

# Matrix variances with projections

Dénes Petz<sup>1</sup> and Géza Tóth<sup>2</sup>

Alfréd Rényi Institute of Mathematics, H-1364 Budapest, POB 127, Hungary

Department of Theoretical Physics,  
The University of the Basque Country, P.O. Box 644, E-48080 Bilbao, Spain

## Abstract

The quantum variance of a self-adjoint operator depends on a density matrix whose particular example is a pure state (formulated by a projection). A general variance can be obtained from certain variances of pure states. This is very different from the probabilistic case.

**Keywords:** Density matrix, pure state, projection. variance.

**AMS 2010 Subject Classification:** Primary 81Q50; Secondary 40C05

By a density matrix  $D \in M_n(\mathbb{C})$  we mean  $D \geq 0$  and  $\text{Tr } D = 1$ . In quantum information theory the traditional variance is

$$\text{Var}_D(A) = \text{Tr } DA^2 - (\text{Tr } DA)^2 \quad (1)$$

when  $D$  is a density matrix and  $A \in M_n(\mathbb{C})$  is a self-adjoint operator [3, 4]. This is the straightforward analogy of the variance in probability theory [2], a standard notation is  $\langle A^2 \rangle - \langle A \rangle^2$  in both formalism. It is rather different from probability theory that this variance can be strictly positive even in the case when  $D$  has rank 1. If  $D$  has rank 1, then it is an orthogonal projection of rank 1 and it is also called as pure state.

It is easy to show that

$$\text{Var}_D(A + \lambda I) = \text{Var}_D(A) \quad \text{for } \lambda \in \mathbb{R}.$$

The variance functional  $D \mapsto \text{Var}_D(A)$  is concave:

$$\text{Var}_D(A) \geq \sum_i \lambda_i \text{Var}_{D_i}(A) \quad \text{if } D = \sum_i \lambda_i D_i.$$

(Here  $\lambda_i \geq 0$  and  $\sum_i \lambda_i = 1$ .)

---

<sup>1</sup>E-mail: petz@math.bme.hu

<sup>2</sup>E-mail: toth@alumni.nd.edu

If we change the basis of the  $n$ -dimensional space such that  $D = \text{Diag}(p_1, p_2, \dots, p_n)$ , then we have

$$\text{Var}_D(A) = \sum_{i,j} \frac{p_i + p_j}{2} |A_{ij}|^2 - \left( \sum_{i=1}^n p_i A_{ii} \right)^2. \quad (2)$$

In the projection example  $P = \text{Diag}(1, 0, \dots, 0)$ , formula (2) gives

$$\text{Var}_P(A) = \sum_{i \neq 1} |A_{1i}|^2$$

and this can be strictly positive.

**Theorem.** *Let  $D$  be a density matrix. Take all the decompositions such that*

$$D = \sum_i q_i Q_i, \quad (3)$$

where  $Q_i$  are pure states and  $(q_i)$  is a probability distribution. Then

$$\text{Var}_D(A) = \sup \left( \sum_i q_i (\text{Tr } Q_i A^2 - (\text{Tr } Q_i A)^2) \right), \quad (4)$$

where the supremum is over all decompositions (3).

The proof will be an application of matrix theory. In the first lemma contains a trivial computation on block matrices.

**Lemma 1** *Assume that*

$$D = \begin{bmatrix} D^\wedge & 0 \\ 0 & 0 \end{bmatrix}, \quad D_i = \begin{bmatrix} D_i^\wedge & 0 \\ 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} A^\wedge & B \\ B^* & C \end{bmatrix}$$

and

$$D = \sum_i \lambda_i D_i, \quad D^\wedge = \sum_i \lambda_i D_i^\wedge.$$

Then

$$\begin{aligned} & (\text{Tr } D^\wedge (A^\wedge)^2 - (\text{Tr } D^\wedge A^\wedge)^2) - \sum_i \lambda_i (\text{Tr } D_i^\wedge (A^\wedge)^2 - (\text{Tr } D_i^\wedge A^\wedge)^2) \\ &= (\text{Tr } D A^2 - (\text{Tr } D A)^2) - \sum_i \lambda_i (\text{Tr } D_i A^2 - (\text{Tr } D_i A)^2). \end{aligned}$$

This lemma shows that if  $D \in M_n(\mathbb{C})$  has a rank  $k < n$ , then the computation of a variance  $\text{Var}_D(A)$  can be reduced to  $k \times k$  matrices. The equality in (4) is rather obvious for a rank 2 density matrix and due to the previous lemma the computation will be with  $2 \times 2$  matrices.

**Lemma 2** For a rank 2 matrix  $D$  the equality holds in (4).

*Proof:* Due to Lemma 1 we can make the computation with  $2 \times 2$  matrices. We can assume that

$$D = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}, \quad A = \begin{bmatrix} a_1 & b \\ \bar{b} & a_2 \end{bmatrix}.$$

Then

$$\text{Tr } DA^2 = p(a_1^2 + |b|^2) + (1-p)(a_2^2 + |b|^2).$$

We can assume that

$$\text{Tr } DA = pa_1 + (1-p)a_2 = 0.$$

Let

$$Q_1 = \begin{bmatrix} p & ce^{-i\varphi} \\ ce^{i\varphi} & 1-p \end{bmatrix},$$

where  $c = \sqrt{p(1-p)}$ . This is a projection and

$$\text{Tr } Q_1 A = a_1 p + a_2(1-p) + bce^{-i\varphi} + \bar{b}ce^{i\varphi} = 2c \text{Re } be^{-i\varphi}.$$

We choose  $\varphi$  such that  $\text{Re } be^{-i\varphi} = 0$ . Then  $\text{Tr } Q_1 A = 0$  and

$$\text{Tr } Q_1 A^2 = p(a_1^2 + |b|^2) + (1-p)(a_2^2 + |b|^2) = \text{Tr } DA^2.$$

Let

$$Q_2 = \begin{bmatrix} p & -ce^{-i\varphi} \\ -ce^{i\varphi} & 1-p \end{bmatrix}.$$

Then

$$D = \frac{1}{2}Q_1 + \frac{1}{2}Q_2$$

and we have

$$\frac{1}{2}(\text{Tr } Q_1 A^2 + \text{Tr } Q_2 A^2) = p(a_1^2 + |b|^2) + (1-p)(a_2^2 + |b|^2) = \text{Tr } DA^2.$$

Therefore we have an equality. □

We denote by  $r(D)$  the rank of an operator  $D$ . The idea of the proof is to reduce the rank and the block-diagonal formalism will be used [1].

**Lemma 3** Let  $D$  be a density matrix and  $A = A^*$  be an observable. Assume the block-matrix forms

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 \\ A_2^* & A_3 \end{bmatrix}.$$

and  $r(D_1), r(D_2) > 1$ . We construct

$$D' := \begin{bmatrix} D_1 & X^* \\ X & D_2 \end{bmatrix}$$

such that

$$\text{Tr } DA = \text{Tr } D'A, \quad D' \geq 0, \quad r(D') < r(D).$$

*Proof:* The  $\text{Tr } DA = \text{Tr } D'A$  condition is equivalent with  $\text{Tr } XA_2 + \text{Tr } X^*A_2^* = 0$  and this holds if and only if  $\text{Re } \text{Tr } XA_2 = 0$ .

We can have unitaries  $U$  and  $W$  such that  $UD_1U^*$  and  $WD_2W^*$  are diagonal:

$$UD_1U^* = \text{Diag}(0, \dots, 0, a_1, \dots, a_k), \quad WD_2W^* = \text{Diag}(b_1, \dots, b_l, 0, \dots, 0)$$

where  $a_i, b_j > 0$ . Then  $D$  has the same rank as the matrix

$$\begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} D \begin{bmatrix} U^* & 0 \\ 0 & W^* \end{bmatrix} = \begin{bmatrix} UD_1U^* & 0 \\ 0 & WD_2W^* \end{bmatrix},$$

the rank is  $k + l$ . A possible modification of this matrix is

$$\begin{aligned} Y &: = \begin{bmatrix} \text{Diag}(0, \dots, 0, a_1, \dots, a_{k-1}) & 0 & 0 & 0 \\ 0 & a_k & \sqrt{a_k b_1} & 0 \\ 0 & \sqrt{a_k b_1} & b_1 & 0 \\ 0 & 0 & 0 & \text{Diag}(b_2, \dots, b_l, 0, \dots, 0) \end{bmatrix} \\ &= \begin{bmatrix} UD_1U^* & M \\ M & WD_2W^* \end{bmatrix} \end{aligned}$$

and  $r(Y) = k + l - 1$ . So  $Y$  has a smaller rank than  $D$ . Next we take

$$\begin{bmatrix} U^* & 0 \\ 0 & W^* \end{bmatrix} Y \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} = \begin{bmatrix} D_1 & U^* M W \\ W^* M U & D_2 \end{bmatrix}$$

which has the same rank as  $Y$ . If  $X_1 := W^* M U$  is multiplied with  $e^{i\alpha}$  ( $\alpha > 0$ ), then the positivity condition and the rank remain. On the other hand, we can choose  $\alpha > 0$  such that  $\text{Re } \text{Tr } e^{i\alpha} X_1 A_2 = 0$ . Then  $X := e^{i\alpha} X_1$  is the matrix we wanted,  $\square$

**Lemma 4** *Let  $D$  be a density matrix of rank  $m > 0$  and  $A = A^*$  be an observable. We claim the existence of a decomposition*

$$D = pD_- + (1 - p)D_+, \quad (5)$$

such that  $r(D_-) < m$ ,  $r(D_+) < m$ , and

$$\text{Tr } AD_+ = \text{Tr } AD_- = \text{Tr } DA. \quad (6)$$

*Proof:* By unitary transformation we can get to the formalism of the previous lemma:

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 \\ A_2^* & A_3 \end{bmatrix}.$$

We choose

$$D_+ = D' = \begin{bmatrix} D_1 & X^* \\ X & D_2 \end{bmatrix}, \quad D_- = \begin{bmatrix} D_1 & -X^* \\ -X & D_2 \end{bmatrix}.$$

Then

$$D = \frac{1}{2}D_- + \frac{1}{2}D_+$$

and the requirements  $\text{Tr } AD_+ = \text{Tr } AD_- = \text{Tr } DA$  also hold.  $\square$

*Proof of the Theorem:* For rank-2 states, it is true because of Lemma 2. Any state with a rank larger than 2 can be decomposed into the mixture of lower rank states, according to Lemma 4, that have the same expectation value for  $A$ , as the original state has. The lower rank states can then be decomposed into the mixture of states with an even lower rank, until we reach rank-2 states. Thus, any state  $D$  can be decomposed into the mixture of

$$D = \sum p_k Q_k \tag{7}$$

such that  $\text{Tr } AQ_k = \text{Tr } AD$ . Hence the statement of the theorem follows.  $\square$

The above theorem is relevant to the paper [5] and to the book [4].

## References

- [1] R. Bhatia, *Matrix Analysis*, Springer, 1997.
- [2] W. Feller, *An introduction to probability theory and its applications I*, Oxford, England: Wiley. (1950).
- [3] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, North-Holland, Amsterdam, 1982.
- [4] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer-Verlag, Heidelberg, 2008.
- [5] G. Tóth and D. Petz, Optimal generalized variance and quantum Fisher information, to be published, <http://fr.arxiv.org/abs/1109.2831>.