–94 Zbl 1500.46052 MR 4393903

[45] W. L. Paschke, Inner product modules over B^*-algebras. *Trans. Amer. Math. Soc.* **182** (1973), 443–468 Zbl 0239.46062 MR 355613

[46] V. Paulsen, *Completely bounded maps and operator algebras*. Cambridge Stud. Adv. Math. 78, Cambridge University Press, Cambridge, 2002 Zbl 1029.47003 MR 1976867

[47] D. Petz, *Quantum information theory and quantum statistics*. Theoret. and Math. Phys., Springer, Berlin, 2008 Zbl 1145.81002 MR 2363070

[48] R. Pluta, *Ranges of bimodule projections and conditional expectations*. Cambridge Scholars Publishing, Newcastle upon Tyne, 2013 Zbl 1306.46001 MR 3052312

[49] V. A. Rohlin, On the fundamental ideas of measure theory. *Amer. Math. Soc. Translation* **1952** (1952), no. 71, 55 MR 0047744

[50] W. Rudin, *Real and complex analysis*. 3rd edn., McGraw-Hill Book, New York, 1987 Zbl 0925.00005 MR 924157

[51] S. Sakai, C^*-*algebras and W^*-algebras*. Ergeb. Math. Grenzgeb. 60, Springer, New York, 1971 Zbl 0219.46042 MR 0442701

[52] Z. Semadeni, *Monads and their Eilenberg–Moore algebras in functional analysis*. Queen's Papers in Pure and Appl. Math. 33, Queen's University, Kingston, Ontario, 1973 Zbl 0272.46049 MR 0367674

[53] W. F. Stinespring, Positive functions on C^*-algebras. *Proc. Amer. Math. Soc.* **6** (1955), 211–216 Zbl 0064.36703 MR 69403

[54] E. Størmer, *Positive linear maps of operator algebras*. Springer Monogr. Math., Springer, Heidelberg, 2013 MR 3012443

[55] T. Świrszcz, Monadic functors and convexity. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **22** (1974), 39–42 Zbl 0276.46036 MR 390019

[56] M. Wendt, The category of disintegration. *Cahiers Topologie Géom. Différentielle Catég.* **35** (1994), no. 4, 291–308 Zbl 0818.18003 MR 1307262

[57] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned. *Nature* **299** (1982), no. 5886, 802–803 Zbl 1369.81022

Arthur J. Parzygnat
Institut des Hautes Études Scientifiques, 35 Route de Chartres, 91440 Bures-sur-Yvette, France;
current address: Department of Mathematics, Massachusetts Institute of Technology, Cambridge,
MA 02139, USA; parzygnat@ihes.fr, arthurjp@mit.edu

Benjamin P. Russo
Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge,
TN 37831, USA; russobp@ornl.gov