Advances in Mathematics **141**, 282–321 (1999)
Article ID aima.1998.1780, available online at http://www.idealibrary.com on

Some Minimization Problems for the Free Analogue of the Fisher Information

Alexandru Nica*

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
E-mail: anica@math.uwaterloo.ca

Dimitri Shlyakhtenko†

Department of Mathematics, University of California at Los Angeles, Los Angeles, California 90095-1555

and

Roland Speicher‡

Institut für Angewandte Mathematik, Universität Heidelberg, D-69120 Heidelberg, Germany
E-mail: roland.speicher@urz.uni-heidelberg.de

Received January 1, 1998; accepted September 19, 1998

We consider the free non-commutative analogue Φ^*, introduced by D. Voiculescu, of the concept of Fisher information for random variables. We determine the minimal possible value of $\Phi^*(a, a^*)$, if a is a non-commutative random variable subject to the constraint that the distribution of a^*a is prescribed. More generally, we obtain the minimal possible value of $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d})$, if $\{a_{ij}\}_{1 \leqslant i, j \leqslant d}$ is a family of non-commutative random variables such that the distribution of A^*A is prescribed, where A is the matrix $(a_{ij})_{i, j=1}^d$. The $d \times d$-generalization is obtained from the case $d=1$ via a result of independent interest, concerning the minimal value of $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d})$ when the matrix $A=(a_{ij})_{i, j=1}^d$ and its adjoint have a given joint distribution. (A version of this result describes the minimal value of $\Phi^*(\{b_{ij}\}_{1 \leqslant i, j \leqslant d})$ when the matrix $B=(b_{ij})_{i, j=1}^d$ is selfadjoint and has a given distribution.)

* Research supported by a grant from the Natural Sciences and Engineering Research Council, Canada.

† Supported in part by a Graduate Research Fellowship of the National Science Foundation, USA.

‡ Supported by a Heisenberg Fellowship of the DFG, Germany.

0001-8708/99 $30.00

We then show how the minimization results obtained for Φ^* lead to maximization results concerning the free entropy χ^*, also defined by Voiculescu.

1. INTRODUCTION

In this paper we determine the minimal possible value of the free Fisher information $\Phi^*(a, a^*)$, if a is a non-commutative random variable subject to the constraint that the distribution of a^*a is prescribed. More generally, we obtain the minimal possible value of $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d})$, if $\{a_{ij}\}_{1 \leqslant i, j \leqslant d}$ is a family of non-commutative random variables such that the distribution of A^*A is prescribed, where A is the matrix $(a_{ij})_{i,j=1}^d$. The $d \times d$-generalization is obtained via a result of independent interest on the minimal free Fisher information of a family of matrix entries, when the distribution/ $*$-distribution of the matrix itself is given.

The framework we will consider is the one of a W^*-probability space $(\mathscr{A}, \varphi)$, with φ a faithful trace (i.e. $-\mathscr{A}$ is a W^*-algebra, and $\varphi: \mathscr{A} \to \mathbf{C}$ is a normal faithful trace-state). An element $a \in \mathscr{A}$ will be referred to as a "non-commutative random variable," and $\varphi(a)$ will be called "the expectation of a." If $a = a^* \in \mathscr{A}$, then the unique probability measure with compact support μ on $\mathbf{R}$ which has $\int_{-\infty}^{\infty} t^n \, d\mu(t) = \varphi(a^n)$, $\forall n \geqslant 0$, is called the distribution of a. An element $a = a^* \in \mathscr{A}$ is said to be semicircular of radius $r > 0$ if its distribution is absolutely continuous with respect to the Lebesgue measure, with density $\rho(t) = 2(\pi r^2)^{-1} \sqrt{r^2 - t^2}$ on $[-r, r]$.

A fundamental concept used throughout the paper is the one of freeness for a family of subsets of $\mathscr{A}$. For the definition and basic properties of freeness, we refer the reader to [9], Chapter 2.

The free analogues of entropy and of Fisher information for random variables were introduced and studied in a series of papers of D. Voiculescu ([4]–[8]), in connection to the isomorphism problem for the von Neumann algebras associated to free groups. Free analogues for some well-known inequalities concerning the Fisher information were obtained in this way. In particular, one has a "free Cramer-Rao inequality," which says the following: if $(x_1, ..., x_n)$ is an n-tuple of selfadjoint elements of $\mathscr{A}$ such that the total variance $\varphi(x_1^2 + \cdots + x_n^2)$ is prescribed, then the free Fisher information $\Phi^*(x_1, ..., x_n)$ is minimized when the x_j's are semicircular of equal radii, and free (see [8], Proposition 6.9). In the particular case $n = 2$, if one sets $a := x_1 + ix_2$ and works with a, a^* instead of x_1, x_2, then the free Cramer–Rao inequality can also be formulated like this: let a be a non-commutative random variable, such that the expectation of a^*a is prescribed; then the free Fisher information $\Phi^*(a, a^*)$ is minimized when a is a circular element (which means, by definition, that the real and imaginary part of a are free and have semicircular distributions of equal radii).

In the present paper we examine a similar minimization problem, where not only the expectation, but the whole distribution (i.e. the moments of all orders) of a^*a are prescribed. More precisely: given a probability measure ν with compact support on $[0, \infty)$, what can be said about

$$\inf\{\Phi^*(a, a^*) \mid a^*a \text{ has distribution } \nu\}? \tag{1.1}$$

One cannot of course hope to have the infimum in (1.1) achieved by a circular element; this is simply because, given ν as in (1.1), there does not exist in general a circular element a such that a^*a has distribution ν. (In fact: if a is circular, then the distribution of a^*a can only be of the form $2(\alpha\pi)^{-1}\sqrt{(\alpha-t)/t}\,dt$ on $[0, \alpha]$ for some $\alpha > 0$—see [9], Section 5.1.)

A remarkable family of relatives of the circular element is provided by the so-called "R-diagonal elements", introduced in [1]. There are several possible descriptions for the fact that an element $a \in \mathscr{A}$ is R-diagonal. The one taken as starting point in [1] is that the R-transform—i.e. free analogue for the log of the Fourier transform—of the pair (a, a^*) has a special form, which is in a certain sense "diagonal"; this is in fact where the name of "R-diagonal" comes from. In the present paper we will use an equivalent characterization of R-diagonality, described as follows: a is R-diagonal if and only if the $*$-distribution of a (i.e., the family of expectations of words in a and a^*) coincides with the $*$-distribution of an element of the form up, where u is a unitary distributed according to the Haar measure on the circle, $p = p^*$, and $\{u, u^*\}$ is free from $\{p\}$. The equivalence between the two characterizations of an R-diagonal element is shown in [1]. The circular element is R-diagonal, e.g. because its polar decomposition is known to be of the form up, with u Haar unitary such that $\{u, u^*\}$ is free from $\{p\}$ (see [9], Section 5.1).

Now, given a probability measure ν, with compact support on $[0, \infty)$, there always exists an R-diagonal element a such that a^*a has distribution ν. This a is "unique up to isomorphism," in the sense that the $*$-distribution of a is completely determined (which in turn determines the unital W^*-algebra generated by a); see Remark 3.3 below. The result we obtain is that the R-diagonal element attains the infimum considered in (1.1). Moreover, finding the actual value of the infimum is reduced to the calculation of a free Fisher information $\Phi^*(\mu)$, where μ is a symmetric distribution naturally associated to ν; and for $\Phi^*(\mu)$ one can use an explicit formula established in [4]. To summarize, we have:

1.1. THEOREM. *Let ν be a probability measure with compact support on $[0, \infty)$. Let μ be the symmetric probability measure on $\mathbf{R}$ determined by the fact that $\mu(S) = \nu(S^2)$ for every symmetric Borel set $S \subset \mathbf{R}$. Then*

$$\min\{\Phi^*(a, a^*) \mid a^*a \text{ has distribution } \nu\} = 2\Phi^*(\mu), \tag{1.2}$$

and the minimum is attained when a is R-diagonal. If in particular ν is absolutely continuous, with density ρ, then the quantities in (1.2) *equal*:

$$\tfrac{4}{3} \cdot \int_0^\infty t\rho(t)^3 \, dt \in [0, \infty]. \tag{1.3}$$

The facts stated in Theorem 1.1 are discussed in more detail (and proved) in the Section 3 of the paper.

A natural question which arises in connection to Theorem 1.1 is the following: if the minimum discussed in the theorem is finite, is it also possible to reach it as $\Phi^*(a, a^*)$ for an element a which is not R-diagonal? Up to present we were not able to settle this problem. What we can show is its (non-trivial) equivalence to another problem, also open, of deciding if a certain freeness condition is implied by the equality of two free Fisher informations with respect to subalgebras; see Sections 3.10, 3.11 below.

It is interesting that one can formulate a "matrix version" of the Theorem 1.1 —i.e. a version where "a" becomes a $d \times d$-matrix over a W^*-probability space. The possibility of making such a generalization is created by the following result, which is of independent interest:

1.2. THEOREM. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, and let d be a positive integer. Then*:

1. *For every matrix $A = (a_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$ we have*:

$$\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}) \geqslant d^3 \Phi^*(A, A^*); \tag{1.4}$$

moreover, (1.4) *holds with equality if $\{A, A^*\}$ is free from the subalgebra of "scalar matrices" $M_d(\mathbf{C}I) \subseteq M_d(\mathcal{A})$ (with I = the unit of $\mathcal{A}$).*

2. *For every selfadjoint matrix $B = (b_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$ we have*:

$$\Phi^*(\{b_{ij}\}_{1 \leqslant i, j \leqslant d}) \geqslant d^3 \Phi^*(B); \tag{1.5}$$

and (1.5) *holds with equality if B is free from $M_d(\mathbf{C}I) \subseteq M_d(\mathcal{A})$.*

It is easy to see that the freeness conditions appearing in Theorem 1.2 can indeed be fulfilled, in the context where the $*$-distribution of A (in 1) and the distribution of B (in 2) are prescribed—see the discussion preceding Proposition 4.1 in Section 4.

The conditions under which equality is reached in (1.4), (1.5) have again to do with the more general concept of free Fisher information with respect to a subalgebra. For instance, the fact standing behind the statement of

Theorem 1.2.1 is the following: if in addition to the family $\{a_{ij}\}_{1\leqslant i,\, j\leqslant d}$ we also consider a unital W^*-subalgebra $\mathcal{B}\subseteq\mathcal{A}$, then:

$$\Phi^*(\{a_{ij}, a_{ij}^*\}_{1\leqslant i,\, j\leqslant d}\colon \mathcal{B}) = d^3\Phi^*(\{A, A^*\}\colon \mathcal{M}_d(\mathcal{B})); \tag{1.6}$$

in the particular case when $\mathcal{B}=\mathbf{C}I$, this leads to

$$\Phi^*(\{a_{ij}, a_{ij}^*\}_{1\leqslant i,\, j\leqslant d}) = d^3\Phi^*(\{A, A^*\}\colon \mathcal{M}_d(\mathbf{C}I)) \geqslant d^3\Phi^*(\{A, A^*\}),$$

which is (1.4) (see Proposition 4.1 below, and the comment following to it).

By combining the results of the Theorems 1.1 and 1.2.1, one obtains the above mentioned generalization of 1.1:

1.3. Theorem. *Let ν and μ be as in Theorem* 1.1, *and let d be a positive integer. Then*

$$\min\left\{\Phi^*(\{a_{ij}, a_{ij}^*\}_{1\leqslant i,\, j\leqslant d}) \;\middle|\; \begin{array}{l} A := (a_{ij})_{i,\, j=1}^d \in M_d(\mathcal{A}) \text{ is such} \\ \text{that } A^*A \text{ has distribution } \nu \end{array}\right\}$$
$$= 2d^3\Phi^*(\mu). \tag{1.7}$$

The minimum is attained if the matrix $A=(a_{ij})_{i,\, j=1}^d$ is an R-diagonal element of $M_d(\mathcal{A})$, and if $\{A, A^*\}$ is free from the algebra of scalar matrices $M_d(\mathbf{C}I)\subseteq M_d(\mathcal{A})$.

It is easy to see that minimization problems for Φ^* correspond to maximization problems for the concept of free entropy χ^*, which was also defined (in terms of Φ^*) in Voiculescu's work [8]. We will conclude the paper by spelling out the maximization results for χ^* which follow from the theorems presented above. The counterpart of Theorem 1.3 is:

1.4. Theorem. *Let ν and μ be as in Theorem* 1.1, *and let d be a positive integer. Then*

$$\max\left\{\chi^*(\{a_{ij}, a_{ij}^*\}_{1\leqslant i,\, j\leqslant d}) \;\middle|\; \begin{array}{l} A := (a_{ij})_{i,\, j=1}^d \in M_d(\mathcal{A}) \text{ is such} \\ \text{that } A^*A \text{ has distribution } \nu \end{array}\right\}$$
$$= 2d^2\left(\chi^*(\mu) - \frac{\log d}{2}\right). \tag{1.8}$$

The maximum is attained if the matrix $A=(a_{ij})_{i,\, j=1}^d$ is an R-diagonal element of $M_d(\mathcal{A})$, and if $\{A, A^*\}$ is free from the algebra of scalar matrices $M_d(\mathbf{C}I)\subseteq M_d(\mathcal{A})$.

The Theorem 1.4 is obtained from its particular case $d=1$ via a maximization result for the free entropy of a family of matrix entries, which is an analogue of Theorem 1.2:

1.5. THEOREM. *Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, and let d be a positive integer. Then:*

1. *For every matrix $A=(a_{ij})_{i,j=1}^d \in M_d(\mathscr{A})$ we have:*

$$\chi^*(\{a_{ij}, a_{ij}^*\}_{1\leqslant i, j\leqslant d}) \leqslant d^2(\chi^*(A, A^*) - \log d); \tag{1.9}$$

moreover, (1.9) *holds with equality if $\{A, A^*\}$ is free from the subalgebra of scalar matrices $M_d(\mathbf{C}I) \subseteq M_d(\mathscr{A})$.*

2. *For every selfadjoint matrix $B=(b_{ij})_{i,j=1}^d \in M_d(\mathscr{A})$ we have:*

$$\chi^*(\{b_{ij}\}_{1\leqslant i, j\leqslant d}) \leqslant d^2\left(\chi^*(B) - \frac{\log d}{2}\right); \tag{1.10}$$

and (1.10) *holds with equality if B is free from $M_d(\mathbf{C}I) \subseteq M_d(\mathscr{A})$.*

It is tempting to believe that the results obtained about χ^* in this way remain true if "χ^*" is replaced by "χ", the free entropy defined via approximations with matrices which was studied in [5]–[7]. But at the moment it is not proved (though it might very well be true) that χ and χ^* coincide; and consequently, when we replace χ^* by χ in our maximization results, we just obtain some statements for which proofs are needed. We hope to discuss these statements about χ (and supply their proofs) in a future work.

The paper is organized as follows: after reviewing the concept of free Fisher information in Section 2, we will prove the Theorem 1.1 in Section 3, the Theorems 1.2, 1.3 in Section 4, and the Theorems 1.4, 1.5 in Section 5.

2. REVIEW OF THE CONCEPT OF FREE FISHER INFORMATION

For general "free probabilistic" terminology and basic results, we refer the reader to the monograph [9].

2.1. *Notations.* Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ a faithful trace.

1. $L^2(\mathscr{A}, \varphi)$ will denote the Hilbert space obtained by completing $\mathscr{A}$ with respect to the norm $\|a\|_2 := \sqrt{\varphi(a^*a)}$, $a \in \mathscr{A}$.

2. For d a positive integer, we will denote by $M_d(\mathscr{A})$ the W^*-algebra of $d\times d$-matrices over $\mathscr{A}$. Also, we will denote: $\varphi_d := tr\otimes\varphi : M_d(\mathscr{A})\to \mathbf{C}$, where tr is the normalized trace on $M_d(\mathbf{C})$. In other words, φ_d is the faithful trace-state which acts by the formula

$$\varphi_d(A)=\frac{1}{d}\sum_{i=1}^{d}\varphi(a_{ii}),\qquad \text{for}\quad A=(a_{ij})_{i,j=1}^{d}\in M_d(\mathscr{A}). \tag{2.1}$$

3. An immediate consequence of (2.1) is that:

$$\|A\|^2_{L^2(\varphi_d)}=\frac{1}{d}\sum_{i,j=1}^{d}\|a_{ij}\|^2_{L^2(\varphi)},\ \forall A=(a_{ij})_{i,j=1}^{d}\in M_d(\mathscr{A}). \tag{2.2}$$

Thus if we fix a pair of indices $k, l\in\{1, ..., d\}$, then we get:

$$\|a_{k,l}\|_{L^2(\varphi)}\leqslant\sqrt{d}\,\|A\|_{L^2(\varphi_d)},\ \forall A=(a_{ij})_{i,j=1}^{d}\in M_d(\mathscr{A});$$

and consequently, the map $A\mapsto a_{k,l}$ extends by continuity to a bounded linear map "$\text{Entry}_{k,l}$" from $L^2(M_d(\mathscr{A}),\varphi_d)$ to $L^2(\mathscr{A},\varphi)$. Eq. (2.2) can then be extended by continuity to:

$$\|X\|^2_{L^2(\varphi_d)}=\frac{1}{d}\sum_{i,j=1}^{d}\|\text{Entry}_{i,j}(X)\|^2_{L^2(\varphi)},\ \forall X\in L^2(M_d(\mathscr{A}),\varphi_d); \tag{2.3}$$

and by using (2.3) it is readily seen that $X\mapsto(\text{Entry}_{i,j}(X))_{i,j=1}^{d}$ is a bijection between $L^2(M_d(\mathscr{A}),\varphi_d)$ and the vector space of $d\times d$-matrices over $L^2(\mathscr{A},\varphi)$. We will identify in what follows the vectors in $L^2(M_d(\mathscr{A}),\varphi_d)$ with matrices over $L^2(\mathscr{A},\varphi)$, via this bijection. It is easily checked that, in this identification, the left and right actions of $M_d(\mathscr{A})$ on $L^2(M_d(\mathscr{A}),\varphi_d)$ become "matrix multiplications"—e.g. we have that:

$$\text{Entry}_{k,l}(XA)=\sum_{m=1}^{d}\text{Entry}_{k,m}(X)\cdot a_{m,l},$$

for every $X\in L^2(M_d(\mathscr{A}),\varphi_d)$, $A=(a_{ij})_{i,j=1}^{d}$, $1\leqslant k, l\leqslant d$. The formulas for the entries of X^*, and for $\varphi_d(X)$, $X\in L^2(M_d(\mathscr{A}),\varphi_d)$, are also obtained by continuity in the obvious way.

The considerations made in this paper revolve around the notion of free Fisher information, which was introduced and studied in [4], [8]. We will next review this notion (Sections 2.2–2.6). A family $\{a_i\}_{i\in I}$ of elements of a W^*-algebra will be called in what follows "selfadjoint" if there exists an involutive bijection $\sigma: I\to I$ such that $a_i^*=a_{\sigma(i)}$ for every $i\in I$.

2.2. Definition. Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace. Let $\{a_i\}_{i \in I}$ be a selfadjoint family of elements of $\mathscr{A}$, and let $\mathscr{B} \subseteq \mathscr{A}$ be a unital W^*-algebra.

1. We say that a family $\{\xi_i\}_{i \in I}$ of vectors in $L^2(\mathscr{A}, \varphi)$ fulfills the conjugate relations for $\{a_i\}_{i \in I}$, with respect to $\mathscr{B}$, if:

$$\varphi(\xi_i b_0 a_{i_1} b_1 \cdots a_{i_n} b_n) = \sum_{m=1}^{n} \delta_{i, i_m} \varphi(b_0 a_{i_1} \cdots a_{i_{m-1}} b_{m-1}) \cdot \varphi(b_m a_{i_{m+1}} \cdots a_{i_n} b_n), \tag{2.4}$$

for every $n \geqslant 0$, $b_0, b_1, ..., b_n \in \mathscr{B}$ and $i, i_1, ..., i_n \in I$.

2. We say that a family $\{\xi_i\}_{i \in I}$ of vectors in $L^2(\mathscr{A}, \varphi)$ is a conjugate system for $\{a_i\}_{i \in I}$ with respect to $\mathscr{B}$ if it satisfies the Eq. (2.4) and if in addition we have that:

$$\xi_i \in \overline{Alg(\{a_j\}_{j \in I} \cup \mathscr{B})}^{\|\cdot\|_2} \subseteq L^2(\mathscr{A}, \varphi), \ \forall i \in I. \tag{2.5}$$

2.3. *Remarks.* 1. The conjugate relations (2.4) can be viewed as a prescription for the inner products in $L^2(\mathscr{A}, \varphi)$ between ξ_i $(i \in I)$ and a monomial $b_0 a_{i_1} b_1 \cdots a_{i_n} b_n$; since the monomials of this form linearly span $Alg(\{a_i\}_{i \in I} \cup \mathscr{B})$, it follows that the conjugate system $\{\xi_i\}_{i \in I}$ for $\{a_i\}_{i \in I}$ with respect to $\mathscr{B}$ is unique, if it exists. Note moreover that the existence of the conjugate system is equivalent to the existence of any family of vectors in $L^2(\mathscr{A}, \varphi)$ which fulfill the conjugate relations (2.4); indeed, if $\{\xi_i\}_{i \in I}$ satisfy (2.4) and if we set η_i to be the projection of ξ_i onto $\overline{Alg(\{a_j\}_{j \in I} \cup \mathscr{B})}^{\|\cdot\|_2}$, $i \in I$, then $\{\eta_i\}_{i \in I}$ will also satisfy (2.4), hence will give the conjugate system.

2. If the family $\{a_i\}_{i \in I}$ from Definition 2.2 has a conjugate system $\{\xi_i\}_{i \in I}$ with respect to $\mathscr{B}$, and if $\sigma: I \to I$ is an involution such that $a_i^* = a_{\sigma(i)}$, $i \in I$, then we necessarily also have:

$$\xi_i^* = \xi_{\sigma(i)}, \ i \in I. \tag{2.6}$$

Indeed, it is easy to see (by using the relations $a_i^* = a_{\sigma(i)}$, $i \in I$, and the properties of the trace-state φ) that if we set $\eta_i = \xi_{\sigma(i)}^*$, $i \in I$, then $\{\eta_i\}_{i \in I}$ will also fulfill the conjugate relations (2.4); therefore $\eta_i = \xi_i$, $i \in I$, by the uniqueness of the conjugate system, and this gives (2.6).

2.4. Definition. Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, let $\{a_i\}_{i \in I}$ be a selfadjoint family of elements of $\mathscr{A}$, and let $\mathscr{B} \subseteq \mathscr{A}$ be a unital W^*-subalgebra. If $\{a_i\}_{i \in I}$ has a conjugate system $\{\xi_i\}_{i \in I}$ with respect to $\mathscr{B}$, then the free Fisher information of $\{a_i\}_{i\in I}$ with respect to $\mathscr{B}$ is:

$$\Phi^*(\{a_i\}_{i\in I} : \mathscr{B}) := \sum_{i\in I} \|\xi_i\|^2. \tag{2.7}$$

If $\{a_i\}_{i\in I}$ has no conjugate system with respect to $\mathscr{B}$, then one takes $\Phi^*(\{a_i\}_{i\in I} : \mathscr{B}) := \infty$.

2.5. DEFINITION. Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace. If $\{a_i\}_{i\in I}$ is a selfadjoint family of elements of $\mathscr{A}$, then we denote:

$$\Phi^*(\{a_i\}_{i\in I}) := \Phi^*(\{a_i\}_{i\in I} : \mathbf{C}I). \tag{2.8}$$

$\Phi^*(\{a_i\}_{i\in I})$ will be simply called "the free Fisher information" of $\{a_i\}_{i\in I}$. Also, if $\{\xi_i\}_{i\in I}$ fulfills the conjugate relations (respectively is a conjugate system) for $\{a_i\}_{i\in I}$ with respect to $\mathbf{C}I$, we will generally omit "with respect to $\mathbf{C}I$" from the formulation.

2.6. *Remarks.* 1. Let $(\mathscr{A}, \varphi)$, $\{a_i\}_{i\in I}$ and $\mathscr{B}$ be as in the Definition 2.4. If a family $\{\xi_i\}_{i\in I}$ in $L^2(\mathscr{A}, \varphi)$ fulfills the conjugate relations for $\{a_i\}_{i\in I}$ with respect to $\mathscr{B}$, but does not necessarily satisfy (2.5), then we still know that:

$$\Phi^*(\{a_i\}_{i\in I} : \mathscr{B}) \leqslant \sum_{i\in I} \|\xi_i\|^2. \tag{2.9}$$

This is a direct consequence of the statement concluding the Remark 2.3.1.

2. Let $(\mathscr{A}, \varphi)$ and $\{a_i\}_{i\in I}$ be as above, and let $\mathscr{B}_1, \mathscr{B}_2$ be W^*-subalgebras of $\mathscr{A}$ such that $I \in \mathscr{B}_1 \subseteq \mathscr{B}_2$. Then

$$\Phi^*(\{a_i\}_{i\in I} : \mathscr{B}_1) \leqslant \Phi^*(\{a_i\}_{i\in I} : \mathscr{B}_2). \tag{2.10}$$

Indeed, if $\Phi^*(\{a_i\}_{i\in I} : \mathscr{B}_2) < \infty$, then the conjugate system for $\{a_i\}_{i\in I}$ with respect to $\mathscr{B}_2$ will fulfill the conjugate relations with respect to $\mathscr{B}_1$; hence (2.10) follows from (2.9).

3. In the particular case of 2 when $\mathscr{B}_1 = \mathbf{C}I$, we obtain the inequality:

$$\Phi^*(\{a_i\}_{i\in I}) \leqslant \Phi^*(\{a_i\}_{i\in I} : \mathscr{B}), \tag{2.11}$$

for every unital W^*-subalgebra $\mathscr{B}$ of $\mathscr{A}$. It is important to record here that, as proved in [8] Proposition 3.6, (2.11) holds with equality whenever $\{a_i\}_{i\in I}$ is free from $\mathscr{B}$.

The problems discussed in the present paper are formulated only in terms of the free information $\Phi^*(\{a_i\}_{i\in I})$ (with respect to the scalars). But however, considerations involving free information with respect to non-trivial subalgebras appear naturally in the solutions. Moreover, in Section 3 we will arrive to use a version of $\Phi^*(\bullet : \mathscr{B})$ where (in addition to $\mathscr{B}$ itself) one also considers a completely positive map $\eta: \mathscr{B} \to \mathscr{B}$. This version of Φ^* was introduced in [3], and is defined as follows.

2.7. Definition. Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, let $x = x^*$ be in $\mathscr{A}$, let $\mathscr{B} \subseteq \mathscr{A}$ be a unital W^*-subalgebra, and let $\eta: \mathscr{B} \to \mathscr{B}$ be a completely positive map.

1. We say that a vector $\xi \in L^2(\mathscr{A}, \varphi)$ fulfills the conjugate relations for x, with respect to $\mathscr{B}$ and η, if:

$$\varphi(\xi b_0 x b_1 \cdots x b_n) = \sum_{m=1}^{n} \varphi(\eta(E_{\mathscr{B}}(b_0 x \cdots x b_{m-1})) \cdot b_m x \cdots x b_n), \tag{2.12}$$

for every $n \geqslant 0$ and every $b_0, b_1, ..., b_n \in \mathscr{B}$, and where $E_{\mathscr{B}}$ denotes the unique trace-preserving conditional expectation from $\mathscr{A}$ onto $\mathscr{B}$.

2. The vector $\xi \in L^2(\mathscr{A}, \varphi)$ is called a conjugate for x, with respect to $\mathscr{B}$ and η, if it satisfies (2.12) and if in addition:

$$\xi \in \overline{Alg(\{x\} \cup \mathscr{B})}^{\|\cdot\|_2}. \tag{2.13}$$

3. The free Fisher information of x with respect to $\mathscr{B}$ and η is defined to be:

$$\Phi^*(x: \mathscr{B}, \eta) := \|\xi\|^2, \tag{2.14}$$

if x has a conjugate vector ξ with respect to $\mathscr{B}$ and η, and $\Phi^*(x: \mathscr{B}, \eta) := \infty$ otherwise.

2.8. *Remarks.* 1. Exactly as in Remark 2.3.1, one sees that the conjugate vector with respect to $\mathscr{B}$ and η is unique, if it exists. (This ensures that the definition of $\Phi^*(x: \mathscr{B}, \eta)$ in (2.14) makes sense.)

2. In the particular case when the completely positive map $\eta: \mathscr{B} \to \mathscr{B}$ is $\eta(b) := \varphi(b) I$, $b \in \mathscr{B}$, one obtains $\Phi^*(x: \mathscr{B}, \eta) = \Phi^*(x: \mathscr{B})$, because (2.12) reduces to (2.4).

3. It is easy to see (exactly as in the Remark 2.6.2) that one has the inequality:

$$\Phi^*(x: \mathscr{B}_1, \eta_1) \leqslant \Phi^*(x: \mathscr{B}_2, \eta_2) \tag{2.15}$$

whenever $\mathscr{B}_1 \subseteq \mathscr{B}_2$ and η_1, η_2 are related by:

$$\eta_2(b) = \eta_1(E_{\mathscr{B}_1}(b)), \forall b \in \mathscr{B}_2. \tag{2.16}$$

It is again important to record that, as proved in [3] Proposition 3.8, (2.15) holds with equality whenever $Alg(\{x\} \cup \mathscr{B}_1)$ is free from $\mathscr{B}_2$, with amalgamation over $\mathscr{B}_1$.

2.9. *Remark.* Let $(\mathscr{A}_1, \varphi_1)$ and $(\mathscr{A}_2, \varphi_2)$ be W^*-probability spaces, with φ_1, φ_2 faithful traces, and let $x_1 = x_1^* \in A_1$, $x_2 = x_2^* \in \mathscr{A}_2$ be elements with identical distributions (i.e., $\varphi_1(x_1^n) = \varphi_2(x_2^n)$, $\forall n \geqslant 0$). Then we must also have that $\Phi^*(x_1) = \Phi^*(x_2)$. Indeed, the coincidence of distributions has as consequence that there exists a unitary operator $U: \overline{Alg(I, x_1)}^{\|\cdot\|_2} \to \overline{Alg(I, x_2)}^{\|\cdot\|_2}$, determined by the relation $U(x_1^n) = x_2^n$, $n \geqslant 0$. It is immediate that U sends a conjugate for x_1 into a conjugate for x_2, and this in turn implies the equality of free Fisher informations.

In particular, if μ is a probability measure with compact support on $\mathbf{R}$, it makes sense to use the notation

$$\Phi^*(\mu) := \Phi^*(x), \tag{2.17}$$

where x is an arbitrary selfadjoint random variable (in some W^*-probability space $(\mathscr{A}, \varphi)$, with φ faithful trace) such that the distribution of x is μ. A detailed discussion about $\Phi^*(\mu)$ is made in [4] (see also Section 2 of [8]); it is in particular shown there that if μ is absolutely continuous with respect to the Lebesgue measure, and has density ρ, then:

$$\Phi^*(\mu) = \tfrac{2}{3} \cdot \int_{-\infty}^{\infty} \rho(t)^3 \, dt. \tag{2.18}$$

3. MINIMIZATION OF $\Phi^*(a, a^*)$, WHEN THE DISTRIBUTION OF a^*a IS PRESCRIBED

Let ν be a probability measure with compact support on $[0, \infty)$. We consider the minimization problem stated in (1.1) of the Introduction, i.e. the problem of determining:

$$\inf\{\Phi^*(a, a^*) \mid a^*a \text{ has distribution } \nu\}$$

where $a \in \mathscr{A}$ and $(\mathscr{A}, \varphi)$ is a W^*-probability space, with φ faithful trace.

In the considerations related to this problem, it is convenient to use the following symmetric measure associated to ν.

3.1. Definition. For ν as above, we will call "symmetric square root of ν" the unique probability measure μ on $\mathbf{R}$ which is symmetric (i.e. $\mu(S)=\mu(-S)$ for every Borel set S), and has the property that $\mu(S)=\nu(\{s^2 \mid s\in S\})$, for every Borel set S such that $S=-S$.

In terms of random variables, the connection between ν and its symmetric square root μ is expressed as follows: a selfadjoint element x in a W^*-probability space $(\mathscr{A}, \varphi)$ has distribution μ if and only if x is even (i.e. $\varphi(x^n)=0$ for n odd), and x^2 has distribution ν.

3.2. Theorem. *Let ν be a probability measure with compact support on $[0, \infty)$, and let μ be the symmetric square root of ν. Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, and let $a\in\mathscr{A}$ be such that a^*a has distribution ν. Then*:

$$\Phi^*(a, a^*)\geqslant 2\Phi^*(\mu). \tag{3.1}$$

Moreover, (3.1) *holds with equality if a is of the form $a=up$, where $u\in\mathscr{A}$ is a unitary with Haar distribution* (*i.e. $\varphi(u^n)=0$ for all $n\in\mathbf{Z}\backslash\{0\}$*), *$p=p^*$ has distribution μ, and $\{p\}$ is free from $\{u, u^*\}$. Thus the infimum considered in* (1.1) *of the Introduction is equal to $2\Phi^*(\mu)$.*

3.3. *Remarks.* 1. If u is a unitary with Haar distribution, $p=p^*$, and $\{p\}$ is free from $\{u, u^*\}$, then the element $a=up$ is said to be R-diagonal ([1]). For such an element, the $*$-distribution of a is completely determined by the distribution of p^2 ([1], Corollary 1.8). This implies that, as far as $*$-distributions are concerned, there is a unique R-diagonal element a such that the distribution of a^*a is a given probability measure ν.

Let us hence notice that in the phrase following Eq. (3.1) (in the statement of Theorem 3.2) we could replace "p has distribution μ" with the apparently more general condition "p^2 has distribution ν." But this wouldn't actually change the $*$-distribution of a—we would still have to do with the same R-diagonal element.

We were in fact unable to determine if the R-diagonal $*$-distribution is the unique one which achieves the minimization of $\Phi^*(a, a^*)$ considered in (1.1). (See also the Sections 3.10, 3.11 below.)

2. The statement of Theorem 3.2 contains the one of Theorem 1.1, with the exception of the formula (1.3). The latter formula follows from Eq. (2.18) of Remark 2.9, combined with the simple observation that μ is absolutely continuous if and only if ν is so, in which case the densities σ of μ and ρ of ν are connected by the relation $\sigma(t)=|t|\,\rho(t^2)$, $t\in\mathbf{R}$.

Our goal in this section is thus to prove Theorem 3.2. Let us set the following:

3.4. *Notations.* ν, μ, $(\mathcal{A}, \varphi)$, $a \in \mathcal{A}$ are fixed from now on, until the end of the section, and are as in the statement of Theorem 3.2. We will consider the space $(M_2(\mathcal{A}), \varphi_2)$ of 2×2 matrices over $(\mathcal{A}, \varphi)$ (as in Notations 2.1.2), and we will give a special attention to the selfadjoint matrix:

$$A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix} \in M_2(\mathcal{A}). \tag{3.2}$$

For $i, j \in \{1, 2\}$ we will denote by V_{ij} the matrix in $M_2(\mathcal{A})$ which has the (i, j)-entry equal to the unit of $\mathcal{A}$, and the other entries equal to 0. Then:

$$\operatorname{span}\{V_{11}, V_{12}, V_{21}, V_{22}\} = M_2(\mathbf{C}I) \subseteq M_2(\mathcal{A});$$

we will also denote

$$\mathcal{D} := \operatorname{span}\{V_{11}, V_{22}\}$$

(the 2-dimensional $*$-subalgebra of $M_2(\mathcal{A})$ consisting of scalar diagonal matrices). We will denote by $E_{\mathcal{M}}$ and $E_{\mathcal{D}}$ the unique trace-preserving conditional expectations from $M_2(\mathcal{A})$ onto $M_2(\mathbf{C}I)$ and $\mathcal{D}$, respectively. For $B = (b_{ij})_{i, j=1}^2 \in M_2(\mathcal{A})$ we have:

$$E_{\mathcal{M}}(B) = \begin{pmatrix} \varphi(b_{11})\, I & \varphi(b_{12})\, I \\ \varphi(b_{21})\, I & \varphi(b_{22})\, I \end{pmatrix}, \qquad E_{\mathcal{D}}(B) = \begin{pmatrix} \varphi(b_{11})\, I & 0 \\ 0 & \varphi(b_{22})\, I \end{pmatrix}. \tag{3.3}$$

3.5. *Remark.* Since A of Equation (3.2) has:

$$A^2 = \begin{pmatrix} aa^* & 0 \\ 0 & a^*a \end{pmatrix},$$

while on the other hand the odd powers of A have 0's on the main diagonal, it is immediate that A is even and that A^2 has distribution ν. Therefore A itself has distribution μ.

3.6. PROPOSITION. *Let $\eta: M_2(\mathbf{C}I) \to M_2(\mathbf{C}I)$ be the completely positive map defined by*:

$$\eta\left(\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\right) := \begin{pmatrix} x_{22} & 0 \\ 0 & x_{11} \end{pmatrix}. \tag{3.4}$$

Then we have:

$$\Phi^*(a, a^*) = 2\Phi^*(A : M_2(\mathbf{C}I), \eta). \tag{3.5}$$

Proof. We first consider the situation when $\Phi^*(a, a^*) < \infty$. In this case there exists $\xi \in \overline{Alg(I, a, a^*)}^{\|\cdot\|_2}$ such that $\{\xi, \xi^*\}$ forms a conjugate system for $\{a, a^*\}$. We define:

$$X := \begin{pmatrix} 0 & \xi^* \\ \xi & 0 \end{pmatrix} \in L^2(M_2(\mathscr{A}), \varphi_2) \tag{3.6}$$

where the identification between vectors in $L^2(M_2(\mathscr{A}), \varphi_2)$ and matrices over $L^2(\mathscr{A}, \varphi)$ is as discussed in the Notations 2.1.3. We will show that X is a conjugate for A, with respect to $M_2(\mathbf{C}I)$ and η. Proving this claim consists in verifying that:

(a) the relation

$$\varphi_2(XB_0AB_1 \cdots AB_n) = \sum_{m=1}^{n} \varphi_2(\eta(E_{\mathscr{M}}(B_0A \cdots AB_{m-1}) \cdot B_mA \cdots AB_n) \tag{3.7}$$

holds for every $n \geqslant 0$ and every $B_0, B_1, ..., B_n \in M_2(\mathbf{C}I)$; and that:

(b) $X \in \overline{Alg(\{A\} \cup M_2(\mathbf{C}I))}^{\|\cdot\|_2} \subseteq L^2(M_2(\mathscr{A}), \varphi_2)$.

Note that once (a) and (b) will be proved, we will have the equality

$$\Phi^*(A: M_2(\mathbf{C}I), \eta) = \|X\|^2_{L^2(\varphi_2)} \overset{(2.3)}{=} \tfrac{1}{2}(\|\xi\|^2 + \|\xi^*\|^2) = \tfrac{1}{2}\Phi^*(a, a^*),$$

which is exactly (3.5) (under the hypothesis $\Phi^*(a, a^*) < \infty$).

Proof of (*a*). Both sides of (3.7) depend multilinearly on $B_0, B_1, ..., B_n$; we can therefore assume without loss of generality that $B_m = V_{i_m j_m}$, $0 \leqslant m \leqslant n$, for some $i_0, j_0, ..., i_n, j_n \in \{1, 2\}$.

By using the trace property of φ_2 we can write the left-hand side of (3.7) as:

$$\varphi_2(V_{j_n j_n} X V_{i_0 j_0} A V_{i_1 j_1} \cdots A V_{i_n j_n}). \tag{3.8}$$

Only the (j_n, j_n)-entry of the matrix product appearing in (3.8) is not 0; this entry equals:

$$(X)_{j_n i_0}(A)_{j_0\, i_1} \cdots (A)_{j_{n-1} i_n},$$

where $(A)_{ij}$, $(X)_{ij}$ stand for the (i, j)-entry of A and X, respectively. Thus the quantity in (3.8) equals:

$$\tfrac{1}{2}\varphi((X)_{j_n i_0}(A)_{j_0 i_1} \cdots (A)_{j_{n-1} i_n}). \tag{3.9}$$

But we know that $(A)_{ij}=0=(X)_{ij}$ if $i=j$; so if we make the convention to denote $\bar{i}:=3-i$ ($=$the number in $\{1, 2\}$ which is not i), for $i\in\{1, 2\}$, then (3.9) becomes:

$$\tfrac{1}{2}\delta_{j_0\bar{i}_1}\delta_{j_1\bar{i}_2}\cdots\delta_{j_{n-1}\bar{i}_n}\delta_{j_n\bar{i}_0}\cdot\varphi((X)_{\bar{i}_0 i_0}(A)_{\bar{i}_1 i_1}\cdots(A)_{\bar{i}_n i_n}). \tag{3.10}$$

In (3.10), $(X)_{\bar{i}_0 i_0}$ is ξ or ξ^*, while every $(A)_{\bar{i}_m i_m}$ is either a or a^*. So the conjugate relations for $\{a, a^*\}$ can be used, to obtain that the quantity in (3.10) equals:

$$\tfrac{1}{2}\delta_{j_0\bar{i}_1}\delta_{j_1\bar{i}_2}\cdots\delta_{j_{n-1}\bar{i}_n}\delta_{j_n\bar{i}_0}$$
$$\times\sum_{m=1}^{n}\delta_{\bar{i}_0 i_m}\varphi((A)_{\bar{i}_1 i_1}\cdots(A)_{\bar{i}_{m-1}i_{m-1}})\cdot\varphi((A)_{\bar{i}_{m+1}i_{m+1}}\cdots(A)_{\bar{i}_n i_n}). \tag{3.11}$$

We now turn to the right-hand side of (3.7). By using the formulas for η and $E_{\mathcal{M}}$ (as in Equations (3.4) and (3.3)) we first see that:

$$\eta(E_{\mathcal{M}}(V_{i_0j_0}A\cdots AV_{i_{m-1}j_{m-1}}))=\delta_{i_0j_{m-1}}\cdot\varphi((A)_{j_0i_1}\cdots(A)_{j_{m-2}i_{m-1}})\cdot V_{\bar{i}_0\bar{i}_0},$$

for $1\leqslant m\leqslant n$. By replacing this into the right-hand side of (3.7), we obtain the expression:

$$\sum_{m=1}^{n}\delta_{i_0j_{m-1}}\cdot\varphi((A)_{j_0i_1}\cdots(A)_{j_{m-2}i_{m-1}})\cdot\varphi_2(V_{\bar{i}_0\bar{i}_0}V_{i_mj_m}A\cdots AV_{i_nj_n}). \tag{3.12}$$

But then a calculation very similar to the ones shown above gives us that the summation in (3.12) coincides, term by term, with the one in (3.11).

Proof of (b). We have

$$\begin{pmatrix} p(a, a^*) & 0\\ 0 & 0\end{pmatrix}\in Alg(\{A\}\cup M_2(\mathbf{C}I)), \tag{3.13}$$

whenever p is a non-commutative polynomial of two variables. (Indeed, it clearly suffices to check the cases $p(a, a^*)=I$ and $p(a, a^*)=a$, when the matrix in (3.13) becomes V_{11} and respectively AV_{21}.) From (3.13) and the fact that $\xi, \xi^*\in Alg(I, a, a^*)$ we infer:

$$\begin{pmatrix} \xi & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix} \xi^* & 0\\ 0 & 0\end{pmatrix}\in\overline{Alg(\{A\}\cup M_2(\mathbf{C}I))}^{\|\cdot\|_2}. \tag{3.14}$$

But $\overline{Alg(\{A\} \cup M_2(\mathbf{C}I))}^{\|\cdot\|_2}$ is invariant under the left/right action of elements from $M_2(\mathbf{C}I)$; so (3.14) implies that:

$$X = V_{21} \begin{pmatrix} \xi & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \xi^* & 0 \\ 0 & 0 \end{pmatrix} V_{12} \in \overline{Alg(\{A\} \cup M_2(\mathbf{C}I))}^{\|\cdot\|_2},$$

as desired.

Hence (3.5) is now proved in the case when $\Phi^*(a, a^*) < \infty$. It remains to show that $\Phi^*(a, a^*) = \infty \Rightarrow \Phi^*(A : M_2(\mathbf{C}I), \eta) = \infty$; or equivalently that $\Phi^*(A : M_2(\mathbf{C}I), \eta) < \infty \Rightarrow \Phi^*(a, a^*) < \infty$.

If $\Phi^*(A : M_2(\mathbf{C}I), \eta) < \infty$, then there exists $X \in L^2(M_2(\mathscr{A}), \varphi_2)$ which satisfies the conjugate relations for A, with respect to $M_2(\mathbf{C}I)$ and η. We identify X with a 2×2 matrix over $L^2(\mathscr{A}, \varphi)$, and denote its (2,1)-entry by ξ; we will show that $\{\xi, \xi^*\}$ satisfy the conjugate relations with respect to $\{a, a^*\}$ (this will entail, as noticed in Remark 2.6.1, that $\Phi^*(a, a^*) < \infty$).

It is in fact sufficient to prove that:

$$\varphi(\xi a_{i_1} \cdots a_{i_n}) = \sum_{m=1}^{n} \delta_{i_m, 1} \varphi(a_{i_1} \cdots a_{i_{m-1}})\, \varphi(a_{i_{m+1}} \cdots a_{i_n}), \tag{3.15}$$

for every $n \geqslant 1$ and every $i_1, ..., i_n \in \{1, 2\}$, where we denoted $a_1 := a$, $a_2 := a^*$. Indeed, the symmetric relation:

$$\varphi(\xi^* a_{i_1} \cdots a_{i_n}) = \sum_{m=1}^{n} \delta_{i_m, 2} \varphi(a_{i_1} \cdots a_{i_{m-1}})\, \varphi(a_{i_{m+1}} \cdots a_{i_n}) \tag{3.16}$$

follows from (3.15) by taking an adjoint and doing a circular permutation under φ. We also have $\varphi(\xi) = 2\varphi_2(XV_{12}) = 0$, by the conjugate relations satisfied by X, and $\varphi(\xi^*) = \overline{\varphi(\xi)} = 0$. Added to (3.15–16), this exhausts the list of conjugate relations for a, a^*.

In order to verify (3.15), we adopt again the conventions of notation used in the "Proof of (a)" above, and we write:

$$\begin{aligned}
&\varphi(\xi a_{i_1} a_{i_2} \cdots {}_{i_n}) \\
&\quad = \varphi((X)_{21} (A)_{i_1 \bar{i}_1} (A)_{i_2 \bar{i}_2} \cdots (A)_{i_n \bar{i}_n}) \\
&\quad = 2\varphi_2(X V_{1 i_1} A V_{\bar{i}_1 i_2} A \cdots V_{\bar{i}_{n-1} i_n} A V_{\bar{i}_n 2}) \\
&\quad = 2 \sum_{m=1}^{n} \varphi_2(\eta(E_{\mathscr{M}}(V_{1 i_1} A V_{\bar{i}_1 i_2} A \cdots V_{\bar{i}_{m-1} i_m})) \cdot V_{\bar{i}_m i_{m+1}} A \cdots V_{\bar{i}_{n-1} i_n} A V_{\bar{i}_n 2})
\end{aligned}$$

(by the conjugate relations for A, with respect to $M_2(\mathbf{C}I)$ and η)

$$=2 \sum_{m=1}^{n} \varphi_2(\delta_{i_m, 1} \varphi((A)_{i_1 \bar{i}_1} \cdots (A)_{i_{m-1} \bar{i}_{m-1}}) V_{22} \cdot V_{\bar{i}_m i_{m+1}} A \cdots V_{\bar{i}_{n-1} i_n} A V_{\bar{i}_n 2})$$

(by writing explicitly how $\eta \circ E_{\mathscr{M}}$ works)

$$=2 \sum_{m=1}^{n} \delta_{i_m, 1} \varphi((A)_{i_1 \bar{i}_1} \cdots (A)_{i_{m-1} \bar{i}_{m-1}}) \cdot \tfrac{1}{2} \delta_{2, \bar{i}_m} \varphi((A)_{i_{m+1} \bar{i}_{m+1}} \cdots (A)_{i_n \bar{i}_n})$$

$$= \sum_{m=1}^{n} \delta_{i_m, 1} \varphi(a_{i_1} \cdots a_{i_{m-1}}) \varphi(a_{i_{m+1}} \cdots a_{i_n}). \qquad \text{Q.E.D.}$$

3.7. Proposition. *Let $\eta_0 : \mathscr{D} \to \mathscr{D}$ be the $*$-automorphism defined by:*

$$\eta_0\left(\begin{pmatrix} x_{11} & 0 \\ 0 & x_{22} \end{pmatrix}\right) := \begin{pmatrix} x_{22} & 0 \\ 0 & x_{11} \end{pmatrix}. \tag{3.17}$$

Then we have:

$$\Phi^*(A : M_2(\mathbf{C}I), \eta) \geqslant \Phi^*(A : \mathscr{D}, \eta_0) = \Phi^*(A) = \Phi^*(\mu). \tag{3.18}$$

Proof. It is immediate that $\eta(B) = \eta_0(E_{\mathscr{D}}(B))$, for every $B \in M_2(\mathbf{C}I)$; thus the inequality appearing in (3.18) is implied by (2.15) of Remark 2.8.3. On the other hand, the equality $\Phi^*(A) = \Phi^*(\mu)$ holds just because A has distribution μ (Remark 3.5). Our main concern in this proof is to show that $\Phi^*(A : \mathscr{D}, \eta_0) = \Phi^*(A)$.

For proving $\Phi^*(A : \mathscr{D}, \eta_0) \geqslant \Phi^*(A)$, we assume the existence of a conjugate X for A, with respect to $\mathscr{D}$ and η_0, and we show that X fulfills the conjugate relations for A with respect to the scalars. The assumption on X is that:

$$\varphi_2(X D_0 A D_1 \cdots A D_n) = \sum_{m=1}^{n} \varphi_2(\eta_0(E_{\mathscr{D}}(D_0 A \cdots A D_{m-1})) \cdot D_m A \cdots A D_n), \tag{3.19}$$

for every $n \geqslant 0$ and every $D_0, D_1, ..., D_n \in \mathscr{D}$. By setting in (3.19) $D_0 = D_1 = \cdots = D_n = I_2$ (the unit of $M_2(\mathscr{A})$), we get:

$$\varphi_2(X A^n) = \sum_{m=1}^{n} \varphi_2(\eta_0(E_{\mathscr{D}}(A^{m-1})) \cdot A^{n-m}), \; n \geqslant 0. \tag{3.20}$$

It is however immediately checked that:

$$\eta_0(D_{\mathscr{D}}(A^k)) = \varphi_2(A^k)\, I_2, k \geqslant 0;$$

hence (3.20) comes to:

$$\varphi_2(XA^n) = \sum_{m=1}^{n} \varphi_2(A^{m-1}) \cdot \varphi_2(A^{n-m}), n \geqslant 0,$$

which says exactly that X fulfills the conjugate relations for A with respect to the scalars.

We now go to the proof of the opposite inequality, $\Phi^*(A : \mathscr{D}, \eta_0) \leqslant \Phi^*(A)$. The method is the same as above (although the calculations will be more complicated): we assume that A has a conjugate vector $X \in L^2(M_2(\mathscr{A}), \varphi_2)$, with respect to the scalars, and we will show that X also fulfills the conjugate relations for A with respect to $\mathscr{D}$ and η_0. We identify the vector X with a matrix over $L^2(\mathscr{A}, \varphi)$ (as in 2.1.3):

$$X = \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix}, \qquad \text{with} \quad \xi_{ij} \in L^2(\mathscr{A}, \varphi). \tag{3.21}$$

Note that

$$A = A^* \text{ in } M_2(\mathscr{A}) \Rightarrow X = X^* \text{ in } L^2(M_2(\mathscr{A}), \varphi_2) \qquad \text{(by Remark 2.3.2)}$$

$$\Rightarrow \xi_{12} = \xi_{21}^* \text{ in } L^2(\mathscr{A}, \varphi).$$

Before doing anything else, let us show that in (3.21) we have $\xi_{11} = \xi_{22} = 0$. To this end we will use "the even half" of the conjugate relations fulfilled by X:

$$\varphi(XA^{2k}) = \sum_{l=1}^{2k} \varphi(A^{l-1}) \cdot \varphi(A^{2k-l}), k \geqslant 0.$$

Every term in the latter sum is 0, because one of A^{l-1} and A^{2k-l} must always have vanishing diagonal entries. So we get $\varphi(XA^{2k}) = 0$, hence $X \perp A^{2k}$ in $L^2(M_2(\mathscr{A}), \varphi_2)$, for every $k \geqslant 0$. Since on the other hand the definition of the conjugate vector contains the fact that

$$X \in \overline{\operatorname{span}}^{\|\cdot\|_2}\{A^n \mid n \geqslant 0\} \subseteq L^2(M_2(\mathscr{A}), \varphi_2),$$

and since (obviously) $A^n \perp A^m$ when n, m have different parities, we infer that actually:

$$X \in \overline{\text{span}}^{\|\cdot\|_2}\{A^{2k+1} \mid k \geqslant 0\} = \overline{\text{span}}^{\|\cdot\|_2}\left\{\begin{pmatrix} 0 & a(a^*a)^k \\ a^*(aa^*)^k & 0 \end{pmatrix} \middle| \, k \geqslant 0\right\}. \tag{3.22}$$

From the discussion in 2.1.3 it is clear that convergence in $L^2(M_2(\mathscr{A}), \varphi_2)$ implies "entry-wise convergence" in $L^2(\mathscr{A}, \varphi)$. Therefore (3.22) has as consequence that $\xi_{11} = \xi_{22} = 0$, as desired, and we can write:

$$X = \begin{pmatrix} 0 & \xi^* \\ \xi & 0 \end{pmatrix}, \tag{3.23}$$

where $\xi := \xi_{21}$ of (3.21).

Besides (3.23), there is another consequence of (3.22) which will be used in the sequel, namely that:

$$\varphi(\xi a(a^*a)^m) = \varphi(\xi^* a^*(aa^*)^m), \quad \forall m \geqslant 0. \tag{3.24}$$

Indeed, for every given $m \geqslant 0$, (3.22) implies:

$$\begin{aligned} \begin{pmatrix} \xi^* a^*(aa^*)^m & 0 \\ 0 & \xi a(a^*a)^m \end{pmatrix} &= XA^{2m+1} \in \overline{\text{span}}^{\|\cdot\|_2}\{A^{2k} \mid k \geqslant m+1\} \\ &= \overline{\text{span}}^{\|\cdot\|_2}\left\{\begin{pmatrix} (aa^*)^k & 0 \\ 0 & (a^*a)^k \end{pmatrix} \middle| \, k \geqslant m+1\right\}. \end{aligned}$$

Then the fact that $\varphi((aa^*)^k) = \varphi((a^*a)^k)$, $k \geqslant m+1$, can be passed through the closed linear span to yield (3.24).

Now, recall that our goal is to prove that X fulfills the conjugate relations for A, with respect to $\mathscr{D}$ and η_0; these relations are exactly as described in (3.19). Since $\mathscr{D} = \text{span}\{V_{11}, V_{22}\}$, it actually suffices to check that:

$$\begin{aligned} &\varphi_2(XV_{i_0 i_0} AV_{i_1 i_1} \cdots AV_{i_n i_n}) \\ &\quad = \sum_{m=1}^{n} \varphi_2(\eta_0(E_{\mathscr{D}}(V_{i_0 i_0} A \cdots AV_{i_{m-1} i_{m-1}})) \cdot V_{i_m i_m} A \cdots AV_{i_n i_n}), \end{aligned} \tag{3.25}$$

for every $n \geqslant 0$ and every $i_0, i_1, ..., i_n \in \{1, 2\}$.

The verification of (3.25) goes on a line similar to the one used for checking Eq. (3.7) in the proof of Proposition 3.6. The left-hand side of (3.25) is evaluated as:

$$\begin{aligned}
&\tfrac{1}{2}\varphi((X)_{i_n i_0}(A)_{i_0 i_1}\cdots(A)_{i_{n-1}i_n})\\
&\quad=\tfrac{1}{2}\delta_{i_0\bar{i}_n}\cdot\delta_{i_0\bar{i}_1}\cdots\delta_{i_{n-1}\bar{i}_n}\cdot\varphi((X)_{\bar{i}_0 i_0}(A)_{i_0\bar{i}_0}\cdots(A)_{i_{n-1}\bar{i}_{n-1}})\\
&\quad=\begin{cases}2^{-1}\varphi(\xi a(a^*a)^k), & \text{if}\quad n=2k+1 \quad\text{and}\\ & (i_0, i_1, ..., i_n)=(1, 2, ..., 1, 2)\\ 2^{-1}\varphi(\xi^* a^*(aa^*)^k), & \text{if}\quad n=2k+1\quad\text{and}\\ & (i_0, i_1, ..., i_n)=(2, 1, ..., 2, 1)\\ 0, & \text{otherwise.}\end{cases}
\end{aligned}\tag{3.26}$$

The general term (indexed by $1\leqslant m\leqslant n$) on the right-hand side of (3.25) is:

$$\begin{aligned}
&\varphi_2(\eta_0(E_{\mathscr{D}}(V_{i_0 i_0}A\cdots AV_{i_{m-1}i_{m-1}}))\cdot V_{i_m i_m}A\cdots AV_{i_n i_n})\\
&\quad=\varphi_2(\eta_0(\delta_{i_0 i_{m-1}}\varphi((A)_{i_0 i_1}\cdots(A)_{i_{m-2}i_{m-1}})\,V_{i_0 i_0})\cdot V_{i_m i_m}A\cdots AV_{i_n i_n})\\
&\quad=\delta_{i_0 i_{m-1}}\varphi((A)_{i_0 i_1}\cdots(A)_{i_{m-2}i_{m-1}})\cdot\varphi_2(V_{\bar{i}_0\bar{i}_0}V_{i_m i_m}A\cdots AV_{i_n i_n})\\
&\quad=\tfrac{1}{2}\delta_{i_0 i_{m-1}}\cdot\delta_{i_0\bar{i}_m}\cdot\delta_{i_0\bar{i}_1}\delta_{i_1\bar{i}_2}\cdots\delta_{i_{n-1}\bar{i}_n}\cdot\varphi((A)_{i_0\bar{i}_0}\cdots(A)_{i_{m-2}\bar{i}_{m-2}})\\
&\quad\quad\cdot\varphi((A)_{i_m\bar{i}_m}\cdots(A)_{i_{n-1}\bar{i}_{n-1}})\\
&\quad=\begin{cases}2^{-1}\varphi((aa^*)^{(m-1)/2})\,\varphi((a^*a)^{(n-m)/2}),\\ \qquad\text{if } m, n \text{ are both odd and}\quad(i_0, i_1, ..., i_n)=(1, 2, ..., 1, 2)\\ 2^{-1}\varphi((a^*a)^{(m-1)/2})\,\varphi((aa^*)^{(n-m)/2}),\\ \qquad\text{if } m, n\text{ are both odd and}\quad(i_0, i_1, ..., i_n)=(2, 1, ..., 2, 1)\\ 0,\quad\text{otherwise.}\end{cases}
\end{aligned}\tag{3.27}$$

By comparing (3.26) with (3.27) (and by also taking (3.24) into account) we see that all it takes in order to obtain (3.25) is:

$$\varphi(\xi a(a^*a)^k)=\sum_{l=0}^{k}\varphi((aa^*)^l)\cdot\varphi((a^*a)^{k-l}),\quad\forall k\geqslant 0.\tag{3.28}$$

Finally, we obtain (3.28) by using "the odd half" of the conjugate relations (with respect to the scalars), which are fulfilled by X:

$$\varphi(XA^{2k+1})=\sum_{l=1}^{2k+1}\varphi_2(A^{l-1})\cdot\varphi_2(A^{2k+1-l}),\qquad k\geqslant 0.\tag{3.29}$$

Indeed, we have:

$$\begin{aligned}\varphi_2(XA^{2k+1}) &= \varphi_2\begin{pmatrix}\xi^* a^*(aa^*)^k & 0\\ 0 & \xi a(a^*a)^k\end{pmatrix}\\ &= \tfrac{1}{2}(\varphi(\xi^* a^*(aa^*)^k) + \varphi(\xi a(a^*a)^k))\\ &= \varphi(\xi a(a^*a)^k), \qquad \text{by (3.24);}\end{aligned}$$

while on the other hand it is immediate that:

$$\begin{aligned}\sum_{l=1}^{2k+1} \varphi_2(A^{l-1}) \cdot \varphi_2(A^{2k+1-l}) &= \sum_{l=0}^{k} \varphi_2(A^{2l}) \cdot \varphi_2(A^{2(k-l)})\\ &= \sum_{l=0}^{k} \varphi((aa^*)^l) \cdot \varphi((a^*a)^{k-l})\end{aligned}$$

(due to the particular form of A). So actually (3.28) reduces to (3.29). Q.E.D.

We now discuss the special property of the R-diagonal element which will ensure the equality in (3.1) of Theorem 3.2.

3.8. PROPOSITION. *In the framework of the Notations* 3.4, *we have that*: $Alg(\{A\} \cup \mathcal{D})$ *is free from* $M_2(\mathbf{C}I)$ *with amalgamation over* $\mathcal{D}$ *if and only if* a *is* R*-diagonal.*

In the proof of the Proposition 3.8 we will use the following lemma.

3.9. LEMMA. *Let* $(\mathcal{M}, \psi)$ *be a* W^**-probability space, and let us denote, for every* $b \in \mathcal{M}$ *and every* $k \geqslant 1$:

$$\begin{cases} w_{11;\,k}(b) = (bb^*)^k - \psi((bb^*)^k)\, I\\ w_{12;\,k}(b) = b(b^*b)^{k-1}\\ w_{21;\,k}(b) = b^*(bb^*)^{k-1}\\ w_{22;\,k}(b) = (b^*b)^k - \psi((b^*b)^k)\, I.\end{cases} \tag{3.30}$$

Then the following statements about an element $b \in \mathcal{M}$ *are equivalent*:

1. b *is* R*-diagonal in* $(\mathcal{M}, \psi)$.
2. *We have that*:

$$\psi(w_{\bar{i}_0 i_1;\,k_1}(b)\, w_{\bar{i}_1 i_2;\,k_2}(b) \cdots w_{\bar{i}_{n-1} i_n;\,k_n}(b)) = 0 \tag{3.31}$$

for every $n \geqslant 1$, $i_0, i_1, ..., i_n \in \{1, 2\}$ *and* $k_1, ..., k_n \geqslant 1$. (*Same as in the preceding propositions, we used in Eq.* (3.31) *the convention of notation* $\bar{i} = 3 - i$, *for* $i \in \{1, 2\}$.)

Proof of Lemma 3.9. $1 \Rightarrow 2$. We can assume that $b = up$ where $u \in \mathcal{M}$ is a Haar unitary, $p = p^*$ is even (i.e. $\psi(p^k) = 0$ for k odd) and $\{u, u^*\}$ is free from $\{p\}$. Thus we have, for every $k \geqslant 1$:

$$w_{11;\,k}(b) = u(p^{2k} - \psi(p^{2k})\, I)\, u^*, \qquad w_{12;\,k}(b) = up^{2k-1},$$
$$w_{21;\,k}(b) = p^{2k-1}u^*, \qquad w_{22;\,k}(b) = p^{2k} - \psi(p^{2k})\, I.$$

Every $w_{ij;\,k}(b)$ can be viewed as a word with 1, 2, or 3 letters over the alphabet:

$$\{u, u^*\} \cup \{p^k - \psi(p^k)\, I \mid k \geqslant 1\}; \tag{3.32}$$

and moreover the letters which form $w_{ij;\,k}(b)$ always come alternatively from $\{u, u^*\}$ and $\{p^k - \psi(p^k)\, I \mid k \geqslant 1\}$.

Given any $n \geqslant 1$, $i_0, i_1, ..., i_n \in \{1, 2\}$ and $k_1, ..., k_n \geqslant 1$, we claim that the product:

$$w := w_{\bar{\imath}_0 i_1;\, k_1}(b)\, w_{\bar{\imath}_1 i_2;\, k_2}(b) \cdots w_{\bar{\imath}_{n-1} i_n;\, k_n}(b) \tag{3.33}$$

still has the same alternance property of the letters, when viewed as a word over the alphabet (3.32). Indeed, for every $1 \leqslant m \leqslant n-1$ there are two possibilities: either $i_m = 1$, in which case $w_{\bar{\imath}_{m-1} i_m;\, k_m}(b)$ ends with u^* and $w_{\bar{\imath}_m i_{m+1};\, k_{m+1}}(b)$ begins with a $p^k - \psi(p^k)\, I$; or $i_m = 2$, in which case $w_{\bar{\imath}_{m-1} i_m;\, k_m}(b)$ ends with a $p^k - \psi(p^k)\, I$ and $w_{\bar{\imath}_m i_{m+1};\, k_{m+1}}(b)$ begins with u. In both cases, the concatenation of $w_{\bar{\imath}_{m-1} i_m;\, k_m}(b)$ and $w_{\bar{\imath}_m i_{m+1};\, k_{m+1}}(b)$ is still alternating.

But if the product w appearing in (3.33) is alternating when viewed as a word with letters from (3.32), then the equality $\psi(w) = 0$ follows from the definition of freeness (since every letter in (3.32) is in the kernel of ψ, and since $\{u, u^*\}$ is free from $\{p\}$).

$2 \Rightarrow 1$. By enlarging the space $(\mathcal{M}, \psi)$ if necessary, we can assume that there exists an R-diagonal element $c \in \mathcal{M}$, such that c^*c has the same distribution as b^*b. We denote $b_1 := b$, $b_2 := b^*$, $c_1 := c$, $c_2 = c^*$. We will show that:

$$\psi(b_{i_1} b_{i_2} \cdots b_{i_n}) = \psi(c_{i_1} c_{i_2} \cdots c_{i_n}), \qquad \forall n \geqslant 1, \quad \forall i_1, ..., i_n \in \{1, 2\}. \tag{3.34}$$

From (3.34) it will follow that b is R-diagonal (since c is so, and (3.34) means that b and c have the same $*$-distribution).

If $w_{ij;\,k}(c) \in \mathcal{M}$ is defined by analogy with Eq. (3.30), for $i, j \in \{1, 2\}$ and $k \geqslant 1$, then the implication $1 \Rightarrow 2$ proved above ensures that:

$$\psi(w_{\bar{\imath}_0 i_1;\, k_1}(c)\, w_{\bar{\imath}_1 i_2;\, k_2}(c) \cdots w_{\bar{\imath}_{n-1} i_n;\, k_n}(c)) = 0 \tag{3.35}$$

for every $n \geqslant 1$, $i_0, i_1, ..., i_n \in \{1, 2\}$ and $k_1, ..., k_n \geqslant 1$. The equality in (3.34) will be obtained by exploiting the similarity between (3.31) and (3.35).

We will prove (3.34) by induction on n. For $n=1$ we have to show that $\psi(b) = \psi(c)$, $\psi(b^*) = \psi(c^*)$. And indeed:

$$\psi(b) = \psi(w_{12;\,1}(b)) \overset{(3.31)}{=} 0 \overset{(3.35)}{=} \psi(w_{12;\,1}(c)) = \psi(c),$$

while $\psi(b^*) = 0 = \psi(c^*)$ can be shown in a similar way.

We consider now an $n \geqslant 2$. We assume that (3.34) is true for $1, 2, ..., n-1$ and we prove it for n. Let us fix some indices $i_1, ..., i_n \in \{1, 2\}$, about which we want to prove that (3.34) holds.

We take the product $b_{i_1} b_{i_2} \cdots b_{i_n}$, and draw a vertical bar between b_{i_m} and $b_{i_{m+1}}$ for every $1 \leqslant m \leqslant n-1$ such that $i_m = i_{m+1}$. (For instance if $b_{i_1} b_{i_2} \cdots b_{i_n}$ were to be $bb^*bbbb^*b^*b$, then our bars would look like this: $bb^*b \mid b \mid bb^* \mid b^*b$.) By examining the sub-products of $b_{i_1} b_{i_2} \cdots b_{i_n}$ which sit between consecutive vertical bars, we find that we have written:

$$b_{i_1} b_{i_2} \cdots b_{i_n} = (w_{\bar{j}_0 j_1;\, k_1}(b) + \lambda_1 I) \cdots (w_{\bar{j}_{s-1} j_s;\, k_s}(b) + \lambda_s I) \tag{3.36}$$

for some $s \geqslant 1$, $j_0, j_1, ..., j_s \in \{1, 2\}$, $k_1, ..., k_s \geqslant 1$ having $k_1 + \cdots + k_s = n$, and $\lambda_1, ..., \lambda_s \in \mathbf{C}$. The number λ_r, $1 \leqslant r \leqslant s$, is determined as follows: if $j_{r-1} = j_r$, then $\lambda_r = 0$; and if $j_{r-1} \neq j_r$, then $\lambda_r = \psi((b^*b)^{k_r})$.

In a similar way we can write:

$$c_{i_1} c_{i_2} \cdots c_{i_n} = (w_{\bar{j}_0 j_1;\, k_1}(c) + \lambda_1 I) \cdots (w_{\bar{j}_{s-1} j_s;\, k_s}(c) + \lambda_s I); \tag{3.37}$$

and moreover, the parameters $s, j_0, j_1, ..., j_s, k_1, ..., k_s, \lambda_1, ..., \lambda_s$ appearing in (3.37) coincide with those from (3.36). Indeed, the values of $s, j_0, j_1, ..., j_s, k_1, ..., k_s$ are determined solely by how the vertical bars are placed between the c_{i_m}'s in $c_{i_1} c_{i_2} \cdots c_{i_n}$, and this is identical to how the vertical bars were placed in $b_{i_1} b_{i_2} \cdots b_{i_n}$. After that, the value of every λ_r is determined as $\delta_{\bar{j}_{r-1},\, j_r} \psi((c^*c)^{k_r})$, which is again the same as in (3.36), due to the fact that b^*b and c^*c have the same distribution.

By applying ψ on both sides of (3.36) and then by expanding the product on the right-hand side, we obtain:

$$\begin{aligned}\psi(b_{i_1} b_{i_2} \cdots b_{i_n}) &= \psi(w_{\bar{j}_0 j_1;\, k_1}(b) \cdots w_{\bar{j}_{s-1} j_s;\, k_s}(b)) \\ &\quad + \sum_{\varnothing \neq A \subseteq \{1, ..., s\}} \left(\prod_{r \in A} \lambda_r \right) \cdot \psi \left(\prod_{r \in \{1, ..., s\} \setminus A} w_{\bar{j}_{r-1} j;\, k}(b) \right).\end{aligned}$$

The corresponding operations done in (3.37) yield an identical formula, where we have c's instead of b's. But we know that

$$\psi(w_{\bar{j}_0 j_1;\, k_1}(b) \cdots w_{\bar{j}_{s-1} j_s;\, k_s}(b)) \overset{(3.31)}{=} 0 \overset{(3.35)}{=} \psi(w_{\bar{j}_0 j_1;\, k_1}(c) \cdots w_{\bar{j}_{s-1} j_s;\, k_s}(c)),$$

while on the other hand the induction hypothesis gives us that:

$$\psi\Bigg(\prod_{r\in\{1,\dots,s\}\setminus A} w_{\bar{j}_{r-1}j;\,k}(b)\Bigg)=\psi\Bigg(\prod_{r\in\{1,\dots,s\}\setminus A} w_{\bar{j}_{r-1}j_r;\,k}(c)\Bigg),$$

for every $\varnothing\neq A\subseteq\{1,\dots,s\}$. These equalities imply in turn that $\psi(b_{i_1}b_{i_2}\cdots b_{i_n})=\psi(c_{i_1}c_{i_2}\cdots c_{i_n})$, as desired. Q.E.D.

Proof of Proposition 3.8. For every $i, j\in\{1,2\}$ and $k\geqslant 1$ we denote by $w_{ij;\,k}(a)$ the element of $\mathscr{A}$ defined by the same recipe as in Eq. (3.30) of Lemma 3.9, and we denote by $W_{ij;\,k}\in M_2(\mathscr{A})$ the matrix which has its (i,j)-entry equal to $w_{ij;\,k}(a)$, and its other entries equal to 0.

It is immediately seen that $Alg(\{A\}\cup\mathscr{D})$ is linearly spanned by the matrices of the form:

$$\begin{pmatrix}(aa^*)^k & 0\\ 0 & 0\end{pmatrix},\ \begin{pmatrix}0 & a(a^*a)^k\\ 0 & 0\end{pmatrix},\ \begin{pmatrix}0 & 0\\ a^*(aa^*)^k & 0\end{pmatrix},\ \begin{pmatrix}0 & 0\\ 0 & (a^*a)^k\end{pmatrix},\ k\geqslant 0;$$

and this implies the formula:

$$\{X\in Alg(\{A\}\cup\mathscr{D})\mid E_{\mathscr{D}}(X)=0\}=\operatorname{span}\{W_{ij;\,k}\mid i,j\in\{1,2\},k\geqslant 1\}. \tag{3.38}$$

On the other hand it is clear that:

$$\{X\in M_2(\mathbf{C}I)\mid E_{\mathscr{D}}(X)=0\}=\operatorname{span}\{V_{12},V_{21}\}. \tag{3.39}$$

From (3.38–39) it follows that $Alg(\{A\}\cup\mathscr{D})$ is free from $M_2(\mathbf{C}I)$ with amalgamation over $\mathscr{D}$ if and only if:

$$\begin{cases} E_{\mathscr{D}}(U'W_{j'_1j''_1;\,k_1}V_{i_1\bar{i}_1}\cdots V_{i_{n-1}\bar{i}_{n-1}}W_{j'_nj''_n;\,k_n}U'')=0,\\ \forall n\geqslant 1,\qquad \forall j'_1,j''_1,\dots,j'_n,j''_n,i_1,\dots,i_{n-1}\in\{1,2\},\\ \forall k_1,\dots,k_n\geqslant 1,\qquad \forall U',U''\in\{V_{11}+V_{22},V_{12},V_{21}\}.\end{cases} \tag{3.40}$$

The matrix product appearing in (3.40) is 0 if it is not true that $j''_1=i_1$, $\bar{i}_1=j'_2,\dots,j''_{n-1}=i_{n-1},\bar{i}_{n-1}=j'_n$. And consequently, (3.40) is equivalent to:

$$\begin{cases} E_{\mathscr{D}}(U'W_{\bar{i}_0i_1;\,k_1}V_{i_1\bar{i}_1}\cdots W_{\bar{i}_{n-2}i_{n-1};\,k_{n-1}}V_{i_{n-1}\bar{i}_{n-1}}W_{\bar{i}_{n-1}i_n;\,k_n}U'')=0,\\ \forall n\geqslant 1,\qquad \forall i_0,i_1,\dots,i_n\in\{1,2\},\\ \forall k_1,\dots,k_n\geqslant 1,\qquad \forall U',U''\in\{V_{11}+V_{22},V_{12},V_{21}\}.\end{cases} \tag{3.41}$$

But now, the matrix product appearing in (3.41) has one entry equal to $w_{\bar{i}_0i_1;\,k_1}(a)\cdots w_{\bar{i}_{n-1}i_n;\,k_n}(a)$ (which can appear on any of the four possible positions, depending on the choices of U' and U''); and has the other three entries equal to 0. This makes it immediate that the condition (3.41) is equivalent to the one presented in the Lemma 3.9.2 (applied here to $(\mathcal{A}, \varphi)$). Q.E.D.

Proof of Theorem 3.2. The inequality (3.1) is obtained by putting together the Equations (3.5) and (3.18) established in the Propositions 3.6, 3.7. From (3.5) and (3.18) it is also clear that (3.1) holds with equality if and only if:

$$\Phi^*(A: M_2(\mathbf{C}I), \eta) = \Phi^*(A: \mathcal{D}, \eta_0). \tag{3.42}$$

As reviewed in the Remark 2.8.3, a sufficient condition for (3.42) to take place is that $Alg(\{A\} \cup \mathcal{D})$ and $M_2(\mathbf{C}I)$ are free with amalgamation over $\mathcal{D}$. But by Proposition 3.8, this sufficient condition is equivalent to the fact that a is an R-diagonal element. Q.E.D.

3.10. *Remark.* If one is only interested in establishing the inequality (3.1), then a substantial short-cut can be taken through the above considerations. The short-cut goes by verifying directly that if (ξ, ξ^*) is a conjugative system for (a, a^*), then $X := \begin{pmatrix} 0 & \xi^* \\ \xi & 0 \end{pmatrix}$ fulfills the conjugate relations for $A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}$, with respect to the scalars; this immediately implies $\Phi^*(a, a^*) \geqslant 2\Phi^*(A) = 2\Phi^*(\mu)$, i.e. (3.1).

The reason for insisting to put into evidence the relations shown in (3.5) and (3.18) is that they also give us a non-trivial necessary and sufficient condition—Eq. (3.42)—for the minimal free Fisher information to be attained. As mentioned in Remark 3.3, the problem of determining what are the $*$-distributions of a which attain the minimal $\Phi^*(a, a^*)$ is open, and seemingly difficult. The condition in (3.42) helps clarifying the nature of this problem, by reducing it to the following one (also open):

3.11. *Problem.* Is it true that if $\Phi^*(A: M_2(\mathbf{C}I), \eta) = \Phi^*(A: \mathcal{D}, \eta_0) < \infty$, then necessarily $Alg(\{A\} \cup \mathcal{D})$ and $M_2(\mathbf{C}I)$ are free with amalgamation over $\mathcal{D}$?

Settling the Problem 3.11 in the affirmative would imply that if the minimal value of $\Phi^*(a, a^*)$ (under the constraint in (1.1)) is finite, then this minimal value can be reached only by an R-diagonal element. While on the other hand, a negative answer in 3.11 would provide examples of situations when the infimum in (1.1) is finite and is reached by non-R-diagonal elements.

4. MINIMIZATION OF FREE FISHER INFORMATION FOR MATRIX ENTRIES

Let d be a fixed positive integer. We will consider here the following two minimization problems:

(a) Determine the minimal possible value of $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d})$, if the family $\{a_{ij}\}_{1 \leqslant i, j \leqslant d}$ (of elements in some W^*-probability space $(\mathcal{A}, \varphi)$, with φ faithful trace) is such that the matrix $A = (a_{ij})_{i, j=1}^d$ has a prescribed $*$-distribution.

(b) Determine the minimal possible value of $\Phi^*(\{b_{ij}\}_{1 \leqslant i, j \leqslant d})$, if the family $\{b_{ij}\}_{1 \leqslant i, j \leqslant d}$ (of elements in some W^*-probability space $(\mathcal{A}, \varphi)$, with φ faithful trace) is such that the matrix $B = (b_{ij})_{i, j=1}^d$ is selfadjoint, and has a prescribed distribution. Note that if $B = B^*$, then $\{b_{ij}\}_{1 \leqslant i, j \leqslant d}$ is a selfadjoint family of elements of $\mathcal{A}$—indeed, the involution $\sigma(i, j) := (j, i)$ has the property that $b_{ij}^* = b_{\sigma(i, j)}$, for every $1 \leqslant i, j \leqslant d$.

The solutions of these problems are provided by the Theorem 1.2 stated in the Introduction. For instance for (b) we have that, given a probability measure μ with compact support on $\mathbf{R}$:

$$\min \left\{ \Phi^*(\{b_{ij}\}_{1 \leqslant i, j \leqslant d}) \,\middle|\, \begin{array}{l} B = (b_{ij})_{i, j=1}^d = B^* \\ \text{has distribution } \mu \end{array} \right\} = d^3 \Phi^*(\mu). \tag{4.1}$$

A similar formula holds in the framework of the problem (a) (but where for the role of μ we must now consider a linear functional on $\mathbf{C}\langle X, X^* \rangle$, which can appear as a $*$-distribution in a tracial W^*-probability space).

In order to infer (4.1) as a consequence of Theorem 1.2.2 (and the corresponding conclusion from Theorem 1.2.1), there is one more detail that needs to be verified—that the freeness conditions appearing in Theorem 1.2 can indeed be fulfilled, in the context where the joint distribution of A and A^* (in 1) and the distribution of B (in 2) are prescribed. We discuss in more detail the selfadjoint case of 2; the non-selfadjoint case is similar.

So, let μ be a fixed probability measure with compact support on $\mathbf{R}$. One can find a W^*-probability space $(\mathcal{M}, \psi)$, with ψ faithful trace, and x and $\{v_{ij}\}_{1 \leqslant i, j \leqslant d}$ in $\mathcal{M}$ such that:

(i) $x = x^*$ has distribution μ;

(ii) the v_{ij}'s form a family of matrix units (i.e., $v_{ij} v_{kl} = \delta_{jk} v_{il}$, $v_{ij}^* = v_{ji}$, $\forall 1 \leqslant i, j, k, l \leqslant d$, and $\sum_{i=1}^d v_{ii} = I$);

(iii) x is free from $\{v_{ij}\}_{1 \leqslant i, j \leqslant d}$.

(An example of such $(\mathcal{M}, \psi)$ is the free product $(L^\infty(\mu), d\mu) \star (M_d(\mathbf{C}), tr)$). Consider the compressed W^*-probability space $(\mathcal{A}, \varphi)$, where $\mathcal{A} := v_{11} \mathcal{M} v_{11}$ and $\varphi(\cdot) := d\psi(\cdot)$ on $\mathcal{A}$; and in $\mathcal{A}$ consider the family of compressions $b_{ij} := v_{1i} x v_{j1}$, $1 \leqslant i, j \leqslant d$. Then the self-adjoint matrix $B = (b_{ij})_{i, j=1}^d$ has distribution μ in $(M_d(\mathcal{A}), \varphi_d)$, and is on the other hand free from $M_d(\mathbf{C}I) \subseteq M_d(\mathcal{A})$. These things happen because the spaces

$(\mathcal{M}, \psi)$ and $(M_d(\mathcal{A}), \varphi_d)$ are isomorphic, via the $*$-isomorphism $\mathcal{M} \ni y \mapsto (v_{1i} y v_{j1})_{i,j=1}^d \in M_d(\mathcal{A})$, which sends x to B and $\operatorname{span}\{v_{ij} \mid 1 \leqslant i, j \leqslant d\}$ onto $M_d(\mathbf{C}I)$.

It thus remains that we prove the Theorem 1.2. We will in fact prove more, namely:

4.1. PROPOSITION. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, and let $\mathcal{B} \subseteq \mathcal{A}$ be a unital W^*-subalgebra. Let d be a positive integer; consider the W^*-probability space $(M_d(\mathcal{A}), \varphi_d)$ (defined as in Notations 2.1.2), and the W^*-subalgebra $M_d(\mathcal{B}) \subseteq M_d(\mathcal{A})$.*

1. *For every $A = (a_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$, we have:*

$$\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d} : \mathcal{B}) = d^3 \Phi^*(\{A, A^*\} : M_d(\mathcal{B})). \tag{4.2}$$

2. *For every $G = (g_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$ such that $G = G^*$, we have:*

$$\Phi^*(\{g_{ij}\}_{1 \leqslant i, j \leqslant d} : \mathcal{B}) = d^3 \Phi^*(G : M_d(\mathcal{B})). \tag{4.3}$$

The name of the selfadjoint matrix appearing in 4.1.2 was changed to G (from B, as was in Theorem 1.2) in order to avoid any confusion with the elements of $M_d(\mathcal{B})$. Note that if in Proposition 4.1 we take $\mathcal{B} = \mathbf{C}I$, then the Eqs. (4.2) and (4.3) become:

$$\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}) = d^3 \Phi^*(\{A, A^*\} : M_d(\mathbf{C}I)), \tag{4.4}$$

and respectively

$$\Phi^*(\{g_{ij}\}_{1 \leqslant i, j \leqslant d}) = d^3 \Phi^*(G : M_d(\mathbf{C}I)). \tag{4.5}$$

The statements of Theorem 1.2 follow immediately from these relations. Indeed, for 1.2.1 we only have to use (4.4) and the fact (reviewed in Remark 2.6) that $\Phi^*(\{A, A^*\} : M_d(\mathbf{C}I)) \geqslant \Phi^*(A, A^*)$, with equality when $\{A, A^*\}$ is free from $M_d(\mathbf{C}I)$; and similarly for 1.2.2.

Proof of Proposition 4.1. The proofs of 4.1.1 and 4.1.2 are similar to each other (and also similar to the proof of Proposition 3.6 from the previous section). For this reason, we will only do 4.1.1, and leave 4.1.2 as an exercise to the reader.

In 4.1.1 we first consider the situation when $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d} : \mathcal{B}) < \infty$. In this case, the family $\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}$ has a conjugate system $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i, j \leqslant d}$ with respect to $\mathcal{B}$. Let us define:

$$X := \frac{1}{d} (\xi_{ji})_{i,j=1}^d \in L^2(M_d(\mathcal{A}), \varphi_d) \tag{4.6}$$

(where the identification discussed in 2.1.3 is used). We will show that $\{X, X^*\}$ is a conjugate system for $\{A, A^*\}$, with respect to $M_d(\mathscr{B})$. This will entail (4.2) (under the hypothesis $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}) < \infty$), because it will give:

$$\begin{aligned}\Phi^*(\{A, A^*\} : M_d(\mathscr{B})) &= \|X\|^2_{L^2(\varphi_d)} + \|X^*\|^2_{L^2(\varphi_d)} \\ &\overset{(2.3)}{=} \frac{1}{d} \sum_{i, j=1}^{d} \left(\left\| \frac{1}{d} \xi_{ji} \right\|^2_{L^2(\varphi)} + \left\| \frac{1}{d} \xi_{ij}^* \right\|^2_{L^2(\varphi)} \right) \\ &= \frac{1}{d^3} \sum_{i, j=1}^{d} (\|\xi_{ij}\|^2_{L^2(\varphi)} + \|\xi_{ij}^*\|^2_{L^2(\varphi)}) \\ &= \frac{1}{d^3} \Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}).\end{aligned}$$

The conjugate relations which we need to verify are:

$$\begin{aligned}&\varphi_d(X B_0 A_{i_1} B_1 \cdots A_{i_n} B_n) \\ &\quad = \sum_{m=1}^{n} \delta_{i_m, 1} \cdot \varphi_d(B_0 A_{i_1} \cdots A_{i_{m-1}} B_{m-1}) \cdot \varphi_d(B_m A_{i_{m+1}} \cdots A_{i_n} B_n), \qquad (4.7)\end{aligned}$$

for $n \geqslant 1$, $B_0, B_1, ..., B_n \in M_d(\mathscr{B})$ and $i_1, i_2, ..., i_n \in \{1, 2\}$, where we denoted $A_1 := A$, $A_2 := A^*$. The list of conjugate relations for $\{A, A^*\}$ with respect to $M_d(\mathscr{B})$ also contains:

$$\begin{aligned}&\varphi_d(X^* B_0 A_{i_1} B_1 \cdots A_{i_n} B_n) \\ &\quad = \sum_{m=1}^{n} \delta_{i_m, 2} \cdot \varphi_d(B_0 A_{i_1} \cdots A_{i_{m-1}} B_{m-1}) \cdot \varphi_d(B_m A_{i_{m+1}} \cdots A_{i_n} B_n), \qquad (4.8)\end{aligned}$$

(for $n \geqslant 1$, $B_0, B_1, ..., B_n \in M_d(\mathscr{B})$, $i_1, i_2, ..., i_n \in \{1, 2\}$), and

$$\varphi_d(XB) = \varphi_d(X^*B) = 0, \forall B \in M_d(\mathscr{B}). \qquad (4.9)$$

However, (4.8) readily follows from (4.7) by taking an adjoint and then doing a circular permutation under φ_d; while (4.9) is a direct consequence of the equations $\varphi(\xi_{ij} b) = 0$, $1 \leqslant i, j \leqslant d$, $b \in \mathscr{B}$ which appear on the list of conjugate relations satisfied by the family $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i, j \leqslant d}$. (Hence indeed, only (4.7) needs to be checked.)

For every $1 \leqslant k, l \leqslant d$ and every $b \in \mathscr{B}$ let us denote by $V_{kl} \otimes b$ the matrix in $M_d(\mathscr{B})$ which has its (k, l)-entry equal to b, and all its other entries equal to 0. By multilinearity we can assume in (4.7) that $B_0 = V_{k_0 l_0} \otimes b_0, ..., B_n = V_{k_n l_n} \otimes b_n$ for some $k_0, l_0, ..., k_n, l_n \in \{1, ..., d\}$ and $b_0, ..., b_n \in \mathscr{B}$. The left-hand side of (4.7) is then equal to:

$$\begin{aligned}
&\varphi_d((V_{l_n l_n} \otimes I)\, X(V_{k_0 l_0} \otimes b_0)\, A_{i_1}(V_{k_1 l_1} \otimes b_1) \cdots A_{i_n}(V_{k_n l_n} \otimes b_n)) \\
&\qquad = \frac{1}{d}\, \varphi((X)_{l_n k_0}\, b_0 (A_{i_1})_{l_0 k_1}\, b_1 \cdots (A_{i_n})_{l_{n-1} k_n} b_n) \\
&\qquad = \frac{1}{d^2}\, \varphi(\xi_{k_0 l_n}\, b_0 (A_{i_1})_{l_0 k_1}\, b_1 \cdots (A_{i_n})_{l_{n-1} k_n} b_n),
\end{aligned} \tag{4.10}$$

where "$(A_{i_1})_{l_0 k_1}$" stands for the (l_0, k_1)-entry of the matrix A_{i_1}, etc (same conventions of notation as in Section 3). By using the conjugate relations satisfied by the family $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i,\, j \leqslant d}$, we can continue (4.10) with:

$$\begin{aligned}
&= \frac{1}{d^2} \sum_{m=1}^{n} \delta_{i_m, 1}\, \delta_{k_0, l_{m-1}}\, \delta_{l_n, k_m} \cdot \varphi(b_0 (A_{i_1})_{l_0 k_1} \cdots (A_{i_{m-1}})_{l_{m-2} k_{m-1}} b_{m-1}) \\
&\qquad \cdot \varphi(b_m (A_{i_{m+1}})_{l_m k_{m+1}} \cdots (A_{i_n})_{l_{n-1} k_n} b_n).
\end{aligned} \tag{4.11}$$

It is straightforward to observe that the summation which appeared in (4.11) is equal to the right-hand side of (4.7).

In order to complete the verification that $\{X, X^*\}$ is the conjugate of $\{A, A^*\}$ with respect to $M_d(\mathscr{B})$, we must also show that:

$$X \in \overline{Alg(\{A, A^*\} \cup M_d(\mathscr{B}))}^{\|\cdot\|_2}. \tag{4.12}$$

For every $1 \leqslant i, j \leqslant d$ let us denote by $A_{ij} \in M_d(\mathscr{A})$ and respectively by $X_{ij} \in L^2(M_d(\mathscr{A}), \varphi_d)$ the matrix which has a_{ij} (respectively ξ_{ij}) on its $(1, 1)$-entry, and 0's on all the other entries. Then:

$$A_{ij} = (V_{1i} \otimes I)\, A(V_{j1} \otimes I) \in Alg(\{A, A^*\} \cup M_d(\mathscr{B})), \forall 1 \leqslant i, j \leqslant d.$$

Among the properties satisfied by $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i,\, j \leqslant d}$ (as conjugate for $\{a_{ij}, a_{ij}^*\}_{1 \leqslant i,\, j \leqslant d}$), we also have that:

$$\xi_{kl} \in \overline{Alg(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i,\, j \leqslant d} \cup \mathscr{B})}^{\|\cdot\|_2}, \qquad \forall 1 \leqslant k, l \leqslant d.$$

Consequently, by forming polynomials with matrices of the form A_{ij}, A_{ij}^* and $V_{11} \otimes b$ $(b \in \mathscr{B})$, and then by taking $\|\cdot\|_2$-limits, we obtain that every X_{kl} $(1 \leqslant k, l \leqslant d)$ belongs to the $\|\cdot\|_2$-closed space indicated in (4.12). This space is invariant under the left/right action of elements from $M_d(\mathscr{B})$, hence we can conclude that it also contains

$$X = \frac{1}{d} \sum_{k, l=1}^{d} (V_{l1} \otimes I)\, X_{kl} (V_{1k} \otimes I),$$

as desired.

Equation (4.2) is now proved in the case when $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d} : \mathscr{B}) < \infty$. It remains to show that $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d} : \mathscr{B}) = \infty \Rightarrow \Phi^*(\{A, A^*\} : M_d(\mathscr{B})) = \infty$; or equivalently, that $\Phi^*(\{A, A^*\} : M_d(\mathscr{B})) < \infty \Rightarrow \Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d} : \mathscr{B}) < \infty$.

If $\Phi^*(\{A, A^*\} : M_d(\mathscr{B})) < \infty$, then there exists $X \in L^2(M_d(\mathscr{A})\, \varphi_d)$ such that $\{X, X^*\}$ fulfills the conjugate relations for $\{A, A^*\}$, with respect to $M_d(\mathscr{B})$. We write X as a $d \times d$-matrix (as in 2.1.3):

$$X = (\eta_{ij})_{i, j=1}^{d}, \qquad \text{with} \quad \eta_{ij} \in L^2(\mathscr{A}, \varphi), \quad 1 \leqslant i, j \leqslant d;$$

and we set $\xi_{ij} := d\eta_{ji}$, $1 \leqslant i, j \leqslant d$. We claim that $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i, j \leqslant d}$ fulfills the conjugate relations for $\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}$, with respect to $\mathscr{B}$. Since the calculation verifying this claim is very similar in spirit with the one which concluded the proof of Proposition 3.6, we will only mention its guiding line, and leave the details to the reader. The generic relation that needs to be proved is of the form:

$$\begin{aligned} &\varphi(\xi_{kl} b_0 (A_{i_1})_{k_1 l_1} b_1 \cdots (A_{i_n})_{k_n l_n} b_n) \\ &\quad = \sum_{m=1}^{n} \delta_{i_m, 1} \delta_{k, k_m} \delta_{l, l_m} \varphi(b_0 (A_{i_1})_{k_1 l_1} \cdots (A_{i_{m-1}})_{k_{m-1} l_{m-1}} b_{m-1}) \\ &\qquad \cdot \varphi(b_m (A_{i_m})_{k_m l_m} \cdots (A_{i_n})_{k_n l_n} b_n), \end{aligned} \tag{4.13}$$

for $n \geqslant 1$, $b_0, ..., b_n \in \mathscr{B}$, $i_1, ..., i_n \in \{1, 2\}$, $k_1, l_1, ..., k_n, l_n \in \{1, ..., d\}$. The line for establishing (4.13) goes by writing its left-hand side as

$$d^2 \varphi_d(X(V_{k, k_1} \otimes b_0)\, A_{i_1}(V_{l_1, k_2} \otimes b_1)\, A_{i_2}(V_{l_2, k_3} \otimes b_2) \cdots A_{i_n}(V_{l_n, l} \otimes b_n)); \tag{4.14}$$

then by using in (4.14) the conjugate relations fulfilled by $\{X, X^*\}$; and finally by evaluating (in a straightforward way) the terms of the summation which is obtained in this manner.

But if $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i, j \leqslant d}$ fulfills the conjugate relations for $\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}$, with respect to $\mathscr{B}$, then it follows that $\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d} : \mathscr{B}) < \infty$, and this concludes the proof.

Q.E.D.

By using the Theorem 1.2, we can now easily prove the generalization of our minimization result for Φ^*, which was stated in Theorem 1.3.

Proof of Theorem 1.3. Let us fix a probability measure ν with compact support on $[0, \infty)$, and a positive integer d. We denote the symmetric square root of ν (defined as in 3.1) by μ.

Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, and let $\{a_{ij}\}_{1 \leqslant i, j \leqslant d}$ be elements of $\mathscr{A}$ such that if we set $A := (a_{ij})_{i,j=1}^{d}$, then A^*A has distribution ν in $(M_d(\mathscr{A}), \varphi_d)$. Then:

$$\Phi^*(A, A^*) \geqslant 2\Phi^*(\mu) \tag{4.15}$$

(by Theorem 1.1); if we combine this with the inequality (1.4) of Theorem 1.2, we get:

$$\Phi^*(\{a_{ij}, a_{ij}^*\}_{1 \leqslant i, j \leqslant d}) \geqslant 2d^3\Phi^*(\mu). \tag{4.16}$$

A discussion similar to the one preceding Proposition 4.1 shows that (in the context where ν and d are prescribed) we can pick the family $\{a_{ij}\}_{1 \leqslant i, j \leqslant d}$ such that in addition to the condition that the distribution of A^*A be ν, we also have:

(i) A is R-diagonal in $M_d(\mathscr{A})$; and

(ii) $\{A, A^*\}$ is free from the algebra of scalar matrices $M_d(\mathbf{C}I) \subseteq M_d(\mathscr{A})$.

The condition (i) implies that (4.15) holds with equality, while (ii) implies equality in (1.4) of Theorem 1.2; hence (i) + (ii) ensure that the lower bound $2d^3\Phi^*(\mu)$ of (4.16) is actually attained. Q.E.D.

In the case when $\Phi^*(\mu) < \infty$, it would be interesting to know if the conditions (i) and (ii) mentioned in the proof of Theorem 1.3 are also necessary for (4.16) to hold with equality. Deciding on this fact would amount to solving the Problem 3.11 (which corresponds to the particular case $d = 1$), and another problem of the same nature—whether the equality $\Phi^*(\{A, A^*\} : M_d(\mathbf{C}I)) = \Phi^*(A, A^*) < \infty$ must imply the freeness of $\{A, A^*\}$ and $M_d(\mathbf{C}I)$.

5. THE CORRESPONDING MAXIMIZATION PROBLEMS FOR THE FREE ENTROPY χ^*

In this section we will consider the concept of free entropy χ^*, defined in [8] in terms of the free information Φ^*. We will treat the questions of maximizing χ^*, under constraints similar to those discussed in the previous sections. The results concerning χ^* will follow from the corresponding results for the free Fisher information.

Let $(\mathscr{A}, \varphi)$ be a W^*-probability space with φ a faithful trace, and consider a selfadjoint family of elements of $\mathscr{A}$ which is given in the form: $\{a_i, a_i^*\}_{1 \leqslant i \leqslant m} \cup \{b_j\}_{1 \leqslant j \leqslant n}$, where $b_j = b_j^*$ for $1 \leqslant j \leqslant n$. By enlarging $(\mathscr{A}, \varphi)$ if necessary, we can assume there exist circular elements

$c_1, ..., c_m \in \mathscr{A}$ and semicircular elements $s_1, ..., s_n \in \mathscr{A}$ such that $\{c_1, c_1^*\}, ..., \{c_m, c_m^*\}, \{s_1\}, ..., \{s_n\}, \{a_1, a_1^*, ..., a_m, a_m^*, b_1, ..., b_n\}$ are free. We will assume in addition that $c_1, ..., c_m$ and $s_1, ..., s_n$ are normalized by their variance (i.e. $\varphi(c_i^* c_i) = 1 = \varphi(s_j^2)$, for every $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$). Then the free entropy $\chi^*(\{a_i, a_i^*\}_{1 \leqslant i \leqslant m} \cup \{b_j\}_{1 \leqslant j \leqslant n}) \in [-\infty, \infty)$ is defined by the formula

$$\begin{aligned}
&\chi^*(\{a_i, a_i^*\}_{1 \leqslant i \leqslant m} \cup \{b_j\}_{1 \leqslant j \leqslant n}) \\
&\qquad = \frac{2m+n}{2} \log(2\pi e) \\
&\qquad\quad + \frac{1}{2} \int_0^\infty \left(\frac{2m+n}{1+t} - \Phi^*(\{a_i + \sqrt{t} c_i, a_i^* + \sqrt{t} c_i^*\}_{1 \leqslant i \leqslant m} \right. \\
&\qquad\quad \left. \cup \{b_j + \sqrt{t} s_j\}_{1 \leqslant j \leqslant n}) \right) dt. \qquad (5.1)
\end{aligned}$$

The integral on the right-hand side of (5.1) makes sense, and takes indeed value in $[-\infty, \infty)$—see Corollary 6.14, Proposition 7.2 in [8]. (In order to apply literally the estimates from [8], one first replaces every pair $\{c_j, c_j^*\}$ with the pair of selfadjoints $\{(c_j + c_j^*)/\sqrt{2}, (c_j - c_j^*)/i\sqrt{2}\}$—this does not affect the integrand on the right-hand side of (5.1).) Moreover, the value of the integral in (5.1) does not depend on the choice of $c_1, ..., c_m$, $s_1, ..., s_n$; in fact it is easy to see that $\chi^*(\{a_i, a_i^*\}_{1 \leqslant i \leqslant m} \cup \{b_j\}_{1 \leqslant j \leqslant n})$ depends only on the joint distribution of $\{a_1, a_1^*, ..., a_m, a_m^*, b_1, ..., b_n\}$ in $(\mathscr{A}, \varphi)$.

If μ is a probability measure with compact support on $\mathbf{R}$, then we will denote (similarly to how we did with Φ^* in Notation 2.9):

$$\chi^*(\mu) := \chi^*(x), \qquad (5.2)$$

where x is an arbitrary selfadjoint random variable with distribution μ. Similarly to the situation for Φ^*, there exists an explicit integral formula for $\chi^*(\mu)$, namely:

$$\chi^*(\mu) = \iint \log |s-t| \, d\mu(s) \, d\mu(t) + \frac{3}{4} + \frac{\log(2\pi)}{2} \qquad (5.3)$$

([5], Proposition 4.5, combined with [8], Proposition 7.6).

We now start towards the proofs of Theorems 1.4 and 1.5. Following the same line which we used for Φ^*, we will first do the Theorem 1.4 in the case $d = 1$. We will use the following freeness result.

5.1. PROPOSITION. *Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace. Let a, c be in $\mathscr{A}$, and assume that c can be factored as $c = up$, where $u \in \mathscr{A}$ is a unitary with Haar distribution, $p = p^* \in \mathscr{A}$ has a symmetric distribution, and $\{u, u^*\}$ is free from $\{p\}$. (In other words, we assume that c is R-diagonal.) If $\{a, a^*\}$ is free from $\{c, c^*\}$ in $(\mathscr{A}, \varphi)$, then the selfadjoint matrices*:

$$A = \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}, S = \begin{pmatrix} 0 & c \\ c^* & 0 \end{pmatrix} \tag{5.4}$$

are free in $(M_2(\mathscr{A}), \varphi_2)$.

Proof. We denote:

$$\mathscr{X} := \{u, u^*\} \cup \{p^k - \varphi(p^k) I \mid k \geqslant 1\}.$$

A word made with letters from the alphabet $\mathscr{X}$ will be called "alternating" if no two consecutive letters of the word are both from $\{u, u^*\}$ or both from $\{p^k - \varphi(p^k) I \mid k \geqslant 1\}$; the set of such alternating words will be denoted by $\mathscr{X}^*_{alt}$. Note that $\mathscr{X}^*_{alt} \subseteq \mathrm{Ker}(\varphi)$; this follows (by using the definition of freeness) from the facts that $\mathscr{X} \subseteq \mathrm{Ker}(\varphi)$ and that $\{u, u^*\}$ is free from $\{p\}$.

Let us consider on the other hand the set:

$$\mathscr{Y} = \mathscr{Y}_{11} \cup \mathscr{Y}_{12} \cup \mathscr{Y}_{21} \cup \mathscr{Y}_{22},$$

where:

$$\begin{aligned} &\mathscr{Y}_{11} = \{(aa^*)^k - \varphi((aa^*)^k) I) \mid k \geqslant 1\}, \\ &\mathscr{Y}_{12} = \{a(a^*a)^k \mid k \geqslant 0\}, \mathscr{Y}_{21} = \{a^*(aa^*)^k \mid k \geqslant 0\}, \\ &\mathscr{Y}_{22} = \{(a^*a)^k - \varphi((a^*a)^k) I) \mid k \geqslant 1\}. \end{aligned}$$

We will look at words of the form

$$w = (y_1 - \lambda_1 I)\, x_1 (y_2 - \lambda_2 I)\, x_2 \cdots (y_n - \lambda_n I)\, x_n, \tag{5.5}$$

where $n \geqslant 1$, $y_1, ..., y_n \in \mathscr{Y}$, $\lambda_1, ..., \lambda_n \in \mathbf{C}$, $x_1, ..., x_n \in \mathscr{X}^*_{alt}$, and where the following rules are obeyed:

$$\begin{cases} \text{if } y_m \in \mathscr{Y}_{11} \cup \mathscr{Y}_{21} (1 \leqslant m \leqslant n), \text{ then } x_m \text{ begins with } u; \\ \text{if } y_m \in \mathscr{Y}_{12} \cup \mathscr{Y}_{22} (1 \leqslant m \leqslant n), \text{ then } x_m \text{ begins with a } p^k - \varphi(p^k) I; \\ \text{if } y_m \in \mathscr{Y}_{11} \cup \mathscr{Y}_{12} (2 \leqslant m \leqslant n), \text{ then } x_{m-1} \text{ ends with } u^*; \\ \text{if } y_m \in \mathscr{Y}_{21} \cup \mathscr{Y}_{22} (2 \leqslant m \leqslant n), \text{ then } x_{m-1} \text{ ends with a } p^k - \varphi(p^k) I; \\ \text{if } y_m \in \mathscr{Y}_{11} \cup \mathscr{Y}_{22} (1 \leqslant m \leqslant n), \text{ then } \lambda_m = 0. \end{cases} \tag{5.6}$$

We will prove the following:

$$\text{Claim: If } w \text{ satisfies (5.6), then } \varphi\ (w)=0. \tag{5.7}$$

The proof of the Claim (5.7) will be done by induction on the number n of x_i's and y_i's entering the word w. For $n=1$, we have:

$$\begin{aligned}\varphi(w) &= \varphi((y_1-\lambda_1)\, x_1) \\ &= \varphi(y_1-\lambda_1)\, \varphi(x_1) && (\text{because } \{a, a^*\} \text{ free from } \{c, c^*\}) \\ &= 0 && (\text{because } \varphi(x_1)=0).\end{aligned}$$

Let us next assume the Claim (5.7) is true for $n-1$, and prove it for n. We first show that:

$$\begin{aligned}&\varphi((y_1-\lambda_1 I)\, x_1(y_2-\lambda_2 I)\, x_2 \cdots (y_n-\lambda_n I)\, x_n) \\ &\quad = \varphi((y_1-\lambda'_1 I)\, x_1(y_2-\lambda'_2 I)\, x_2 \cdots (y_n-\lambda'_n I)\, x_n),\end{aligned} \tag{5.8}$$

for every $y_1, ..., y_n \in \mathscr{Y}$, $\lambda_1, ..., \lambda_n, \lambda'_1, ..., \lambda'_n \in \mathbf{C}$, $x_1, ..., x_n \in \mathscr{X}^*_{alt}$ such that the rules (5.6) are satisfied. Clearly, it suffices to verify (5.8) in the situation when $(\lambda_1, ..., \lambda_n)$ differs from $(\lambda'_1, ..., \lambda'_n)$ on only one position k, $1 \leqslant k \leqslant n$. For that k we must have $y_k \in \mathscr{Y}_{12} \cup \mathscr{Y}_{21}$ (otherwise λ_k and λ'_k are both set to 0 in (5.6)). But in such a case the difference of the two sides of (5.8) equals

$$\begin{aligned}&(\lambda'_k-\lambda_k)\, \varphi((y_1-\lambda_1 I)\, x_1 \cdots (y_{k-1}-\lambda_{k-1} I)\, x_{k-1} x_k (y_{k+1}-\lambda_{k+1} I) \\ &\quad \times x_{k+1} \cdots (y_n-\lambda_n I)\, x_n);\end{aligned} \tag{5.9}$$

and the quantity in (5.9) is indeed equal to 0, due to the induction hypothesis. (The main point, in order to apply the induction hypothesis, is to note that in both the possible cases—$y_k \in \mathscr{Y}_{12}$, $y_k \in \mathscr{Y}_{21}$—we will have $x_{k-1}x_k \in \mathscr{X}^*_{alt}$; this happens because of the four "concatenation" rules stated in (5.6).)

Now, it is immediate that for every word $w=(y_1-\lambda_1 I)\, x_1(y_2-\lambda_2 I)\, x_2 \cdots (y_n-\lambda_n I)\, x_n$ as in (5.5), we can find some new scalars $\lambda'_1, ..., \lambda'_n \in \mathbf{C}$ such that $w'=(y_1-\lambda'_1 I)\, x_1(y_2-\lambda'_2 I)\, x_2 \cdots (y_n-\lambda'_n I)\, x_n$ still satisfies the rules (5.6), and such that in addition:

$$\varphi(y_1-\lambda'_1 I) = \cdots = \varphi(y_n-\lambda'_n I)=0. \tag{5.10}$$

Indeed, if $1 \leqslant m \leqslant n$ is such that $y_m \in \mathscr{Y}_{12} \cup \mathscr{Y}_{21}$, we can take $\lambda'_m=\varphi(y_m)$; while if $1 \leqslant m \leqslant n$ is such that $y_m \in \mathscr{Y}_{11} \cup \mathscr{Y}_{22}$, then the last rule (5.6) imposes $\lambda'_m=0=\lambda_m$—but in this case we also get $\varphi(y_m)=0$ from the definitions of $\mathscr{Y}_{11}, \mathscr{Y}_{22}$. The new word w' satisfies $\varphi(w')=0$; indeed, besides

(5.10) we also have $\varphi(x_1) = \cdots = \varphi(x_n) = 0$ (because $x_1, ..., x_n \in \mathscr{X}^*_{alt}$), and we only need to apply the definition of freeness. Since (5.8) gives us that $\varphi(w) = \varphi(w')$, it follows that $\varphi(w) = 0$, and this concludes the proof of the Claim (5.7).

Let us finally look at the matrices A, S defined in (5.4). In order to verify their freeness, it suffices to check that $\varphi_2(W) = 0$ for every word:

$$W = (A^{k_1} - \varphi_2(A^{k_1})\, I_2)(S^{l_1} - \varphi_2(S^{l_1})\, I_2) \cdots \\ \times (A^{k_n} - \varphi_2(A^{k_n})\, I_2)(S^{l_n} - \varphi_2(S^{l_n})\, I_2), \tag{5.11}$$

with $n, k_1, l_1, ..., k_n, l_n \geqslant 1$. A straightforward calculation shows that both diagonal entries of W in (5.11) are words of the type considered in (5.5)–(5.6). Hence the diagonal entries of W are in $\mathrm{Ker}(\varphi)$, by the Claim (5.7)—and consequently $W \in \mathrm{Ker}(\varphi_2)$, as desired. Q.E.D.

5.2. Proposition (the case $d = 1$ of Theorem 1.4). *Let ν be a probability measure with compact support on $[0, \infty)$, and let μ be the symmetric square root of ν (defined as in 3.1). Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, and let $a \in \mathscr{A}$ be such that a^*a has distribution ν. Then*:

$$\chi^*(a, a^*) \leqslant 2\chi^*(\mu). \tag{5.12}$$

Moreover, (5.12) *holds with equality if a is R-diagonal.*

Proof. We may assume without loss of generality that there exists a circular element $c \in \mathscr{A}$, of variance 1, such that $\{c, c^*\}$ is free from $\{a, a^*\}$. Then, by (5.1):

$$\chi^*(a, a^*) = \frac{1}{2} \int_0^\infty \left(\frac{2}{1+t} - \Phi^*(a + \sqrt{t}\, c, (a + \sqrt{t}\, c)^*) \right) dt + \log(2\pi e). \tag{5.13}$$

Consider on the other hand the space $(M_2(\mathscr{A}), \varphi_2)$ of 2×2-matrices over $(\mathscr{A}, \varphi)$, and the selfadjoint matrices $A, S \in M_2(\mathscr{A})$ defined exactly as in Equation (5.4) of Proposition 5.1. Then A has distribution μ (by Remark 3.5), and is free from S (by Proposition 5.1). From the form of S it is immediate that

$$\varphi_2(S^{2n}) = \varphi((c^*c)^n),\ \varphi_2(S^{2n+1}) = 0,\ \forall n \geqslant 0.$$

It is known that c^*c has the same distribution as the square of a semicircular element (see [9], Section 5.1); this implies that S is semicircular of variance 1.

Now, since S is a normalized semicircular free from A, we can write:

$$\chi^*(\mu) = \chi^*(A) = \frac{1}{2}\int_0^\infty \left(\frac{1}{1+t} - \Phi^*(A+\sqrt{t}\,S)\right) dt + \frac{1}{2}\log(2\pi e). \qquad (5.14)$$

But for every $t \geqslant 0$:

$$A+\sqrt{t}\,S = \begin{pmatrix} 0 & a+\sqrt{t}\,c \\ (a+\sqrt{t}\,c)^* & 0 \end{pmatrix},$$

hence (again by Remark 3.5) the distribution of $A+\sqrt{t}\,S$ is the symmetric square root of the distribution of $(a+\sqrt{t}\,c)^*\,(a+\sqrt{t}\,c)$. When applied to this situation, the Theorem 1.1 gives us that:

$$\Phi^*(a+\sqrt{t}\,c, (a+\sqrt{t}\,c)^*) \geqslant 2\Phi^*(A+\sqrt{t}\,S), \quad \forall t \geqslant 0. \qquad (5.15)$$

The inequality (5.12) is obtained by replacing (5.15) in (5.13), and by comparing the result with (5.14).

If a is R-diagonal, then so is $a+\sqrt{t}\,c$ for every $t \geqslant 0$. Indeed, $\sqrt{t}\,c$ is also R-diagonal, and the sum of two free R-diagonal elements is still R-diagonal (this follows for instance right away from the characterization of R-diagonality in terms of the R-transform—see [1]). But then the Theorem 1.1 implies that (5.15) holds with equality for every $t \geqslant 0$; and consequently, when we replace (5.15) in (5.13) and compare with (5.14), we obtain that (5.12) holds with equality too. Q.E.D.

We now move to the proof of Theorem 1.5. We will use a known freeness result, stated as follows.

5.3. Proposition. *Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace, let $\mathscr{B} \subseteq \mathscr{A}$ be a unital W^*-subalgebra, and let d be a positive integer.*

1. *Let $\{c_{ij}\}_{1 \leqslant i, j \leqslant d}$ be a family of elements of $\mathscr{A}$ such that every c_{ij} is circular of variance 1, and such that $\{c_{11}, c_{11}^*\}, \{c_{12}, c_{12}^*\}, ..., \{c_{dd}, c_{dd}^*\}, \mathscr{B}$ are free. Then the matrix $C = (c_{ij})_{i,j=1}^d$ is a circular element of variance d in $M_d(\mathscr{A})$, and $\{C, C^*\}$ is free from $M_d(\mathscr{B})$.*

2. *Let $\{s_{ij}\}_{1 \leqslant i, j \leqslant d}$ be a family of elements of $\mathscr{A}$ such that: $s_{ij}^* = s_{ji}$, for every $1 \leqslant i, j \leqslant d$; s_{ii} is semicircular of variance 1 for every $1 \leqslant i \leqslant d$; s_{ij} is circular of variance 1 for every $1 \leqslant i < j \leqslant d$; and $\{s_{11}\}, ..., \{s_{dd}\}, \{s_{12}, s_{12}^*\}, ..., \{s_{d-1,d} s_{d-1,d}^*\}, \mathscr{B}$ are free. Then the selfadjoint matrix $S = (s_{ij})_{i,j=1}^d$ is a semicircular element of variance d in $M_d(\mathscr{A})$, and is free from $M_d(\mathscr{B})$.*

For the fact that C of 5.3.1 is circular and that S of 5.3.2 is semicircular, see [9], Section 5.1; for the additional assertions concerning the freeness from $M_d(\mathscr{B})$, see [2]. For the sake of completeness, we indicate a way of proving Proposition 5.3 which, quite amusingly, comes out directly from the considerations of the preceding section. We have:

5.4. LEMMA. *Let $(\mathscr{A}, \varphi)$ be a W^*-probability space, with φ faithful trace. Let $\{s_i\}_{1 \leqslant i \leqslant k}$ be a family of selfadjoint elements of $\mathscr{A}$, and let $\mathscr{B} \subseteq \mathscr{A}$ be a unital W^*-subalgebra. Assume that $\{s_i\}_{1 \leqslant i \leqslant k}$ is its own conjugate with respect to $\mathscr{B}$. Then every $s_i (1 \leqslant i \leqslant k)$ is semicircular of variance* 1, *and $\{s_1\}, ..., \{s_k\}, \mathscr{B}$ are free.*

Proof of Lemma 5.4. This is an immediate consequence of the free Cramer-Rao inequality, as stated in [8], Proposition 6.9. The line of the argument goes as follows. By enlarging $(\mathscr{A}, \varphi)$ if necessary, we can assume that there also exists in $\mathscr{A}$ a family $\{s'_i\}_{1 \leqslant i \leqslant k}$ of semicircular elements of variance 1, such that $\{s'_1\}, ..., \{s'_k\}, \mathscr{B}$ are free. Then $\{s'_i\}_{1 \leqslant i \leqslant k}$ is its own conjugate with respect to $\mathscr{B}$, by Propositions 3.8 and 3.6 of [8]. The hypothesis that $\{s_i\}_{1 \leqslant i \leqslant k}$ is its own conjugate with respect to $\mathscr{B}$ amounts to the fact that

$$\begin{aligned}
&\varphi(s_i b_0 s_{i_1} b_1 \cdots s_{i_n} b_n) \\
&\quad = \sum_{m=1}^{n} \delta_{i, i_m} \varphi(b_0 s_{i_1} \cdots s_{i_{m-1}} b_{m-1}) \cdot \varphi(b_m s_{i_{m+1}} \cdots s_{i_n} b_n) \\
&\quad = \sum_{m=1}^{n} \delta_{i, i_m} \varphi(s_{i_1} b_1 \cdots s_{i_{m-1}} (b_{m-1} b_0)) \cdot \varphi(s_{i_{m+1}} b_{m+1} \cdots s_{i_n} (b_n b_m)),
\end{aligned} \tag{5.16}$$

for every $n \geqslant 0$, $b_0, ..., b_n \in \mathscr{B}$, $1 \leqslant i, i_1, ..., i_n \leqslant k$. Since $\{s'_i\}_{1 \leqslant i \leqslant k}$ also has the property of being its own conjugate with respect to $\mathscr{B}$, (5.16) remains true when we replace s_i by s'_i and s_{i_1} by $s'_{i_1}, ..., s_{i_n}$ by s'_{i_n}. But then an induction argument immediately gives that:

$$\varphi(s_i b_0 s_{i_1} b_1 \cdots s_{i_n} b_n) = \varphi(s'_i b_0 s'_{i_1} b_1 \cdots s'_{i_n} b_n), \tag{5.17}$$

for every $n \geqslant 0$, $b_0, ..., b_n \in \mathscr{B}$, $1 \leqslant i, i_1, ..., i_n \leqslant k$. Finally, from (5.17) and the fact that $s'_1, ..., s'_k$ are normalized semicirculars, with $\{s'_1\}, ..., \{s'_k\}, \mathscr{B}$ free, it follows that $s_1, ..., s_k$ also have these properties. Q.E.D.

Proof of Proposition 5.3. The proofs of 1 and 2 are similar; we will show 1, and leave 2 as an exercise to the reader.

By working with the real and imaginary parts of the elements c_{ij}, and by using Propositions 3.8 and 3.6 of [8], one obtains that the conjugate of

$\{c_{ij}, c_{ij}^*\}_{1 \leqslant i, j \leqslant d}$ with respect to $\mathscr{B}$ is $\{\xi_{ij}, \xi_{ij}^*\}_{1 \leqslant i, j \leqslant d}$, with $\xi_{ij} := c_{ij}^*$, $1 \leqslant i, j \leqslant d$. The Proposition 4.1 from the preceding section (or rather its proof) applies to this situation, and gives that $\{X, X^*\}$ is the conjugate of $\{C, C^*\}$ with respect to $M_d(\mathscr{B})$, where:

$$X := \frac{1}{d} (\xi_{ji})_{i, j=1}^d = \frac{1}{d} C^*. \tag{5.18}$$

From (5.18) it is immediate that if we set $S_1 = (C + C^*)/\sqrt{2d}$, $S_2 = (C - C^*)/i\sqrt{2d}$, then $\{S_1, S_2\}$ is its own conjugate with respect to $M_d(\mathscr{B})$. But then we can use the Lemma 5.4 to infer that S_1, S_2 are semicirculars of variance 1 in $M_d(\mathscr{A})$, such that $\{S_1\}, \{S_2\}, M_d(\mathscr{B})$ are free. This in turn implies that $C = \sqrt{d/2}(S_1 + iS_2)$ is circular of variance d, and free from $M_d(\mathscr{B})$.

Q.E.D.

Proof of Theorem 1.5. The proofs of 1.5.1 and 1.5.2 are similar; in order to offer the reader a variation, we will this time show 2, and leave 1 as an exercise.

The selfadjoint family $\{b_{ij}\}_{1 \leqslant i, j \leqslant d}$ appearing on the left-hand side of (1.10) is to be looked at as $\{b_{ij}, b_{ij}^*\}_{1 \leqslant i < j \leqslant d} \cup \{b_{ii}\}_{1 \leqslant i \leqslant d}$; thus (5.1) applies with $m = d(d-1)/2$, $n = d$, and yields the formula:

$$\chi^*(\{b_{ij}\}_{1 \leqslant i, j \leqslant d}) = \frac{1}{2} \int_0^\infty \left(\frac{d^2}{1+t} - \Phi^*(\{b_{ij} + \sqrt{t}\, s_{ij}\}_{1 \leqslant i, j \leqslant d}) \right) dt + \frac{d^2}{2} \log(2\pi e), \tag{5.19}$$

where the family $\{s_{ij}\}_{1 \leqslant i, j \leqslant d}$ of elements of $\mathscr{A}$ has the following properties: $s_{ij}^* = s_{ji}$, for every $1 \leqslant i, j \leqslant d$; s_{ii} is semicircular of variance 1 for every $1 \leqslant i \leqslant d$; s_{ij} is circular of variance 1 for every $1 \leqslant i < j \leqslant d$; and the sets $\{s_{11}\}, ..., \{s_{dd}\}, \{s_{12}, s_{12}^*\}, ..., \{s_{d-1,d} S_{d-1,d}^*\}, \{b_{ij} \mid 1 \leqslant i, j \leqslant d\}$ are free.

If we denote $S := (s_{ij})_{i,j=1}^d \in M_d(\mathscr{A})$, then $d^{-1/2} S$ is semicircular of variance 1, free from B in $(M_d(\mathscr{A}), \varphi_d)$ (by Proposition 5.3.2, where the choice of W^*-subalgebra $\mathscr{B} \subseteq \mathscr{A}$ is made to be $\mathscr{B} := W^*(\{I\} \cup \{b_{ij}\}_{1 \leqslant i, j \leqslant d})$). We can therefore use $d^{-1/2} S$ in the calculation of the free entropy $\chi^*(B)$; it is in fact more convenient to write the formula for $\chi^*(d^{-1/2} B)$:

$$\chi^*\left(\frac{1}{\sqrt{d}} B\right) = \frac{1}{2} \int_0^\infty \left(\frac{1}{1+t} - \Phi^*\left(\frac{1}{\sqrt{d}} B + \sqrt{\frac{t}{d}}\, S \right) \right) dt + \frac{1}{2} \log(2\pi e). \tag{5.20}$$

The scaling formulas for Φ^* and χ^* are

$$\Phi^*(\lambda x) = \lambda^{-2} \Phi^*(x),\ \chi^*(\lambda x) = \chi^*(x) + \log(\lambda)$$

(for $\lambda > 0$ and x a selfadjoint random variable—see [8], Sections 6.2(b) and 7.8). Thus (5.20) can also be written in the form:

$$\chi^*(B) - \frac{\log d}{2} = \frac{1}{2} \int_0^\infty \left(\frac{1}{1+t} - d\Phi^*(B + \sqrt{t}\, S) \right) dt + \frac{1}{2} \log(2\pi e). \tag{5.21}$$

Now, for every $t \geqslant 0$, the Theorem 1.2.2 gives us:

$$\Phi^*(\{b_{ij} + \sqrt{t} s_{ij}\}_{1 \leqslant i, j \leqslant d}) \geqslant d^3 \Phi^*(B + \sqrt{t} S). \tag{5.22}$$

If we replace (5.22) into (5.19), and compare the result with (5.21), then (1.10) of Theorem 1.5 is obtained.

If B is free from the algebra of scalar matrices $M_d(\mathbf{C}I) \subseteq M_d(\mathscr{A})$, then the same is true for $B + \sqrt{t}\, S$, for every $t \geqslant 0$; this is because (as implied by Proposition 5.3.2) S is free from $M_d(W^*(\{b_{ij}\}_{1 \leqslant i, j \leqslant d}))$, which in turn implies that $\{B, S\}$ is free from $M_d(\mathbf{C}I)$. But in this situation, the Theorem 1.2.2 implies that (5.22) holds with equality, for every $t \geqslant 0$; and the same argument used in the preceding paragraph shows now that (1.10) holds with equality, too. Q.E.D.

Proof of Theorem 1.4. This follows from Proposition 5.2 and Theorem 1.5.1, exactly as in the same way as Theorem 1.3 was obtained from 1.1 and 1.2.1 at the end of Section 4. Q.E.D.

Note added in proof. After the submission of this article we received D. Voiculescu's preprint, "The Analogues of Entropy and of Fisher's Information Measure in Free Probability Theory. VI. Liberation and Mutual Free Information," Berkeley, CA, July 1998. There, in Proposition 5.18(b), the following is shown: If $\Phi^*(X_1, ..., X_m : B) = \Phi^*(X_1, ..., X_m) < \infty$ then $\{X_1, ..., X_m\}$ and B are free. This implies a positive solution to our question on the uniqueness of the minimizing distribution in Theorem 1.2 (see the last paragraph in Section 4). By using ideas similar to those in the preprint of Voiculescu, one can also show that the minimum in Theorem 1.1 (if it is finite) can only be reached by an R-diagonal element. This yields that the minimizing distributions in Theorems 1.1–1.3 and the maximizing distributions in Theorems 1.4 and 1.5 are uniquely determined.

ACKNOWLEDGMENTS

Part of the research reported in this paper was done during a Research in Pairs program (supported by Volkswagen Stiftung) of the Mathematisches Forschungsinstitut Oberwolfach, Germany. We acknowledge the excellent work conditions and very stimulating atmosphere provided by this program.

REFERENCES

1. A. Nica and R. Speicher, R-diagonal pairs—a common approach to Haar unitaries and circular elements, *Fields Inst. Commun.* **12** (1997), 149–188.

2. D. Shlyakhtenko, Some applications of freeness with amalgamation, *J. Reine Angew. Math.* **500** (1998), 191–212.
3. D. Shlyakhtenko, Free entropy with respect to a completely positive map, Preprint May 1998.
4. D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, I., *Commun. Math. Phys.* **155** (1993), 71–92.
5. D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, II., *Invent. Math.* **118** (1994), 411–440.
6. D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, III: The absence of Cartan subalgebras, *Geom. Funct. Anal.* **6** (1996), 172–199.
7. D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, IV: Maximum entropy and freeness, *Fields Inst. Commun.* **12** (1997), 293–302.
8. D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, V: Non-commutative Hilbert transforms, *Invent. Math.* **132** (1998), 189–227.
9. D. V. Voiculescu, K. J. Dykema, and A. Nica, "Free Random Variables," CRM Monograph Series, Vol. 1, American Mathematical Society, Providence, RI, 1992.