COMPUTING EXTREME SINGULAR VALUES OF FREE OPERATORS

EMRE PARMAKSIZ AND RAMON VAN HANDEL

Abstract. A recent development in random matrix theory, the intrinsic freeness principle, establishes that the spectrum of very general random matrices behaves as that of an associated free operator. This reduces the study of such random matrices to the deterministic problem of computing spectral statistics of the free operator. In the self-adjoint case, the spectral edges of the free operator can be computed exactly by means of a variational formula due to Lehner. In this note, we provide variational formulas for the largest and smallest singular values in the non-self-adjoint case.

1. Introduction

Let X be a $d \times d$ random matrix with jointly Gaussian entries. Without loss of generality, such a matrix can always be represented as

$$X = a_0 + \sum_{i=1}^{n} a_i g_i,$$

where $g_1, \ldots, g_n$ are i.i.d. standard Gaussian variables and $a_0, \ldots, a_n \in \mathbb{C}^{d\times d}$ are deterministic matrices. Since no structural assumptions are made, the entries of such matrices can be highly nonhomogeneous and dependent.

Nonetheless, recent advances in random marix theory make it possible to understand the spectrum of such matrices under surprisingly minimal assumptions. To this end, define the deterministic operator

$$x = a_0 \otimes \mathbf{1} + \sum_{i=1}^{n} a_i \otimes s_i, \tag{1.1}$$

where $s_1, \ldots, s_n$ is a free semicircular family (cf. [6] or [1, §4.1]). The *intrinsic freeness principle* [1, 2] establishes that under mild conditions, the spectral statistics of X and x nearly coincide: in the self-adjoint case, the spectra of X and x are close in the Hausdorff distance, while in the general case the same is true for the singular value spectrum. The universality principles of [3] further extend these conclusions to a large family of non-Gaussian random matrices.

The power of these results lies in the fact that they reduce the study of complicated random matrices to the deterministic problem of computing the spectrum of a free operator, which is accessible using free probability theory. In particular, when x is self-adjoint, the following variational principle due to Lehner [5] (see §2.2) provides an explicit formula for the spectral edges of x. In the following, we write $\lambda_{\max}(x) = \sup \mathrm{sp}(x)$ and $\lambda_{\min}(x) = \inf \mathrm{sp}(x) = -\lambda_{\max}(-x)$.

2020 *Mathematics Subject Classification.* 15B52; 46L53; 46L54.

Theorem 1.1 (Lehner)**.** *Let x be as in (1.1) with $a_0, \ldots, a_n \in \mathbb{C}^{d\times d}_{\mathrm{s.a.}}$. Then*

$$\lambda_{\max}(x) = \inf_{z>0} \lambda_{\max}\Bigg(a_0 + z^{-1} + \sum_{i=1}^n a_i z a_i\Bigg),$$

where $z \in \mathbb{C}^{d\times d}_{\mathrm{s.a.}}$. Moreover, the infimum can be restricted to those z such that the matrix in $\lambda_{\max}(\cdots)$ is a multiple of the identity.

Lehner's formula is extremely useful in applications that require a precise understanding of the spectral edges of nonhomogeneous and dependent random matrices. For example, several such applications can be found in [2].

When x is not self-adjoint, one is typically interested in the largest and smallest singular values $\mathrm{s}_{\max}(x) = \sup \mathrm{sp}((xx^*)^{1/2})$ and $\mathrm{s}_{\min}(x) = \inf \mathrm{sp}((xx^*)^{1/2})$, respectively. More generally, in problems related to sample covariance matrices, one is interested in the spectral edges of the operator $xx^* + b \otimes \mathbf{1}$ with $b \in \mathbb{C}^{d\times d}_{\mathrm{s.a.}}$. We emphasize at the outset that the computation of these quantities is in principle fully settled in [5, §5.2]: Lehner provides a recipe for computing the operator norm of *any* noncommutative quadratic polynomial of a free semicircular family with matrix coefficients, of which the above quantities are special cases. However, this general recipe gives rise to complicated and in some cases inexplicit variational formulas that prove to be difficult to use in concrete situations.

The aim of this note is to obtain much simpler explicit formulas, in the spirit of Theorem 1.1, for the spectral edges of the operator $xx^* + b \otimes \mathbf{1}$ in the case that x is centered (that is, $a_0 = 0$ in (1.1)). This setting may be viewed as a matrix-valued analogue of the free Poisson distribution, as will be explained in §5 below. Even though this is a special case of the more general problem considered by Lehner, it is one that arises frequently in applications, and tractable formulas for the spectral edges are essential for the analysis of concrete models.

1.1. **Main result.** The following setting will be considered throughout the remainder of this note. Fix $d, m, n \in \mathbb{N}$, deterministic matrices $a_1, \ldots, a_n \in \mathbb{C}^{d\times m}$, and a self-adjoint deterministic matrix $b \in \mathbb{C}^{d\times d}_{\mathrm{s.a.}}$. Define the operator

$$x = \sum_{i=1}^n a_i \otimes s_i, \tag{1.2}$$

where $s_1, \ldots, s_n$ is a free semicircular family.

Theorem 1.2. *For x as in (1.2), we have*

$$\lambda_{\max}(xx^* + b \otimes \mathbf{1}) = \inf_{\substack{z>0 \\ \sum_j a_j^* z a_j < \mathbf{1}}} \lambda_{\max}\Bigg(b + z^{-1} + \sum_{i=1}^n a_i \Bigg(\mathbf{1} - \sum_{j=1}^n a_j^* z a_j\Bigg)^{-1} a_i^*\Bigg)$$

and

$$\lambda_{\min}(xx^* + b \otimes \mathbf{1}) = \sup_{z<0} \lambda_{\min}\Bigg(b + z^{-1} + \sum_{i=1}^n a_i \Bigg(\mathbf{1} - \sum_{j=1}^n a_j^* z a_j\Bigg)^{-1} a_i^*\Bigg),$$

where $z \in \mathbb{C}^{d\times d}_{\mathrm{s.a.}}$. Moreover, the infimum (supremum) can be restricted to those z such that the matrix in $\lambda_{\max}(\cdots)$ ($\lambda_{\min}(\cdots)$) is a multiple of the identity.

As an immediate corollary, we obtain variational formulas for the largest and smallest singular values $\mathrm{s}_{\max}(x)$ and $\mathrm{s}_{\min}(x)$ of x.

Corollary 1.3. *For x as in (1.2), we have*

$$\mathrm{s}_{\max}(x)^2 = \inf_{\substack{z>0 \\ \sum_j a_j^* z a_j < \mathbf{1}}} \lambda_{\max}\left(z^{-1} + \sum_{i=1}^n a_i \left(\mathbf{1} - \sum_{j=1}^n a_j^* z a_j \right)^{-1} a_i^* \right)$$

and

$$\mathrm{s}_{\min}(x)^2 = \sup_{z<0} \lambda_{\min}\left(z^{-1} + \sum_{i=1}^n a_i \left(\mathbf{1} - \sum_{j=1}^n a_j^* z a_j \right)^{-1} a_i^* \right).$$

Proof. Apply Theorem 1.2 with $b=0$. ☐

1.2. **A reduction principle.** We now state a reduction principle that facilitates the application of Theorem 1.2 to models with symmetries. An analogous result in the setting of Theorem 1.1 appears in [2, Lemma 7.1].

Lemma 1.4. *Let $\mathcal{A}$ be a $*$-subalgebra of $\mathbb{C}^{m\times m}$ and $\mathcal{B}$ be a $*$-subalgebra of $\mathbb{C}^{d\times d}$ such that the following conditions hold:*

$$b \in \mathcal{B}, \qquad \sum_{i=1}^n a_i y a_i^* \in \mathcal{B} \quad \text{for all} \quad y \in \mathcal{A}, \qquad \sum_{j=1}^n a_j^* z a_j \in \mathcal{A} \quad \text{for all} \quad z \in \mathcal{B}.$$

Then for x as in (1.2), all the conclusions of Theorem 1.2 remain valid if the infimum (supremum) is taken only over $z \in \mathcal{B}$.

Let us illustrate this reduction principle in a special case that is important in applications. Let X be a $d\times m$ random matrix with independent centered Gaussian entries $X_{ij} \sim N(0,\sigma_{ij}^2)$. The associated free model x is given by

$$x = \sum_{i=1}^d \sum_{j=1}^m \sigma_{ij}\, e_i e_j^* \otimes s_{ij} \tag{1.3}$$

where $(s_{ij})_{i\in[d],j\in[m]}$ are free semicircular variables and $e_1, e_2, \ldots$ denote the standard basis vectors in $\mathbb{C}^d$ or $\mathbb{C}^m$. Then we have the following.

Corollary 1.5. *Let x have independent entries as in (1.3), and let b be a diagonal matrix with diagonal entries $b_1,\ldots,b_d$. Then we have*

$$\lambda_{\max}(xx^* + b\otimes \mathbf{1}) = \inf_{\substack{\min_j v_j > 0 \\ \max_j \sum_k \sigma_{kj}^2 v_k < 1}} \max_{i\in[d]} \left(b_i + \frac{1}{v_i} + \sum_{j\in[m]} \frac{\sigma_{ij}^2}{1 - \sum_{k\in[d]} \sigma_{kj}^2 v_k} \right)$$

and

$$\lambda_{\min}(xx^* + b\otimes \mathbf{1}) = \sup_{\max_j v_j < 0} \min_{i\in[d]} \left(b_i + \frac{1}{v_i} + \sum_{j\in[m]} \frac{\sigma_{ij}^2}{1 - \sum_{k\in[d]} \sigma_{kj}^2 v_k} \right),$$

where $v\in\mathbb{R}^d$. Moreover, the infimum (supremum) can be restricted to those v such that the vector in the maximum (minimum) is a multiple of the constant vector 1.

Proof. Let $\mathcal{A}$ and $\mathcal{B}$ be the $*$-algebras of diagonal matrices in $\mathbb{C}^{m\times m}$ and $\mathbb{C}^{d\times d}$, respectively. Then it is readily verified that the assumptions of Lemma 1.4 are satisfied, and the conclusion follows by restricting the variational principles of Theorem 1.2 to diagonal matrices z with diagonal entries $v_1,\ldots,v_d$. $\square$

An application of Corollary 1.5 to phase transitions of nonhomogeneous sample covariance matrices is developed in [2, §3.6 and §8.5].

1.3. **Organization of this note.** In §2, we recall some results of matrix algebra and the formulas of Lehner that will be used in the proofs. The first variational formula of Theorem 1.2 is proved in §3, and the second variational formula is proved in §4. The last statement of Theorem 1.2 and Lemma 1.4 are proved in §5.

2. Preliminaries

2.1. **Matrix algebra.** The following two theorems of matrix algebra will play an important role in our proofs. The first is a classical result on Schur complements.

Lemma 2.1 (Schur complements)**.** *Let $A\in\mathbb{C}^{d_1\times d_1}_{\mathrm{s.a.}}$, $B\in\mathbb{C}^{d_1\times d_2}$, $D\in\mathbb{C}^{d_2\times d_2}_{\mathrm{s.a.}}$, and*

$$M=\begin{bmatrix} A & B\\ B^* & D\end{bmatrix},$$

and assume that D is invertible. Define the Schur complement

$$M/D=A-BD^{-1}B^*.$$

Then

$$M>0 \qquad \text{if and only if} \qquad D>0 \quad \text{and} \quad M/D>0.$$

Moreover, if M and M/D are invertible, we have

$$M^{-1}=\begin{bmatrix} (M/D)^{-1} & -(M/D)^{-1}BD^{-1}\\ -D^{-1}B^*(M/D)^{-1} & D^{-1}+D^{-1}B^*(M/D)^{-1}BD^{-1}\end{bmatrix}.$$

The first statement can be found in [8, Theorem 1.12] and the second statement can be found in [8, Theorem 1.2].

The second result is a simple matrix inversion identity that is closely related to the Schur complement. Its correctness is readily verified algebraically.

Lemma 2.2 (Matrix inversion lemma)**.** *Let $B\in\mathbb{C}^{d_1\times d_2}$ and $D\in\mathbb{C}^{d_2\times d_2}_{\mathrm{s.a.}}$. Then*

$$(\mathbf{1}-BD^{-1}B^*)^{-1}=\mathbf{1}+B(D-B^*B)^{-1}B^*,$$

provided that all the inverses in this identity exist.

2.2. **Lehner's formulas.** Theorem 1.1 is essentially [5, Corollary 1.5], but it is phrased there in a slightly different manner. For completeness, we spell out the straightforward translation to the form given here.

Proof of Theorem 1.1. Choose $c>0$ so that $a_0+c\mathbf{1}\geq 0$ and $x+c\mathbf{1}\geq 0$ (this is possible since x is a bounded operator). By [5, Corollary 1.5], we have

$$\|x+c\mathbf{1}\|=\inf_{z>0}\left\|a_0+c\mathbf{1}+z^{-1}+\sum_{i=1}^n a_i z a_i\right\|,$$

where the infimum may also be restricted to those z such that the matrix inside the norm is a multiple of the identity. The conclusion follows by applying to both sides of this identity the elementary fact, that if y is self-adjoint and $y + c\mathbf{1} \geq 0$, then $\|y + c\mathbf{1}\| = \lambda_{\max}(y + c\mathbf{1}) = \lambda_{\max}(y) + c$. $\square$

For the proof of the second formula of Theorem 1.2, we will need a more general form Theorem 1.1. To state it, we must first recall a standard construction of a free semicircular family on the free Fock space, cf. [6, pp. 102–108]. Let

$$\mathcal{F}(\mathbb{C}^n) := \bigoplus_{k=0}^{\infty} (\mathbb{C}^n)^{\otimes k}$$

be the free Fock space over $\mathbb{C}^n$, where by convention $(\mathbb{C}^n)^{\otimes 0}$ denotes the one-dimensional Hilbert space spanned by the vacuum state Ω. For $i = 1, \ldots, n$, we define the creation operator l_i on $\mathcal{F}(\mathbb{C}^n)$ by

$$l_i(x_1 \otimes \cdots \otimes x_n) := e_i \otimes x_1 \otimes \cdots \otimes x_n,$$

where $e_1, \ldots, e_n$ is the coordinate basis in $\mathbb{C}^n$, and we define

$$s_i := l_i + l_i^*.$$

Then $s_1, \ldots, s_n$ is a free semicircular family. Moreover, we have the identities

$$l_i^* l_j = 1_{i=j}\mathbf{1}, \qquad \sum_{i=1}^{n} l_i l_i^* = \mathbf{1} - p_\Omega, \tag{2.1}$$

where p_Ω denotes the orthogonal projection on $(\mathbb{C}^n)^{\otimes 0}$.

We now state a generalization of Theorem 1.1 to self-adjoint linear combinations of the operators l_i, p_Ω with matrix coefficients.

Theorem 2.3 (Lehner)**.** *Let $a_0, b \in \mathbb{C}^{d\times d}_{\text{s.a.}}$ and $a_1, \ldots, a_n \in \mathbb{C}^{d\times d}$. Then*

$$\lambda_{\max}\left(a_0 \otimes \mathbf{1} + b \otimes p_\Omega + \sum_{i=1}^{n} \left(a_i \otimes l_i + a_i^* \otimes l_i^*\right)\right) = \inf_{\substack{z>0 \\ z^{-1}>b}} \lambda_{\max}\left(a_0 + z^{-1} + \sum_{i=1}^{n} a_i^* z a_i\right),$$

where $z \in \mathbb{C}^{d\times d}_{\text{s.a.}}$. Moreover, the infimum can be restricted to those z such that the matrix in $\lambda_{\max}(\cdots)$ is a multiple of the identity.

Proof. This follows from [5, Theorem 1.8] as in the above proof of Theorem 1.1 and with the substitution $z + b \leftarrow z^{-1}$. $\square$

We finally recall a classical dilation result that will be used in conjunction with Lehner's formulas; cf. [1, Lemma 4.9].

Lemma 2.4. *Let $y : H_1 \to H_2$ be a bounded linear operator between Hilbert spaces H_1 and H_2, and let $\tilde{y}$ be the bounded operator on $H_1 \oplus H_2$ defined by*

$$\tilde{y} = \begin{bmatrix} 0 & y^* \\ y & 0 \end{bmatrix}.$$

Then $\operatorname{sp}(\tilde{y}) \cup \{0\} = \operatorname{sp}((yy^*)^{1/2}) \cup -\operatorname{sp}((yy^*)^{1/2}) \cup \{0\}$.

3. The upper edge

The aim of this section is to prove the first variational formula of Theorem 1.2. This formula will be deduced from Theorem 1.1 by a linearization argument. As we will see, however, the reduction to the simple form given in Theorem 1.2 requires a careful analysis of the resulting optimization problem.

Let us note at the outset that as $\lambda_{\max}(y + c\mathbf{1}) = \lambda_{\max}(y) + c$ for any $c \in \mathbb{R}$, the variational formula in Theorem 1.2 is unchanged if we replace b by $b + c\mathbf{1}$. By choosing c large enough, we may assume without loss of generality that $b > 0$. This assumption will henceforth be made throughout this section. In particular, this enables us to write $b = a_0 a_0^*$ for some $a_0 \in \mathbb{C}^{d\times d}$.

We begin with a reduction to the setting of Theorem 1.1.

Lemma 3.1. *Let $r = m + 2d$ and*

$$\tilde{x} = \begin{bmatrix} 0 & 0 & a_0^* \otimes \mathbf{1} \\ 0 & 0 & x^* \\ a_0 \otimes \mathbf{1} & x & 0 \end{bmatrix} = \tilde{a}_0 + \sum_{i=1}^n \tilde{a}_i \otimes s_i,$$

where for $i = 1, \ldots, n$ we define the $r \times r$ matrices

$$\tilde{a}_0 = \begin{bmatrix} 0 & 0 & a_0^* \\ 0 & 0 & 0 \\ a_0 & 0 & 0 \end{bmatrix}, \qquad \tilde{a}_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a_i^* \\ 0 & a_i & 0 \end{bmatrix}.$$

Then $\lambda_{\max}(\tilde{x}) = \lambda_{\max}(xx^* + b \otimes \mathbf{1})^{1/2}$.

Proof. This follows directly from Lemma 2.4. $\square$

Theorem 1.1 immediately yields the following.

Corollary 3.2. *Define*

$$f(\tilde{z}) = \tilde{a}_0 + \tilde{z}^{-1} + \sum_{i=1}^n \tilde{a}_i \tilde{z} \tilde{a}_i.$$

Then we have

$$\lambda_{\max}(xx^* + b \otimes \mathbf{1})^{1/2} = \inf_{\tilde{z}>0} \lambda_{\max}(f(\tilde{z})) = \inf_{\lambda>0} \inf_{\substack{\tilde{z}>0 \\ f(\tilde{z})=\lambda\mathbf{1}}} \lambda_{\max}(f(\tilde{z})),$$

where the infimum is taken over $\tilde{z} \in \mathbb{C}^{r\times r}_{\text{s.a.}}$.

Next, we identify some consequences of the identity $f(\tilde{z}) = \lambda\mathbf{1}$.

Lemma 3.3. *Write $\tilde{z} \in \mathbb{C}^{r\times r}_{\text{s.a.}}$ in the block decomposition of Lemma 3.1 as*

$$\tilde{z} = \begin{bmatrix} p & v_1 & v_2 \\ v_1^* & q & w \\ v_2^* & w^* & z \end{bmatrix}.$$

Let $\lambda > 0$, and suppose that $\tilde{z} > 0$ and $f(\tilde{z}) = \lambda\mathbf{1}$. Then

$$z > 0, \qquad \sum_{j=1}^n a_j^* z a_j < \lambda\mathbf{1}, \qquad \lambda^{-1} b + z^{-1} + \sum_{i=1}^n a_i \left(\lambda\mathbf{1} - \sum_{j=1}^n a_j^* z a_j\right)^{-1} a_i^* \leq \lambda\mathbf{1}.$$

Proof. We begin by noting that

$$f(\tilde{z}) = \begin{bmatrix} 0 & 0 & a_0^* \\ 0 & 0 & 0 \\ a_0 & 0 & 0 \end{bmatrix} + \tilde{z}^{-1} + \sum_{i=1}^{n} \begin{bmatrix} 0 & 0 & 0 \\ 0 & a_i^* z a_i & a_i^* w^* a_i^* \\ 0 & a_i w a_i & a_i q a_i^* \end{bmatrix}.$$

To compute $\tilde{z}^{-1}$, we express

$$\tilde{z} = \begin{bmatrix} p & V \\ V^* & Z \end{bmatrix}, \qquad Z = \begin{bmatrix} q & w \\ w^* & z \end{bmatrix}, \qquad V = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$$

and apply Lemma 2.1. Using that $f(\tilde{z}) = \lambda\mathbf{1}$ yields

$$\begin{aligned} &(\tilde{z}/Z)^{-1} = \lambda\mathbf{1}, \\ &\begin{bmatrix} 0 & a_0^* \end{bmatrix} - (\tilde{z}/Z)^{-1} V Z^{-1} = 0, \\ &Z^{-1} + Z^{-1} V^* (\tilde{z}/Z)^{-1} V Z^{-1} + \sum_{i=1}^{n} \begin{bmatrix} a_i^* z a_i & a_i^* w^* a_i^* \\ a_i w a_i & a_i q a_i^* \end{bmatrix} = \lambda\mathbf{1}. \end{aligned}$$

Plugging the first two equations into the last equation yields

$$Z^{-1} + \frac{1}{\lambda} \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix} + \sum_{i=1}^{n} \begin{bmatrix} a_i^* z a_i & a_i^* w^* a_i^* \\ a_i w a_i & a_i q a_i^* \end{bmatrix} = \lambda\mathbf{1},$$

where we used that $b = a_0 a_0^*$. We note that the above equations are all well defined, as $\tilde{z} > 0$ ensures that $Z > 0$ and $\tilde{z}/Z > 0$ by Lemma 2.1.

Next, we compute Z^{-1} using Lemma 2.1 and plug the resulting expression into the last equation display. Then we obtain

$$(Z/z)^{-1} + \sum_{i=1}^{n} a_i^* z a_i = \lambda\mathbf{1}, \tag{3.1}$$

$$z^{-1} + z^{-1} w^* (Z/z)^{-1} w z^{-1} + \lambda^{-1} b + \sum_{i=1}^{n} a_i q a_i^* = \lambda\mathbf{1}. \tag{3.2}$$

Moreover, as $Z > 0$, we have $z > 0$ and $Z/z > 0$. This yields the desired inequality $\lambda\mathbf{1} - \sum_i a_i^* z a_i = (Z/z)^{-1} > 0$ by (3.1). Finally, by (3.1) and the definition of Z/z,

$$q \geq q - w z^{-1} w^* = Z/z = \left(\lambda\mathbf{1} - \sum_{i=1}^{n} a_i^* z a_i \right)^{-1},$$

and thus the last inequality in the statement follows from (3.2). □

The idea of the proof is now to show that given any $\tilde{z}$ as in Lemma 3.3, we can find a matrix with a smaller value for the variational principle of Corollary 3.2.

Lemma 3.4. *Let $\lambda > 0$ and $z \in \mathbb{C}^{d\times d}_{\text{s.a.}}$ such that $z > 0$ and $\sum_j a_j^* z a_j < \lambda\mathbf{1}$. Define*

$$g(\lambda; z) = \begin{bmatrix} \lambda^{-1}\mathbf{1} + \lambda^{-2} a_0^* z a_0 & 0 & \lambda^{-1} a_0^* z \\ 0 & \left(\lambda\mathbf{1} - \sum_j a_j^* z a_j\right)^{-1} & 0 \\ \lambda^{-1} z a_0 & 0 & z \end{bmatrix}.$$

Then $g(\lambda;z)>0$ and

$$f(g(\lambda;z))=\begin{bmatrix}\lambda\mathbf{1} & 0 & 0\\ 0 & \lambda\mathbf{1} & 0\\ 0 & 0 & \lambda^{-1}b+z^{-1}+\sum_i a_i\big(\lambda\mathbf{1}-\sum_j a_j^*za_j\big)^{-1}a_i^*\end{bmatrix}.$$

In particular, if $\tilde{z}$ is as in Lemma 3.3, then $f(g(\lambda;z))\le f(\tilde{z})$.

Proof. Express $\hat{z}=g(\lambda;z)$ in block form as

$$\hat{z}=\begin{bmatrix}\lambda^{-1}\mathbf{1}+\lambda^{-2}a_0^*za_0 & W\\ W^* & Y\end{bmatrix}$$

with

$$Y=\begin{bmatrix}\big(\lambda\mathbf{1}-\sum_j a_j^*za_j\big)^{-1} & 0\\ 0 & z\end{bmatrix},\qquad W=\begin{bmatrix}0 & \lambda^{-1}a_0^*z\end{bmatrix}.$$

Then $Y>0$ and $\hat{z}/Y=\lambda^{-1}\mathbf{1}>0$, so that $\hat{z}>0$ by Lemma 2.1.

We can further use Lemma 2.1 to compute

$$\hat{z}^{-1}=\begin{bmatrix}\lambda\mathbf{1} & 0 & -a_0^*\\ 0 & \lambda\mathbf{1}-\sum_j a_j^*za_j & 0\\ -a_0 & 0 & \lambda^{-1}b+z^{-1}\end{bmatrix}.$$

The identity for $f(\hat{z})$ now follows readily. Finally, if $\tilde{z}$ is as in Lemma 3.3, we obtain $f(\hat{z})\le\lambda\mathbf{1}=f(\tilde{z})$ using the last inequality of Lemma 3.3. $\square$

We can now prove the first part of Theorem 1.2.

Proposition 3.5. *For x as in (1.2), we have*

$$\lambda_{\max}(xx^*+b\otimes\mathbf{1})=\inf_{\substack{z>0\\ \sum_j a_j^*za_j<\mathbf{1}}}\lambda_{\max}\left(b+z^{-1}+\sum_{i=1}^n a_i\left(\mathbf{1}-\sum_{j=1}^n a_j^*za_j\right)^{-1}a_i^*\right).$$

Proof. Corollary 3.2 and Lemma 3.4 yield

$$\begin{aligned}\lambda_{\max}(xx^*+b\otimes\mathbf{1})^{1/2}&=\inf_{\lambda>0}\inf_{\substack{\tilde{z}>0\\ f(\tilde{z})=\lambda\mathbf{1}}}\lambda_{\max}(f(\tilde{z}))\\ &\ge\inf_{\lambda>0}\inf_{\substack{z>0\\ \sum_j a_j^*za_j<\lambda\mathbf{1}}}\lambda_{\max}(f(g(\lambda;z)))\\ &\ge\inf_{\tilde{z}>0}\lambda_{\max}(f(\tilde{z}))=\lambda_{\max}(xx^*+b\otimes\mathbf{1})^{1/2}.\end{aligned}$$

Therefore, the expression for $f(g(\lambda;z))$ in Lemma 3.4 yields

$$\lambda_{\max}(xx^*+b\otimes\mathbf{1})^{1/2}=\inf_{\lambda>0}\inf_{\substack{z>0\\ \sum_j a_j^*za_j<\lambda\mathbf{1}}}\lambda_{\max}(f(g(\lambda;z)))=$$

$$\inf_{\lambda>0}\inf_{\substack{z>0\\ \sum_j a_j^*za_j<\lambda\mathbf{1}}}\max\left\{\lambda,\ \lambda_{\max}\left(\lambda^{-1}b+z^{-1}+\sum_i a_i\left(\lambda\mathbf{1}-\sum_{j=1}^n a_j^*za_j\right)^{-1}a_i^*\right)\right\}.$$

The conclusion follows readily by replacing $z\leftarrow\lambda z$ in the inner infimum and using that $\inf_{\lambda>0}\max\{\lambda,\lambda^{-1}c\}=c^{1/2}$ for $c\ge0$. $\square$

4. The lower edge

The aim of this section is to prove the second variational formula of Theorem 1.2. The basic structure of the proof is similar to that of the first variational formula. However, the complication that arises in the present case is that we must rely on a more involved linearization argument that appears in [5, §5.2].

To this end, define the operator

$$L = \begin{bmatrix} \mathbf{1}\otimes p_\Omega & \mathbf{1}\otimes l_1^* & \mathbf{1}\otimes l_1 & \cdots & \mathbf{1}\otimes l_n^* & \mathbf{1}\otimes l_n \end{bmatrix}$$

in $\mathbb{C}^{1\times(2n+1)}\otimes\mathbb{C}^{d\times d}\otimes B(\mathcal{F}(\mathbb{C}^n))$, and define the matrix

$$T = \begin{bmatrix} \frac{a_0a_0^*}{n+1} & 0 & \cdots & 0 \\ 0 & \frac{a_0a_0^*}{n+1} - a_1a_1^* & & -a_1a_n^* \\ 0 & -a_1a_1^* & & -a_1a_n^* \\ \vdots & & \ddots & \vdots \\ 0 & -a_na_1^* & & -a_na_n^* \\ 0 & -a_na_1^* & \cdots & \frac{a_0a_0^*}{n+1} - a_na_n^* \end{bmatrix}$$

in $\mathbb{C}^{(2n+1)\times(2n+1)}\otimes\mathbb{C}^{d\times d}$, where $a_0\in\mathbb{C}^{d\times d}$ will be chosen shortly. Then (2.1) yields

$$L(T\otimes\mathbf{1})L^* = a_0a_0^*\otimes\mathbf{1} - xx^*.$$

Now recall, as was explained at the beginning of §3, that the variational formula that we aim to prove is unchanged if we replace b by $b+c\mathbf{1}$ for any $c\in\mathbb{R}$. By choosing c to be sufficiently negative, we may assume without loss of generality that we can write $-b = a_0a_0^* > 0$ for some $a_0\in\mathbb{C}^{d\times d}$, and that moreover T is positive definite. This is assumed henceforth throughout this section.

Since T is positive definite, it can be expressed as $T = RR^*$ where

$$R = \begin{bmatrix} \frac{a_0}{\sqrt{n+1}} & 0 \\ 0 & B \end{bmatrix}, \qquad B = \begin{bmatrix} M_1 \\ N_1 \\ \vdots \\ M_n \\ N_n \end{bmatrix}$$

for some matrix $B\in\mathbb{C}^{2nd\times 2nd}$ that we expressed in block form with blocks $M_i, N_i\in\mathbb{C}^{d\times 2nd}$. Then we obtain the following analogue of Corollary 3.2.

Lemma 4.1. *Let $s = 2(n+1)d$, define the $s\times s$ matrices*

$$A_0 = \frac{1}{\sqrt{n+1}}\begin{bmatrix} 0 & 0 & a_0^* \\ 0 & 0 & 0 \\ a_0 & 0 & 0 \end{bmatrix}, \qquad A_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & M_i^* \\ 0 & N_i & 0 \end{bmatrix}$$

for $i = 1,\ldots,n$, and define

$$F(Z) = Z^{-1} + \sum_{i=1}^n A_i^* Z A_i.$$

Then

$$\big(-\lambda_{\min}(xx^* + b\otimes \mathbf{1})\big)^{1/2} = \inf_{\substack{Z>0\\ Z^{-1}>A_0}} \lambda_{\max}(F(Z)) = \inf_{\lambda>0} \inf_{\substack{Z>0\\ Z^{-1}>A_0\\ F(Z)=\lambda\mathbf{1}}} \lambda_{\max}(F(Z)).$$

Proof. Note that we can write

$$A_0\otimes p_\Omega + \sum_{i=1}^n \big(A_i\otimes l_i + A_i^*\otimes l_i^*\big) = \begin{bmatrix} 0 & (R\otimes \mathbf{1})^*L^* \\ L(R\otimes\mathbf{1}) & 0\end{bmatrix},$$

so that Lemma 2.4 yields

$$\lambda_{\max}\Bigg(A_0\otimes p_\Omega + \sum_{i=1}^n \big(A_i\otimes l_i + A_i^*\otimes l_i^*\big)\Bigg) = \\ \lambda_{\max}\big(L(T\otimes\mathbf{1})L^*\big)^{1/2} = \big(-\lambda_{\min}(xx^*+b\otimes\mathbf{1})\big)^{1/2}.$$

The conclusion now follows from Theorem 2.3. □

We can now proceed in a similar manner as in §3.

Lemma 4.2. *Write $Z\in\mathbb{C}^{s\times s}_{\mathrm{s.a.}}$ in the block decomposition of Lemma 4.1 as*

$$Z = \begin{bmatrix} P & V_1 & V_2 \\ V_1^* & Q & W \\ V_2^* & W^* & z\end{bmatrix}.$$

Let $\lambda>0$, and suppose that $Z>0$, $Z^{-1}>A_0$, and $f(Z)=\lambda\mathbf{1}$. Then

$$z^{-1} > \frac{1}{\lambda}\frac{a_0a_0^*}{n+1}, \qquad \sum_{j=1}^n N_j^* z N_j < \lambda\mathbf{1},$$

and

$$z^{-1} + \sum_{i=1}^n M_i\Bigg(\lambda\mathbf{1} - \sum_{j=1}^n N_j^*zN_j\Bigg)^{-1} M_i^* \le \lambda\mathbf{1}.$$

Proof. We begin by noting that

$$F(Z) = Z^{-1} + \sum_{i=1}^n \begin{bmatrix} 0 & 0 & 0 \\ 0 & N_i^*zN_i & N_i^*W^*M_i^* \\ 0 & M_iWN_i & M_iQM_i^*\end{bmatrix}.$$

To compute Z^{-1}, we express

$$Z = \begin{bmatrix} P & V \\ V^* & Z_0\end{bmatrix}, \qquad Z_0 = \begin{bmatrix} Q & W \\ W^* & z\end{bmatrix}, \qquad V = \begin{bmatrix} V_1 & V_2\end{bmatrix}$$

and apply Lemma 2.1. Using that $F(Z)=\lambda\mathbf{1}$ yields $V=0$, $P=\lambda^{-1}\mathbf{1}$, and

$$Z_0^{-1} + \sum_{i=1}^n \begin{bmatrix} N_i^*zN_i & N_i^*W^*M_i^* \\ M_iWN_i & M_iQM_i^*\end{bmatrix} = \lambda\mathbf{1},$$

where we note that $Z>0$ implies that $Z_0>0$.

Next, computing Z_0^{-1} using Lemma 2.1, we obtain

$$(Z_0/z)^{-1} + \sum_{j=1}^{n} N_j^* z N_j = \lambda \mathbf{1}, \tag{4.1}$$

$$z^{-1} + z^{-1} W^* (Z_0/z)^{-1} W z^{-1} + \sum_{i=1}^{n} M_i Q M_i^* = \lambda \mathbf{1}. \tag{4.2}$$

with $z > 0$ and $Z_0/z > 0$. This yields the desired inequality $\lambda \mathbf{1} - \sum_{j=1}^{n} N_j^* z N_j = (Z_0/z)^{-1} > 0$ by (4.1). Moreover, by (4.1) and the definition of Z_0/z,

$$Q \geq Q - W z^{-1} W^* = Z_0/z = \left(\lambda \mathbf{1} - \sum_{j=1}^{n} N_j^* z N_j \right)^{-1},$$

and thus the last inequality in the statement follows from (4.2).

Finally, note that as $P = \lambda^{-1}\mathbf{1}$ and $V = 0$, the condition $Z^{-1} > A_0$ becomes

$$\begin{bmatrix} \lambda \mathbf{1} & C^* \\ C & Z_0^{-1} \end{bmatrix} > 0, \qquad C = -\frac{1}{\sqrt{n+1}} \begin{bmatrix} 0 \\ a_0 \end{bmatrix},$$

so that Lemma 2.1 implies that $\lambda \mathbf{1} - C^* Z_0 C > 0$. Thus $\frac{a_0^* z a_0}{n+1} < \lambda \mathbf{1}$. As we assumed that $a_0 a_0^* > 0$, the matrix a_0 is invertible, and thus taking the inverse of both sides and rearranging yields the desired inequality $z^{-1} > \frac{1}{\lambda} \frac{a_0 a_0^*}{n+1}$. □

The above lemma gives rise to the following value reduction principle.

Lemma 4.3. *Let $\lambda > 0$, $z \in \mathbb{C}^{d\times d}_{\text{s.a.}}$ with $z^{-1} > \frac{1}{\lambda}\frac{a_0 a_0^*}{n+1}$, $\sum_{j=1}^{n} N_j^* z N_j < \lambda \mathbf{1}$. Define*

$$G(\lambda; z) = \begin{bmatrix} \lambda^{-1}\mathbf{1} & 0 & 0 \\ 0 & \left(\lambda \mathbf{1} - \sum_{j=1}^{n} N_j^* z N_j\right)^{-1} & 0 \\ 0 & 0 & z \end{bmatrix}.$$

Then $G(\lambda; z) > 0$, $G(\lambda, z)^{-1} > A_0$, and

$$F(G(\lambda; z)) = \begin{bmatrix} \lambda \mathbf{1} & 0 & 0 \\ 0 & \lambda \mathbf{1} & 0 \\ 0 & 0 & z^{-1} + \sum_{i=1}^{n} M_i \left(\lambda \mathbf{1} - \sum_{j=1}^{n} N_j^* z N_j\right)^{-1} M_i^* \end{bmatrix}.$$

In particular, if Z is as in Lemma 4.2, then $F(G(\lambda; z)) \leq F(Z)$.

Proof. The only conclusion of the lemma that does not follow immediately from the definitions is $G(\lambda, z)^{-1} > A_0$, that is, that

$$\begin{bmatrix} \lambda \mathbf{1} & 0 & -\frac{a_0^*}{\sqrt{n+1}} \\ 0 & \lambda \mathbf{1} - \sum_{j=1}^{n} N_j^* z N_j & 0 \\ -\frac{a_0}{\sqrt{n+1}} & 0 & z^{-1} \end{bmatrix} > 0.$$

By Lemma 2.1, this is equivalent to

$$\lambda \mathbf{1} - \frac{1}{n+1} \begin{bmatrix} 0 & a_0^* \end{bmatrix} \begin{bmatrix} \lambda \mathbf{1} - \sum_{j=1}^{n} N_j^* z N_j & 0 \\ 0 & z^{-1} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ a_0 \end{bmatrix} > 0,$$

which is further equivalent to the assumption that $z^{-1} > \frac{1}{\lambda}\frac{a_0 a_0^*}{n+1}$ by the argument given at the end of the proof of Lemma 4.2. ☐

We readily deduce the following.

Corollary 4.4. *For x as in (1.2), we have*

$$-\lambda_{\min}(xx^* + b\otimes \mathbf{1}) = \inf_{\substack{z^{-1} > \frac{a_0 a_0^*}{n+1} \\ \sum_j N_j^* z N_j < \mathbf{1}}} \lambda_{\max}\left(z^{-1} + \sum_{i=1}^n M_i \left(\mathbf{1} - \sum_{j=1}^n N_j^* z N_j\right)^{-1} M_i^*\right).$$

Proof. Apply Lemmas 4.1 and 4.3 exactly as in the proof of Proposition 3.5. ☐

Unlike in §3, the complication that now arises is that Corollary 4.4 depends only implicitly on the matrices $a_1, \ldots, a_n, b$ that define the operator on the left-hand side. We must therefore further simplify the variational principle to arrive at the explicit form that is stated in Theorem 1.2. To this end, define the matrices

$$N = \begin{bmatrix} N_1 \\ \vdots \\ N_n \end{bmatrix}, \qquad A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

in $\mathbb{C}^{n\times 1} \otimes \mathbb{C}^{d\times 2nd}$ and $\mathbb{C}^{n\times 1}\otimes \mathbb{C}^{d\times m}$, respectively. Then we have the following.

Lemma 4.5. *Let $z^{-1} > \frac{a_0 a_0^*}{n+1}$. Then $\sum_j N_j^* z N_j < \mathbf{1}$ and*

$$\sum_{i=1}^n M_i \left(\mathbf{1} - \sum_{j=1}^n N_j^* z N_j\right)^{-1} M_i^* =$$

$$\frac{n}{n+1} a_0 a_0^* - \sum_{i=1}^n a_i \left(\mathbf{1} + \sum_{j=1}^n a_j^* \left(z^{-1} - \frac{a_0 a_0^*}{n+1}\right)^{-1} a_j\right)^{-1} a_i^*.$$

Proof. We first note that $M_i N_j^* = -a_i a_j^*$ and $M_i M_j^* = N_i N_j^* = 1_{i=j}\frac{a_0 a_0^*}{n+1} - a_i a_j^*$ by the definition of M_i, N_i. We can therefore write

$$\mathbf{1}\otimes z^{-1} - NN^* = \mathbf{1}\otimes\left(z^{-1} - \frac{a_0 a_0^*}{n+1}\right) + AA^* > 0,$$

which yields the desired inequality

$$\begin{aligned}\sum_{j=1}^n N_j^* z N_j &= N^*(\mathbf{1}\otimes z)N = N^*\big(\mathbf{1}\otimes z^{-1} - NN^* + NN^*\big)^{-1} N \\ &= \mathbf{1} - \big(\mathbf{1} + N^*\big(\mathbf{1}\otimes z^{-1} - NN^*\big)^{-1} N\big)^{-1} < \mathbf{1}\end{aligned}$$

by Lemma 2.2. Next, we rearrange the previous equation display as

$$\left(\mathbf{1} - \sum_{j=1}^n N_j^* z N_j\right)^{-1} = \mathbf{1} + N^*\big(\mathbf{1}\otimes z^{-1} - NN^*\big)^{-1} N.$$

As $M_i N^* = a_i A^*$ and $M_i M_i^* = \frac{a_0 a_0^*}{n+1} - a_i a_i^*$, we can compute

$$\sum_{i=1}^n M_i \left(\mathbf{1} - \sum_{j=1}^n N_j^* z N_j\right)^{-1} M_i^* =$$
$$\frac{n}{n+1} a_0 a_0^* + \sum_{i=1}^n a_i \left[A^* \left(\mathbf{1} \otimes \left(z^{-1} - \frac{a_0 a_0^*}{n+1} \right) + AA^* \right)^{-1} A - \mathbf{1} \right] a_i^*.$$

Applying Lemma 2.2 again yields the conclusion. □

We can now prove the second part of Theorem 1.2.

Proposition 4.6. *For x as in (1.2), we have*

$$\lambda_{\min}(xx^* + b \otimes \mathbf{1}) = \sup_{z<0} \lambda_{\min}\left(b + z^{-1} + \sum_{i=1}^n a_i \left(\mathbf{1} - \sum_{j=1}^n a_j^* z a_j \right)^{-1} a_i^* \right).$$

Proof. This follows readily from Corollary 4.4, Lemma 4.5, and $a_0 a_0^* = -b$, where we make the substitution $z^{-1} - \frac{a_0 a_0^*}{n+1} \leftarrow -z^{-1}$. □

5. The matrix Cauchy transform

It remains to prove the last part of Theorem 1.2, viz., that the variational principles can be restricted to z such that the matrix on the right-hand side is a multiple of the identity. This will be deduced from another result that is useful in its own right. Lemma 1.4 will follow as a byproduct of the proof.

Define an analytic function

$$G : \mathbb{C} \backslash \mathrm{sp}(xx^* + b \otimes \mathbf{1}) \to \mathbb{C}^{d \times d},$$

the *matrix Cauchy transform* of $xx^* + b \otimes \mathbf{1}$, by

$$G(\lambda) := (\mathrm{id} \otimes \tau)\Big[\big(\lambda - xx^* - b \otimes \mathbf{1} \big)^{-1} \Big].$$

Here τ is the trace on the C^*-probability space $(\mathcal{A}, \tau)$ in which the free semicircular family $s_1, \ldots, s_n$ is defined. The following is the main result of this section.

Theorem 5.1. *For every $\lambda \in \mathbb{C} \backslash \mathrm{conv}(\mathrm{sp}(xx^* + b \otimes \mathbf{1}))$, we have*

$$b + G(\lambda)^{-1} + \sum_{i=1}^n a_i \left(\mathbf{1} - \sum_{j=1}^n a_j^* G(\lambda) a_j \right)^{-1} a_i^* = \lambda \mathbf{1}.$$

In particular, both inverses in this equation exist.

In the case that $d = 1$ (that is, when the coefficients a_i, b are scalar), the equation in Theorem 5.1 reduces to the well known equation for the Cauchy transform of the free Poisson distribution [6, pp. 203–206]. This explains the assertion made in the introduction that the operators considered in this note may be viewed as matrix-valued analogues of the free Poisson distribution.

Proof of Theorem 5.1. When $|\lambda|$ is sufficiently large that $g(\lambda) = (\lambda - b \otimes \mathbf{1})^{-1}$ satisfies $\|g(\lambda)\|\|xx^*\| < 1$, $G(\lambda)$ has a convergent power series expansion

$$G(\lambda) = \sum_{k=0}^{\infty} (\mathrm{id} \otimes \tau)\Big[g(\lambda)\big(xx^* g(\lambda)\big)^k\Big].$$

Let us fix such a λ until further notice, and define

$$C_k = (\mathrm{id} \otimes \tau)\Big[\big(xx^* g(\lambda)\big)^k\Big],$$
$$D_k = (\mathrm{id} \otimes \tau)\Big[x^* g(\lambda)\big(xx^* g(\lambda)\big)^k x\Big].$$

Reasoning as in the proof of [7, Lemma 4.4] and using that the trace of a product of an odd number of semicircular variables vanishes, we obtain the recursions

$$C_{k+1} = \sum_{i=1}^{n} a_i a_i^* (\lambda - b)^{-1} C_k + \sum_{l=0}^{k-1} \sum_{i=1}^{n} a_i D_l a_i^* (\lambda - b)^{-1} C_{k-1-l},$$
$$D_k = \sum_{i=1}^{n} a_i^* (\lambda - b)^{-1} C_k a_i + \sum_{l=0}^{k-1} \sum_{i=1}^{n} a_i^* (\lambda - b)^{-1} C_l a_i D_{k-1-l}$$

for $k \geq 0$ with initial condition $C_0 = \mathbf{1}$. Summing these recursions over k yields

$$(\lambda - b)G(\lambda) = \mathbf{1} + \sum_{i=1}^{n} a_i H(\lambda) a_i^* G(\lambda), \tag{5.1}$$
$$H(\lambda) = \mathbf{1} + \sum_{i=1}^{n} a_i^* G(\lambda) a_i H(\lambda), \tag{5.2}$$

where we define

$$H(\lambda) := \mathbf{1} + \sum_{k=0}^{\infty} D_k = \mathbf{1} + (\mathrm{id} \otimes \tau)\Big[x^* \big(\lambda - xx^* - b \otimes \mathbf{1}\big)^{-1} x\Big].$$

We have therefore established the validity of (5.1)–(5.2) for all $\lambda \in \mathbb{C}$ with $|\lambda|$ sufficiently large. However, since both G and H are analytic, the validity of these equations extends to every $\lambda \in \mathbb{C} \backslash \mathrm{sp}(xx^* + b \otimes \mathbf{1})$.

Now note that as $\mathrm{Im}\, \frac{1}{\lambda - t} = -\frac{\mathrm{Im}\,\lambda}{|\lambda - t|^2}$ for $t \in \mathbb{R}$, it follows from the definitions of $G(\lambda)$ and $H(\lambda)$ that $\mathrm{Im}\, G(\lambda)$ and $\mathrm{Im}\, H(\lambda)$ are negative definite when $\mathrm{Im}\,\lambda > 0$ and positive definite when $\mathrm{Im}\,\lambda < 0$. Thus $G(\lambda)$ and $H(\lambda)$ are invertible whenever $\mathrm{Im}\,\lambda \neq 0$ by [4, Lemma 3.1]. Consequently, (5.2) yields

$$H(\lambda) = \left(\mathbf{1} - \sum_{j=1}^{n} a_j^* G(\lambda) a_j\right)^{-1}, \tag{5.3}$$

and substituting this expression into (5.1) and multiplying on the right by $G(\lambda)^{-1}$ yields the conclusion in the case that $\mathrm{Im}\,\lambda \neq 0$.

It remains to consider the cases where λ is real. If $\lambda > \lambda_{\max}(xx^* + b \otimes \mathbf{1})$, then $G(\lambda)$ and $H(\lambda)$ are positive definite and hence invertible by their definition, and the proof is concluded as above. If $\lambda < \lambda_{\min}(xx^* + b \otimes \mathbf{1})$, then $G(\lambda)$ is negative definite and thus $\mathbf{1} - \sum_j a_j^* G(\lambda) a_j$ is positive definite. Together with (5.2), this implies that $G(\lambda)$ and $H(\lambda)$ are invertible, and we again conclude as above. $\square$

We can now conclude the proof of Theorem 1.2.

Proof of Theorem 1.2. The two variational principles were proved above in Propositions 3.5 and 4.6, respectively. It is convenient to write

$$h(z) = b + z^{-1} + \sum_{i=1}^{n} a_i \left(\mathbf{1} - \sum_{j=1}^{n} a_j^* z a_j \right)^{-1} a_i^*,$$

so that the two variational principles may by expressed as

$$\lambda_{\max}(xx^* + b \otimes \mathbf{1}) = \inf_{\substack{z>0 \\ \sum_j a_j^* z a_j < \mathbf{1}}} \lambda_{\max}(h(z)),$$
$$\lambda_{\min}(xx^* + b \otimes \mathbf{1}) = \sup_{z<0} \lambda_{\min}(h(z)).$$

It remains to prove the last assertion of the theorem.

Let us first consider the second variational principle. Clearly

$$\lambda_{\min}(xx^* + b \otimes \mathbf{1}) = \sup_{z<0} \lambda_{\min}(h(z)) \geq \sup_{\lambda \in \mathbb{R}} \sup_{\substack{z<0 \\ h(z)=\lambda\mathbf{1}}} \lambda_{\min}(h(z)),$$

since we restrict the supremum to a smaller set. On the other hand, fix any $\varepsilon > 0$ and let $\mu = \lambda_{\min}(xx^* + b \otimes \mathbf{1}) - \varepsilon$. Then $G(\mu) < 0$ and $h(G(\mu)) = \mu\mathbf{1}$ by the definition of the matrix Cauchy transform and Theorem 5.1, respectively. Therefore

$$\sup_{\lambda \in \mathbb{R}} \sup_{\substack{z<0 \\ h(z)=\lambda\mathbf{1}}} \lambda_{\min}(h(z)) \geq \lambda_{\min}(h(G(\mu))) = \lambda_{\min}(xx^* + b \otimes \mathbf{1}) - \varepsilon.$$

Letting $\varepsilon \downarrow 0$ shows that

$$\lambda_{\min}(xx^* + b \otimes \mathbf{1}) = \sup_{\lambda \in \mathbb{R}} \sup_{\substack{z<0 \\ h(z)=\lambda\mathbf{1}}} \lambda_{\min}(h(z)),$$

which is the desired conclusion.

For the first variational principle, let $\varepsilon > 0$ and $\mu = \lambda_{\max}(xx^* + b \otimes \mathbf{1}) + \varepsilon$. Then $G(\mu) > 0$ and $h(G(\mu)) = \mu\mathbf{1}$ by the definition of the matrix Cauchy transform and Theorem 5.1. Moreover, that $\sum_j a_j^* G(\mu) a_j < \mathbf{1}$ follows from (5.3) since $H(\mu) > 0$ by its definition. The proof for the first variational principle can now be completed in the identical manner as for the second variational principle. $\square$

We conclude by proving Lemma 1.4. A similar result in the setting of Theorem 1.1 appears in [2, Lemma 7.1], but the proof given here is entirely different.

Proof of Lemma 1.4. Under the assumptions of Lemma 1.4, it is readily verified from the recursions in the proof of Theorem 5.1 that $C_k \in \mathcal{B}$ and $D_k \in \mathcal{A}$ for all $k \geq 0$. Thus $G(\lambda) \in \mathcal{B}$ and $H(\lambda) \in \mathcal{A}$ for all $\lambda \in \mathbb{C} \backslash \mathrm{sp}(xx^* + b \otimes \mathbf{1})$. The rest of the proof now proceeds exactly as in the proof of the last part of Theorem 1.2. $\square$

Acknowledgments. RvH was supported in part by NSF grant DMS-2347954.

References

[1] A. S. Bandeira, M. T. Boedihardjo, and R. van Handel. Matrix concentration inequalities and free probability. *Invent. Math.*, 234(1):419–487, 2023.

[2] A. S. Bandeira, G. Cipolloni, D. Schröder, and R. van Handel. Matrix concentration inequalities and free probability II. Two-sided bounds and applications, 2024. arxiv:2406.11453.

[3] T. Brailovskaya and R. van Handel. Universality and sharp matrix concentration inequalities. *Geom. Funct. Anal.*, 34(6):1734–1838, 2024.

[4] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2):711–775, 2005.

[5] F. Lehner. Computing norms of free operators with matrix coefficients. *Amer. J. Math.*, 121(3):453–486, 1999.

[6] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *LMS Lecture Note Series*. Cambridge University Press, Cambridge, 2006.

[7] R. van Handel. The strong convergence phenomenon. In *Current Developments in Mathematics 2026*. arxiv:2507.00346.

[8] F. Zhang, editor. *The Schur complement and its applications*, volume 4 of *Numer. Methods Algorithms*. Springer-Verlag, New York, 2005.

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Email address: `yep@mit.edu`

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
Email address: `rvan@math.princeton.edu`