

# Jensen's operator inequality for functions of several variables

Huzihiro Araki\* and Frank Hansen†

September 14, 1998

## Abstract

The operator convex functions of several variables are characterized in terms of a non-commutative generalization of Jensen's inequality, extending previous results of F. Hansen and G.K. Pedersen for functions of one variable and of F. Hansen for functions of two variables.

## 1 Introduction and the main result

Let  $I_1, \dots, I_k$  be intervals in  $\mathbf{R}$  and let  $f : I_1 \times \dots \times I_k \rightarrow \mathbf{R}$  be a real function defined on the product of the intervals. Let  $x = (x_1, \dots, x_k)$  be a  $k$ -tuple of bounded selfadjoint operators on Hilbert spaces  $H_1, \dots, H_k$  such that the spectrum of  $x_i$  is contained in  $I_i$  for  $i = 1, \dots, k$ . We say that such a  $k$ -tuple is in the domain of  $f$ . If

$$x_i = \int_{I_i} \lambda_i E_i(d\lambda_i)$$

is the spectral resolution of  $x_i$  for  $i = 1, \dots, k$ , then we define

$$f(x) = f(x_1, \dots, x_k) = \int_{I_1 \times \dots \times I_k} f(\lambda_1, \dots, \lambda_k) E_1(d\lambda_1) \otimes \dots \otimes E_k(d\lambda_k)$$

as a bounded selfadjoint operator on  $H_1 \otimes \dots \otimes H_k$ . The above function  $f : I_1 \times \dots \times I_k \rightarrow \mathbf{R}$  is said to be operator convex, if the operator inequality

$$(1.1) \quad f(\lambda x_1 + (1 - \lambda)y_1, \dots, \lambda x_k + (1 - \lambda)y_k) \leq \lambda f(x_1, \dots, x_k) + (1 - \lambda)f(y_1, \dots, y_k)$$

holds for any  $\lambda \in [0, 1]$ , any Hilbert spaces  $H_1, \dots, H_k$  and any selfadjoint operators  $x_1, \dots, x_k$  and  $y_1, \dots, y_k$  on  $H_1, \dots, H_k$  contained in the domain of  $f$ . The definition is meaningful since

---

\*Department of Mathematics, Faculty of Science and Technology, The Science University of Tokyo, Noda, Chiba-ken 278-8510, Japan.

†Institute of Economics, University of Copenhagen, Studiestraede 6, 1455 Copenhagen K, Denmark.

also the spectrum of  $\lambda x_i + (1 - \lambda)y_i$  is contained in the interval  $I_i$  for each  $i = 1, \dots, k$ . If the inequality (1.1) holds just for Hilbert spaces of finite dimensions  $(n_1, \dots, n_k)$ , then we say that  $f$  is matrix convex of order  $(n_1, \dots, n_k)$  or just order  $n$  if  $n_i = n$  for  $i = 1, \dots, k$ .

**Definition 1.1** An  $l$ -tuple of bounded linear operators  $a = (a_1, \dots, a_l)$  on a Hilbert space  $H$  is called a contraction row, if

$$(1.2) \quad a_1 a_1^* + \dots + a_l a_l^* \leq 1.$$

It is called a unitary row, if there exists a unitary operator  $U$  on the direct sum of  $l$  copies of  $H$  such that  $(a_1, \dots, a_l)$  is the first row in the  $l \times l$  block matrix representation of  $U$ .

If  $a = (a_1, \dots, a_l)$  is a contraction row, it follows that the norm  $\|a_i\| \leq 1$  for  $i = 1, \dots, l$ . A necessary but in general not sufficient condition for an  $l$ -tuple  $(a_1, \dots, a_l)$  to be a unitary row is that

$$(1.3) \quad a_1 a_1^* + \dots + a_l a_l^* = 1.$$

A unitary row is in particular a contraction row. If one of the operators in a contraction row is normal, then (1.3) is a sufficient condition for the row to be a unitary row, cf. Remark 2.3 below. This result was obtained in [5] for  $l = 2$ . In particular, any  $l$ -tuple of orthogonal projections  $(p_1, \dots, p_l)$  with sum  $p_1 + \dots + p_l = 1$  is a unitary row. This also follows explicitly, because the operator represented by the block matrix

$$(1.4) \quad U = \left( p_{j-i \pmod{l}+1} \right)_{i,j=1}^l = \begin{pmatrix} p_1 & p_2 & \cdots & p_l \\ p_l & p_1 & \cdots & p_{l-1} \\ \vdots & \vdots & & \vdots \\ p_2 & p_3 & \cdots & p_1 \end{pmatrix}$$

is unitary on  $H \otimes \mathbf{C}^l$ . Any extension of a unitary row of length  $l$  with zero operators to an  $m$ -tuple  $(a_1, \dots, a_l, 0, \dots, 0)$  is a unitary row of length  $m$ , since

$$U' = \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix}_{m \times m}$$

is unitary on  $H \otimes \mathbf{C}^m$ , if  $U$  is unitary on  $H \otimes \mathbf{C}^l$ . It is a non-trivial problem to determine whether a general contraction row satisfying (1.3) is a unitary row.

We are going to present Jensen's operator inequality for functions of several variables. It generalizes previous results for functions of one variable and for functions of two variables, cf. [7] and [5].

**Theorem 1.2** Let  $f$  be a continuous real valued function defined on  $I_1 \times \dots \times I_k$  where  $I_i = [0, \alpha_i[$  with  $\alpha_i \leq \infty$  for  $i = 1, \dots, k$ . The following statements are equivalent (the statement (iii)<sub>l</sub> is for each natural number  $l \geq 2$ ):

- (i)  $f$  is operator convex and  $f(r_1, \dots, r_k) \leq 0$  if  $r_i = 0$  for some  $i = 1, \dots, k$ .
- (ii) For each natural number  $l \geq 2$  and for each  $j = 0, 1, \dots, l - 1$  the operator inequality

$$(1.5) \quad \begin{aligned} & \text{diag}\left(f(a_{s_1 1}^* x_1 a_{s_1 1}, \dots, a_{s_k k}^* x_k a_{s_k k})\right)_{|s|=j \pmod{l}} \\ & \leq \left((a_{t_1 1}^* \otimes \dots \otimes a_{t_k k}^*)f(x_1, \dots, x_k)(a_{s_1 1} \otimes \dots \otimes a_{s_k k})\right)_{|t|=|s|=j \pmod{l}} \end{aligned}$$

is valid for all unitary rows  $a_i = (a_{1i}, \dots, a_{li})$  of length  $l$  acting on a Hilbert space  $H_i$  for  $i = 1, \dots, k$  and all  $k$ -tuples  $(x_1, \dots, x_k)$  of selfadjoint operators in the domain of  $f$  acting on  $H_1, \dots, H_k$ . The indices  $s, t$  are multi-indices of the form  $s = (s_1, \dots, s_k)$ , where  $s_i = 1, \dots, l$  for  $i = 1, \dots, k$  with weight  $|s| = s_1 + \dots + s_k$ .

- (iii)<sub>l</sub> For each  $j = 0, 1, \dots, l - 1$  the operator inequality

$$(1.6) \quad \begin{aligned} & \text{diag}\left(f(p_{s_1 1} x_1 p_{s_1 1}, \dots, p_{s_k k} x_k p_{s_k k})\right)_{|s|=j \pmod{l}} \\ & \leq \left((p_{t_1 1} \otimes \dots \otimes p_{t_k k})f(x_1, \dots, x_k)(p_{s_1 1} \otimes \dots \otimes p_{s_k k})\right)_{|t|=|s|=j \pmod{l}} \end{aligned}$$

is valid for all partitions of unity  $p_{1i} + \dots + p_{li} = 1$  on  $H_i$  by orthogonal projections for each  $i = 1, \dots, k$  and all  $k$ -tuples  $(x_1, \dots, x_k)$  of selfadjoint operators in the domain of  $f$  acting on  $H_1, \dots, H_k$ . The index notation  $s$  and  $t$  for rows and columns of the block matrices are the same as in the statement (ii) above.

**Remark 1.3** The statements in (ii) and (iii)<sub>l</sub> extended to  $l = 1$  are trivially true with equality for any  $f$ . To get (i) from (iii)<sub>l</sub> for any fixed  $l \geq 2$ , it is enough to assume the inequality in (iii)<sub>l</sub> for  $j = 0$  and identical partitions of the unity acting on the same Hilbert space  $H$  for  $i = 1, \dots, k$ .

**Corollary 1.4** Let  $f$  be a continuous real valued function defined on  $I_1 \times \dots \times I_k$  where  $I_i = [0, \alpha_i[$  with  $\alpha_i \leq \infty$  for  $i = 1, \dots, k$ . If  $f$  is operator convex and  $f(r_1, \dots, r_k) \leq 0$  if  $r_i = 0$  for some  $i = 1, \dots, k$ , then the operator inequality

$$(1.7) \quad \begin{aligned} & \text{diag}\left(f(a_{s_1 1}^* x_1 a_{s_1 1}, \dots, a_{s_k k}^* x_k a_{s_k k})\right)_{|s|=j \pmod{m}} \\ & \leq \left((a_{t_1 1}^* \otimes \dots \otimes a_{t_k k}^*)f(x_1, \dots, x_k)(a_{s_1 1} \otimes \dots \otimes a_{s_k k})\right)_{|t|=|s|=j \pmod{m}} \end{aligned}$$

is valid for all natural numbers  $l \geq 1$ ,  $m > l$ ,  $j = 0, 1, \dots, m - 1$ , all contraction rows  $a_i = (a_{1i}, \dots, a_{li})$  of length  $l$  acting on a Hilbert space  $H_i$  for  $i = 1, \dots, k$  and all  $k$ -tuples  $(x_1, \dots, x_k)$  of selfadjoint operators in the domain of  $f$  acting on  $H_1, \dots, H_k$ .

**Proof:** We apply lemma 2.1 to construct unitary rows of length  $m$  from the given contraction rows  $a_i = (a_{1i}, \dots, a_{li})$  and consider Jensen's operator inequality for  $k$  and  $m$ . Take  $j = 0, 1, \dots, m-1$  and consider the submatrix corresponding to indices  $s, t$  taking values only in  $\{1, \dots, l\}^k$  and having weights  $j \pmod{m}$ . **QED**

**Remark 1.5** *If we set  $k = 1, l = 1$  and  $m = 2$  in Corollary 1.4, then we get the inequality*

$$f(a^*xa) \leq a^*f(x)a$$

*valid for all contractions  $a$  and selfadjoint  $x$  in the domain of  $f$ . This is Jensen's operator inequality for functions of one variable [7]. If we set  $k = 2, l = 1$ , and  $m = 2$  in the corollary and choose projections  $p$  and  $q$  as contraction rows of length 1, then we get Aujla's inequality [2]*

$$f(pxp, qqy) \leq (p \otimes q)f(x, y)(p \otimes q)$$

*which characterizes separately operator convex functions of two variables.*

**Remark 1.6** *If we put  $l = k$  and  $j = 0$  in (ii) of Theorem 1.2 and take the submatrix corresponding to coinciding indices  $s_1 = \dots = s_k = i$  for  $i = 1, \dots, k$ , then we obtain the inequality*

$$(1.8) \quad \text{diag}\left(f(a_{i1}^*x_1a_{i1}, \dots, a_{ik}^*x_k a_{ik})\right)_{i=1}^k \leq \left(\left(a_{i1}^* \otimes \dots \otimes a_{ik}^*\right)f(x_1, \dots, x_k)(a_{j1} \otimes \dots \otimes a_{jk})\right)_{i,j=1}^k$$

*for all unitary rows  $a_i = (a_{1i}, \dots, a_{ki})$  acting on a Hilbert space  $H_i$  for  $i = 1, \dots, k$  and all  $k$ -tuples  $(x_1, \dots, x_k)$  of selfadjoint operators in the domain of  $f$  acting on  $H_1, \dots, H_k$ .*

**Remark 1.7** *If we let all the operators in (1.8) act on the same Hilbert space and only consider one common unitary row, which with a change of notation we denote  $(a_1, \dots, a_k)$ , then we obtain the inequality*

$$(1.9) \quad \text{diag}\left(f(a_i^*x_1a_i, \dots, a_i^*x_k a_i)\right)_{i=1}^k \leq \left(\left(a_i^* \otimes \dots \otimes a_i^*\right)f(x_1, \dots, x_k)(a_j \otimes \dots \otimes a_j)\right)_{i,j=1}^k.$$

*This is for  $k = 2$  the generalization of Jensen's operator inequality to functions of two variables obtained in [5], except that in the reference it was obtained only for unitary rows of a certain type.*

It appears that the inequality (1.9) is too weak to imply operator convexity of  $f$  except in the cases  $k = 1, 2$  which were already settled in the literature [7, 5].

## 2 Preliminaries

**Lemma 2.1** *A contraction row  $(a_1, \dots, a_l)$  of length  $l$  on a Hilbert space  $H$  can for any  $m > l$  be extended to a unitary row of length  $m$ . More specifically, the row  $(a_1, \dots, a_l, a)$ , where*

$$a = \left( 1 - \sum_{i=1}^l a_i a_i^* \right)^{1/2},$$

*as well as  $(a_1, \dots, a_l, a, 0, \dots, 0)$  for any number of zeros are unitary rows.*

**Proof:** The block matrix with operators on  $H$  as its elements

$$A = \begin{pmatrix} a_1 & \cdots & a_l \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{l \times l}$$

is by the assumption a contraction on  $H \otimes \mathbf{C}^l$ . Indeed, the norm  $\|AA^*\| = \|\sum a_i a_i^*\| \leq 1$ . The  $2l \times 2l$  block matrix

$$U = \begin{pmatrix} A & (1 - AA^*)^{1/2} \\ -(1 - A^*A)^{1/2} & A^* \end{pmatrix}$$

is unitary [3, Solution 177]. When explicitly written out, this block matrix splits as a direct sum of two unitary maps (a block matrix of size  $l + 1$  and the identity matrix of size  $l - 1$ ), one of which has the first row  $(a_1, \dots, a_l, a)$ . **QED**

In fact, the unitary matrix corresponding to a unitary row of length  $l + 1$  appearing in the proof of Lemma 2.1 can be computed explicitly as

$$\begin{pmatrix} a_1 & \cdots & a_l & a \\ & & a_1^* & \\ & -(1 - A^*A)^{1/2} & \vdots & \\ & & a_l^* & \end{pmatrix} = \begin{pmatrix} a_1 & \cdots & a_l & a \\ v_{11} & \cdots & v_{1l} & a_1^* \\ \vdots & & \vdots & \vdots \\ v_{l1} & \cdots & v_{ll} & a_l^* \end{pmatrix},$$

where  $v_{ij} = a_i^*(1 + a)^{-1}a_j - \delta_{ij}$  for  $i, j = 1, \dots, l$ .

**Corollary 2.2** *Let  $(a_1, \dots, a_l)$  be a contraction row and suppose that  $\dim \ker a_i = \dim \ker a_i^*$  for at least one  $i = 1, \dots, l$ . Then the row is a unitary row, if and only if (1.3) holds.*

**Proof:** The necessity of (1.3) is immediate as mentioned earlier. To prove the sufficiency of (1.3) we may without loss of generality assume that  $\dim \ker a_l = \dim \ker a_l^*$ . Let  $a_l = u|a_l|$  be the polar decomposition of  $a_l$ . Then  $u$  is a partial isometry with  $u^*u$  and  $uu^*$  being the self-adjoint projections on  $(\ker a_l)^\perp$  and  $(\ker a_l^*)^\perp$ . By assumption, there exists a partial isometry

$v$  with  $v^*v$  and  $vv^*$  being the self-adjoint projections on  $\ker a_l$  and  $\ker a_l^*$ . Therefore  $w = u + v$  is unitary, and since the support of  $|a_l|$  is  $(\ker a_l)^\perp$  we obtain  $a_l = w|a_l|$ . Set  $b_j = w^*a_j$  for  $j = 1, \dots, k-1$ . Then by (1.3)

$$1 - \sum_{j=1}^{l-1} b_j b_j^* = w^* \left( 1 - \sum_{j=1}^{l-1} a_j a_j^* \right) w = w^* a_l a_l^* w = |a_l|^2.$$

Hence  $(b_1, \dots, b_{l-1}, |a_l|)$  is a unitary row by Lemma 2.1, thus being the first row of a unitary block matrix  $V$ . Finally,  $(a_1, \dots, a_{l-1}, a_l) = (wb_1, \dots, wb_{l-1}, w|a_l|)$  is the first row of the unitary block matrix  $(w \otimes E_l)V$ , where  $E_l$  denotes the identity matrix of order  $l$ . **QED**

**Remark 2.3** *If one of the operators in a contraction row is normal and (1.3) is satisfied, then the contraction row is a unitary row. This follows from Corollary 2.2, because the kernel of a normal operator coincides with the kernel of its adjoint.*

The root of unity  $\beta = e^{2\pi i/l}$  is a simple root of the polynomial  $X^l - 1$ . In the following we make repeated use of the identities  $\beta^l = 1$ ,  $\bar{\beta} = \beta^{-1}$ , and  $1 + \beta^j + \beta^{2j} + \dots + \beta^{(l-1)j} = 0$  for  $j = 1, \dots, l-1$ . We consider the operator  $R_i$  on  $H_i \otimes \mathbf{C}^l$  given by the diagonal block matrix representation

$$R_i = \text{diag}(\beta^p)_{p=1, \dots, l}.$$

For  $i = 1, \dots, k$  and any unitary row  $a_i = (a_{1i}, \dots, a_{li})$  acting on a Hilbert space  $H_i$  with associated unitary  $U_i$  on  $H_i \otimes \mathbf{C}^l$  we set

$$U_{mi} = U_i R_i^m = U_i \cdot \text{diag}(\beta^{pm})_{p=1, \dots, l}$$

which defines a unitary operator on  $H_i \otimes \mathbf{C}^l$ . Relative to these unitary operators we define the average  $\Delta_{U_i}(X)$  of an operator  $X_i$  acting on  $H_i \otimes \mathbf{C}^l$  by setting

$$\Delta_{U_i}(X_i) = \frac{1}{l} \sum_{m=1}^l U_{mi}^* X_i U_{mi}.$$

**Lemma 2.4** *For any bounded operator  $x_i$  on  $H_i$ , consider the block matrix*

$$D_l(x_i) = \begin{pmatrix} x_i & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{l \times l}$$

*as an operator on  $H_i \otimes \mathbf{C}^l$ . Then for  $i = 1, \dots, k$  we have*

$$\Delta_{U_i}(D_l(x_i)) = \text{diag} \left( a_{pi}^* x_i a_{pi} \right)_{p=1}^l$$

*for any unitary row  $a_i = (a_{1i}, \dots, a_{li})$  on  $H_i$  with associated unitary  $U_i$ .*

**Proof:** The entry at the  $p$ -th row and  $q$ -th column of  $Y R_i^m$  is just  $\beta^{mq} y_{pq}$  for any operator  $Y$  acting on  $H_i \otimes \mathbb{C}^l$  with block matrix representation  $Y = (y_{pq})_{p,q=1,\dots,l}$ . Likewise, the entry

$$[(R_i^m)^* Y R_i^m]_{pq} = \beta^{(q-p)m} y_{pq}$$

and consequently

$$\frac{1}{l} \sum_{m=1}^l (R_i^m)^* Y R_i^m$$

is a diagonal block matrix composed of the diagonal of  $Y$ . Since the diagonal of  $U_i^* D_l(x_i) U_i$  is given by  $(a_{1i}^* x_i a_{1i}, \dots, a_{li}^* x_i a_{li})$ , the statement follows. **QED**

### 3 Proof of Jensen's operator inequality

**Proof of Theorem 1.2:**

(i)  $\Rightarrow$  (ii) : We consider unitary rows  $a_i = (a_{1i}, \dots, a_{li})$  for  $i = 1, \dots, k$  with associated unitaries  $U_1, \dots, U_k$ . The  $l^k \times l^k$  block matrix

$$\text{diag} \left( f(a_{s_1 1}^* x_1 a_{s_1 1}, \dots, a_{s_k k}^* x_k a_{s_k k}) \right)_{s_1, \dots, s_k=1}^l$$

is naturally identified with

$$f \left( \left( \begin{array}{ccc} a_{11}^* x_1 a_{11} & & \\ & \ddots & \\ & & a_{l1}^* x_1 a_{l1} \end{array} \right), \dots, \left( \begin{array}{ccc} a_{1k}^* x_k a_{1k} & & \\ & \ddots & \\ & & a_{lk}^* x_k a_{lk} \end{array} \right) \right)$$

under the block matrix representation of tensor products. Applying lemma 2.4 and the convexity of  $f$ , we thus obtain

$$\begin{aligned} & \text{diag} \left( f(a_{s_1 1}^* x_1 a_{s_1 1}, \dots, a_{s_k k}^* x_k a_{s_k k}) \right)_{s_1, \dots, s_k=1}^l \\ &= f \left( \Delta_{U_1}(D_l(x_1)), \dots, \Delta_{U_k}(D_l(x_k)) \right) \\ &\leq \frac{1}{l} \sum_{m=1}^l f \left( U_{m1}^* D_l(x_1) U_{m1}, \dots, U_{mk}^* D_l(x_k) U_{mk} \right) \\ &= \frac{1}{l} \sum_{m=1}^l (U_{m1} \otimes \dots \otimes U_{mk})^* f \left( D_l(x_1), \dots, D_l(x_k) \right) (U_{m1} \otimes \dots \otimes U_{mk}) \\ &\leq \frac{1}{l} \sum_{m=1}^l (U_{m1} \otimes \dots \otimes U_{mk})^* D_l^k \left( f(x_1, \dots, x_k) \right) (U_{m1} \otimes \dots \otimes U_{mk}). \end{aligned}$$

The last inequality is a consequence of the assumption that  $f(r_1, \dots, r_k) \leq 0$  whenever  $r_i = 0$  for some  $i = 1, \dots, k$ . The matrix elements of each term of the last sum are calculated as

$$\begin{aligned} & \left[ (U_{m1} \otimes \cdots \otimes U_{mk})^* D_{lk} \left( f(x_1, \dots, x_k) \right) (U_{m1} \otimes \cdots \otimes U_{mk}) \right]_{ts} \\ &= \beta^{m(|s|-|t|)} \left[ (U_1 \otimes \cdots \otimes U_k)^* D_{lk} \left( f(x_1, \dots, x_k) \right) (U_1 \otimes \cdots \otimes U_k) \right]_{ts}. \end{aligned}$$

The average over  $m$  gives rise to

$$\frac{1}{l} \sum_{m=1}^l \beta^{m(|s|-|t|)} = \delta_{|s||t|}$$

and this automatically splits the above inequality into  $l$  parts. Take  $j = 0, 1, \dots, l-1$  and let  $Q_j$  denote the projection on the subspace corresponding to the indices  $s = (s_1, \dots, s_k)$  of a fixed weight  $|s| = s_1 + \cdots + s_k = j \pmod{l}$ . It follows that

$$\begin{aligned} & Q_j \operatorname{diag} \left( f(a_{s_{11}}^* x_1 a_{s_{11}}), \dots, f(a_{s_{kk}}^* x_k a_{s_{kk}}) \right)_{s_1, \dots, s_k=1}^l Q_j \\ & \leq Q_j (U_1 \otimes \cdots \otimes U_k)^* D_{lk} \left( f(x_1, \dots, x_k) \right) (U_1 \otimes \cdots \otimes U_k) Q_j. \end{aligned}$$

The left-hand side is a diagonal matrix with non-trivial diagonal entries for  $|s| = j \pmod{l}$  and  $D_{lk}(f(x_1, \dots, x_k))$  is a diagonal matrix which has its only non-trivial element at  $1(k) = (1, \dots, 1)$ . Because of

$$[(U_1 \otimes \cdots \otimes U_k) Q_j]_{1(k)_s} = a_{s_{11}} \otimes \cdots \otimes a_{s_{kk}},$$

we conclude that the above inequality, when restricted to the image subspace of  $Q_j$ , is Jensen's operator inequality (1.5).

(ii)  $\Rightarrow$  (iii)<sub>l</sub> : Obvious specialization.

(iii)<sub>l</sub>  $\Rightarrow$  (i) : We shall only use (iii)<sub>l</sub> for  $j = 0$ . The operator  $V$  defined by the block matrix

$$V = (v_{pq})_{p,q=1}^l = l^{-1/2} \begin{pmatrix} 1 & \beta & \beta^2 & \cdots & \beta^{l-1} \\ 1 & \beta^2 & \beta^4 & \cdots & \beta^{2(l-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \beta^{l-1} & \beta^{2(l-1)} & \cdots & \beta^{(l-1)(l-1)} \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix},$$

where  $v_{pq} = l^{-1/2} \beta^{p(q-1)}$ , is unitary. We introduce the following block matrices with operator elements:

$$X_i = \operatorname{diag}(x_{ji})_{j=1, \dots, l} \quad \text{and} \quad Z_i = V^* X_i V$$



for  $1 = 1, \dots, k$ . The diagonal elements of  $Z_i$  are given by

$$[Z_i]_{jj} = \sum_{s=1}^l \bar{v}_{sj} v_{sj} x_{si} = \frac{x_{1i} + \dots + x_{li}}{l}$$

and consequently we obtain

$$P_j Z_i P_j = \frac{x_{1i} + \dots + x_{li}}{l} P_j$$

for  $i = 1, \dots, k$  and  $j = 1, \dots, l$ , where  $P_i$  is the projection given by the block matrix  $P_i = \text{diag}(\delta_{ji})_{j=1, \dots, l}$  and  $\delta$  is Kronecker's symbol. The statement  $(iii)_l$  of Theorem 1.2 now entails the inequality

$$\begin{aligned} & \text{diag}\left(f(P_{s_1} Z_1 P_{s_1}, \dots, P_{s_k} Z_k P_{s_k})\right)_{|s|=0 \pmod{l}} \\ (3.1) \quad & \leq \left( (P_{t_1} \otimes \dots \otimes P_{t_k}) f(Z_1, \dots, Z_k) (P_{s_1} \otimes \dots \otimes P_{s_k}) \right)_{|t|=|s|=0} \\ & = \left( (P_{t_1} V^* \otimes \dots \otimes P_{t_k} V^*) f(X_1, \dots, X_k) (V P_{s_1} \otimes \dots \otimes V P_{s_k}) \right)_{|t|=|s|=0 \pmod{l}} \end{aligned}$$

which we shall examine in more detail. The block matrix  $V P_j$  has vanishing matrix elements except for the  $j$ -th column which is equal to the  $j$ -th column of  $V$ . We can therefore calculate the matrix elements

$$\begin{aligned} & [V P_{s_1} \otimes \dots \otimes V P_{s_k}]_{pq} \\ & = [V P_{s_1}]_{p_1 q_1} \dots [V P_{s_k}]_{p_k q_k} \\ & = \delta_{q_1 s_1} l^{-1/2} \beta^{p_1(s_1-1)} \dots \delta_{q_k s_k} l^{-1/2} \beta^{p_k(s_k-1)} \\ & = l^{-k/2} \delta_{qs} \beta^{p \cdot s - |p|} \end{aligned}$$

for  $p = (p_1, \dots, p_k)$ ,  $q = (q_1, \dots, q_k)$  and  $p_1, \dots, p_k, q_1, \dots, q_k = 1, \dots, l$ . We observe in particular that the only non zero column in  $V P_{s_1} \otimes \dots \otimes V P_{s_k}$  is the  $s$ -th column. The right hand side of (3.1) is a  $l^{k-1} \times l^{k-1}$  block matrix where each block is a  $l^k \times l^k$  block matrix of operators. Let us calculate the  $(p, q)$  entry of this  $l^k \times l^k$  block matrix.

$$\begin{aligned} & \left[ \left( P_{t_1} V^* \otimes \dots \otimes P_{t_k} V^* \right) f(X_1, \dots, X_k) \left( V P_{s_1} \otimes \dots \otimes V P_{s_k} \right) \right]_{pq} \\ & = \sum_{u_1, \dots, u_k=1}^l \left[ \left( P_{t_1} V^* \otimes \dots \otimes P_{t_k} V^* \right) f(X_1, \dots, X_k) \right]_{pu} l^{-k/2} \delta_{qs} \beta^{u \cdot s - |u|} \\ & = l^{-k} \sum_{u_1, \dots, u_k=1}^l \delta_{pt} \beta^{|u| - u \cdot t} f(x_{u_1 1}, \dots, x_{u_k k}) \delta_{qs} \beta^{u \cdot s - |u|} \\ & = l^{-k} \delta_{pt} \delta_{qs} \sum_{u_1, \dots, u_k=1}^l f(x_{u_1 1}, \dots, x_{u_k k}) \beta^{u \cdot (s-t)} \end{aligned}$$

where the calculation was facilitated because  $f(X_1, \dots, X_k)$  is a diagonal matrix with  $u$ -th diagonal entry  $f(x_{u_1 1}, \dots, x_{u_k k})$ . We discard all rows and columns with indices different from  $t$  and  $s$  in each block in (3.1). Since this operation preserves the inequality we obtain

$$\begin{aligned} & f\left(\frac{x_{11} + \dots + x_{l1}}{l}, \dots, \frac{x_{1k} + \dots + x_{lk}}{l}\right) E_{l^{k-1}} \\ & \leq l^{-k} \sum_{u_1, \dots, u_k=1}^l f(x_{u_1 1}, \dots, x_{u_k k}) \left(\beta^{u \cdot (s-t)}\right)_{|t|=|s|=0 \pmod{l}} \end{aligned}$$

where  $E_{l^{k-1}}$  denotes the identity matrix of order  $l^{k-1}$ . If we set

$$\Pi_u = l^{-(k-1)} \left(\beta^{u \cdot (s-t)}\right)_{|t|=|s|=0 \pmod{l}}$$

then the inequality can be written as

$$f\left(\frac{x_{11} + \dots + x_{l1}}{l}, \dots, \frac{x_{1k} + \dots + x_{lk}}{l}\right) E_{l^{k-1}} \leq \frac{1}{l} \sum_{u_1, \dots, u_k=1}^l f(x_{u_1 1}, \dots, x_{u_k k}) \Pi_u.$$

It is an easy calculation to show that the matrices  $\Pi_u$  are self-adjoint projections. Because of

$$\sum_{u_1, \dots, u_k=1}^l \beta^{u \cdot (s-t)} = l^k \delta_{ts},$$

it follows that

$$\sum_{u_1, \dots, u_k=1}^l \Pi_u = l E_{l^{k-1}}.$$

Due to  $|t| = |s|$  we note that  $\Pi_u = \Pi_{u+j(k) \pmod{l}}$  for  $j = 0, 1, \dots, l-1$  where  $j(k) = (j, \dots, j)$  and  $u+j(k) \pmod{l}$  are different indices for different  $j$ . We introduce an equivalence relation  $u \sim v$  by setting  $u \sim v$ , if  $u - v = j(k) \pmod{l}$  for any natural number  $j$  and denote the equivalence class of  $u$  by  $\langle u \rangle$ . Each equivalence class contains exactly  $l$  distinct members of the form  $u + j(k)$  for  $j = 0, 1, \dots, l-1$ . We denote the common projection corresponding to each equivalence class  $\langle u \rangle$  by  $\Pi_{\langle u \rangle}$  and obtain

$$\sum_{\langle u \rangle} \Pi_{\langle u \rangle} = E_{l^{k-1}}.$$

Hence the projections  $\Pi_{\langle u \rangle}$  are mutually orthogonal, and setting  $u = 1(k)$  we obtain in particular

$$\begin{aligned} & f\left(\frac{x_{11} + \dots + x_{l1}}{l}, \dots, \frac{x_{1k} + \dots + x_{lk}}{l}\right) \Pi_{1(k)} \\ & \leq \frac{1}{l} \left( f(x_{11}, \dots, x_{1k}) + \dots + f(x_{l1}, \dots, x_{lk}) \right) \Pi_{1(k)}. \end{aligned}$$

Therefore  $f$  is operator convex. The rest of statement (i) follows by examining diagonal elements in (3.1), which at  $t$  for  $|t| = 0 \pmod{l}$  gives

$$\begin{aligned} & f(P_{t_1} Z_1 P_{t_1}, \dots, P_{t_k} Z_k P_{t_k}) \\ & \leq (P_{t_1} \otimes \dots \otimes P_{t_k}) f(Z_1, \dots, Z_k) (P_{t_1} \otimes \dots \otimes P_{t_k}). \end{aligned}$$

If we look at the  $s$ -th diagonal entry of this inequality for  $s \neq t$ , then

$$f\left(\frac{x_{11} + \dots + x_{l1}}{l} \delta_{t_1 s_1}, \dots, \frac{x_{1k} + \dots + x_{lk}}{l} \delta_{t_k s_k}\right) \leq 0$$

and we conclude that  $f(r_1, \dots, r_k) \leq 0$  if  $r_i = 0$  for at least one  $i = 1, \dots, k$ . **QED**

## References

- [1] T. Ando. Concavity of certain maps of positive definite matrices and applications to Hadamard products. *Linear Algebra Appl.*, 26:203–241, 1979.
- [2] J.S. Aujla. Matrix convexity of functions of two variables. *Linear Algebra and Its Applications*, 194:149–160, 1993.
- [3] P.R. Halmos. *A Hilbert space problem book*. American Book, 1967.
- [4] F. Hansen. An operator inequality. *Math. Ann.*, 246:249–250, 1980.
- [5] F. Hansen. Jensen’s operator inequality for functions of two variables. *Proc. Amer. Math. Soc.*, 125:2093–2102, 1997.
- [6] F. Hansen. Operator convex functions of several variables. *Publ. RIMS*, 33:443–463, 1997.
- [7] F. Hansen and G.K. Pedersen. Jensen’s inequality for operators and Löwner’s theorem. *Math. Ann.*, 258:229–241, 1982.
- [8] A. Korányi. On some classes of analytic functions of several variables. *Trans Amer. Math. Soc.*, 101:520–554, 1961.
- [9] G.K. Pedersen. Extreme  $n$ -tuples of elements in  $c^*$ -algebras. *Bull. London Math. Soc.*, 19:264–270, 1987.