

IDENTIFIABILITY OF PARAMETRIC RANDOM MATRIX MODELS

TOMOHIRO HAYASE

ABSTRACT. We investigate parameter identifiability of spectral distributions of random matrices. In particular, we treat compound Wishart type and signal-plus-noise type. We show that each model is identifiable up to some kind of rotation of parameter space. Our method is based on free probability theory.

CONTENTS

1. Introduction	1
2. Related Works	3
3. Preliminary	3
3.1. Freeness	3
3.2. Random Matrix Models and Free Deterministic Equivalents	5
4. Identifiability	6
4.1. Identifiability of CW Model	6
4.2. Identifiability of SPN Model	6
5. Acknowledgement	16
References	16

1. INTRODUCTION

Identifiability analysis is fundamental in a theoretical understanding of statistical models, for example, log-likelihood maximization. A parametric statistical model $(P_\vartheta)_{\vartheta \in \Theta}$, a parametric family of probability measures, is said to be *identifiable* if the map $\vartheta \mapsto P_\vartheta$ is injective. For a statistical model, its identifiability is necessary for its regularity. Under regularity condition, then maximal likelihood estimator has a good behavior such as asymptotic normality. In general, a geometry of log-likelihood is determined by the Fisher information matrix (see [2]), which is expected Hessian of log-likelihood with respect to parameters. If a statistical model is non-identifiable, then the Fisher information matrix is singular, and the eigenspace for the zero eigenvalue is determined by non-identifiable parameters. Therefore, determining non-identifiable parameters is important in non-identifiable models.

In this paper, we investigate identifiability of statistical models introduced for parameter estimation of random matrix models. In [8], two typical random matrix models, the

Date: December 31, 2018.

Key words and phrases. Identifiability, Random Matrix Theory, Free Probability Theory, Statistical Models.

compound Wishart model W_{CW} and the signal-plus-noise model W_{SPN} are treated. They are defined as the following:

$$\begin{aligned} W_{\text{CW}}(D) &= Z^* D Z, \\ W_{\text{SPN}}(A, \sigma) &= (A + \sigma Z)^* (A + \sigma Z), \end{aligned}$$

where Z is $p \times d$ matrix of independent and identically distributed Gaussian random variables with mean zero and variance $1/d$, D and A are deterministic matrices, and $\sigma \in \mathbb{R}$. For any self-adjoint matrix W , let us denote by μ_W the eigenvalue distribution defined as

$$\mu_W = \frac{1}{d} \sum_{k=1}^d \delta_{\lambda_k},$$

where $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_d$ are eigenvalues of W . The parameter estimation method introduced in [8] is minimizing modified KL-divergence between a statistical model

$$\begin{aligned} &\mu_{\text{CW}}^\square(D), \quad D \in M_p(\mathbb{C}) \\ &\text{(resp. } \mu_{\text{SPN}}^\square(A, \sigma), \quad A \in M_{p,d}(\mathbb{C}), \quad \sigma \in \mathbb{R}) \end{aligned}$$

and a sample of the empirical eigenvalue distribution $\mu_{W_{\text{CW}}(D_0)}$ (resp. $\mu_{W_{\text{SPN}}(A_0, \sigma_0)}$), where true parameters D_0, A_0, σ_0 are unknown. The definition of the statistical models μ_x^\square is based on *free deterministic equivalent*. The free deterministic equivalent is introduced by [14], which is a deterministic and infinite-dimensional approximation of random matrices based on a central limit theorem of the eigenvalue distribution.

It directly follows from the definition of μ_x^\square that

$$\begin{aligned} \mu_D = \mu_{D'} &\implies \mu_{\text{CW}}^\square(D) = \mu_{\text{CW}}^\square(D'), \\ \mu_A = \mu_{A'}, \quad \sigma^2 = \sigma'^2 &\implies \mu_{\text{SPN}}^\square(A, \sigma) = \mu_{\text{SPN}}^\square(A', \sigma'). \end{aligned}$$

In particular, these statistical models are not identifiable. For the CW model, it is easy to show that the converse also holds:

$$\mu_D = \mu_{D'} \iff \mu_{\text{CW}}^\square(D) = \mu_{\text{CW}}^\square(D').$$

In other words, if we replace the parameter set by the set of eigenvalue distributions then this model becomes identifiable. Note that there is a bijection between the set of eigenvalue distributions and

$$\{v \in \mathbb{R}^p \mid v_k = v_{\pi(k)}, \forall k = 1, \dots, p, \quad \forall \pi \in S_p\},$$

where S_p is the permutation group of p elements. However, it is not clear that the converse holds for the SPN model.

The main theorem of this paper is as follows.

Theorem 1.1. *Let $p, d \in \mathbb{N}$ with $p \geq d$. For $A, B \in M_{p,d}(\mathbb{C})$ and $\sigma, \rho \in \mathbb{R}$, the following holds:*

$$\mu_{\text{SPN}}^\square(A, \sigma) = \mu_{\text{SPN}}^\square(B, \rho) \iff \begin{cases} \mu_{A^*A} = \mu_{B^*B}, \\ \sigma^2 = \rho^2. \end{cases}$$

In particular, if we replace the parameter space by the direct product of singular value distribution and the nonzero real numbers, then this statistical model becomes identifiable. Note that there is a bijective between the direct product and

$$\{v \in \mathbb{R}^d \mid v_{\pi(k)} = v_k \geq 0 \ \forall k = 1, \dots, d, \forall \pi \in S_d\} \times \{v \in \mathbb{R} \mid v \geq 0\}.$$

Our proof consists of an analytic part based on operator-valued analytic free additive subordination [3] and a combinatorial part based on free multiplicative deconvolution [11, 12].

2. RELATED WORKS

The compound Wishart random matrix was introduced by [13]. It appears as sample covariance matrices of correlated samplings [4, 5, 7]. The signal-plus-noise random matrix appears in signal precessing [11, 6, 15].

Free probability is invented by Voiculescu [16]. In free probability theory, motivated by solving a problem in operator algebras, some infinite-dimensional operators are described as infinite-dimensional limit of random matrices. The approximation is based on a central limit theorem, which is called the *free central limit theorem*, of eigenvalue distribution of random matrices [17]. Conversely, the purpose of free deterministic equivalent is to approximate fixed-size but large random matrix models by deterministic operators.

For analysis of non-identifiable models, generic identifiability was introduced in [1].

3. PRELIMINARY

3.1. Freeness.

First, we summarize some definitions from operator algebras and free probability theory. See [9] for the detail.

Definition 3.1.

- (1) A C^* -probability space is a pair (\mathfrak{A}, τ) satisfying followings.
 - (a) The set \mathfrak{A} is a unital C^* -algebra, that is, a possibly non-commutative subalgebra of the algebra $B(\mathcal{H})$ of bounded \mathbb{C} -linear operators on a Hilbert space \mathcal{H} over \mathbb{C} satisfying the following conditions:
 - (i) it is stable under the adjoint $*$: $a \rightarrow a^*$, $a \in \mathfrak{A}$,
 - (ii) it is closed under the topology of the operator norm of $B(\mathcal{H})$,
 - (iii) it contains the identity operator $\text{id}_{\mathcal{H}}$ as the unit $1_{\mathfrak{A}}$ of \mathfrak{A} .
 - (b) The function τ on \mathfrak{A} is a faithful tracial state, that, is a \mathbb{C} -valued linear functional with
 - (i) $\tau(a) \geq 0$ for any $a \geq 0$, and the equality holds if and only if $a = 0$,
 - (ii) $\tau(1_{\mathfrak{A}}) = 1$,
 - (iii) $\tau(ab) = \tau(ba)$ for any $a, b \in \mathfrak{A}$.
- (2) A subalgebra \mathfrak{B} of a C^* -algebra \mathfrak{A} is called a $*$ -subalgebra if it is stable under the adjoint operator $*$. Moreover, it is called a unital C^* -subalgebra if the $*$ -subalgebra is closed under the operator norm topology and contains $1_{\mathfrak{A}}$ as its unit.
- (3) Two unital C^* -algebras are called $*$ -isomorphic if there is a bijective linear map between them which preserves the $*$ -operation and the multiplication.
- (4) Let us denote by $\mathfrak{A}_{\text{s.a.}}$ the set of self-adjoint elements, that is, $a = a^*$ of \mathfrak{A} .
- (5) Write $\text{Re } a := (a + a^*)/2$ and $\text{Im } a := (a - a^*)/2i$ for any $a \in \mathfrak{A}$.

(6) The *distribution* of $a \in \mathfrak{A}_{\text{s.a.}}$ is the probability measure $\mu_a \in \mathcal{B}_c(\mathbb{R})$ determined by

$$\int x^k \mu_a(dx) = \tau(a^k), \quad k \in \mathbb{N}.$$

(7) For $a \in \mathfrak{A}_{\text{s.a.}}$, we define its Cauchy transform G_a by $G_a(z) := \tau[(z-a)^{-1}]$ ($z \in \mathbb{C} \setminus \mathbb{R}$), equivalently, $G_a := G_{\mu_a}$.

Definition 3.2. A family of $*$ -subalgebras $(\mathfrak{A}_j)_{j \in J}$ of \mathfrak{A} is said to be *free* if the following factorization rule holds: for any $n \in \mathbb{N}$ and indexes $j_1, j_2, \dots, j_n \in J$ with $j_1 \neq j_2 \neq j_3 \neq \dots \neq j_n$, and $a_l \in \mathfrak{A}_{j_l}$ with $\tau(a_l) = 0$ ($l = 1, \dots, n$), it holds that

$$\tau(a_1 \cdots a_n) = 0.$$

Let $(x_j)_{j \in J}$ be a family of self-adjoint elements $x_j \in \mathfrak{A}_{\text{s.a.}}$. For $j \in J$, let \mathfrak{A}_j be the $*$ -subalgebra of polynomials of x_j . Then $(x_j)_{j \in J}$ is said to be free if \mathfrak{A}_j is free.

We introduce special elements in a non-commutative probability space.

Definition 3.3. Let (\mathfrak{A}, τ) be a C^* -probability space.

(1) An element $s \in \mathfrak{A}_{\text{s.a.}}$ is called *standard semicircular* if its distribution is given by the standard semicircular law;

$$\mu_s(dx) = \frac{\sqrt{4-x^2}}{2\pi} \mathbf{1}_{[-2,2]}(x) dx,$$

where $\mathbf{1}_S$ is the indicator function for any subset $S \subseteq \mathbb{R}$.

(2) Let $v > 0$. An element $c \in \mathfrak{A}$ is called *circular of variance v* if

$$c = \sqrt{v} \frac{s_1 + i s_2}{\sqrt{2}},$$

where (s_1, s_2) is a pair of free standard semicircular elements. In addition, c is called *standard circular element* if $v = 1$.

(3) A *$*$ -free circular family* (resp. standard *$*$ -free circular family*) is a family $\{c_j \mid j \in J\}$ of circular elements $c_j \in \mathfrak{A}$ such that $\bigcup_{j \in J} \{\text{Re } c_j, \text{Im } c_j\}$ is free (resp. and each element is of variance 1).

Definition 3.4. Let (\mathfrak{A}, τ) be a C^* -probability space and \mathfrak{B} be a unital C^* -subalgebra of \mathfrak{A} . Recall that they share the unit: $I_{\mathfrak{A}} = I_{\mathfrak{B}}$.

(1) Then a linear operator $E: \mathfrak{A} \rightarrow \mathfrak{B}$ is called a *conditional expectation onto \mathfrak{B}* if it satisfies following conditions;

- (a) $E[b] = b$ for any $b \in \mathfrak{B}$,
- (b) $E[b_1 a b_2] = b_1 E[a] b_2$ for any $a \in \mathfrak{A}$ and $b_1, b_2 \in \mathfrak{B}$,
- (c) $E[a^*] = E[a]^*$ for any $a \in \mathfrak{A}$.

(2) We write $\mathbb{H}^+(\mathfrak{B}) := \{W \in \mathfrak{B} \mid \text{there is } \varepsilon > 0 \text{ such that } \Im W \geq \varepsilon I_{\mathfrak{A}}\}$ and $\mathbb{H}^-(\mathfrak{B}) := -\mathbb{H}^+(\mathfrak{B})$.

(3) Let $E: \mathfrak{A} \rightarrow \mathfrak{B}$ be a conditional expectation. For $a \in \mathfrak{A}_{\text{s.a.}}$, we define a *E -Cauchy transform* as the map $G_a^E: \mathbb{H}^+(\mathfrak{B}) \rightarrow \mathbb{H}^-(\mathfrak{B})$, where

$$G_a^E(Z) := E[(Z - a)^{-1}], \quad Z \in \mathbb{H}^+(\mathfrak{B}).$$

If there is no confusion, we also call E a \mathfrak{B} -valued Cauchy transform.

Definition 3.5. (Operator-valued Freeness) Let (\mathfrak{A}, τ) be a C^* -probability space, and $E : \mathfrak{A} \rightarrow \mathfrak{B}$ be a conditional expectation. Let $(\mathfrak{B}_j)_{j \in J}$ be a family of $*$ -subalgebras of \mathfrak{A} such that $\mathfrak{B} \subseteq \mathfrak{B}_j$. Then $(\mathfrak{B}_j)_{j \in J}$ is said to be E -free if the following factorization rule holds: for any $n \in \mathbb{N}$ and indexes $j_1, j_2, \dots, j_n \in J$ with $j_1 \neq j_2 \neq j_3 \neq \dots \neq j_n$, and $a_l \in \mathfrak{B}_l$ with $E(a_l) = 0$ ($l = 1, \dots, n$), it holds that

$$E(a_1 \cdots a_n) = 0.$$

In addition, a family of elements $X_j \in \mathfrak{A}_{\text{s.a.}}$ ($j \in J$) is called E -free if the family of $*$ -subalgebra of the \mathfrak{B} -coefficient polynomials of X_j is E -free.

3.2. Random Matrix Models and Free Deterministic Equivalents.

Definition 3.6. Fix a probability measure space $(\Omega, \mathfrak{F}, \mathbb{P})$. Write $\mathbb{E}[\cdot] = \int \cdot \mathbb{P}(d\omega)$. Let $p, d \in \mathbb{N}$. Then real (resp. complex) $p \times d$ Ginibre random matrix of variance $v > 0$ is defined as $p \times d$ matrix of independent and identically distributed real (resp. complex) Gaussian random variables Z_{ij} ($i = 1, \dots, p, j = 1, \dots, d$) such that

$$\mathbb{E}[Z_{ij}] = 0, \mathbb{E}[\bar{Z}_{ij} Z_{ij}] = v.$$

Definition 3.7. Let $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$). Let us denote by Z the real (resp. complex) $p \times d$ Ginibre random matrix of variance $1/d$.

- (1) A real (resp. complex) *compound Wishart model* (CW model for short) of type (p, d) is defined as a parametric family W_{CW} , where

$$W_{\text{CW}}(D) := Z^* D Z, \quad D \in M_p(\mathbb{K}).$$

- (2) A real (resp. complex) *signal-plus-noise model* (SPN model for short) of type (p, d) is defined as a parametric family W_{SPN} , where

$$W_{\text{SPN}}(A, \sigma) := (A + \sigma Z)^* (A + \sigma Z), \quad A \in M_{p,d}(\mathbb{K}), \quad \sigma \in \mathbb{R}.$$

Here we introduce free deterministic equivalent of each random matrix model. Note that the free deterministic equivalent does not depend on the choice of the field \mathbb{R} or \mathbb{C} .

Definition 3.8. Let $p, d \in \mathbb{N}$. Fix a C^* -probability space (\mathfrak{A}, τ) . Let us denote by C the $p \times d$ matrix of $*$ -free circular elements in (\mathfrak{A}, τ) so that

$$\tau(C_{ij}) = 0, \quad \tau(C_{ij}^* C_{ij}) = 1/d.$$

- (1) The free deterministic equivalent of CW model (FDECW model, for short) of type (p, d) is defined as a parametric family W_{CW}^\square , where

$$W_{\text{CW}}^\square(D) = C^* D C, \quad D \in M_p(\mathbb{C}).$$

In addition, we denote by $\mu_{\text{CW}}(D)$ the distribution of $W_{\text{CW}}^\square(D)$ in the C^* -probability space $(M_d(\mathfrak{A}), \text{tr}_d \otimes \tau)$:

$$\mu_{\text{CW}}^\square(D) = \mu_{W_{\text{CW}}^\square(D)}.$$

- (2) The free deterministic equivalent of SPN model (FDESPN model, for short) of type (p, d) is defined as a parametric family W_{SPN}^\square , where

$$W_{\text{SPN}}^\square(A, \sigma) = (A + \sigma C)^* (A + \sigma C), \quad A \in M_{p,d}(\mathbb{C}), \quad \sigma \in \mathbb{R}.$$

In addition we denote by $\mu_{\text{SPN}}(A, \sigma)$ the distribution of $W_{\text{SPN}}^{\square}(A, \sigma)$ in the C^* -probability space $(M_d(\mathfrak{A}), \text{tr}_d \otimes \tau)$, that is,

$$\mu_{\text{SPN}}^{\square}(A, \sigma) = \mu_{W_{\text{SPN}}^{\square}(A, \sigma)}.$$

4. IDENTIFIABILITY

4.1. Identifiability of CW Model.

First, we quickly check the identifiability of the CW model. Fix $p, d \in \mathbb{N}$. Let $D, D' \in M_p(\mathbb{C})]_{\text{s.a.}}$ and $v = (v_1 \leq v_2 \leq \dots \leq v_p), v' = (v'_1 \dots v'_d) \in \mathbb{R}^p$ be the vectors of eigenvalues of D, D' respectively. Assume that

$$\mu_{\text{CW}}^{\square}(D) = \mu_{\text{CW}}^{\square}(D'). \quad (4.1)$$

Now since $\mu_{\text{CW}}^{\square}(D)$ is a compound free Poisson law (see [10]), the \mathcal{R} -transform of $\mu_{\text{CW}}^{\square}(D)$ is given by the following.

$$\mathcal{R}(b, v) = \frac{1}{d} \sum_{k=1}^p \frac{v_k}{1 - v_k b}, \quad b \in \mathbb{H}^-(\mathbb{C}).$$

By the assumption (4.1), it holds that

$$\mathcal{R}(b, v) = \mathcal{R}(b, v'), \quad b \in \mathbb{H}^-(\mathbb{C}).$$

Since all polos of $\mathcal{R}(\cdot, v)$ are order one, v and v' are equal up to permutation of entries, that is, there is a permutation $\pi \in S_p$ such that

$$v_{\pi(k)} = v'_k, \quad k = 1, \dots, p.$$

Equivalently, we have

$$\mu_D = \mu_{D'}.$$

4.2. Identifiability of SPN Model.

Next, we work on the SPN model. We prove the following identifiability of the statistical model $\mu_{\text{SPN}}^{\square}$ for the random matrix model W_{SPN} . The proof is divided into an analytic part and a combinatorial one.

Theorem 4.1. *Let $p, d \in \mathbb{N}$ with $p \geq d$, $A, B \in M_{p,d}(\mathbb{C})$, and $\sigma, \rho \in \mathbb{R}$. Then $\mu_{\text{SPN}}^{\square}(A, \sigma) = \mu_{\text{SPN}}^{\square}(B, \rho)$ if and only if $\mu_{A^*A} = \mu_{B^*B}$ and $\sigma^2 = \rho^2$.*

The proof is postponed to Section 4.2.5.

4.2.1. Analytic Part.

Write

$$\mathfrak{D}_2 = \left\{ \left[\begin{array}{cc} z_1 I_d & 0 \\ 0 & z_2 I_p \end{array} \right] \mid z_1, z_2 \in \mathbb{C} \right\} \subseteq M_{p+d}(\mathbb{C}) \subseteq M_{p+d}(\mathfrak{A}).$$

We identify \mathfrak{D}_2 and \mathbb{C}^2 via the following isomorphism $\mathfrak{D}_2 \simeq \mathbb{C}^2$:

$$\left[\begin{array}{cc} z_1 I_d & 0 \\ 0 & z_2 I_p \end{array} \right] \mapsto \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

We define a conditional expectation $E: M_{p+d}(\mathfrak{A}) \rightarrow \mathbb{C}^2$ by

$$E(X) = \begin{bmatrix} \text{tr}_d \otimes \tau(X_{+,+}) \\ \text{tr}_p \otimes \tau(X_{-,-}) \end{bmatrix},$$

where $X_{+,+} \in M_d(\mathfrak{A})$ is the $d \times d$ -upper left corner of $X \in M_{p,d}(\mathfrak{A})$ and $X_{-,-} \in M_p(\mathfrak{A})$ is the $p \times p$ -lower right corner of X . For $X \in M_{p+d}(\mathfrak{A})$ and $z \in \mathbb{H}^+(\mathbb{C}^2) = \{(z_1, z_2) \in \mathbb{C}^2 \mid \Im z_1, \Im z_2 > 0\}$, we write

$$\begin{aligned} G_X(z) &= E[(z - X)^{-1}], \\ h_X(z) &= G_X(z)^{-1} - z. \end{aligned}$$

For any rectangular matrix $Y \in M_{p,d}(\mathfrak{A})$, write

$$\Lambda(Y) = \begin{bmatrix} 0 & Y^* \\ Y & 0 \end{bmatrix}.$$

Let $z = (\alpha, \beta) \in \mathbb{C}^2$. Then we have

$$(z - \Lambda(Y))^{-1} = \begin{bmatrix} \alpha I_d & -Y^* \\ -Y & \beta I_p \end{bmatrix}^{-1} = \begin{bmatrix} \beta(\alpha\beta I_d - Y^*Y)^{-1} & Y^*(\alpha\beta I_p - YY^*)^{-1} \\ (\alpha\beta I_p - YY^*)^{-1}Y & \alpha(\alpha\beta I_p - YY^*)^{-1} \end{bmatrix}.$$

Applying E , we have

$$G_{\Lambda(Y)}(z) = \begin{bmatrix} \beta \text{tr}_d \otimes \tau[(\alpha\beta I_d - Y^*Y)^{-1}] \\ \alpha \text{tr}_p \otimes \tau[(\alpha\beta I_p - YY^*)^{-1}] \end{bmatrix}.$$

In particular, $G_{\Lambda(Y)}$ is determined by μ_{Y^*Y} . Let $C \in M_{p,d}(\mathfrak{A})$ be a matrix of $*$ -free standard circular elements. By [8, Proposition 5.30], $\Lambda(C)$ is a \mathbb{C}^2 -valued semicircular element (see [9, Section 9.1] for the definition) with the following variance mapping $\eta: \mathbb{C}^2 \rightarrow \mathbb{C}^2$:

$$\eta\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} (p/d)y \\ x \end{bmatrix}.$$

Hence the following equations hold for any $z \in \mathbb{H}^+(\mathbb{C}^2)$:

$$\begin{aligned} G_{\sigma\Lambda(C)}(z)^{-1} &= z - \sigma^2 \eta(G_{\sigma\Lambda(C)}(z)), \\ h_{\sigma\Lambda(C)}(z) &= \sigma^2 \eta(G_{\sigma\Lambda(C)}(z)). \end{aligned} \tag{4.2}$$

Next, to prove a key lemma, we refer to an analytic free additive subordination formula based on [3].

Corollary 4.2. *Set $a := \Lambda(A)$ and $s := \sigma\Lambda(C)$. Then there exists a pair of Fréchet analytic (equivalently, holomorphic) mappings $\psi_1, \psi_2 \in \text{Hol}(\mathbb{H}^+(\mathbb{C}^2))$ so that for all $z \in \mathbb{H}^+(\mathbb{C}^2)$,*

$$\begin{aligned} \Im \psi_j(z) &\geq \Im z, \forall j \in \{1, 2\}, \\ h_a(\psi_1(z)) + z &= \psi_2(z), \\ h_s(\psi_2(z)) + z &= \psi_1(z), \end{aligned} \tag{4.3}$$

$$G_{a+s}(z) = G_a(\psi_1(z)), \text{ and,} \tag{4.4}$$

$$G_{a+s}(z) = G_s(\psi_2(z)). \tag{4.5}$$

Proof. By [8, Proposition 5.30], the pair (a, s) is E -free. Then the assertion follows from [3, Theorem 2.7]. \square

Lemma 4.3. *Let $p, d \in \mathbb{N}$ with $p \geq d$. Let $A \in M_{p,d}(\mathbb{C})$ and $\sigma \in \mathbb{R}$. Then we have the following equation between holomorphic mappings on $\mathbb{H}^+(\mathbb{C}^2)$:*

$$G_{\Lambda(A+\sigma C)}(z) = G_{\Lambda(A)}[\sigma^2 \eta(G_{\Lambda(A+\sigma C)}(z)) + z], \forall z \in \mathbb{H}^+(\mathbb{C}^2).$$

Proof. Set $a := \Lambda(A)$ and $s := \Lambda(C)$. Pick same holomorphic mappings ψ_1 and ψ_2 as in Corollary 4.2. Then for any $z \in \mathbb{H}^+(\mathbb{C}^2)$,

$$\begin{aligned} G_{a+s}(z) &= G_a(\psi_1(z)) && \text{(by (4.4))} \\ &= G_a(h_s(\psi_2(z)) + z) && \text{(by (4.3))} \\ &= G_a(\sigma^2 \eta(G_s(\psi_2(z))) + z) && \text{(by (4.2))} \\ &= G_a(\sigma^2 \eta(G_{a+s}(z)) + z). && \text{(by (4.5))} \end{aligned}$$

\square

Now we have prepared to prove the first key lemma.

Lemma 4.4. *Fix $p, d \in \mathbb{N}$ with $p \geq d$. Let $A, B \in M_{p,d}(\mathbb{C})$ and $\sigma \in \mathbb{R}$. If $\mu_{\text{SPN}}^\square(A, \sigma) = \mu_{\text{SPN}}^\square(B, 0)$ then $\sigma = 0$.*

Proof. Assume that $\mu_{\text{SPN}}^\square(A, \sigma) = \mu_{\text{SPN}}^\square(B, 0)$. Then $G_{\Lambda(A+\sigma C)} = G_{\Lambda(B)}$ since $G_{\Lambda(Y)}$ is determined by μ_{Y^*Y} for any $Y \in M_{p,d}(\mathfrak{A})$.

In the case $B = 0$, it holds that $(A+\sigma C)^*(A+\sigma C) = 0$. Thus $A = -\sigma C$ and $A^*A = \sigma^2 C^*C$. Since μ_{C^*C} has no atom and μ_{A^*A} is a sum of delta measures, we have $\sigma = 0$.

Consider the case $B \neq 0$. Write $\beta := \|B^*B\|^{1/2} > 0$. Now for any $z \in \mathbb{H}^+(\mathbb{C}^2)$, by the assumption and Lemma 4.3, the following holds:

$$G_{\Lambda(A)}[\sigma^2 \eta(G_{\Lambda(B)}(z)) + z] = G_{\Lambda(B)}(z), z \in \mathbb{H}^+(\mathbb{C}^2). \quad (4.6)$$

Let

$$\begin{aligned} g(z) &:= G_{\Lambda(B)}(z), \\ f(z) &:= (z_2 z_1 - \beta^2) G_{\Lambda(B)}(z). \end{aligned}$$

Then

$$\lim_{\gamma \rightarrow +0} f(\beta + i\gamma, \beta + i\gamma) = \left(\frac{m\beta}{d}, \frac{m\beta}{p} \right) \neq 0, \quad (4.7)$$

where $m \geq 1$ is the multiplicity of the eigenvalue β of $\sqrt{B^*B}$. Let $a_1 \leq \dots \leq a_d$ be eigenvalues of $\sqrt{A^*A}$. Then

$$G_{\Lambda(A)}(z_1, z_2) = \left(\frac{z_2}{d} \sum_{k=1}^d \frac{1}{z_2 z_1 - a_k^2}, \frac{z_1}{p} \sum_{k=1}^d \frac{1}{z_2 z_1 - a_k^2} + \frac{p-d}{pz_2} \right).$$

Now for any $k = 1, \dots, d$ and $j = 1, 2$,

$$\frac{g(z)_j}{g(z)_1 g(z)_2 - a_k^2} = \frac{f(z)_j}{f(z)_1 f(z)_2 - a_k^2 (z_1 z_2 - \beta^2)^2} (z_1 z_2 - \beta^2). \quad (4.8)$$

Let $\gamma > 0$ and $z_1 = z_2 = \beta + i\gamma$. Then (4.8) converges to 0 as $\gamma \rightarrow +0$ by (4.7).

Assume that $\sigma \neq 0$, then by (4.8), it holds that

$$\lim_{\substack{z=(\beta+i\gamma, \beta+i\gamma) \\ \gamma \rightarrow +0}} G_{\Lambda(A)} [\sigma^2 \eta(G_{\Lambda(B)}(z)) + z] = (0, \frac{p-d}{p\beta}).$$

In particular,

$$\lim_{\substack{z=(\beta+i\gamma, \beta+i\gamma) \\ \gamma \rightarrow +0}} (z_1 z_2 - \beta^2) G_{\Lambda(A)} [\sigma^2 \eta(G_{\Lambda(B)}(z)) + z] = 0.$$

By (4.6), this contradicts (4.7). Therefore $\sigma = 0$. \square

4.2.2. Combinatorial Part.

We use the free multiplicative deconvolution introduced by [12, 11]. We quickly review the deconvolution.

First, we introduce a family of formal power series, since the deconvolution is defined as an operation between moment power series. Let us denote by Ξ the set of formal power series without the constant term of the form

$$f(z) = \sum_{n=1}^{\infty} \alpha_n z^n, \quad (4.9)$$

with $\alpha_n \in \mathbb{C} (\forall n \in \mathbb{N})$. Let $f \in \Xi$ be as in (4.9). For every $n \in \mathbb{N}$ we denote

$$\text{Cf}_n(f) = \alpha_n.$$

Second, we introduce Kreweras complement and boxed convolution. Here we only need one-dimensional boxed convolution. See [10, Lecture 17, 18] for the detail. Let $n \in \mathbb{N}$ and $\pi \in \text{NC}(n)$. Write $[n] = \{1, 2, \dots, n\}$ and consider the discriminant union $[n] \amalg [n]$. We write the elements from the second entry as \bar{k} ($k \in [n]$), and write $[\bar{n}] = \{\bar{1}, \bar{2}, \dots, \bar{n}\}$. We define an order as follows:

$$1 \leq \bar{1} \leq 2 \leq \bar{2} \dots \leq n \leq \bar{n}.$$

Then the set $[n] \amalg [n]$ is a totally ordered set. Let $\pi \in \text{NC}(n)$ and

$$J := \{\rho \in \text{NC}([\bar{n}]) \mid \pi \cup \rho \in \text{NC}([n] \cup [n])\}.$$

Then J has the biggest element with respect to the following partially order of $\text{NC}(n)$: for ρ and $\pi \in \text{NC}(n)$, $\rho \leq \pi$ if $\forall V_1, V_2 \in \rho, \exists W \in \pi$ such that $V_1 \cup V_2 \subseteq W$. The *Kreweras complement* of π , denoted by $K(\pi)$ is defined as

$$K(\pi) := \max J. \quad (4.10)$$

For $n \in \mathbb{N}$ and $\text{NC}(n)$, we denote

$$\text{Cf}_{n;\pi}(f) := \prod_{V \in \pi} \text{Cf}_{|V|}(f),$$

where $|V|$ is the number of elements in V . For $f, g \in \Xi$, the *one dimensional boxed convolution* (boxed convolution, for short), denoted by $f \boxtimes g$ is defined as

$$(f \boxtimes g)(z) := \sum_{m=1}^{\infty} \sum_{\pi \in \text{NC}(m)} \text{Cf}_{n;\pi}(f) \text{Cf}_{n;K(\pi)}(g) z^m,$$

where $K(\pi)$ is the Kreweras complement (4.10). One has the operation \boxtimes is associative and commutative [10, Proposition 17.5, Corollary 17.10]. In addition, let us denote by Δ the series in Ξ defined as

$$\Delta(z) = z.$$

Then Δ is the unit of (Ξ, \boxtimes) [10, Proposition 17.5]. We denote by Ξ^\times the set of invertible elements in Ξ with respect to \boxtimes . For $f \in \Xi$, we denote by f^{-1} its inverse with respect to \boxtimes . Then by [10, Proposition 17.7],

$$\Xi^\times = \{f \in \Xi \mid \text{Cf}_1(f) \neq 0\}.$$

Third, we define the *Zeta function* as

$$\text{Zeta}(z) := \sum_{n=1}^{\infty} z^n.$$

Clearly $\text{Zeta} \in \Xi^\times$. Then we define the R-transform of formal power series.

Definition 4.5. (R-transform) Let $f \in \Xi$. Let us define the *R-transform* of f as

$$R_f := f \boxtimes \text{Zeta}^{-1}.$$

For any probability measure μ on \mathbb{R} with all moments finite, we denote by M_μ its moment formal power series:

$$M[\mu](z) = \sum_{n=1}^{\infty} m_n(z) z^n.$$

Let (\mathfrak{A}, φ) be a C^* -probability space, and let a be an element of \mathfrak{A} . The *moment power series* of a , denote by M_a , is a formal power series defined as

$$M[a](z) = \sum_{n=1}^{\infty} \varphi(a^n) z^n.$$

We simply write

$$\begin{aligned} R[\mu] &= R_{M[\mu]}, \\ R[a] &= R_{M[a]}. \end{aligned} \tag{4.11}$$

Usually R-transform of $a \in \mathfrak{A}$ is defined as formal power series whose coefficients are free cumulants (see [10]). The compatibility of our definition (4.11) and usual definition is proven in [10, Proposition 17.4]. In addition, the following holds.

Lemma 4.6. *Let (\mathfrak{A}, φ) be a C^* -probability space and $a, b \in \mathfrak{A}$. Assume that (a, b) is free. Then*

$$R[ab] = R[a] \boxtimes R[b].$$

Proof. This is a direct consequence of [10, Proposition 17.2]. \square

Lastly, note that it holds that for $f \in \Xi$,

$$f \in \Xi^\times \text{ if and only if } R_f \in \Xi^\times,$$

since $\text{Cf}_1(R_f) = \text{Cf}_1(f)$. Now we have prepared to define the free multiplicative deconvolution.

Definition 4.7. (free multiplicative deconvolution) For $f \in \Xi$ and $g \in \Xi^\times$, the *free multiplicative deconvolution* of f with g is defined as

$$f \boxminus g := (R_f \boxtimes R_g^{-1}) \boxtimes \text{Zeta}.$$

Equivalently, $f \boxminus g$ is the unique formal power series in Ξ determined by

$$R_f = R_g \boxtimes R_{(f \boxminus g)}.$$

Example 4.8. Let $\beta \in \mathbb{R}$ and δ_β be the delta measure on \mathbb{R} whose support is $\{\beta\} \subseteq \mathbb{R}$. Then

$$M[\delta_\beta](z) = \sum_{n=1}^{\infty} \beta^n z^n = [\text{Zeta} \boxtimes (\beta \Delta)](z),$$

since

$$\text{Cf}_{n;K(\pi)}(\beta \Delta) = \begin{cases} \beta^n & ; \pi = \{\{1, 2, \dots, n\}\}, \\ 0 & ; \text{otherwise.} \end{cases}$$

Note that $K(\{\{1, 2, \dots, n\}\}) = \{\{1\}, \{2\}, \dots, \{n\}\}$. Hence

$$R[\delta_\beta] = \beta \Delta.$$

Then for any $f \in \Xi$, we have

$$R_{f(\beta \cdot)}[z] = \sum_{n=1}^{\infty} \text{Cf}_n(R_f) \beta^n z^n = R_f \boxtimes R[\delta_\beta](z).$$

In particular, if $f \in \Xi^\times$, it holds that

$$f(\beta \cdot) \boxminus f = M[\delta_\beta]. \quad (4.12)$$

In the case $f = M[a]$ with $a \in \mathfrak{A}$, it is easy to show that

$$M[\beta a] \boxminus M[a] = M[\beta] = M[\delta_\beta],$$

since each scalar is free from any element of \mathfrak{A} .

Definition 4.9. Let $f, g \in \Xi$. Then their *free additive convolution*, denoted by $f \boxplus g \in \Xi$, is defined as

$$f \boxplus g := (R_f + R_g) \boxtimes \text{Zeta}.$$

Equivalently, $f \boxplus g$ is the unique formal power series in Ξ determined by

$$R_{f \boxplus g} = R_f + R_g.$$

Notation 4.10. Let (\mathfrak{A}, φ) be a C^* -probability space. Let $q \in \mathfrak{A}$ be a non-zero projection, that is, $q = q^* = q^2$. Then

$$(q \mathfrak{A} q, \frac{1}{\varphi(q)} \varphi)$$

becomes a C^* -probability space. For $a \in \mathfrak{A}$, we denote by $M^{q \mathfrak{A} q}[q a q]$ the moment power series of $q a q$ ($a \in \mathfrak{A}$) in $(q \mathfrak{A} q, \varphi(q)^{-1} \varphi)$:

$$M^{q \mathfrak{A} q}[q a q] = \sum_{n=1}^{\infty} \frac{1}{\varphi(q)} \varphi[(q a q)^n] z^n.$$

Proposition 4.11. *Let (\mathfrak{A}, φ) be a C^* -probability space. Assume that $a, c, p \in \mathfrak{A}$ satisfies the following conditions:*

- (1) $a^* = a$,
- (2) c is a circular element, that is,

$$c = \sigma \frac{s_1 + is_2}{\sqrt{2}},$$

where (s_1, s_2) is a pair of free standard semicircular elements in (\mathfrak{A}, φ) and $\sigma \in \mathbb{R}$,

- (3) q is a projection, and,
- (4) $(\{c, c^*\}, \{a, q\})$ is a pair of free families.

Set $\lambda := \varphi(q)$ and

$$f_\lambda(z) := \sum_{n=1}^{\infty} \lambda^{n-1} z^n.$$

Then we have

$$M^{q\mathfrak{A}q}[q(a+c)^*(a+c)q] \boxtimes f_\lambda = (M^{q\mathfrak{A}q}[qa^*aq] \boxtimes f_\lambda) \boxplus (M^{q\mathfrak{A}q}[qc^*cq] \boxtimes f_\lambda).$$

Proof. This is a direct consequence of [12, Theorem 3.4]. \square

4.2.3. Free Poisson Distribution.

The formal power series f_λ in Proposition 4.11 is R-transform of a free Poisson distribution. We review on the free Poisson distribution.

Definition 4.12. (Free Poisson Distribution) Let $\lambda > 0$, $\alpha \in \mathbb{R}$. Then the *free Poisson distribution* with rate λ and jump size α is defined as the probability measure on \mathbb{R} determined by

$$R[\nu_{\lambda, \alpha}] = \lambda \sum_{n=1}^{\infty} \alpha^n z^n.$$

Usually free Poisson law is defined as the limit law of free version of law of small numbers [10, Definition 12.12]. The compatibility between our definition and usual definition is given by [10, Proposition 12.11]. Note that $\nu_{\lambda, \alpha}$ is, in fact, a compactly supported probability measure. Note that

$$f_\lambda = R[\nu_{\lambda^{-1}, \lambda}].$$

Lemma 4.13. *Let (\mathfrak{A}, φ) be a C^* -probability space, $a \in \mathfrak{A}$, and $q \in \mathfrak{A}$ be a non-zero projection free from a . Then it holds that*

$$R^{q\mathfrak{A}q}[qaq](z) = \lambda^{-1} R[a](\lambda z),$$

where $\lambda := \varphi(q)$.

This is well-known, but for the reader's convenience, we sketch the proof.

Proof. Note that $M^{q\mathfrak{A}q}[qaq] = \lambda^{-1} M[qaq]$. By the tracial condition and Lemma 4.6,

$$M[qaq] = M[aq] = R[aq] \boxtimes \text{Zeta} = R[a] \boxtimes R[q] \boxtimes \text{Zeta} = R[a] \boxtimes M[q] = R[a] \boxtimes (\lambda \text{Zeta}).$$

By definition of the boxed convolution, we have

$$R[a] \boxed{\star} (\lambda \text{Zeta})(z) = \sum_{n=1}^{\infty} \sum_{\pi \in \text{NC}(n)} \text{Cf}_{n;\pi}(R[a]) \lambda^{\#K(\pi)} z^n.$$

Since $\#\pi + \#K(\pi) = n + 1$, this is equal to

$$\lambda \sum_{n=1}^{\infty} \sum_{\pi \in \text{NC}(n)} \text{Cf}_{n;\pi} \left(\frac{1}{\lambda} R[a] \right) \lambda z^n = \lambda [(\lambda^{-1} R[a](\lambda \cdot)) \boxed{\star} \text{Zeta}](z).$$

Thus

$$\begin{aligned} M^{q\mathfrak{A}q}[qaq] &= (\lambda^{-1} R[a](\lambda \cdot)) \boxed{\star} \text{Zeta}, \\ R^{q\mathfrak{A}q}[qaq] &= \lambda^{-1} R[a](\lambda \cdot). \end{aligned}$$

□

Example 4.14. Let $q, c \in \mathfrak{A}$ and q be a nonzero-projection. Assume that $(\{q\}, \{c, c^*\})$ is free pair in (\mathfrak{A}, τ) and c is a standard circular element. Then by Lemma 4.13,

$$R^{q\mathfrak{A}q}[qc^*cq](z) = \lambda^{-1} R[c^*c](\lambda z) = \lambda^{-1} \sum_{n=1}^{\infty} \lambda^n z^n = R[\nu_{\lambda^{-1}, \lambda}](z) = f_{\lambda}(z).$$

4.2.4. Second Lemma.

In this section, we convert the model to an operator of the form qaq where q is a projection. Let (\mathfrak{A}, φ) be a C^* -probability space. Let $p, d \in \mathbb{N}$ with $p \geq d$ and write $\lambda = d/p$. In this section and in next one, we denote by $C^{p,d}$ be a $p \times d$ matrix of \star -free circular elements with

$$\varphi[(C_{ij}^{p,d})^* C_{ij}^{p,d}] = \frac{1}{d}.$$

Recall that

$$W_{\text{SPN}}^{\square}(A, \sigma) = (A + \sigma C^{p,d})^* (A + \sigma C^{p,d}).$$

Now we identify $C^{p,d}$ with $d \times d$ upper-left corner of $C^{p,p}$ with a normalization as the following:

$$C_{ij}^{p,p} = \sqrt{\lambda} C_{ij}^{p,d}, \quad \forall i \in \{1, 2, \dots, p\}, \quad \forall j \in \{1, 2, \dots, d\}.$$

Recall that a family $\{C_{ij}^{p,p} \mid 1 \leq i, j \leq p\}$ is a \star -free family of circular elements such as

$$\varphi[(C_{ij}^{p,p})^* C_{ij}^{p,p}] = \frac{1}{p}.$$

We write

$$\mathfrak{C} := M_p(\mathfrak{A}), \quad \tau := \text{tr}_p \otimes \varphi.$$

Then $C^{p,p}$ is a circular element in (\mathfrak{C}, τ) , and it is standard, that is,

$$\text{Cf}_n(R_{(C^{p,p})^* C^{p,p}}) = 1.$$

We define a projection $\Pi \in M_p(\mathbb{C}) \subseteq \mathfrak{C}$ as

$$\Pi_{ij} = \begin{cases} 1, & \text{if } i = j \leq d, \\ 0, & \text{otherwise.} \end{cases}$$

One has $\tau(\Pi) = \lambda$. For a $p \times d$ -matrix A , let us denote by \tilde{A} be the $p \times p$ -square matrix obtained by adding zeros to A ;

$$\tilde{A} := \begin{bmatrix} A & O_{p,d} \end{bmatrix}.$$

Now by definition, we have

$$\Pi \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right)^* \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right) \Pi = \begin{bmatrix} (A + \sigma C^{p,d})^* (A + \sigma C^{p,d}) & O_{d,p-d} \\ O_{p-d,d} & O_{p-d,p-d} \end{bmatrix}.$$

Therefore, for any $m \in \mathbb{N}$,

$$\begin{aligned} \frac{1}{d} \text{Tr}_d \otimes \varphi [W_{\text{SPN}}^\square(A, \sigma)^m] &= \frac{1}{\lambda} \frac{1}{p} \text{Tr}_p \otimes \varphi \left\{ \left[\Pi \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right)^* \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right) \Pi \right]^m \right\}, \\ \text{tr}_d \otimes \varphi [W_{\text{SPN}}^\square(A, \sigma)^m] &= \frac{1}{\text{tr}_p \otimes \varphi(\Pi)} \text{tr}_p \otimes \varphi \left\{ \left[\Pi \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right)^* \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right) \Pi \right]^m \right\}. \end{aligned}$$

Equivalently, we have

$$M[W_{\text{SPN}}^\square(A, \sigma)] = M^{\Pi \epsilon \Pi} \left[\Pi \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right)^* \left(\tilde{A} + \frac{\sigma}{\sqrt{\lambda}} C^{p,p} \right) \Pi \right]. \quad (4.13)$$

Recall that

$$M^{\Pi \epsilon \Pi}[\Pi X \Pi](z) = \sum_{n=1}^{\infty} \frac{1}{\tau(\Pi)} \tau[(\Pi X \Pi)^n] z^n.$$

Lemma 4.15. *Let $\alpha \in \mathbb{R}$. Then*

$$M^{\Pi \epsilon \Pi}[\alpha \Pi (C^{p,p})^* C^{p,p} \Pi](z) \boxtimes M[\nu_{\lambda^{-1}, \lambda}] = M[\delta_\alpha],$$

where δ_α is the delta measure on \mathbb{R} whose support is $\{\alpha\}$.

Proof. Now $\{C^{p,p}\}$ and $\{\tilde{A}, \Pi\}$ is $*$ -free in (\mathfrak{C}, τ) , since the entries of A and Π are scalar. By Lemma 4.13,

$$\begin{aligned} R^{\Pi \epsilon \Pi}[\Pi (C^{p,p})^* C^{p,p} \Pi](z) &= \lambda^{-1} R[(C^{p,p})^* C^{p,p}](\lambda z) \\ &= \lambda^{-1} \sum_{n=1}^{\infty} (\lambda z)^n \\ &= R[\nu_{\lambda^{-1}, \lambda}]. \end{aligned}$$

Hence by (4.12),

$$M^{\Pi \epsilon \Pi}[\alpha \Pi (C^{p,p})^* C^{p,p} \Pi](z) \boxtimes M[\nu_{\lambda^{-1}, \lambda}] = M[\nu_{\lambda^{-1}, \lambda}](\alpha \cdot) \boxtimes M[\nu_{\lambda^{-1}, \lambda}] = M[\delta_\alpha]. \quad \square$$

Corollary 4.16. *Let $p, d \in \mathbb{N}$, $A \in M_{p,d}(\mathbb{C})$, $\sigma \in \mathbb{R}$. Assume that $p \geq d$ and set $\lambda := d/p$. Then*

$$M[W_{\text{SPN}}^\square(A, \sigma)] \boxtimes f_\lambda = (M[A^* A] \boxtimes f_\lambda) \boxplus M[\delta_{\sigma^2/\lambda}].$$

Proof. By (4.13) and Proposition 4.11, the left-hand side is equal to

$$(M^{\Pi \epsilon \Pi}[\Pi \tilde{A}^* \tilde{A} \Pi] \boxtimes f_\lambda) \boxplus \left(M^{\Pi \epsilon \Pi} \left[\frac{\sigma^2}{\lambda} \Pi (C^{p,p})^* C^{p,p} \Pi \right] \boxtimes f_\lambda \right).$$

Now

$$M^{\Pi\epsilon\Pi}[\Pi\tilde{A}^*\tilde{A}\Pi] = \frac{1}{\tau(\Pi)}M[\tilde{A}^*\tilde{A}] = \frac{1}{\tau(\Pi)}\frac{d}{p}M[A^*A] = M[A^*A].$$

By Lemma 4.15, it holds that

$$M^{\Pi\epsilon\Pi}\left[\frac{\sigma^2}{\lambda}\Pi(C^{p,p})^*C^{p,p}\Pi\right] \boxtimes f_\lambda = M[\delta_{\sigma^2/\lambda}].$$

Hence the assertion holds. \square

Lemma 4.17. *Assume that $\alpha, \beta \in \mathbb{R}$, and $f, g \in \Xi$ satisfy*

$$f \boxplus M[\delta_\alpha] = g \boxplus M[\delta_\beta]. \quad (4.14)$$

Then

$$f \boxplus M[\delta_{\alpha-\beta}] = g. \quad (4.15)$$

Proof. Apply \boxtimes Zeta to both hand side of (4.14), then

$$\begin{aligned} R_f(z) + \alpha z &= R_g(z) + \beta z, \\ R_f(z) + (\alpha - \beta)z &= R_g(z). \end{aligned}$$

Applying \boxtimes Zeta to both hand side, we have (4.15). \square

Now we prove the second key lemma.

Lemma 4.18. *Let $p, d \in \mathbb{N}$, $\sigma, \rho \in \mathbb{R}$, and, A and $B \in M_{p,d}(\mathbb{C})$. Assume that $\sigma^2 \geq \rho^2$ and*

$$\mu_{\text{SPN}}^\square(A, \sigma) = \mu_{\text{SPN}}^\square(B, \rho).$$

Then

$$\mu_{\text{SPN}}^\square(A, \sqrt{\sigma^2 - \rho^2}) = \mu_{\text{SPN}}^\square(B, 0).$$

Proof. By Corollary 4.16 and the assumption, we have

$$(M_{A^*A} \boxtimes f_\lambda) \boxplus M[\delta_{\sigma^2/\lambda}] = (M_{B^*B} \boxtimes f_\lambda) \boxplus M[\delta_{\rho^2/\lambda}].$$

Thus by Lemma 4.17, it holds that

$$(M[A^*A] \boxtimes f_\lambda) \boxplus M[\delta_{(\sigma^2 - \rho^2)/\lambda}] = M[B^*B] \boxtimes f_\lambda.$$

By using Corollary 4.16 again, we have

$$M[\mu_{\text{SPN}}^\square(A, \sqrt{\sigma^2 - \rho^2})] \boxtimes f_\lambda = M[\mu_{\text{SPN}}^\square(B, 0)] \boxtimes f_\lambda.$$

Equivalently,

$$R[\mu_{\text{SPN}}^\square(A, \sqrt{\sigma^2 - \rho^2})] \boxtimes R[f_\lambda]^{-1} = R[\mu_{\text{SPN}}^\square(B, 0)] \boxtimes R[f_\lambda]^{-1}.$$

Applying \boxtimes $R[f_\lambda]$ \boxtimes Zeta to the both hand sides, we have

$$M[\mu_{\text{SPN}}^\square(A, \sqrt{\sigma^2 - \rho^2})] = M[\mu_{\text{SPN}}^\square(B, 0)].$$

Since any compactly supported probability measure is determined by its moments, the assertion holds. \square

4.2.5. Proof of Identifiability.

proof of Theorem 4.1. Without loss of generality, we may assume that $\sigma^2 \geq \rho^2$. Let $\mu_{\text{SPN}}^\square(A, \sigma) = \mu_{\text{SPN}}^\square(B, \rho)$. First, by Lemma 4.18, we have

$$\mu^\square(A, \sqrt{\sigma^2 - \rho^2}) = \mu^\square(B, 0).$$

Second, Lemma 4.4 implies $\sqrt{\sigma^2 - \rho^2} = 0$. Then $\mu_{A^*A} = \mu_{B^*B}$, which completes the proof. \square

5. ACKNOWLEDGEMENT

We would like to thank Hiroaki Yoshida for discussions. We appreciate Yuichi Ike's valuable comments on our manuscript.

REFERENCES

- [1] E. S. Allman, C. Matias, J. A. Rhodes, et al. Identifiability of parameters in latent structure models with many observed variables. *Ann. Stat.*, 37(6A):3099–3132, 2009.
- [2] S. Amari. *Information geometry and its applications*. Springer, 2016.
- [3] S. T. Belinschi, T. Mai, and R. Speicher. Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. *J. Reine Angew. Math.*, 2013.
- [4] Z. Burda, A. Jarosz, M. A. Nowak, J. Jurkiewicz, G. Papp, and I. Zahed. Applying free random variables to random matrix analysis of financial data. part I: The gaussian case. *Quant. Fin.*, 11(7):1103–1124, 2011.
- [5] R. Couillet, M. Debbah, and J. W. Silverstein. A deterministic equivalent for the analysis of correlated mimo multiple access channels. *IEEE Trans. on Inform. Theory*, 57(6):3493–3514, 2011.
- [6] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. Large information plus noise random matrix models and consistent subspace estimation in large sensor networks. *Random Matrices Theory Appl.*, 1(02):1150006, 2012.
- [7] A. Hasegawa, N. Sakuma, and H. Yoshida. Random matrices by MA models and compound free poisson laws. *Probab. Math. Statist*, 33(2):243–254, 2013.
- [8] T. Hayase. Cauchy noise loss for stochastic optimization of random matrix models via free deterministic equivalents. [arXiv:1804.03154 \[stat.ML\]](https://arxiv.org/abs/1804.03154), 2018.
- [9] J. A. Mingo and R. Speicher. *Free probability and random matrices*, volume 35. Springer, 2017.
- [10] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge Univ. Press, 2006.
- [11] Ø. Ryan and M. Debbah. Free deconvolution for signal processing applications. In *IEEE Trans. Inform. Theory*, pages 1846–1850. IEEE, 2007.
- [12] Ø. Ryan and M. Debbah. Multiplicative free convolution and information-plus-noise type matrices. [arXiv:math/0702342 \[math.PR\]](https://arxiv.org/abs/math/0702342), 2007.
- [13] R. Speicher. *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, volume 627. American Math. Soc., 1998.
- [14] R. Speicher and C. Vargas. Free deterministic equivalents, rectangular random matrix models, and operator-valued free probability theory. *Random Matrices Theory Appl.*, 1(02):1150008, 2012.
- [15] P. Vallet, P. Loubaton, and X. Mestre. Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case. *IEEE Tran. Inf. Theory*, 58(2):1043–1068, 2012.
- [16] D. V. Voiculescu. Symmetries of some reduced free product C^* -algebras. In *Operator algebras and their connections with topology and ergodic theory*, volume 1132, pages 556–588. Springer, Berlin, 1985.
- [17] D. V. Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF TOKYO, 3-8-1 KOMABA, MEGURO-KU, TOKYO, 153-8914, JAPAN

E-mail address: hayase@ms.u-tokyo.ac.jp