

Chapter 8

Generalized Semimodularity: Order Statistics



Iosif Pinelis

Abstract A notion of generalized n -semimodularity is introduced, which extends that of (sub/super)modularity in four ways at once. The main result of this paper, stating that every generalized $(n: 2)$ -semimodular function on the n th Cartesian power of a distributive lattice is generalized n -semimodular, may be considered a multi/infinite-dimensional analogue of the well-known Muirhead lemma in the theory of Schur majorization. This result is also similar to a discretized version of the well-known theorem due to Lorentz, which latter was given only for additive-type functions. Illustrations of our main result are presented for counts of combinations of faces of a polytope; one-sided potentials; multiadditive forms, including multilinear ones—in particular, permanents of rectangular matrices and elementary symmetric functions; and association inequalities for order statistics. Based on an extension of the FKG inequality due to Rinott & Saks and Aharoni & Keich, applications to correlation inequalities for order statistics are given as well.

Keywords Semimodularity · Submodularity · Supermodularity · FKG-type inequalities · Association inequalities · Correlation inequalities

2010 Mathematics Subject Classification Primary 06D99, 26D15, 26D20, 60E15; Secondary 05A20, 05B35, 06A07, 60C05, 62H05, 62H10, 82D99, 90C27

8.1 Summary and Discussion

As pointed out e.g. in [3, 4], the notion of submodularity has become useful in various areas: combinatorial optimization, with many applications in operations research; machine learning; computer vision; electrical networks; signal processing;

I. Pinelis (✉)

Department of Mathematical Sciences, Michigan Technological University, Houghton, MI, USA

e-mail: ipinelis@mtu.edu

several areas of theoretical computer science, such as matroid theory; economics. One may also note the use of this notion in potential theory [6], as a capacity is a submodular function.

Let L be any distributive lattice; for definitions and facts pertaining to lattices, see e.g. [10].

A function $\lambda: L \rightarrow \mathbb{R}$ is called submodular if

$$\lambda(f) + \lambda(g) \geq \lambda(f \vee g) + \lambda(f \wedge g) \tag{8.1}$$

for all f and g in L . A function λ is called supermodular if the function $-\lambda$ is submodular, and λ is called modular if it is both submodular and supermodular. See e.g. [4, 9, 18, 19, 24, 25]. Let us say that a function μ is log-submodular if $\ln \mu$ is submodular. The log-submodularity condition and the corresponding log-supermodularity condition were referred to in Karlin and Rinott [13, 14] as the multivariate total positivity of order 2 (MTP_2) and the multivariate reverse rule of order 2 (MRR_2), respectively. As noted by Choquet [6, §14.3], a nondecreasing function λ is alternating of order 2 iff it satisfies inequality (8.1), that is, λ is submodular; it was also shown in [6] that the classical Newtonian capacity is such a function.

The log-supermodularity condition is the condition under which the famous Fortuin–Kasteleyn–Ginibre (FKG) correlation inequality [8] holds. Therefore, using inequality (8.17) together with the FKG inequality and its generalizations, we will be able to obtain the corresponding applications, in Corollaries 8.2.11 and 8.2.12.

More generally, let \mathcal{R} be any set, endowed with a transitive relation \bowtie , so that for any a, b, c in \mathcal{R} one has the implication $a \bowtie b \ \& \ b \bowtie c \implies a \bowtie c$. For any natural n , let us say that a function $\Lambda: L^n \rightarrow \mathcal{R}$ is *generalized n -semimodular* if

$$\Lambda(f_1, \dots, f_n) \bowtie \Lambda(f_{n:1}, \dots, f_{n:n})$$

for all $f = (f_1, \dots, f_n) \in L^n$, where $f_{n:1}, \dots, f_{n:n}$ are the “order statistics” for f defined by the formula

$$f_{n:j} = \bigwedge \left\{ \bigvee_{i \in J} f_i : J \in \binom{[n]}{j} \right\} \tag{8.2}$$

for $j \in [n] := \overline{1, n}$, with $\binom{[n]}{j}$ denoting the set of all subsets J of the set $[n]$ such that the cardinality of J is j . Here and in the sequel we use the notation $\overline{\alpha, \beta} := \{j \in \mathbb{Z} : \alpha \leq j \leq \beta\}$. In particular, $f_{n:1} = f_1 \wedge \dots \wedge f_n$ and $f_{n:n} = f_1 \vee \dots \vee f_n$.

For any function $\lambda: L \rightarrow \mathbb{R}$, let the function $\Lambda_\lambda: L^2 \rightarrow \mathbb{R}$ be given by the formula $\Lambda_\lambda(f, g) := \lambda(f) + \lambda(g)$ for f and g in L . Then, obviously, λ is submodular or supermodular or modular if and only if Λ_λ is generalized 2-semimodular with the relation “ \bowtie ” being “ \geq ” or “ \leq ” or “ $=$ ”, respectively.

Thus, the notion of generalized n -semimodularity extends that of (sub/super)modularity in four ways at once: (1) the function Λ may be a function of any natural

number n of arguments, whereas λ is a function of only one argument; (2) in contrast with a general form of dependence of $\Lambda(f_1, \dots, f_n)$ on f_1, \dots, f_n , the function Λ_λ of two arguments is of the special form, linear in $\lambda(f)$ and $\lambda(g)$; (3) whereas the values of λ are real numbers, those of Λ may be in any set \mathcal{R} ; and (iv) we now have an arbitrary transitive relation \bowtie over \mathcal{R} instead of one of the three particular relations “ \geq ” or “ \leq ” or “ $=$ ” over \mathbb{R} .

For any $k \in [n]$, let us say that a function $\Lambda: L^n \rightarrow \mathbb{R}$ is *generalized $(n:k)$ -semimodular* if for each $j \in \overline{0, n-k}$ and each $(n-k)$ -tuple $(f_i: i \in [n] \setminus j+1, j+k) \in L^{n-k}$ the function $L^k \ni (f_{j+1}, \dots, f_{j+k}) \mapsto \Lambda(f_1, \dots, f_n)$ is generalized k -semimodular. In particular, Λ is generalized $(n:n)$ -semimodular if and only if it is generalized n -semimodular.

Whenever the relation “ \bowtie ” is denoted as “ \geq ” or “ \leq ” or “ $=$ ”, let us replace “semi” in the above definitions by “sub”, “super”, and “”, respectively. For instance, “generalized n -modular” will stand for “generalized n -semimodular” with the relation “ \bowtie ” being “ $=$ ”.

The main result of this note is

Theorem 8.1.1 *Again, let L be any distributive lattice. If a function $\Lambda: L^n \rightarrow \mathcal{R}$ is generalized $(n:2)$ -semimodular, then it is generalized n -semimodular.*

The necessary proofs will be given in Sect. 8.3.

As will be seen from the proof of Theorem 8.1.1, the condition that the function Λ be generalized $(n:2)$ -semimodular can be relaxed to the following: for each $j \in \overline{1, n-1}$ and each $f = (f_1, \dots, f_n) \in L^n$ such that $f_1 \leq \dots \leq f_j$, one has $L(f_1, \dots, f_n) \bowtie L(f_1, \dots, f_{j-1}, f_j \wedge f_{j+1}, f_j \vee f_{j+1}, f_{j+2}, \dots, f_n)$.

Remark 8.1.2 Theorem 8.1.1 will not hold in general if the lattice L is not assumed to be distributive. For instance, let L be defined by the set $[5] = \{1, 2, 3, 4, 5\}$ with the partial order being the subset of the natural order \leq on the set $[5]$ with elements 2, 3, 4 now considered non-comparable with one another, so that the resulting order relation is the set $\{(f, f): f \in [5]\} \cup \{(1, 2), (1, 3), (1, 4), (2, 5), (3, 5), (4, 5), (1, 5)\}$; then, in particular, $2 \wedge 3 = 1$ and $2 \vee 3 = 5$. This lattice is one of the simplest examples of non-distributive lattices. It is isomorphic to the diamond lattice \mathbf{M}_3 —see e.g. [10, p. 110]. Let $n = 3$, $\mathcal{R} = \mathbb{R}$, and define the function $\Lambda: L^3 \rightarrow \mathbb{R}$ by the formula $\Lambda(f_1, f_2, f_3) := 12f_1f_2 + 3f_2f_3 + 5f_1f_3$ for all $f = (f_1, f_2, f_3) \in L^3$. Then one can verify directly—by a straightforward but tedious calculation consisting in checking $2 \times 5^3 = 250$ inequalities, two inequalities for each $f = (f_1, f_2, f_3) \in [5]^3$ —that this function Λ is generalized $(3:2)$ -submodular. However, Λ is not generalized 3-submodular, because for $f = (2, 3, 4)$ one has $(f_{3:1}, f_{3:2}, f_{3:3}) = (1, 5, 5)$ and $\Lambda(f_1, f_2, f_3) = \Lambda(2, 3, 4) = 148 \not\geq 160 = \Lambda(1, 5, 5) = \Lambda(f_{3:1}, f_{3:2}, f_{3:3})$. \square

Remark 8.1.3 A well-known fact, which will be crucial in the proof of Theorem 8.1.1, is the representation theorem due to Birkhoff and Stone stating that any distributive lattice L is isomorphic to a lattice of subsets of (and hence to a lattice of nonnegative real-valued functions on) a certain set S , depending on L (see e.g. [10, Theorem 119]). For such a lattice of functions, the “order statistics” $f_{n:1}, \dots, f_{n:n}$

are uniquely determined by the condition that

$$f_{n:1}(s) \leq \dots \leq f_{n:n}(s) \quad \text{and} \quad \{\{f_{n:1}(s), \dots, f_{n:n}(s)\}\} = \{\{f_1(s), \dots, f_n(s)\}\} \tag{8.3}$$

for each $s \in S$, where the double braces are used to denote multisets, with appropriate multiplicities. To quickly see why this is true, one may reason as follows: Let us now use condition (8.3) to *define* $f_{n:1}, \dots, f_{n:n}$. Note that the value of the right-hand side (rhs) of (8.2) at any point $s \in S$ is invariant with respect to all permutations of the values $f_1(s), \dots, f_n(s)$. So, the value of the rhs of (8.2) at s will not change if one replaces there f_1, \dots, f_n by $f_{n:1}, \dots, f_{n:n}$, and this value will equal $f_{n:j}(s)$. Thus, the definition of $f_{n:1}, \dots, f_{n:n}$ by means of formula (8.3) is equivalent to the one given by (8.2), if the lattice L is already a lattice of real-valued functions on S . Moreover, it is clear now that, if the lattice L is distributive, then definition (8.2) can be rewritten in the dual form, as

$$f_{n:j} = \bigvee \left\{ \bigwedge_{i \in J} f_i : J \in \binom{[n]}{n+1-j} \right\} \tag{8.4}$$

for all $j \in [n]$.

On the other hand, it can be seen that, if L is not distributive, then this duality can be lost and each of the definitions (8.2) and (8.4) of $f_{n:j}$ can be rather unnaturally skewed up or down. For instance, in the counterexample given in Remark 8.1.2, for $f = (2, 3, 4)$ we had $(f_{3:1}, f_{3:2}, f_{3:3}) = (1, 5, 5)$ according to definition (8.2), but we would have $(f_{3:1}, f_{3:2}, f_{3:3}) = (1, 1, 5)$ according to (8.4).

However, one may note that the right-hand side of (8.4) is always \leq than that of (8.2); this follows because for any $J \in \binom{[n]}{n+1-j}$ and any $K \in \binom{[n]}{j}$ there is some $k \in J \cap K$, and then $\bigwedge_{i \in J} f_i \leq f_k \leq \bigvee_{i \in K} f_i$. □

In view of the lattice representation theorem cited in Remark 8.1.3, Theorem 8.1.1 may be considered a multi/infinite-dimensional analogue of the well-known Muirhead lemma in the theory of Schur majorization (cf. e.g. [17, Lemma 2.B.1, p. 32]), which may be stated as follows: for vectors x and y in \mathbb{R}^n such that $x \prec y$ (that is, x is majorized by y), there exist finitely many vectors x_0, \dots, x_m in \mathbb{R}^n such that $x = x_0 \prec \dots \prec x_m = y$ and for each $j \in \overline{0, m-1}$ the vectors x_j and x_{j+1} differ only in two coordinates. However, no direct multi-dimensional extension of the Muirhead lemma seems to exist, even in two dimensions (see e.g. [20, p. 11]).

For functions that are “infinite-dimensional” counterparts of the “ m -dimensional” function $\Lambda : L^m \rightarrow \mathbb{R}$ given by the formula of the additive form

$$\Lambda(g_1, \dots, g_m) = \sum_{j=1}^m \lambda_j(g_j), \tag{8.5}$$

Lorentz [16] obtained a result similar to Theorem 8.1.1; for readers' convenience, let us reproduce it here: For each $j \in [n]$, let f_j^* denote the equimeasurable decreasing rearrangement [11] of a function $f_j: (0, 1) \rightarrow \mathbb{R}$. Let a real-valued expression $\Phi(x, u_1, \dots, u_n)$ be continuous in $(x, u_1, \dots, u_n) \in (0, 1) \times [0, \infty) \times \dots \times [0, \infty)$. Then the inequality

$$\int_0^1 \Phi(x, f_1(x), \dots, f_n(x)) dx \leq \int_0^1 \Phi(x, f_1^*(x), \dots, f_n^*(x)) dx \tag{8.6}$$

holds for all bounded positive measurable functions f_1, \dots, f_n from $(0, 1)$ to \mathbb{R} if and only if the following two conditions hold:

$$\Phi(u_i + h, u_j + h) - \Phi(u_i + h, u_j) - \Phi(u_i, u_j + h) + \Phi(u_i, u_j) \geq 0 \tag{8.7}$$

and

$$\int_0^\delta [\Phi(x-t, u_i+h) - \Phi(x-t, u_i) - \Phi(x+t, u_i+h) + \Phi(x+t, u_i)] dt \geq 0 \tag{8.8}$$

for all $h > 0, x \in (0, 1), \delta \in (0, x \wedge (1 - x)), (u_1, \dots, u_n) \in [0, \infty)^n$, and i, j in $[n]$ such that $i < j$; here, in each of inequalities (8.7) and (8.8), the arguments of Φ that are the same for all the four instances of Φ are omitted, for brevity.

To establish the connection between Lorentz's result and our Theorem 8.1.1, suppose e.g. that each of the functions f_1, \dots, f_n in [16] is a step function, constant on each of the intervals $(\frac{j-1}{m}, \frac{j}{m}]$ for $j \in [m]$, and then let $g_j(s) := f_s(\frac{j}{m})$ for $j \in [m]$ and $s \in S := [n]$. In fact, in the proof in [16] the result is first established for such step functions f_1, \dots, f_n . It is also shown in [16] that, for such "infinite-dimensional" counterparts of the functions given by the "additive" formula (8.5), the sufficient condition is also necessary. In turn, as pointed out in [16], the result there generalizes an inequality in [23]. Another proof of a special case of the result in [16] was given in [5].

8.2 Illustrations and Applications

8.2.1 A General Construction of Generalized n -Submodular Functions from Submodular Ones

Recall here some basics of majorization theory [17]. For $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in \mathbb{R}^n , write $x < y$ if $x_1 + \dots + x_n = y_1 + \dots + y_n$ and $x_{n:1} + \dots + x_{n:k} \geq y_{n:1} + \dots + y_{n:k}$ for all $k \in [n]$. For any $D \subseteq \mathbb{R}^n$, a function $F: D \rightarrow \mathbb{R}$ is called Schur-concave if for any x and y in D such that $x < y$ one has $F(x) \geq F(y)$. If $D = I^n$ for some open interval $I \subseteq \mathbb{R}$ and the function

F is continuously differentiable then, by Schur’s theorem [17, Theorem A.4], F is Schur-concave iff $(\frac{\partial F}{\partial x_i} - \frac{\partial F}{\partial x_j})(x_i - x_j) \leq 0$ for all $x = (x_1, \dots, x_n) \in D$.

Proposition 8.2.1 *Suppose that a real-valued function λ defined on a distributive lattice L is submodular and nondecreasing, and a function $\mathbb{R}^n \ni x = (x_1, \dots, x_n) \rightarrow F(x_1, \dots, x_n)$ is nondecreasing in each of its n arguments and Schur-concave. Then the function $\Lambda = \Lambda_{\lambda, F}: L^n \rightarrow \mathbb{R}$ defined by the formula*

$$\Lambda(f_1, \dots, f_n) := \Lambda_{\lambda, F}(f_1, \dots, f_n) := F(\lambda(f_1), \dots, \lambda(f_n)) \tag{8.9}$$

for $(f_1, \dots, f_n) \in L^n$ is generalized $(n: 2)$ -submodular and hence generalized n -submodular.

A rather general construction of submodular functions on rings of sets is provided by [6, §23.2], which implies that \cup -homomorphisms preserve the property of being alternating of a given order, and the proposition at the end of [6, §23.1], which describes general \cup -homomorphisms as maps of the form

$$S \supseteq A \mapsto G(A) := \{t \in T : (s, t) \in G \text{ for some } s \in A\},$$

where S and T are sets and $G \subseteq S \times T$; in the case when G is (the graph of) a map, the above notation $G(A)$ is of course consistent with that for the image of a set A under the map G ; according to the definition in the beginning of [6, §23], a \cup -homomorphism is a map φ of set rings defined by the condition $\varphi(A \cup B) = \varphi(A) \cup \varphi(B)$ for all relevant sets A and B .

Therefore and because an additive function on a ring of sets is modular and hence submodular, we conclude that functions of the form

$$A \mapsto \mu(G(A)) \tag{8.10}$$

are submodular, where μ is a measure or, more generally, an additive function (say on a discrete set, to avoid matters of measurability). From this observation, one can immediately obtain any number of corollaries of Proposition 8.2.1 such as the following:

Corollary 8.2.2 *Let P be a polytope of dimension d . For each $\alpha \in \overline{0, d}$, let \mathcal{F}_α denote the set of all α -faces (that is, faces of dimension α) of P . For any distinct α, β, γ in $\overline{0, d}$, let $G = G_{\alpha, \beta, \gamma}$ be the set of all pairs $(f_\alpha, (f_\beta, f_\gamma)) \in \mathcal{F}_\alpha \times (\mathcal{F}_\beta \times \mathcal{F}_\gamma)$ such that $f_\alpha \cap f_\beta \neq \emptyset$, $f_\alpha \cap f_\gamma \neq \emptyset$, and $f_\beta \cap f_\gamma \neq \emptyset$. Let L be a lattice of subsets of \mathcal{F}_α . Let a function $\mathbb{R}^n \ni x = (x_1, \dots, x_n) \rightarrow F(x_1, \dots, x_n)$ be nondecreasing in each of its n arguments and Schur-concave. Then the function $\Lambda = \Lambda_{\alpha, \beta, \gamma}: L^n \rightarrow \mathbb{R}$ defined by the formula*

$$\Lambda(A_1, \dots, A_n) := F(\text{card } G(A_1), \dots, \text{card } G(A_n))$$

for $(A_1, \dots, A_n) \in L^n$ is generalized $(n: 2)$ -submodular and hence generalized n -submodular.

For readers' convenience, here is a direct verification of the fact that maps of the form (8.10) are submodular: noting that $G(A \cup B) = G(A) \cup G(B)$ and $G(A \cap B) \subseteq G(A) \cap G(B)$ and using the additivity of μ , we have

$$\mu(G(A \cup B)) + \mu(G(A \cap B)) \leq \mu(G(A) \cup G(B)) + \mu(G(A) \cap G(B)) = \mu(G(A)) + \mu(G(B))$$

for all relevant sets A and B .

8.2.2 Generalized One-Sided Potential

Let here L be the lattice of all measurable real-valued functions on a measure space (S, Σ, μ) , with the pointwise lattice operations \vee and \wedge . Consider the function $\Lambda : L^n \rightarrow \mathcal{R}$ given by the formula

$$\Lambda(f_1, \dots, f_n) := \Lambda_{\varphi, \psi}(f_1, \dots, f_n) := \sum_{j,k=1}^n \Psi(f_j - f_k) \tag{8.11}$$

for all $f = (f_1, \dots, f_n) \in L^n$, where

$$\Psi(g) := \psi \left(\int_S (\varphi \circ g) \, d\mu \right) \tag{8.12}$$

for all $g \in L$, $\varphi : \mathbb{R} \rightarrow [0, \infty]$ is a nondecreasing or nonincreasing function, and $\psi : [0, \infty] \rightarrow (-\infty, \infty]$ is a concave function. Thus, the function $\Lambda = \Lambda_{\varphi, \psi}$ may be referred to as a generalized one-sided potential, since the function φ is assumed to be monotonic.

Proposition 8.2.3 *The function $\Lambda = \Lambda_{\varphi, \psi}$ defined by formula (8.11) is generalized $(n : 2)$ -submodular and hence generalized n -submodular.*

8.2.3 Symmetric Sums of Nonnegative Multiadditive Functions

Let k be a natural number. Let L be a sublattice of the lattice \mathbb{R}^S of all real-valued functions on a set S . Let us say that the lattice L is *complementable* if $f \setminus g := f - f \wedge g \in L$ for any f and g in L , so that $f = f \wedge g + f \setminus g$. Assuming that L is complementable, let us say that a function $m : L \rightarrow \mathbb{R}$ is *additive* if

$$m(f) = m(f \wedge g) + m(f \setminus g)$$

for all f and g in L ; further, let us say that a function $m : L^k \rightarrow \mathbb{R}$ is *multiadditive* or, more specifically, *k-additive* if m is additive in each of its k arguments, that

is, if for each $j \in [k]$ and each $(k - 1)$ -tuple $(f_i : i \in [k] \setminus \{j\})$ the function $L \ni f_j \mapsto m(f_1, \dots, f_k)$ is additive.

To state the main result of this subsection, we shall need the following notation: for any set J , let Π_k^J denote the set of all k -permutations of J , that is, the set of all injective maps of the set $[k]$ to J .

Proposition 8.2.4 *Suppose that k and n are natural numbers such that $k \leq n$, L is a complementable sublattice of \mathbb{R}^S , and $m : L^k \rightarrow \mathbb{R}$ is a nonnegative multiadditive function. Then the function $\Lambda_m : L^n \rightarrow \mathbb{R}$ defined by the formula*

$$\Lambda_m(f_1, \dots, f_n) := \sum_{\pi \in \Pi_k^{[n]}} m(f_{\pi(1)}, \dots, f_{\pi(k)}) \tag{8.13}$$

for $(f_1, \dots, f_n) \in L^n$ is generalized $(n : 2)$ -submodular and hence generalized n -submodular.

Formula (8.13) can be rewritten in the following symmetrized form:

$$\Lambda_m(f_1, \dots, f_n) = k! \sum_{I \in \binom{[n]}{k}} \bar{m}(f_I), \tag{8.14}$$

where, for $I = \{i_1, \dots, i_k\}$ with $1 \leq i_1 < \dots < i_k \leq n$,

$$\bar{m}(f_I) := \bar{m}(f_{i_1}, \dots, f_{i_k}) := \frac{1}{k!} \sum_{\pi \in \Pi_k^I} m(f_{\pi(1)}, \dots, f_{\pi(k)}); \tag{8.15}$$

note that the so-defined function $\bar{m} : L^k \rightarrow \mathbb{R}$ is multiadditive and nonnegative, given that m is so. Also, \bar{m} is permutation-symmetric in the sense that $\bar{m}(f_{\pi(1)}, \dots, f_{\pi(k)}) = \bar{m}(f_1, \dots, f_k)$ for all $(f_1, \dots, f_k) \in L^k$ and all permutations $\pi \in \Pi_k^{[k]}$.

Example 8.2.5 If V is a vector sublattice of the lattice \mathbb{R}^S and L is the lattice of all nonnegative functions in V then, clearly, L is complementable and the restriction to L^k of any multilinear function from V^k to \mathbb{R} is multiadditive.

In particular, if μ is a measure on a σ -algebra Σ over S , V is a vector sublattice of $L^k(S, \Sigma, \mu)$, and L is the lattice of all nonnegative functions in V , then the function $m : L^k \rightarrow \mathbb{R}$ given by the formula

$$m(f_1, \dots, f_k) := \int_S f_1 \cdots f_k \, d\mu$$

for $(f_1, \dots, f_k) \in L^k$ is multiadditive.

So, by Proposition 8.2.4, the functions Λ_m corresponding to the functions m presented above in this example are generalized $(n : 2)$ -submodular and hence generalized n -submodular.

Let now $B = (b_{i,j})$ be a $d \times p$ matrix with $d \leq p$ and nonnegative entries $b_{i,j}$. The permanent of B is defined by the formula

$$\text{perm } B := \sum_{J \in \binom{[p]}{d}} \text{perm } B_{\cdot J},$$

where $B_{\cdot J}$ the square submatrix of B consisting of the columns of B with column indices in the set $J \in \binom{[p]}{d}$; and for a square $d \times d$ matrix $B = (b_{i,j})$,

$$\text{perm } B := \sum_{\pi \in \Pi_d^{[d]}} b_{1,\pi(1)} \cdots b_{d,\pi(d)}.$$

So, $\text{perm } B$ is a multilinear function of the d -tuple $(b_{1,\cdot}, \dots, b_{d,\cdot})$ of the rows of B . Also, if $d = p$, then $\text{perm } B$ is a multilinear function of the d -tuple $(b_{\cdot,1}, \dots, b_{\cdot,d})$ of the columns of B . If $d \geq p$, then $\text{perm } B$ may be defined by the requirement that the permanent be invariant with respect to transposition.

Thus, from Proposition 8.2.4 we immediately obtain

Corollary 8.2.6 *Assuming that the entries $b_{i,j}$ of the $d \times p$ matrix B are nonnegative, $\text{perm } B$ is a generalized d -submodular function of the d -tuple $(b_{1,\cdot}, \dots, b_{d,\cdot})$ of its rows and a generalized p -submodular function of the p -tuple $(b_{\cdot,1}, \dots, b_{\cdot,p})$ of its columns (with respect to the standard lattice structures on $\mathbb{R}^{1 \times p}$ and $\mathbb{R}^{d \times 1}$, respectively):*

$$\text{perm} \begin{pmatrix} b_{d:1,\cdot} \\ \vdots \\ b_{d:d,\cdot} \end{pmatrix} \leq \text{perm} \begin{pmatrix} b_{1,\cdot} \\ \vdots \\ b_{d,\cdot} \end{pmatrix} [= \text{perm } B],$$

$$\text{perm}(b_{\cdot,p:1}, \dots, b_{\cdot,p:p}) \leq \text{perm}(b_{\cdot,1}, \dots, b_{\cdot,p}) [= \text{perm } B].$$

Note that the condition $d \leq p$ is not needed or assumed in Corollary 8.2.6.

Yet another way in which multilinear and hence multiadditive functions may arise is via the elementary symmetric polynomials. Let n be any natural number, and let $k \in [n]$. The elementary symmetric polynomials are defined by the formula

$$e_k(x_1, \dots, x_n) := \sum_{J \in \binom{[n]}{k}} \prod_{j \in J} x_j.$$

In particular, $e_1(x_1, \dots, x_n) := \sum_{j \in [n]} x_j$ and $e_n(x_1, \dots, x_n) := \prod_{j \in [n]} x_j$.

Let $f = (f_1, \dots, f_n)$ be the vector of measurable functions f_1, \dots, f_n defined on a measure space (S, Σ, μ) with values in the interval $[0, \infty)$. Then it is not hard to see that the ‘‘order statistics’’ are nonnegative measurable functions as well. As usual, let $\mu(h) := \int_S h \, d\mu$.

If the measure μ is a probability measure, then the functions f_1, \dots, f_n are called random variables (r.v.'s) and, in this case, $f_{n:1}, \dots, f_{n:n}$ will indeed be what is commonly referred to as the order statistics based on the “random sample” $f = (f_1, \dots, f_n)$; cf. e.g. [7]. In contrast with settings common in statistics, in general we do not impose any conditions on the joint or individual distributions of the r.v.'s f_1, \dots, f_n —except that these r.v.'s be nonnegative.

Then we have the following.

Corollary 8.2.7

$$e_k(\mu(f_1), \dots, \mu(f_n)) \geq e_k(\mu(f_{n:1}), \dots, \mu(f_{n:n})). \quad (8.16)$$

In particular,

$$\mu(f_1) \cdots \mu(f_n) \geq \mu(f_{n:1}) \cdots \mu(f_{n:n}). \quad (8.17)$$

This follows immediately from Proposition 8.2.4 and formula (8.14), since the product $\mu(f_1) \cdots \mu(f_k)$ is clearly multilinear and hence multiadditive in (f_1, \dots, f_k) .

To deal with cases when some of the $\mu(f_j)$'s (or the $\mu(f_{n:j})$'s) equal 0 and other ones equal ∞ , let us assume here the convention $0 \cdot \infty := 0$. One may note that, if the nonnegative functions f_1, \dots, f_n are scalar multiples of one another or, more generally, if $f_{\pi(1)} \leq \dots \leq f_{\pi(n)}$ for some permutation π of the set $[n]$, then inequality (8.16) turns into the equality.

As mentioned above, in Corollary 8.2.7 it is not assumed that f_1, \dots, f_n are independent r.v.'s. However, if μ is a probability measure and the r.v.'s f_1, \dots, f_n are independent (but not necessarily identically distributed), then $\mu(f_1) \cdots \mu(f_n) = \mu(f_1 \cdots f_n) = \mu(f_{n:1} \cdots f_{n:n})$ by the second part of (8.3), and so, (8.17) can then be rewritten as the following positive-association-type inequality for the order statistics:

$$\mu(f_{n:1} \cdots f_{n:n}) \geq \mu(f_{n:1}) \cdots \mu(f_{n:n}). \quad (8.18)$$

Let now ψ be any monotone (that is, either nondecreasing or nonincreasing) function from $[0, \infty]$ to $[0, \infty]$. For $f = (f_1, \dots, f_n)$ as before, let

$$\psi \bullet f := (\psi \circ f_1, \dots, \psi \circ f_n).$$

Then for $j \in [n]$ one has $(\psi \bullet f)_{n:j} = \psi \circ f_{n:j}$ if ψ is nondecreasing and $(\psi \bullet f)_{n:j} = \psi \circ f_{n:n+1-j}$ if ψ is nonincreasing. Thus, we have the following ostensibly more general forms of (8.17) and (8.18):

Corollary 8.2.8

$$\mu(\psi \circ f_1) \cdots \mu(\psi \circ f_n) \geq \mu((\psi \bullet f)_{n:1}) \cdots \mu((\psi \bullet f)_{n:n}). \quad (8.19)$$

If μ is a probability measure and the r.v.'s f_1, \dots, f_n are independent, then

$$\mu((\psi \bullet f)_{n:1} \cdots (\psi \bullet f)_{n:n}) \geq \mu((\psi \bullet f)_{n:1}) \cdots \mu((\psi \bullet f)_{n:n}). \quad (8.20)$$

The property of the order statistics $f_{n:1}, \dots, f_{n:n}$ given by inequality (8.20) may be called the diagonal positive orthant dependence—cf. e.g. Definition 2.3 in [12] of the negative orthant dependence.

Immediately from Theorem 8.1.1 or from inequality (8.19) in Corollary 8.2.8, one obtains

Corollary 8.2.9 *Take any $p \in \mathbb{R} \setminus \{0\}$. Then*

$$\mu(f_1^p)^r \cdots \mu(f_n^p)^r \geq \mu(f_{n:1}^p)^r \cdots \mu(f_{n:n}^p)^r \quad (8.21)$$

for any $r \in (0, \infty)$, and

$$\mu(f_1^p)^r \cdots \mu(f_n^p)^r \leq \mu(f_{n:1}^p)^r \cdots \mu(f_{n:n}^p)^r \quad (8.22)$$

for any $r \in (-\infty, 0)$. Here we use the conventions $0^t := \infty$ and $\infty^t := 0$ for $t \in (-\infty, 0)$. We also the following conventions: $0 \cdot \infty := 0$ concerning (8.21) and $0 \cdot \infty := \infty$ concerning (8.22).

Consider now the special case of Corollary 8.2.9 with $r = 1/p$. Letting then $p \rightarrow \infty$, we see that (8.21) will hold with the $\mu(f_j^p)^r$'s and $\mu(f_{n:j}^p)^r$'s replaced there by μ -ess sup f_j and μ -ess sup $f_{n:j}$, respectively, where μ -ess sup denotes the essential supremum with respect to measure μ . This follows because $\mu(h^p)^{1/p} \xrightarrow{p \rightarrow \infty} \mu$ -ess sup h . Similarly, letting $p \rightarrow -\infty$, we see that (8.22) will hold with the $\mu(f_j^p)^r$'s and $\mu(f_{n:j}^p)^r$'s replaced there by μ -ess inf f_j and μ -ess inf $f_{n:j}$, respectively, where μ -ess inf denotes the essential infimum with respect to μ . Moreover, considering (say) the counting measures μ on finite subsets of the set S and noting that $\sup h = \sup_S h$ coincides with the limit of the net $(\max_J h)$ over the filter of all finite subsets J of S , we conclude that (8.21) will hold with the $\mu(f_j^p)^r$'s and $\mu(f_{n:j}^p)^r$'s replaced there by $\sup f_j$ and $\sup f_{n:j}$, respectively. (The statement about the limit can be spelled out as follows: $\sup_S h \geq \max_J h$ for all finite $J \subseteq S$, and for each real c such that $c < \sup h$ there is some finite set $J_c \subseteq S$ such that for all finite sets J such that $J_c \subseteq J \subseteq S$ one has $\max_J h > c$.) Similarly, (8.22) will hold with the $\mu(f_j^p)^r$'s and $\mu(f_{n:j}^p)^r$'s replaced there by $\inf f_j$ and $\inf f_{n:j}$, respectively. Thus, we have

Corollary 8.2.10

$$(\sup f_1) \cdots (\sup f_n) \geq (\sup f_{n:1}) \cdots (\sup f_{n:n}) \quad (8.23)$$

and

$$(\inf f_1) \cdots (\inf f_n) \leq (\inf f_{n:1}) \cdots (\inf f_{n:n}). \quad (8.24)$$

Here we use the following conventions: $0 \cdot \infty := 0$ concerning (8.23) and $0 \cdot \infty := \infty$ concerning (8.24).

Alternatively, one can obtain (8.23) and (8.24) directly from Theorem 8.1.1.

Also, of course there is no need to assume in Corollary 8.2.10 that the functions f_1, \dots, f_n are measurable.

The special cases of inequalities (8.22) and (8.24) for $n = 2$ mean that the functions $h \mapsto \mu(h^p)^r$ and $h \mapsto \inf h$ are log-supermodular functions on the distributive lattice (say \mathcal{L}_Σ) of all nonnegative Σ -measurable functions on S and on the distributive lattice (say \mathcal{L}) of all nonnegative functions on S , respectively.

At this point, let us recall the famous Fortuin–Kasteleyn–Ginibre (FKG) correlation inequality [8], which states that for any log-supermodular function ν on a finite distributive lattice L and any nondecreasing functions F and G on L we have

$$\nu(FG)\nu(1) \geq \nu(F)\nu(G),$$

where $\nu(F) := \sum_{f \in L} \nu(f)$.

Then we immediately obtain

Corollary 8.2.11 *Let \mathcal{L}_Σ° be any finite sub-lattice of the lattice \mathcal{L}_Σ , and let F and G be nondecreasing functions from \mathcal{L}_Σ° to \mathbb{R} . Then*

$$\left(\sum_{h \in \mathcal{L}_\Sigma^\circ} F(h)G(h)\mu(h)^r \right) \left(\sum_{h \in \mathcal{L}_\Sigma^\circ} \mu(h)^r \right) \geq \left(\sum_{h \in \mathcal{L}_\Sigma^\circ} F(h)\mu(h)^r \right) \left(\sum_{h \in \mathcal{L}_\Sigma^\circ} G(h)\mu(h)^r \right)$$

for any $r \in (-\infty, 0)$. Similarly, let \mathcal{L}° be any finite sub-lattice of the lattice \mathcal{L} , and let F and G be nondecreasing functions from \mathcal{L}° to \mathbb{R} . Then

$$\left(\sum_{h \in \mathcal{L}_\Sigma^\circ} F(h)G(h) \inf h \right) \left(\sum_{h \in \mathcal{L}_\Sigma^\circ} \inf h \right) \geq \left(\sum_{h \in \mathcal{L}_\Sigma^\circ} F(h) \inf h \right) \left(\sum_{h \in \mathcal{L}_\Sigma^\circ} G(h) \inf h \right).$$

As shown by Ahlswede and Daykin [2, pp. 288–289], their inequality [2, Theorem 1] almost immediately implies, and is in a sense sharper than, the FKG inequality. Furthermore, Rinott and Saks [21, 22] and Aharoni and Keich [1] independently obtained a more general inequality “for n -tuples of nonnegative functions on a distributive lattice, of which the Ahlswede–Daykin inequality is the case $n = 2$.” More specifically, in notation closer to that used in the present paper, [1, Theorem 1.1] states the following:

Let $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ be nonnegative functions defined on a distributive lattice L such that

$$\prod_{j=1}^n \alpha_j(f_j) \leq \prod_{j=1}^n \beta_j(f_{n:j})$$

for all f_1, \dots, f_n in L . Then for any finite subsets F_1, \dots, F_n of L

$$\prod_{j=1}^n \sum_{f_j \in F_j} \alpha_j(f_j) \leq \prod_{j=1}^n \sum_{g_j \in F_{n:j}} \beta_j(g_j),$$

where

$$F_{n:j} := \{f_{n:j} : f = (f_1, \dots, f_n) \in F_1 \times \dots \times F_n\}.$$

Note that the definition of the “order statistics” used in [1] is different from (8.2) in that their “order statistics” go in the descending, rather than ascending, order; also, the term “order statistics” is not used in [1].

In view of this result of [1] and our Corollaries 8.2.9 and 8.2.10, one immediately obtains the following statement, which generalizes and strengthens Corollary 8.2.11:

Corollary 8.2.12 *Let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be any finite subsets of the lattice \mathcal{L}_Σ . For each $j \in [n]$, let*

$$\mathcal{F}_{n:j} := \{f_{n:j} : f = (f_1, \dots, f_n) \in \mathcal{F}_1 \times \dots \times \mathcal{F}_n\}.$$

Then

$$\prod_{j=1}^n \sum_{f_j \in \mathcal{F}_j} \mu(f_j)^r \leq \prod_{j=1}^n \sum_{h_j \in \mathcal{F}_{n:j}} \mu(h_j)^r \quad (8.25)$$

for any $r \in (-\infty, 0)$.

Similarly, let now $\mathcal{F}_1, \dots, \mathcal{F}_n$ be any finite subsets of the lattice \mathcal{L} . Then

$$\prod_{j=1}^n \sum_{f_j \in \mathcal{F}_j} \inf f_j \leq \prod_{j=1}^n \sum_{h_j \in \mathcal{F}_{n:j}} \inf h_j.$$

Comparing inequalities (8.21) and (8.22) in Corollary 8.2.9 or inequalities (8.23) and (8.24) in Corollary 8.2.10, one may wonder whether the FKG-type inequalities stated in Corollaries 8.2.11 and 8.2.12 for the functions $h \mapsto \mu(h)^r$ with $r < 0$ and $h \mapsto \inf h$ admit of the corresponding reverse analogues for the functions $h \mapsto \mu(h)^r$ with $r > 0$ and $h \mapsto \sup h$. However, it is not hard to see that such FKG-type inequalities are not reversible in this sense, a reason being that the sets $\mathcal{F}_{n:j}$ may be much larger than the sets \mathcal{F}_j .

E.g., suppose that $n = 2$, $S = \mathbb{R}$, μ is a Borel probability measure on \mathbb{R} , $0 < \varepsilon < \delta < 1$, N is a natural number, \mathcal{F}_1 is the set of N pairwise distinct constant functions f_1, \dots, f_N on \mathbb{R} such that $1 - \varepsilon < f_j < 1 + \varepsilon$ for all $j \in [n]$, and $\mathcal{F}_2 = \{g_1, \dots, g_N\}$, where $g_j := (1 - \delta)\mathbf{1}_{(-\infty, j]} + (1 + \delta)\mathbf{1}_{(j, \infty)}$ and $\mathbf{1}_A$ denotes

the indicator of a set A . Then it is easy to see that each of the sets $\mathcal{F}_{2:1}$ and $\mathcal{F}_{2:2}$ is of cardinality N^2 . So, letting $\delta \downarrow 0$ (so that $\varepsilon \downarrow 0$ as well), we see that, for any real r , the right-hand side of (8.25) goes to N^4 whereas its left-hand side goes to N^2 , which is much less than N^4 if N is large.

Example 8.2.13 Closely related to Example 8.2.5 is as follows. Suppose that (S, Σ) is a measurable space, μ is a measure on the product σ -algebra $\Sigma^{\otimes k}$, and L is a subring of Σ . Then L is complementable and the function $m: L^k \rightarrow \mathbb{R}$ given by the formula

$$m(A_1, \dots, A_k) := \mu(A_1 \times \dots \times A_k) \quad (8.26)$$

for $(A_1, \dots, A_k) \in L^k$ is multiadditive.

A particular case of formula (8.26) is

$$m(A_1, \dots, A_k) := \text{card}(G \cap (A_1 \times \dots \times A_k)), \quad (8.27)$$

where card stands for the cardinality and G is an arbitrary subset of S^k . If G is symmetric in the sense that $(s_1, \dots, s_k) \in G$ iff $(s_{\pi(1)}, \dots, s_{\pi(k)}) \in G$ for all permutations π of the set $[k]$, then G represents the set (say E) of all hyperedges of a k -uniform hypergraph over S , in the sense that $(s_1, \dots, s_k) \in G$ iff $\{s_1, \dots, s_k\} \in E$.

We now have another immediate corollary of Proposition 8.2.4:

Corollary 8.2.14 *Suppose that k and n are natural numbers such that $k \leq n$, (S, Σ) is a measurable space, μ is a measure on the product σ -algebra $\Sigma^{\otimes k}$, and L is a subring of Σ . Then*

$$\sum_{\pi \in \Pi_k^{[n]}} \mu(A_{n:\pi(1)} \times \dots \times A_{n:\pi(k)}) \leq \sum_{\pi \in \Pi_k^{[n]}} \mu(A_{\pi(1)} \times \dots \times A_{\pi(k)}) \quad (8.28)$$

for all $(A_1, \dots, A_n) \in L^n$.

8.3 Proofs

One may note that formula (8.31) in the proof of Theorem 8.1.1 below defines a step similar to a step in the process of the so-called insertion search (cf. e.g. [15, Section 5.2.1] (also called the sifting or sinking technique)—except that here we do the pointwise comparison of functions (rather than numbers) and therefore we do not stop when the right place of the value $f_{n+1}(s)$ of the “new” function f_{n+1} among the already ordered values $f_{n:1}(s), \dots, f_{n:n}(s)$ at a particular point $s \in S$ has been found, because this place will in general depend on s . So, the proof that (8.31) implies (8.34) may be considered as (something a bit more than) a rigorous proof of the validity of the insertion search algorithm, avoiding such informal, undefined terms as swap, moving, and interleaving.

Proof of Theorem 8.1.1 Let us prove the theorem by induction in n . For $n = 1$, the result is trivial. To make the induction step, it suffices to prove the following: For any natural $n \geq 2$, if the function $\Lambda: L^n \rightarrow \mathcal{R}$ is generalized $(n:2)$ -semimodular and the function $L^{n-1} \ni (f_1, \dots, f_{n-1}) \mapsto \Lambda(f_1, \dots, f_n)$ is generalized $(n-1)$ -semimodular for each $f_n \in L$, then Λ is generalized n -semimodular. Thus, we are assuming that the function $\Lambda: L^n \rightarrow \mathcal{R}$ is generalized $(n:2)$ -semimodular and

$$\Lambda(f_1, \dots, f_n) \bowtie \Lambda(f_{n-1:1}, \dots, f_{n-1:n-1}, f_n) \quad (8.29)$$

for all $(f_1, \dots, f_n) \in L^n$, where $f_{n-1:1}, \dots, f_{n-1:n-1}$ are the “order statistics” based on (f_1, \dots, f_{n-1}) .

Take indeed any $(f_1, \dots, f_n) \in L^n$. Define the rectangular array of functions $(g_{k,j}: k \in \overline{0, n-1}, j \in [n])$ recursively, as follows:

$$(g_{0,1}, \dots, g_{0,n-1}, g_{0,n}) := (f_{n-1:1}, \dots, f_{n-1:n-1}, f_n) \quad (8.30)$$

and, for $k \in \overline{1, n-1}$ and $j \in [n]$,

$$g_{k,j} := \begin{cases} g_{k-1,j} & \text{if } j \in \overline{1, n-k-1} \cup \overline{n-k+2, n}, \\ g_{k-1,n-k} \wedge g_{k-1,n-k+1} & \text{if } j = n-k, \\ g_{k-1,n-k} \vee g_{k-1,n-k+1} & \text{if } j = n-k+1. \end{cases} \quad (8.31)$$

By (8.29) and (8.30),

$$\Lambda(f_1, \dots, f_n) \bowtie \Lambda(g_{0,1}, \dots, g_{0,n-1}, g_{0,n}). \quad (8.32)$$

Moreover, for each $k \in \overline{1, n-1}$,

$$\Lambda(g_{k-1,1}, \dots, g_{k-1,n}) \bowtie \Lambda(g_{k,1}, \dots, g_{k,n}), \quad (8.33)$$

since Λ is generalized $(n:2)$ -semimodular.

It follows from (8.32) and (8.33) that

$$\Lambda(f_1, \dots, f_n) \bowtie \Lambda(g_{n-1,1}, \dots, g_{n-1,n}).$$

It remains to verify the identity

$$(g_{n-1,1}, \dots, g_{n-1,n}) \stackrel{(?)}{=} (f_{n:1}, \dots, f_{n:n}). \quad (8.34)$$

In accordance with Remark 8.1.3, we may and shall assume that the distributive lattice L is a lattice of nonnegative real-valued functions on a set S , so that (8.3) holds for each $s \in S$.

In the remainder of the proof, fix any $s \in S$. Then

$$\{\{g_{0,1}(s), \dots, g_{0,n}(s)\}\} = \{\{f_1(s), \dots, f_n(s)\}\},$$

by (8.30) and the second part of (8.3) used with $n - 1$ in place of n ; also, for each $k \in \overline{1, n - 1}$,

$$\{\{g_{k,1}(s), \dots, g_{k,n}(s)\}\} = \{\{g_{k-1,1}(s), \dots, g_{k-1,n}(s)\}\},$$

by (8.31). So,

$$\{\{g_{n-1,1}(s), \dots, g_{n-1,n}(s)\}\} = \{\{f_1(s), \dots, f_n(s)\}\}.$$

Therefore, to complete the proof of (8.34) and thus that of Theorem 8.1.1, it remains to show that

$$g_{n-1,1}(s) \stackrel{(?)}{\leq} \cdots \stackrel{(?)}{\leq} g_{n-1,n}(s), \quad (8.35)$$

which will follow immediately from

Lemma 8.3.1 *For each $k \in \overline{1, n - 1}$, the following assertion is true for all $s \in S$:*

$$\begin{aligned} g_{k,j}(s) \leq g_{k,j+1}(s) \text{ for all } j \in \overline{1, n - k - 2} \cup \overline{n - k, n - 1}; \\ \text{also, } g_{k,n-k-1}(s) \leq g_{k,n-k+1}(s) \text{ if } k \leq n - 2. \end{aligned} \quad (A_k)$$

Indeed, (8.35) is the first clause in assertion (A_k) with $k = n - 1$. Thus, what finally remains to prove Theorem 8.1.1 is to present the following.

Proof of Lemma 8.3.1 For simplicity, let us be dropping (s) —thus writing $g_{k,j}$, f_n , \dots in place of $g_{k,j}(s)$, $f_n(s)$, \dots . We shall prove Lemma 8.3.1 by induction in $k \in \overline{1, n - 1}$. Assertion (A_1) means that $g_{1,1} \leq \cdots \leq g_{1,n-2}$, $g_{1,n-1} \leq g_{1,n}$, and $g_{1,n-2} \leq g_{1,n}$ if $1 \leq n - 2$. So, in view of (8.31) and (8.30), (A_1) can be rewritten as follows: $f_{n-1:1} \leq \cdots \leq f_{n-1:n-2}$, $f_{n-1:n-1} \wedge f_n \leq f_{n-1:n-1} \vee f_n$, and $f_{n-1:n-2} \leq f_{n-1:n-1} \vee f_n$; all these inequalities are obvious. So, (A_1) holds.

Take now any $k \in \overline{2, n - 1}$ and suppose that (A_{k-1}) holds. We need to show that then (A_k) holds.

For all $j \in \overline{1, n - k - 2} \cup \overline{n - k + 2, n - 1}$, we have $j + 1 \in \overline{1, n - k - 1} \cup \overline{n - k + 2, n}$, whence, by (8.31) and the first clause of (A_{k-1}) , $g_{k,j} = g_{k-1,j} \leq g_{k-1,j+1} = g_{k,j+1}$. So,

$$g_{k,j} \leq g_{k,j+1} \quad \text{for } j \in \overline{1, n - k - 2} \cup \overline{n - k + 2, n - 1}. \quad (8.36)$$

If $j = n - k$ then, by (8.31), $g_{k,j} = g_{k-1,n-k} \wedge g_{k-1,n-k+1} \leq g_{k-1,n-k} \vee g_{k-1,n-k+1} = g_{k,j+1}$.

If $j = n - k + 1$ then the condition $k \in \overline{2, n-1}$ implies $j \leq n - 1$, and so, by (8.31) and the second and first clauses of (A_{k-1}) , $g_{k,j} = g_{k-1,n-k} \vee g_{k-1,n-k+1} \leq g_{k-1,n-k+2} = g_{k-1,j+1} = g_{k,j+1}$.

Thus, in view of (8.36), the first clause of (A_k) holds. Also, if $k \leq n - 2$ then, by (8.31) and the first clause of (A_{k-1}) , $g_{k,n-k-1} = g_{k-1,n-k-1} \leq g_{k-1,n-k} \leq g_{k-1,n-k} \vee g_{k-1,n-k+1} = g_{k,n-k+1}$, so that the second clause of (A_k) holds as well. This completes the proof of Lemma 8.3.1. \square

Thus, Theorem 8.1.1 is proved. \square

Proof of Proposition 8.2.1 Take any $(f_1, \dots, f_n) \in L^n$. Corollary B.3 in [17] states that $x < y$ iff x is in the convex hull of the set of all points obtained by permuting the coordinates of the vector y . Also, since the function λ is nondecreasing, we have $\lambda(f_1 \vee f_2) \geq \lambda(f_1) \vee \lambda(f_2)$. For any real a, b, c such that $c \geq a \vee b$, we have $(a, b) = (1-t)(a+b-c, c) + t(c, a+b-c)$ for $t = \frac{c-b}{2c-a-b} \in [0, 1]$ if $c > (a+b)/2$ and for any $t \in [0, 1]$ otherwise (that is, if $a = b = c$). So, the point (a, b) is a convex combination of points $(a+b-c, c)$ and $(c, a+b-c)$. Using this fact for $a = \lambda(f_1)$, $b = \lambda(f_2)$, $c = \lambda(f_1 \vee f_2)$, we see that

$$(\lambda(f_1), \dots, \lambda(f_n)) < (\lambda(f_1) + \lambda(f_2) - \lambda(f_1 \vee f_2), \lambda(f_1 \vee f_2), \lambda(f_3), \dots, \lambda(f_n)).$$

Also, $\lambda(f_1 \wedge f_2) \leq \lambda(f_1) + \lambda(f_2) - \lambda(f_1 \vee f_2)$, by the submodularity of λ . Therefore and because F is nondecreasing (in each of its n arguments) and Schur-concave, we conclude that

$$\begin{aligned} F(\lambda(f_1 \wedge f_2), \lambda(f_1 \vee f_2), \lambda(f_3), \dots, \lambda(f_n)) \\ \leq F(\lambda(f_1) + \lambda(f_2) - \lambda(f_1 \vee f_2), \lambda(f_1 \vee f_2), \lambda(f_3), \dots, \lambda(f_n)) \\ \leq F(\lambda(f_1), \dots, \lambda(f_n)). \end{aligned}$$

Quite similarly,

$$\begin{aligned} F(\lambda(f_1), \dots, \lambda(f_{i-1}), \lambda(f_i \wedge f_{i+1}), \lambda(f_i \vee f_{i+1}), \lambda(f_{i+2}), \dots, \lambda(f_n)) \\ \leq F(\lambda(f_1), \dots, \lambda(f_n)) \end{aligned}$$

for all $i \in \overline{1, n-1}$, so that the function F is indeed generalized $(n:2)$ -submodular and hence, by Theorem 8.1.1, generalized n -submodular. \square

Proof of Proposition 8.2.3 In view of Theorem 8.1.1, it is enough to show that the function $\Lambda = \Lambda_{\varphi, \psi}$ is generalized $(n:2)$ -submodular. Without loss of generality (w.l.o.g.), we may and shall assume that the function φ is nondecreasing, since $\Lambda_{\varphi^-, \psi} = \Lambda_{\varphi, \psi}$, where $\varphi^-(u) := \varphi(-u)$ for all real u . Also, w.l.o.g. $\psi(0) = 0$ and hence $\Psi(0) = 0$.

Take any $f = (f_1, \dots, f_n) \in L^n$. Then, letting

$$\tilde{\Psi}(g) := \Psi(g) + \Psi(-g) \tag{8.37}$$

for $g \in L$, one has

$$\Lambda(f_1, f_2, f_3, \dots, f_n) = \tilde{\Psi}(f_1 - f_2) + \sum_{j=3}^n (\tilde{\Psi}(g_j) + \tilde{\Psi}(h_j)) + R, \quad (8.38)$$

where $g_j := f_1 - f_j$, $h_j := f_2 - f_j$, and $R := \sum_{3 \leq j < k \leq n} \tilde{\Psi}(f_j - f_k)$. Since $f_1 \wedge f_2 - f_1 \vee f_2 = -|f_1 - f_2|$, one similarly has

$$\Lambda(f_1 \wedge f_2, f_1 \vee f_2, f_3, \dots, f_n) = \tilde{\Psi}(|f_1 - f_2|) + \sum_{j=3}^n (\tilde{\Psi}(g_j \wedge h_j) + \tilde{\Psi}(g_j \vee h_j)) + R. \quad (8.39)$$

Next,

$$\tilde{\Psi}(f_1 - f_2) = \psi \left(\int_S \varphi \circ (f_1 - f_2) \, d\mu \right) + \psi \left(\int_S \varphi \circ (f_2 - f_1) \, d\mu \right), \quad (8.40)$$

$$\tilde{\Psi}(|f_1 - f_2|) = \psi \left(\int_S \varphi \circ |f_1 - f_2| \, d\mu \right) + \psi \left(\int_S \varphi \circ (-|f_2 - f_1|) \, d\mu \right), \quad (8.41)$$

$\varphi \circ (f_1 - f_2) + \varphi \circ (f_2 - f_1) = \varphi \circ |f_1 - f_2| + \varphi \circ (-|f_1 - f_2|)$ and hence

$$\int_S \varphi \circ (f_1 - f_2) \, d\mu + \int_S \varphi \circ (f_2 - f_1) \, d\mu = \int_S \varphi \circ |f_1 - f_2| \, d\mu + \int_S \varphi \circ (-|f_2 - f_1|) \, d\mu. \quad (8.42)$$

Also, since φ is nondecreasing, $\varphi \circ (f_1 - f_2) \vee \varphi \circ (f_2 - f_1) \leq \varphi \circ |f_1 - f_2|$ and hence

$$\int_S \varphi \circ (f_1 - f_2) \, d\mu \vee \int_S \varphi \circ (f_2 - f_1) \, d\mu \leq \int_S \varphi \circ |f_1 - f_2| \, d\mu. \quad (8.43)$$

Since the function ψ is convex, it follows from (8.40)–(8.43) that

$$\tilde{\Psi}(f_1 - f_2) \leq \tilde{\Psi}(|f_1 - f_2|). \quad (8.44)$$

Further, take any $j \in \overline{3, n}$. Then $\varphi \circ g_j + \varphi \circ h_j = \varphi \circ (g_j \wedge h_j) + \varphi \circ (g_j \vee h_j)$. So,

$$\int_S (\varphi \circ g_j) \, d\mu + \int_S (\varphi \circ h_j) \, d\mu = \int_S \varphi \circ (g_j \wedge h_j) \, d\mu + \int_S \varphi \circ (g_j \vee h_j) \, d\mu.$$

Moreover, since φ is nondecreasing, $\int_S \varphi \circ (g_j \vee h_j) \, d\mu$ is no less than each of the integrals $\int_S (\varphi \circ g_j) \, d\mu$ and $\int_S (\varphi \circ h_j) \, d\mu$. So, in view of (8.12) and the convexity

of the function ψ , one has $\Psi(g_j) + \Psi(h_j) \leq \Psi(g_j \wedge h_j) + \Psi(g_j \vee h_j)$. Similarly, because $\int_S \varphi \circ (-g_j \wedge h_j) d\mu$ is no less than each of the integrals $\int_S \varphi \circ (-g_j) d\mu$ and $\int_S \varphi \circ (-h_j) d\mu$, one has $\Psi(-g_j) + \Psi(-h_j) \leq \Psi(-g_j \wedge h_j) + \Psi(-g_j \vee h_j)$. So, by (8.37), $\tilde{\Psi}(g_j) + \tilde{\Psi}(h_j) \leq \tilde{\Psi}(g_j \wedge h_j) + \tilde{\Psi}(g_j \vee h_j)$.

Therefore, by (8.38), (8.39), and (8.44), $\Lambda(f_1, f_2, f_3, \dots, f_n) \leq \Lambda(f_1 \wedge f_2, f_1 \vee f_2, f_3, \dots, f_n)$. Similarly, $\Lambda(f_1, \dots, f_{j-1}, f_j, f_{j+1}, f_{j+2}, \dots, f_n) \leq \Lambda(f_1, \dots, f_{j-1}, f_j \wedge f_{j+1}, f_j \vee f_{j+1}, f_{j+2}, \dots, f_n)$ for all $j \in \overline{1, n-1}$.

Thus, the function Λ is generalized $(n:2)$ -supermodular, and so, by Theorem 8.1.1, it is generalized n -supermodular. \square

Proof of Proposition 8.2.4 Fix any $(f_1, \dots, f_n) \in L^n$. Then, in view of the permutation symmetry of \bar{m} defined by (8.15),

$$\frac{1}{k!} \Lambda_m(f_1, \dots, f_n) = \lambda_2(f_{n-1}, f_n) + \lambda_1(f_{n-1}) + \lambda_1(f_n) + \lambda_0, \quad (8.45)$$

where

$$\begin{aligned} \lambda_2(f, g) &:= \sum_{1 \leq i_1 < \dots < i_{k-2} \leq n-2} \bar{m}(f_{i_1}, \dots, f_{i_{k-2}}, f, g), \\ \lambda_1(f) &:= \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n-2} \bar{m}(f_{i_1}, \dots, f_{i_{k-1}}, f), \\ \lambda_0 &:= \sum_{1 \leq i_1 < \dots < i_k \leq n-2} \bar{m}(f_{i_1}, \dots, f_{i_k}), \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{1}{k!} \Lambda_m(f_1, \dots, f_{n-2}, f_{n-1} \wedge f_n, f_{n-1} \vee f_n) &= \lambda_2(f_{n-1} \wedge f_n, f_{n-1} \vee f_n) \\ &+ \lambda_1(f_{n-1} \wedge f_n) + \lambda_1(f_{n-1} \vee f_n) + \lambda_0. \end{aligned} \quad (8.46)$$

Note that the function $\lambda_2: L^2 \rightarrow \mathbb{R}$ is 2-additive and permutation-symmetric, and the function $\lambda_1: L^2 \rightarrow \mathbb{R}$ is additive. Take any f and g in L . Then $(f \vee g) \wedge f = f$ and $(f \vee g) \setminus f = g \setminus f$. So, by the additivity of λ_1 we have $\lambda_1(f \vee g) = \lambda_1(f) + \lambda_1(g \setminus f)$, whereas $\lambda_1(f \wedge g) + \lambda_1(g \setminus f) = \lambda_1(g)$. So,

$$\lambda_1(f \wedge g) + \lambda_1(f \vee g) = \lambda_1(f \wedge g) + \lambda_1(f) + \lambda_1(g \setminus f) = \lambda_1(f) + \lambda_1(g). \quad (8.47)$$

By the 2-additivity and permutation symmetry of λ_2 and because the function λ_2 is 2-additive, permutation-symmetric, and nonnegative, we have

$$\begin{aligned}
 \lambda_2(f \wedge g, f \vee g) &= \lambda_2(f \wedge g, f \setminus g) + \lambda_2(f \wedge g, g) \\
 &= \lambda_2(f \wedge g, f \setminus g) + \lambda_2(f, g) - \lambda_2(f \setminus g, g) \\
 &= \lambda_2(f \wedge g, f \setminus g) + \lambda_2(f, g) - \lambda_2(f \setminus g, g \wedge f) - \lambda_2(f \setminus g, g \setminus f) \\
 &= \lambda_2(f, g) - \lambda_2(f \setminus g, g \setminus f) \\
 &\leq \lambda_2(f, g).
 \end{aligned}
 \tag{8.48}$$

It follows from (8.45), (8.46), (8.47), and (8.48) (with $f = f_{n-1}$ and $g = f_n$) that

$$\Lambda_m(f_1, \dots, f_{n-2}, f_{n-1} \wedge f_n, f_{n-1} \vee f_n) \leq \Lambda_m(f_1, \dots, f_n).$$

Therefore, being permutation-symmetric, the function Λ_m is indeed generalized ($n:2$)-submodular. Hence, by Theorem 8.1.1, Λ_m is generalized n -submodular. \square

References

1. R. Aharoni, U. Keich, A generalization of the Ahlswede-Daykin inequality. *Discrete Math.* **152**(1–3), 1–12 (1996)
2. R. Ahlswede, D.E. Daykin, Inequalities for a pair of maps $S \times S \rightarrow S$ with S a finite set. *Math. Z.* **165**(3), 267–289 (1979)
3. F. Bach, Learning with submodular functions: a convex optimization perspective. *Found. Trends® Mach. Learn.* **6**(2–3), 145–373 (2013)
4. F. Bach, Submodular functions: from discrete to continuous domains. *Math. Program.* **175**(1–2), 419–459 (2019)
5. C. Borell, A note on an inequality for rearrangements. *Pac. J. Math.* **47**, 39–41 (1973)
6. G. Choquet, Theory of capacities. *Ann. Inst. Fourier* **5**, 131–295 (1954)
7. H.A. David, H.N. Nagaraja, *Order Statistics*. Wiley Series in Probability and Statistics, 3rd edn. (Wiley, Hoboken, 2003)
8. C.M. Fortuin, P.W. Kasteleyn, J. Ginibre, Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* **22**, 89–103 (1971)
9. S. Fujishige, *Submodular Functions and Optimization*. *Annals of Discrete Mathematics*, vol. 47 (North-Holland, Amsterdam, 1991)
10. G. Grätzer, *Lattice Theory: Foundation* (Birkhäuser, Basel, 2011)
11. G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities*, 2nd edn. (Cambridge University Press, Cambridge, 1952)
12. K. Joag-Dev, F. Proschan, Negative association of random variables, with applications. *Ann. Stat.* **11**(1), 286–295 (1983)
13. S. Karlin, Y. Rinott, Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10**(4), 467–498 (1980)
14. S. Karlin, Y. Rinott, Classes of orderings of measures and related correlation inequalities. II. Multivariate reverse rule distributions. *J. Multivariate Anal.* **10**(4), 499–516 (1980)

15. D.E. Knuth, *The Art of Computer Programming*, vol. 3, 2nd edn. (Addison-Wesley, Reading, 1998). Sorting and searching [MR0445948]
16. G.G. Lorentz, An inequality for rearrangements. *Am. Math. Mon.* **60**, 176–179 (1953)
17. A.W. Marshall, I. Olkin, B.C. Arnold, *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics, 2nd edn. (Springer, New York, 2011)
18. P. Milgrom, J. Roberts, Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica* **58**(6), 1255–1277 (1990)
19. H. Narayanan, *Submodular Functions and Electrical Networks*. Annals of Discrete Mathematics, vol. 54 (North-Holland, Amsterdam, 1997)
20. I. Pinelis, Optimal binomial, Poisson, and normal left-tail domination for sums of nonnegative random variables. *Electron. J. Probab.* **21**(20), 19 (2016)
21. Y. Rinott, M. Saks, On FKG-type and permanental inequalities, in *Stochastic Inequalities (Seattle, WA, 1991)*. IMS Lecture Notes Monograph Series, vol. 22 (Institute of Mathematical Statistics, Hayward, 1992), pp. 332–342
22. Y. Rinott, M. Saks, Correlation inequalities and a conjecture for permanents. *Combinatorica* **13**(3), 269–277 (1993)
23. H.D. Ruderman, Two new inequalities. *Am. Math. Mon.* **59**, 29–32 (1952)
24. D.M. Topkis, Equilibrium points in nonzero-sum n -person submodular games. *SIAM J. Control Optim.* **17**(6), 773–787 (1979)
25. D.M. Topkis, *Supermodularity and Complementarity*. Frontiers of Economic Research (Princeton University Press, Princeton, 1998).