# A General Approach to Sparse Basis Selection: Majorization, Concavity, and Affine Scaling

*UCSD Department of Electrical and Computer Engineering*

CENTER FOR INFORMATION ENGINEERING

Report No. UCSD-CIE-97-7-1

Kenneth Kreutz-Delgado
*Center for Information Engineering*
*Department of Electrical and Computer Engineering*

Bhaskar D. Rao
*Center for Wireless Communication*
*Department of Electrical and Computer Engineering*

UNIVERSITY OF CALIFORNIA, SAN DIEGO

July 15, 1997

# A General Approach to Sparse Basis Selection: Majorization, Concavity, and Affine Scaling

K. Kreutz-Delgado and B.D. Rao
Electrical and Computer Engineering Department
University of California, San Diego
9500 Gilman Drive, Mail Code 0407
La Jolla, California 92093-0407
{kreutz, brao}@ucsd.edu

## Abstract

*Measures for sparse best–basis selection are analyzed and shown to fit into a general framework based on majorization, Schur-concavity, and concavity. This framework facilitates the analysis of algorithm performance and clarifies the relationships between existing proposed concentration measures useful for sparse basis selection. It also allows one to define new concentration measures, and several general classes of measures are proposed and analyzed in this paper. Admissible measures are given by the Schur-concave functions, which are the class of functions consistent with the so-called Lorentz ordering (a partial ordering on vectors also known as majorization). In particular, concave functions form an important subclass of the Schur-concave functions which attain their minima at sparse solutions to the best basis selection problem. A general affine scaling optimization algorithm obtained from a special factorization of the gradient function is developed and proved to converge to a sparse solution for measures chosen from within this subclass.*

**Keywords:** Best Basis Selection; Sparse Basis Selection; Sparsity; Sparsity Measures; Diversity Measures; Concave Minimization; Schur-Concavity; Affine Scaling Methods; Model Reduction; Complexity Reduction.

# 1 INTRODUCTION

The problem of best basis selection has many important applications to signal representation [19, 81], biomagnetic imaging [43, 36], speech coding [76], and function approximation [18, 61], among others [16, 44, 39, 54]. Of particular interest in this paper are approaches that select basis vectors by minimizing concentration measures subject to the linear constraint

$$Ax = b, \tag{1}$$

where $A$ is an $m \times n$ matrix formed using the vectors from an overdetermined dictionary of basis vectors, $m < n$, and it is assumed that $\text{rank}(A) = m$ [17, 1]. The system of equations (1) has infinitely many solutions, and the solution set is a linear variety denoted by $LV(A, b) = x_p + \mathcal{N}(A)$, where $x_p$ is any particular solution to (1) and $\mathcal{N}(A) =$ Nullspace of $A$ [49]. Constrained minimization of concentration measures results in sparse solutions consistent with membership in $LV(A, b)$. Sparse solutions refer to *basic solutions*, solutions with $m$ nonzero entries, and *degenerate* basic solutions, solutions with less than $m$ nonzero entries [29]. The degenerate basic solutions, if they exist, are more desirable from a concentration objective. The nonzero entries of a sparse solution indicate the basis vectors (columns of $A$) selected. Popular concentration measures used in this context are the Shannon Entropy, the Gaussian Entropy, and the $\ell_{(p\leq 1)}$ ($p$-norm-like) concentration measures, $p \leq 1$ [19, 81, 27, 69].

In our earlier paper [69], an affine scaling methodology was proposed to obtain sparse solutions to (1) by minimizing the Gaussian entropy and the $\ell_{(p\leq 1)}$ ($p$-norm-like) concentration measures for $p \leq 1$ (including $p$ negative). The resulting algorithms for the $\ell_{(p\leq 1)}$ concentration measures generalize the FOCUSS (**FOC**al **U**nderdetermined **S**ystem **S**olver) class of algorithms first developed in [36, 37]. It is also shown in [69] that the algorithm for $\ell_{(p=0)}$ is well-defined in a certain sense and yields precisely the same algorithm and solution as for the Gaussian entropy—a result that is consistent with the definition of the $p = 0$ case commonly considered in the literature on mathematical inequalities [38, 57, 9]. The case $p = 0$ corresponds to using the numerosity measure proposed in [27] (see also [69]). Interestingly, in [69] the algorithm corresponding to the Shannon entropy measure proposed in [19, 81] was shown to not converge to a fully sparse solution, although an increase in concentration was seen to occur. In this paper we carefully analyze these and other sparsity measures and discuss desirable properties for good concentration measures to have.

Broadly speaking, the contributions of this paper are three-fold. First we discuss the concepts of majorization, Schur-concavity, and concavity, and describe their relevance to constructing measures of sparsity and diversity. Next, utilizing this framework, we generalize the concentration measures mentioned above by introducing the class of signomial measures and the Renyi entropy, along with several useful variants. A detailed analysis of these measures is conducted, ascertaining their suitability as concentration measures. Finally, using a special factorization of the gradient of the concentration measure, we develop a class of convergent algorithms for sparse basis selection. Specifically, the affine scaling methodology of [69] is shown to be extensible to deal with the broader classes of concentration measures proposed herein, and a general convergence analysis is conducted. The general classes of measures covered by the convergence analysis eliminates the need for a convergence analysis on a case-by-case basis, as would otherwise be necessary.

The subject of concentration measures for best basis selection was first discussed in [19, 81, 27]. In [81], the Shannon entropy and the $\ell_{(p\leq 1)}$, $0 < p \leq 1$, measures, both evaluated on the "probability" $\tilde{x} = |x|^2/\|x\|_2^2 \in \mathsf{R}^n$, are analyzed at length.[1] It is shown that these functions are consistent with concentration as measured by the partial sums of the decreasing rearrangement of the elements of $\tilde{x}$. Ordering of vectors according to their partial sums is known as *majorization* and many results relating majorization to functional inequalities exist that can be exploited to more fully understand the relationship between majorization and measures of concentration [38, 13, 9, 57, 53, 5, 65, 47].

Inspired by the insightful discussion given in Chapter 8 of [81], we have been motivated to analyze and develop concentration measures from the perspective of majorization theory and to consider measures drawn from the general class of Schur-concave functions, which are precisely the functions consistent with the partial order induced by majorization. In this paper, we argue that concentration measures should be drawn from the class of Schur-concave functions [38, 53, 5] and, in particular, that good concentration measures are a subclass of concave functions. We fully investigate the properties of functions drawn from this class, with a special emphasis on the subclass of concave functions, and construct several general classes of possible concentration measures taking as our point of departure the analysis begun in [81].

There is a long history of exploiting the properties of majorization, Schur-concavity, and concavity for obtaining good concentration measures in economics and social science [6, 75, 23, 74, 15, 30, 32, 58, 7]. Any measure of concentration is also a measure of equality or diversity, so researchers looking for good measures of economic concentration or social equality use many of the same mathematical constructs. Similar ideas have been used in ecology [66, 67, 77, 64], physics [2], and computer science [26, 68, 62]. Recently, [46] has discussed the utility of Schur-concave selection functionals (i.e., concentration measures) for obtaining multiscale signal representations basis vectors. This rich vein of past and present research is a good source for many ideas and possible concentration functions that can be exploited to produce good sparse basis algorithms.

[1] We use the notation where $|x|$, $x^2$, $x^{\frac{1}{2}}$, $x \geq 0$, etc., are defined component-wise for $x \in \mathsf{R}^n$.

Because we consider functions that are *minimized* in order to *increase* concentration (i.e., sparsity), as in much of the economics and ecology literature we refer to these functions as *diversity measures* (i.e., we speak of minimizing diversity in order to increase concentration). Since, as mentioned in the previous paragraph, a measure of concentration also provides a measure of diversity, we can interchangeably refer to the sparsity measures discussed in this paper either as concentration measure (consistent with earlier work such as [19, 81, 69]) or as diversity functions, with the latter terminology being preferred and used predominantly in the remainder of this paper.[2]

The availability of provably globally convergent algorithms for minimization of diversity measures is limited. This is because the subject of optimization theory largely deals with the minimization of *convex* or locally convex functions, whereas the cost functions considered here are usually *concave*. The minimization of concave functions is well known to be difficult because of the existence of multiple minima [73, 63, 42, 41, 12]. In our algorithmic study we further explore the class of affine scaling methods developed in our earlier work [69], and based on a particular factorization of the gradient function, which naturally extends to the general classes of measures analyzed in this paper. The insights provided by the majorization-based framework for diversity measures enables a convergence analysis that includes a wide class of measures.

The outline of the paper is as follows. In Section 2 we discuss majorization, Schur-concavity, concavity, and the properties of Schur-concave and concave functions. In Section 3 we develop and analyze a variety of diversity measures, including those described in [19, 81, 27, 69]. In Section 4 we present gradient factorization–based Affine Scaling Transformation (AST) algorithms guaranteed to locally minimize useful classes of diversity measures, and thereby provide sparse solutions to the best basis selection problem. Conclusions are given in Section 5.

# 2 THE MEASUREMENT OF DIVERSITY

In this section we first discuss majorization, a partial ordering on vectors. We then describe the property of Schur-concavity as a reasonable necessary condition for a measure of diversity. Finally, we discuss the class of separable concave functions and their importance as good measures of diversity. This discussion, which covers important known results and provides some new results (Theorems 3 and 8) and definitions (Definition 6), is motivated by two factors. Firstly, these results are relevant to the development of the paper. Secondly, the review provides an opportunity to highlight results that may not be well known to the signal

[2] Reference [27] refers to “measures of anti-sparsity.”

processing community, and which are useful in understanding the subject of sparsity and its application to the basis selection problem. A detailed background on majorization and the Lorentz order is readily available from references [53, 5, 4, 65, 47].

For this discussion, a basic familiarity with the concepts of convex sets and convex and concave functions is assumed. The necessary background material can be found in one of the many fine introductory expositions, including the references [73, 72, 65, 80]. In this paper an overbar denotes absolute value, $\bar{x} = |x|$, while a "tilde" denotes a function of $x$ *specified by context*: $\tilde{x} = |x|$, $|x|^2$, $|x|/\|x\|_1$, or $|x|^2/\|x\|_2^2$ . $\|\cdot\|_p$ denotes the standard $p$-norm, $\|x\|_p^p = \sum_i |x|^p$. The $i^{th}$ component of a vector $x$ is given by $x[i]$ or, at times, $x_i$, with a particular choice made to improve readability of detailed equations. We will also denote the $k^{th}$ vector in a sequence by $x_k$; the distinction between $x_k$ as an element and $x_k$ as a vector should be clear from context. The bold-face vector $\mathbf{e}_i$ denotes the canonical unit vector with a 1 in the $i^{th}$ position and zeros elsewhere. The bold-face vector $\mathbf{1} \in \mathsf{R}^n$ denotes the vector with 1 in every position. $\mathcal{Q}_l \subset \mathsf{R}^n$, $l \leq \cdots \leq 2^n$, denote the $2^n$ orthants of $\mathsf{R}^n$. $\mathcal{Q}_1$ is the positive orthant with $x \in \mathcal{Q}_1$ iff $x \geq 0$, where the inequality is defined component-wise, $x[i] \geq 0$. The function $\log(x)$ denotes the natural logarithm of $x$.

## 2.1 Majorization and Schur-Concavity

To simplify the discussion, in this section we restrict our discussion to the positive orthant $\mathcal{Q}_1 \subset \mathsf{R}^n$. A preordering[3] on $\mathcal{Q}_1$ is defined for $x, y \in \mathcal{Q}_1 \subset \mathsf{R}^n$ by

$$x \prec y \quad \text{iff} \quad \sum_{i=1}^{k} x_{\lfloor i \rfloor} \leq \sum_{i=1}^{k} y_{\lfloor i \rfloor}\,, \quad \sum_{i=1}^{n} x_{\lfloor i \rfloor} = \sum_{i=1}^{n} y_{\lfloor i \rfloor} \tag{2}$$

where $x_{\lfloor 1 \rfloor} \geq \cdots \geq x_{\lfloor n \rfloor}$ denotes the *decreasing rearrangement*[4] of the elements of $x$. (E.g., $x_{\lfloor 1 \rfloor} = \max_i x[i]$, $x_{\lfloor n \rfloor} = \min_i x[i]$, etc.) If $x \prec y$, we say that $y$ majorizes $x$, or that $x$ is majorized by $y$. If, as in [81], we denote the sequence of partial sums in (2) as

$$S_x[k] = \sum_{i=1}^{k} x_{\lfloor i \rfloor}$$

then the basic definition (2) can be stated as follows.

**Definition 1** (Majorization of $x$ by $y$)

$$x \prec y \quad \text{iff} \quad S_x[k] \leq S_y[k]\,, \quad S_x[n] = S_y[n]\,. \tag{3}$$

[3] A partial ordering on a space obeys the properties of reflexivity, transitivity, and antisymmetry [3]. A preordering has only the properties of reflexivity and transitivity. A total ordering is a partial ordering for which every element of the space can be ordered with respect to any other element. If we agree to identify vectors modulo rearrangements of their elements, then majorization defines a partial ordering on $\mathcal{Q}_1$.

[4] Many authors use *increasing* rearrangements, $x^{\lceil 1 \rceil} \leq \cdots \leq x^{\lceil n \rceil}$, so some care in reading the literature is required. Also the physics community *reverses* the symbol definition in (2) to read $y \prec x$, calling $x$ "more mixed" than $y$ [2].

Frequently $S_x[n]$ is normalized to one, $S_x[n] = 1$.

A plot of $S_x[k]$ versus $k$ is known as a *Lorentz curve* [48], $\mathcal{L}_x$, and $x \prec y$ iff $\mathcal{L}_y$ is everywhere above the curve $\mathcal{L}_x$ (Figure 1). When $x \prec y$, the curve $\mathcal{L}_x$ graphically shows greater equality, or diversity, for the values of the elements of $x \in \mathcal{Q}_1$ than is the case for $\mathcal{L}_y$. The elements of $y$ are more concentrated in value, or less diverse, than the elements of $x$. In Figure 1, the curve $\mathcal{L}_E$ corresponds to the vector with maximum equality or diversity, viz. the vector with all components having equal value, while the curve $\mathcal{L}_I$ is the curve of minimum diversity, or maximum concentration, associated with a vector having only one nonzero element. This graphical representation explains why majorization is also known as the Lorentz order. Lorentz curves that intersect correspond to vectors that cannot be ordered by majorization.

Doubly stochastic matrices play an important role in majorization theory. A real $n \times n$ matrix $M$ is said to be *doubly stochastic* if all entries are nonnegative and each column and each row sum to one. It is well known (*Birkhoff's Theorem*) that every doubly stochastic matrix $M$ is the convex combination of permutation matrices $P_i$, $M = \sum_i \alpha_i P_i$, $\sum_i \alpha_i = 1$, $\alpha_j \geq 0$ [13, 53, 2, 4]. Let $x = My$ for a doubly stochastic matrix $M$. Then $x$ is a convex sum of permutations (rearrangements) of $y$, so that $x$ is seen to be a smoothed or averaged version of $y$. The following theorem says that smoothing $y$ in this manner results in greater diversity in the sense of the Lorentz ordering.

**Theorem 1** (Smoothing and Majorization [2, 4, 53]) *Let $x, y \in \mathcal{Q}_1$. Then $x \prec y$ iff $x = My$ for some doubly stochastic matrix $M$.*

When $x \prec y$, we say that $x$ is less concentrated (more diverse) than $y$ or, equivalently, that $y$ is more concentrated (less diverse) than $x$.

It is natural to ask which functions from $\mathsf{R}^n$ to $\mathsf{R}$ preserve majorization. By definition, these functions belong to the class of *Schur-concave* functions.

**Definition 2** (Permutation Invariance) *A function $\phi(\cdot)$ is called* permutation invariant *iff it is invariant with respect to all permutations of its argument $x$, i.e. $\phi(x) = \phi(Px)$ for any permutation matrix $P$.*

**Definition 3** (Schur-Concavity) *A function $\phi(\cdot) : \mathsf{R}^n \to \mathsf{R}$ is said to be* Schur-concave *if $\phi(x) \geq \phi(y)$ whenever $x \prec y$, and* strictly *Schur-concave if in addition $\phi(x) > \phi(y)$ when $x$ is not a permutation of $y$.*

A Schur-concave function *must be invariant* with respect to permutations of the elements of the vector $x$.

A reasonable necessary condition for a function to be a good measure of diversity is that it preserve the Lorentz ordering, i.e., be Schur-concave. Thus, the class of Schur-concave functions are candidates for measures of diversity. Not surprisingly, Schur-concave functions have been extensively studied as measures of economic equality/concentration [6, 75, 23, 74, 15, 30, 32, 58], ecological diversity [66, 67, 77, 64], and ergodic mixing [2]. For $\phi(\cdot)$ Schur-concave, it is natural to consider $x$ to be more diverse, or less concentrated, than $y$ if $\phi(x) \geq \phi(y)$ [53, 64, 5, 58]. A reasonable approach to sparse basis selection might then be based on minimizing diversity, as measured by a Schur-concave function $\phi(\cdot)$, subject to the constraint (1).

The following two theorems are useful for identifying Schur-concave functions.

**Theorem 2** (Derivative Test for Schur-Concavity [53, 5]) *A function $\phi(\cdot)$ is Schur-concave on $\mathcal{Q}_1$ iff it is permutation invariant and satisfies the* Schur condition*,*

$$(x[i]-x[j])\left(\frac{\partial\phi(x)}{\partial x[i]}-\frac{\partial\phi(x)}{\partial x[j]}\right)\leq 0\,,\quad \forall x\in\mathcal{Q}_1\,,\quad \forall i,j=1,\cdots,n\,. \tag{4}$$

*Furthermore, because of the assumed permutation invariance of $\phi(x)$, one only need verify (4) for a single set of specific values for the pair $(i,j)$.*

We now extend this result to show that Schur-Concavity is preserved even when the variable is normalized using the 1-norm. This will be found to be valuable when we examine different variants of a diversity measure in Section 3.

**Theorem 3** (1-Normalization Preserves Schur-Concavity) *If $\phi(\cdot)$ is Schur-concave on the interior of $\mathcal{Q}_1$, then the scale invariant function $\psi$ defined by $\psi(x)=\phi(x/\|x\|_1)$ is also Schur-concave on the interior of $\mathcal{Q}_1$.*

**Proof.** To prove Theorem 3, let $\tilde{x}=x/\|x\|_1$ and note that on the interior of $\mathcal{Q}_1$,

$$\frac{\partial\psi(x)}{\partial x[i]}=\sum_k\frac{\partial\phi(\tilde{x})}{\partial\tilde{x}[k]}\frac{\partial\tilde{x}[k]}{\partial x[i]}=\frac{1}{\|x\|_1}\left(\frac{\partial\phi(\tilde{x})}{\partial\tilde{x}[i]}-\sum_k\tilde{x}[k]\frac{\partial\phi(\tilde{x})}{\partial\tilde{x}[k]}\right)\,.$$

This follows from the fact that on the positive orthant, $x\in\mathcal{Q}_1$,

$$\frac{\partial\tilde{x}[k]}{\partial x[i]}=\frac{1}{\|x\|_1}\left(\delta_{k,i}-\tilde{x}[k]\right)\,.$$

Therefore on the interior of $\mathcal{Q}_1$,

$$(x[i]-x[j])\left(\frac{\partial\psi(x)}{\partial x[i]}-\frac{\partial\psi(x)}{\partial x[j]}\right)=(\tilde{x}[i]-\tilde{x}[j])\left(\frac{\partial\phi(\tilde{x})}{\partial\tilde{x}[i]}-\frac{\partial\phi(\tilde{x})}{\partial\tilde{x}[j]}\right)\,.$$

Theorem 3 now follows directly from Theorem 2. ■

We can ask if there are subclasses of the set of Schur-concave functions that have especially desirable properties as measures of diversity. In the next subsection we show that concave functions provide a set of good concentration functions in that minimizing a concave diversity measure is guaranteed to result in a sparse solution.

## 2.2 Concave Functions as Measures of Diversity

A subset $\mathcal{C}$ of $\mathsf{R}^n$ is said to be *convex* iff for all $0 \leq \lambda \leq 1$ and for all $x, y \in \mathcal{C}$ one has $(1-\lambda)x + \lambda y \in \mathcal{C}$. Note that every orthant, including the positive orthant $\mathcal{Q}_1$, is convex. A function $\phi(\cdot) : \mathsf{R}^n \to \mathsf{R}$ is *concave* on the convex subset $\mathcal{C}$ iff $\phi((1-\lambda)x + \lambda y) \geq (1-\lambda)\phi(x) + \lambda\phi(y)$ and *strictly concave* if strict inequality holds for $\lambda \in (0,1)$ and $x \neq y$. A function concave over $\mathcal{C}$ is continuous on the interior of $\mathcal{C}$ [72]. Recall that the function $\phi(\cdot)$ is called permutation invariant if it is invariant with respect to permutations (rearrangements) of its argument, i.e., if $\phi(x) = \phi(Px)$ for all $P =$ permutation matrix. The set $\mathcal{C}$ is said to be *permutation symmetric* iff $x \in \mathcal{C}$ implies $Px \in \mathcal{C}$ for every permutation matrix $P$.

**Theorem 4** (Permutation Invariant, Concave Functions are Schur-Concave [53, 2]) *Let $x, y$ belong to a permutation symmetric, convex set $\mathcal{C} \subset \mathsf{R}^n$. Then $x \prec y$ iff $\phi(x) \geq \phi(y)$ for all permutation invariant and concave functions $\phi(\cdot) : \mathcal{C} \to \mathsf{R}$.*

A particularly useful and tractable set of diversity measures is provided by the subclass of separable concave functions.

**Definition 4** *A function $\phi(\cdot) : \mathsf{R}^n \to \mathsf{R}$ is* separable *if there exists a scalar function $g(\cdot) : \mathsf{R} \to \mathsf{R}$ such that $\phi(x) = \sum_{i=1}^n g(x[i])$.*

**Theorem 5** (Separable, Concave Functions are Schur-Concave [53, 2]) *Let $x, y$ belong to a permutation symmetric, convex set $\mathcal{C} \subset \mathsf{R}^n$. Then $x \prec y$ iff $\sum_{i=1}^n g(x[i]) \geq \sum_{i=1}^n g(y[i])$ for every concave function $g : \mathcal{C} \to \mathsf{R}$.*

**Proof of Theorems 4 and 5** (Partial). Theorems 4 and 5 are proved by first noting that the necessity of Theorem 4 implies the necessity of Theorem 5 while sufficiency of Theorem 5 implies sufficiency of Theorem 4, since a separable, concave function is permutation invariant and concave. To show necessity of Theorem 4, let $x \prec y$ and let $\phi(\cdot)$ be permutation invariant and concave. Then, based on our discussion of the previous subsection, for some doubly stochastic matrix $M = \sum_i \lambda_i P_i$, $\sum_i \lambda_i = 1$, with $\lambda_j \geq 0$ and each $P_i$ a permutation matrix, we have $\phi(x) = \phi(My) = \phi(\sum_i \lambda_i P_i y) \geq \sum_i \lambda_i \phi(P_i y) = \sum_i \lambda_i \phi(y) = \phi(y)$. Sufficiency of Theorem 5 is proved in references [53, 2]. ■

The utility of certain concave measures of diversity for obtaining sparse solutions to the best basis problem comes from their highly desirable property of attaining their minima on the boundary of their convex domain of definition.

**Theorem 6** (Optimality of Boundary Points) *Let $\phi(\cdot) : \mathcal{C} \to \mathsf{R}$ be strictly concave and bounded from below on a closed convex set $\mathcal{C} \subset \mathsf{R}^n$. The function $\phi(\cdot)$ attains its local minima (and hence its global minima) at boundary points of $\mathcal{C}$.*

**Proof.** If $x^*$ is assumed to yield a local minima in the interior of $\mathcal{C}$, then there exist interior points $y, z$ arbitrarily close to $x^*$ and $\lambda > 0$ such that $x^* = (1-\lambda)y + \lambda z$ yielding $\phi(x^*) > (1-\lambda)\phi(y) + \lambda\phi(z) \geq \min\{\phi(y), \phi(z)\}$, contradicting the putative local optimality of $x^*$. Thus $x^*$ must be on the boundary of $\mathcal{C}$. ■

If $\mathcal{C}$ is closed and convex, its boundary contains the *extreme points* of $\mathcal{C}$, which are those points that cannot be written as a convex combination of any other points contained in $\mathcal{C}$ [73]. The following theorem holds when $\phi(\cdot)$ is concave, but not necessarily strictly concave.

**Theorem 7** (Optimality of Extreme Points [73, 72]) *Let $\phi(\cdot) : \mathcal{C} \to \mathsf{R}$ be concave on a closed convex set $\mathcal{C} \subset \mathsf{R}^n$ which contains no lines. If $\phi(\cdot)$ attains a global minimum somewhere on $\mathcal{C}$, it is also attained at an extreme point of $\mathcal{C}$.*

Theorems 4–7 show that permutation invariant concave diversity measures are Schur-concave and that if the boundary points correspond to sparse solutions, then minimizing such measures can yield sparse solutions to the best basis selection problem. This is indeed possible by choosing a diversity measure from a proper subclass of functions which are also *sign invariant.*

**Definition 5** (Sign Invariance) *A function $\phi(\cdot)$ is said to be* sign invariant *if $\phi(x) = \phi(\bar{x}), \forall x \in \mathsf{R}^n$, where $\bar{x} = |x| \in \mathcal{Q}_1$ .*

Recall that the concept of a *basic solution* to the system (1) was described earlier in the introduction.

**Theorem 8** (Global Optimality of a Basic Solution) *Let $\phi(\cdot) : \mathsf{R}^n \to \mathsf{R}$ be permutation invariant, sign invariant, and concave on the positive orthant $\mathcal{Q}_1$. Then the global minimum of $\phi(x)$ subject to the linear constraints of (1) is attained at a basic solution.*

**Proof.** Because $\phi(x)$ is sign invariant and concave on $\mathcal{Q}_1$, it is concave on each of the orthants $\mathcal{Q}_l, 1 \leq l \leq 2^n$. This does not mean however that $\phi(x)$ is concave on $\mathsf{R}^n$ (cf. Figure 4). Recalling that the solution set to (1) is the linear variety $LV(A, b) = \{y : Ay = b\}$, the minimization of $\phi(x)$ over $LV(A, b)$ can be rewritten as

$$\min_{LV(A,b)} \phi(x) = \min_{1 \leq l \leq 2^n} \left( \min_{\mathcal{Q}_l \cap LV(A,b)} \phi(x) \right) . \tag{5}$$

Since $\mathcal{Q}_l \cap LV(A, b)$ is a convex set, and $\phi(x)$ is concave on $\mathcal{Q}_l$, from theorem 7 the solution to $\min_{\mathcal{Q}_l \cap LV(A,b)} \phi(x)$ is attained at a basic solution, or degenerate basic solution, or (1) because the subset of the basic solutions/degenerate basic solutions are the extreme points of $\mathcal{Q}_l \cap LV(A, b)$. Therefore from (5), the global minimum is attained at a basic or degenerate basic solution. Furthermore, the local minima are at basic or degenerate basic solutions. ■

Theorems 4–8 show that the permutation and sign invariant concave functions are particularly good measures of diversity, if our intent is to obtain sparse solutions to (1) by minimizing diversity measures. It is of interest, then, to be able to identify concave functions. We state two theorems relating to concave functions which we have found useful in this regard.

**Theorem 9** (Derivative Test for Concavity [73, 72, 8, 50, 51]) *Let $\mathcal{C} \subset \mathsf{R}^n$ be an open convex set and let $\phi(\cdot) : \mathcal{C} \to \mathsf{R}$ be differentiable on $\mathcal{C}$. Then $\phi(\cdot)$ is concave on $\mathcal{C}$ iff for any $x \in \mathcal{C}$ we have*

$$\nabla\phi(x)^T(y - x) \geq \phi(x) - \phi(y)\,, \quad \forall y \in \mathcal{C}\,. \tag{6}$$

*Furthermore $\phi(\cdot)$ is strictly concave iff the inequality is strict for every $y \neq x$.*

**Theorem 10** (Hessian Test for Concavity [73, 72, 8, 50, 51]) *Let $\mathcal{C} \subset \mathsf{R}^n$ be an open convex set and let $\phi(\cdot) : \mathcal{C} \to \mathsf{R}$ be twice differentiable on $\mathcal{C}$. Let $H(x)$ denote the Hessian matrix of second partial derivatives of $\phi(\cdot)$ evaluated at the point $x \in \mathcal{C}$. The function $\phi(\cdot)$ is concave on $\mathcal{C}$ iff for any $x \in \mathcal{C}$ $H(x)$ is negative semidefinite. Furthermore $\phi(\cdot)$ is strictly concave on $\mathcal{C}$ if $H(x)$ is negative definite for all $x \in \mathcal{C}$.*

## 2.3 Almost Concave Functions

We will latter consider diversity measures that are globally Schur-concave, and hence preserve the Lorentz order, on the $n$-dimensional positive orthant $\mathcal{Q}_1 \subset \mathsf{R}^n$, and which are concave on a (possibly position dependent) $(n-1)$-dimensional subspace in an open neighborhood of every point in the interior of $\mathcal{Q}_1$. Given a randomly initialized recursive algorithm to minimize such a measure subject to (1), for $n \geq 2$ the algorithm generically will converge to a sparse solution located on the boundary of $\mathcal{Q}_1$.

**Definition 6** (Almost Concave Function) *Let $\mathcal{C} \subset \mathsf{R}^n$ be an open convex set and the function $\phi(\cdot) : \mathcal{C} \to \mathsf{R}$ be Schur-concave on $\mathcal{C}$. The function $\phi(\cdot)$ is said to be* Almost Concave *(respectively,* Almost Strictly Concave*) on the interior of $\mathcal{C}$ if, for every point $x$ in the interior of $\mathcal{C}$, in a open neighborhood of $x$ it is concave (respectively, strictly concave) on an $(n-1)$-dimensional subspace with origin located at $x$.*

If $\mathrm{B}(x) \subset \mathcal{C}$ is an open ball centered at $x \in \mathcal{C}$ and there exists an $(n-1)$-dimensional linear variety $\mathcal{V}_x$ containing $x$ for which either the derivative test (Theorem 9) or the Hessian test (Theorem 10) holds for all $y \in \mathrm{B}(x) \cap \mathcal{V}_x$, then $\phi(\cdot)$ is concave on $\mathrm{B}(x) \cap \mathcal{V}_x$. If this is true for all $x$ in the interior of $\mathcal{C}$, then $\phi(\cdot)$ is almost concave on the interior of $\mathcal{C}$. Obviously, every concave function is almost concave. Also, note that an almost concave function must be Schur-concave by definition.

# 3 SCALAR MEASURES OF DIVERSITY

A general diversity measure is denoted by $d(\cdot) : \mathsf{R}^n \to \mathsf{R}$. The diversity measures considered in this paper are assumed to be both permutation invariant and sign invariant (as defined in Section 2). Note that the measures considered previously in [69]—the Shannon Entropy, the Gaussian Entropy, and the $\ell_{(p\leq 1)}$ concentration measure—satisfy this property. Motivated by Theorems 4–8, in this section we examine the properties of Schur-concavity and concavity on the positive orthant $\mathcal{Q}_1$. Because of the assumed sign invariance, $d(x) = d(|x|)$, Schur-concavity or concavity (or lack thereof) over $\mathcal{Q}_1$ corresponds to Schur-concavity or concavity (or lack thereof) respectively over any other orthant $\mathcal{Q}_l$. Note, however, that this does not guarantee Schur-concavity or concavity *across* orthants, and in general this property will not be true.

## 3.1 Signomial Measures

In this subsection, we present a general class of separable concave (and hence Schur-concave) functions that include as a special case the class of $\ell_{(p\leq 1)}$ diversity measures defined by

$$d_p(x) = \operatorname{sgn}(p) \sum_{i=1}^{n} |x[i]|^p, \quad p \leq 1, \tag{7}$$

and described in [81, 27, 78, 69].[5]

### 3.1.1 $\mathcal{S}$-functions

The generalization we are interested in is the subclass of *signomials* [10, 11] (also referred to as *generalized polynomials* [28, 72] or *algebraic functions* [73]) given by the separable function

$$\begin{aligned}
d_{\text{sig}}(x) &= \sum_{i=1}^{n} \mathsf{S}(|x[i]|) = \sum_{j=1}^{q} \omega_j \; d_{p_j}(x)\,, \\
d_{p_j}(x) &= \operatorname{sgn}(p_j) \sum_{i=1}^{n} |x[i]|^{p_j}, \;\; p_j \leq 1\,, \\
\mathsf{S}(s) &= \operatorname{sgn}(p_1)\,\omega_1\, s^{p_1} + \cdots + \operatorname{sgn}(p_q)\,\omega_q\, s^{p_q}\,, \quad s \in \mathsf{R}_+\,, \\
\text{where} \quad p_j &< 1\,, \quad p_j \neq 0\,, \quad \text{and} \quad \omega_j \geq 0\,, \\
\text{or} \quad p_j &= 0, 1\,, \quad \text{and} \quad \omega_j \in \mathsf{R}\,.
\end{aligned} \tag{8}$$

Note that, unlike a regular polynomial, $\mathsf{S}(s)$ has fractional and possibly negative powers, $p_j \leq 1$. A general signomial, unlike the subclass defined by (8), has no constraints on its powers or coefficients. If the powers $p_k$ are also constrained to be positive, the coefficients in (8) are all positive and $d_{\text{sig}}(x)$ is then a special case of a posynomial [28, 82, 10], a signomial

[5] As discussed later in this subsection, reference [81] first normalizes $x$ to one in the 2-norm sense.

constrained to have positive coefficients. With no loss of generality, in (8) we can take $\sum_j \omega_j = 1$. Henceforth, we will refer to functions of the form (8) as *$\mathcal{S}$-functions*.

It is readily shown that $d_{\text{sig}}(x)$ has a diagonal, negative semidefinite Hessian for $x \in \mathcal{Q}_1$. Therefore, from Theorem 10 we know that $d_{\text{sig}}(x)$ is concave on the interior of the positive orthant $\mathcal{Q}_1 \subset \mathsf{R}^n$. Furthermore, if there exists $j$ such that $p_j < 1$, $p_j \neq 0$, then the Hessian is negative definite and $d_{\text{sig}}(x)$ is strictly concave on the interior of the positive orthant $\mathcal{Q}_1$. By construction, $d_{\text{sig}}(x)$ is separable. Thus $d_{\text{sig}}(x)$ is Schur-concave and can be designed to be strictly concave, thereby insuring that a sparse solution can be obtained to the sparse basis selection problem by searching for a minimum of the function $d_{\text{sig}}(x)$. Summarizing our results, we have the following theorems.

**Theorem 11** ($\mathcal{S}$-Functions are Schur-Concave) *Let $x, y$ belong to a symmetric, convex set $\mathcal{C} \subset \mathcal{Q}_1$. Then $x \prec y$ only if $d_{\text{sig}}(x) \geq d_{\text{sig}}(y)$ for every $\mathcal{S}$-function $d_{\text{sig}} : \mathcal{C} \to \mathsf{R}$.*

**Theorem 12** ($\mathcal{S}$-Functions are Concave) *Every $\mathcal{S}$-function $d_{\text{sig}}$ is concave on the interior of $\mathcal{Q}_1$. Furthermore, any $\mathcal{S}$-function such that there exists $j$ for which $p_j < 1$, $p_j \neq 0$, is strictly concave on the interior of $\mathcal{Q}_1$.*

It can be shown that for $p > 1$, $d_p(x)$ of (7) is *not* Schur-concave and hence not concave. Indeed, it is well known (and readily demonstrated) that $d_p(x)$ is *convex* over $\mathcal{Q}_1$ (and a metric for the space $\mathsf{R}^n$) for $p > 1$. Figure 2 shows graphs and contour plots of $d_p(x)$ evaluated on the positive quandrant of $\mathsf{R}^2$ for $p = 1.0$ and $p = 0.1$, values of $p$ for which $d_p(x)$ is a well-defined diversity measure.

### 3.1.2 1-Normalized $\mathcal{S}$-functions

From the class of $\mathcal{S}$-functions, one can define the *1-norm normalized $\mathcal{S}$-functions* by taking

$$d_{\text{sig}}^{(1)}(x) \triangleq d_{\text{sig}}(\tilde{x}), \quad \tilde{x} = |x|/\|x\|_1 .$$

Note that $d_{\text{sig}}^{(1)}(x)$ is not separable. Of particular interest is the subclass of 1-normalized $p$-norm-like diversity measures obtained from (7), $d_p^{(1)}(x) \triangleq d_p(\tilde{x})$. Note that

$$d_{\text{sig}}^{(1)}(x) = \sum_{j=1}^{q} \omega_j \, d_{p_j}^{(1)}(x) = \sum_{j=1}^{q} \omega_j \, d_{p_j}(\tilde{x}), \quad \tilde{x} = |x|/\|x\|_1, \quad p_j \leq 1 .$$

It is an immediate consequence of Theorem 3 that the 1-normalized $\mathcal{S}$-functions are Schur-concave on the interior of $\mathcal{Q}_1$:

**Theorem 13** (1-Normalized $\mathcal{S}$-Functions are Schur-Concave) *Let $x, y$ belong to a permutation symmetric, convex set $\mathcal{C} \subset \mathcal{Q}_1$. Then $x \prec y$ only if $d_{\text{sig}}^{(1)}(x) \geq d_{\text{sig}}(y)$ for every 1-normalized $\mathcal{S}$-function $d_{\text{sig}}^{(1)} : \mathcal{C} \to \mathsf{R}$.*

A closer examination of the 1-normalized $\mathcal{S}$-functions reveal that a stronger result is possible. They can be shown to be almost concave.

**Theorem 14** (1-Normalized $\mathcal{S}$-Functions are Almost Concave) *The 1-normalized $\mathcal{S}$-function $d^{(1)}_{\text{sig}}(x) = d_{\text{sig}}(\tilde{x})$, $\tilde{x} = |x|/\|x\|_1$, is* Almost Concave *on the interior of $\mathcal{Q}_1$. If in addition $p_j < 1$ for at least one $p_j$, then $d^{(1)}_{\text{sig}}(x)$ is* Almost Strictly Concave *on the interior of $\mathcal{Q}_1$.*

**Proof.** From Theorem 10, the function $d^{(1)}_{\text{sig}}(x)$ is concave on $\mathcal{Q}_1$ iff the Hessian $\mathcal{H}^{(1)}_{\text{sig}}(x)$ is negative semidefinite on $\mathcal{Q}_1$, where

$$\mathcal{H}^{(1)}_{\text{sig}}(x) = \frac{\partial^2}{\partial x^2} d^{(1)}_{\text{sig}}(x) = \sum_{j=1}^{q} |p_j|\, \omega_j\, \mathcal{H}^{(1)}_{p_j}(x) = \sum_{j=1}^{q} |p_j|\, \omega_j\, \mathcal{H}_{p_j}(\tilde{x})\,, \quad \tilde{x} = \frac{|x|}{\|x\|_1}\,, \quad x \in \mathcal{Q}_1\,.$$

The Hessian $\mathcal{H}^{(1)}_{\text{sig}}(x)$ will be negative semidefinite if $\mathcal{H}^{(1)}_{p_j}(x)$, the Hessian of $d^{(1)}_{p_j}(x)$, is negative semidefinite on $\mathcal{Q}_1$ for each $j$. Furthermore, $d^{(1)}_{\text{sig}}(x)$ is strictly concave if in addition $\mathcal{H}^{(1)}_{p_j}(x)$ is negative definite for at least one $j$. It is shown in the appendix that the Hessian $\mathcal{H}^{(1)}_{p}(x)$ for $x \in \mathcal{Q}_1$ is given by

$$\begin{aligned} \mathcal{H}^{(1)}_{p}(x) &= \frac{\partial^2}{\partial x^2} d^{(1)}_{p}(x) \qquad (9) \\ &= -\frac{|p|}{\|x\|_1^2} \sum_{i=1}^{n} \left\{ \frac{p}{\tilde{x}[i]^{1-p}} \left( \mathbf{e}_i \mathbf{1}^T + \mathbf{1} \mathbf{e}_i^T \right) + \frac{1-p}{\tilde{x}[i]^{2-p}}\, \mathbf{e}_i \mathbf{e}_i^T - (1+p)\, \tilde{x}[i]^p\, \mathbf{1}\mathbf{1}^T \right\}. \end{aligned}$$

Let $\mathbf{1}^\perp \subset \mathsf{R}^n$ be the $n-1$ dimensional subspace of vectors perpendicular to the vector $\mathbf{1}$. It is straightforward to show that $y^T \mathcal{H}^{(1)}_p(x) y \le 0$, for all nonzero $y \in \mathbf{1}^\perp \subset \mathsf{R}^n$ and all $x$ in the interior of $\mathcal{Q}_1$ when $p \le 1$, with $y^T \mathcal{H}^{(1)}_p(x) y < 0$ for $p < 1$, proving the theorem. A key factor that makes this manipulation possible is the structured form in which the Hessian has been expressed. The details of the proof are given in the appendix. ∎

### 3.1.3 2-Normalized $\mathcal{S}$-functions

As proposed in [81], one can also form the *2-norm normalized* $p$-norm-like diversity measures obtained from (7),

$$d^{(2)}_p(x) \triangleq d_p(\tilde{x})\,, \quad \tilde{x} = |x|^2/\|x\|_2^2\,.$$

The measures $d^{(2)}_p(x)$ are a special case of the *2-normalized* $\mathcal{S}$-functions formed from the class of $\mathcal{S}$-functions by taking

$$d^{(2)}_{\text{sig}}(x) \triangleq d_{\text{sig}}(\tilde{x}) = \sum_{j=1}^{n} \omega_j d^{(2)}_{p_j}(x) = \sum_{j=1}^{n} \omega_j d_{p_j}(\tilde{x})\,, \quad \tilde{x} = |x|^2/\|x\|_2^2\,.$$

Compared to the 1-Normalized $\mathcal{S}$-Functions, a greater restriction on the range of $p$ has to be placed the 2-Normalized $\mathcal{S}$-Functions if Schur-concavity is to be preserved. It is found that $p \le \frac{1}{2}$ leads to desirable properties.

**Theorem 15** (2-Normalized $\mathcal{S}$-Functions are Schur-Concave for $p \leq \frac{1}{2}$) *Let $x, y$ belong to a symmetric, convex set $\mathcal{C} \subset \mathcal{Q}_1$. Then $x \prec y$ implies that $d^{(2)}_{\rm sig}(x) \geq d_{\rm sig}(y)$ for every 2-normalized $\mathcal{S}$-function $d^{(2)}_{\rm sig} : \mathcal{C} \to \mathsf{R}$ with $p \leq 1/2$. Furthermore, for $p > \frac{1}{2}$, $d^{(2)}_{\rm sig}(x)$ is* not Schur-concave *(and, therefore, not concave) over the interior of $\mathcal{Q}_1$.*

**Proof.** Because $d^{(2)}_{\rm sig}(x) = \sum_{j=1}^{n} \omega_j d^{(2)}_{p_j}(x)$, it is enough to show that the theorem is true for the simpler case of $d^{(2)}_p(x)$. This is done by making use of Theorem 2 and equation (4). The details are given in the appendix. ■

Actually, for $p \leq 1/2$, even more can be said:

**Theorem 16** (2-Normalized $\mathcal{S}$-Functions are Almost Strictly Concave for $p \leq \frac{1}{2}$) *Let $p_j \leq \frac{1}{2}$ for every $p_j$ in (8). Then the 2-normalized $\mathcal{S}$-function $d^{(2)}_{\rm sig}(x) = d_{\rm sig}(\tilde{x})$, $\tilde{x} = |x|^2/\|x\|_2^2$, is* Almost Strictly Concave *on the interior of $\mathcal{Q}_1$.*

**Proof.** The function $d^{(2)}_{\rm sig}(x)$ is strictly concave on $\mathcal{Q}_1$ iff the Hessian $\mathcal{H}^{(2)}_{\rm sig}(x)$ is negative definite on $\mathcal{Q}_1$, where

$$\mathcal{H}^{(2)}_{\rm sig}(x) = \frac{\partial^2}{\partial x^2} d^{(2)}_{\rm sig}(x) = \sum_{j=1}^{q} |p_j| \, \omega_j \, \mathcal{H}^{(2)}_{p_j}(x) \, , \quad \tilde{x} = \frac{|x|^2}{\|x\|_2^2} \, , \quad x \in \mathcal{Q}_1 \, .$$

$\mathcal{H}^{(2)}_{\rm sig}(x)$ is negative definite if $\mathcal{H}^{(2)}_{p_j}(x)$ is negative definite on $\mathcal{Q}_1$ for every $j$. It is shown in the appendix that the Hessian $\mathcal{H}^{(2)}_p(x)$ over $\mathcal{Q}_1 \subset \mathsf{R}^n$ is given by

$$\begin{aligned} \mathcal{H}^{(2)}_p(x) &= \frac{\partial^2}{\partial x^2} d^{(2)}_p(x) \\ &= -\frac{2|p|}{\|x\|_2^2} \sum_{i=1}^{n} \left\{ \frac{2p}{\tilde{x}[i]^{\frac{1}{2}-p}} \left( \mathbf{e}_i \tilde{x}^{\frac{T}{2}} + \tilde{x}^{\frac{1}{2}} \mathbf{e}_i^T \right) \right. \qquad (10) \\ &\quad \left. + \left( \frac{1-2p}{\tilde{x}[i]^{1-p}} + |d^{(2)}_p(x)| \right) \mathbf{e}_i \mathbf{e}_i^T - 2\tilde{x}[i]^p \, (1+p) \, \tilde{x}^{\frac{1}{2}} \tilde{x}^{\frac{T}{2}} \right\} . \end{aligned}$$

The remainder of the proof is very similar to that of Theorem 14. It is based on the readily verified fact that for each $x$ in the interior of $\mathcal{Q}_1$ the matrix $\mathcal{H}^{(2)}_p(x)$ is negative definite on the $(n-1)$–dimensional subspace of $\mathsf{R}^n$ perpendicular to the vector $\tilde{x}^{\frac{1}{2}}$. Again, the form in which the Hessian is expressed plays an important role in reducing the complexity of the proof. The details of the proof are given in the appendix. ■

## 3.2 Entropy Measures

There has been a large interest in utilizing entropy measures as measures of diversity or concentration in biology, economics, computer science, and physics [79, 40, 67, 2, 31, 22, 68] as well as for sparse basis selection [19, 81, 27, 78, 69]. Here, we discuss some of these measures at length and show the close relationship between the entropy measures and the $p$-norm-like measures discussed in the previous subsection.

### 3.2.1 Gaussian Entropy

References [81, 78] propose the use of the "logarithm of energy" function

$$H_G(x) = \sum_{i=1}^{n} \log |x[i]|^2 \tag{11}$$

as a measure of concentration. Reference [81] points out that this can be interpreted as the Shannon entropy of a Gauss-Markov process [21]; for this reason in this paper, and in reference [69], we refer to (11) as the *Gaussian entropy* measure of diversity.

Note that $H_G$ is separable. It is straightforward to demonstrate that the Hessian of $H_G$ is everywhere positive definite on the positive orthant $\mathcal{Q}_1$, showing that $H_G$ is strictly concave on the interior of $\mathcal{Q}_1$ and hence Schur-concave. The Gaussian entropy is therefore a good measure of concentration and we expect that minimizing $H_G$ will result in sparse solutions to the best basis selection problem.

In [69], an algorithm is presented to minimize $H_G$ that indeed shows very good performance in obtaining sparse solutions. It is also shown that the algorithm to minimize $H_G$ is the same as the algorithm that minimizes (7) for $p = 0$ and can therefore be given the interpretation of optimizing the numerosity ($p = 0$) measure described by [27]. The interpretation of $H_G$ as a $p = 0$ measure follows naturally from the literature on inequalities where Exp($H_G$)= $(\prod_i |x[i]|)^2$ is shown to be intimately related to the $p = 0$ norm (e.g., see [9], page 16; reference [57], page 74; or [38] page 15). In fact,

$$\text{Exp}(\frac{1}{2n} H_G(x)) = \lim_{p \to 0} \left( \frac{1}{n} d_p(x) \right)^{\frac{1}{p}} . \tag{12}$$

### 3.2.2 Shannon Entropy

References [19, 81, 27] have proposed the use of the Shannon entropy function as a measure of diversity appropriate for sparse basis selection. This entropy has also long been used by economists and ecologists as a measure of concentration and diversity [79, 40, 67, 2, 31, 22].

Given a probability distribution, the Shannon entropy is well defined. However starting from $x$, there is some freedom in how one precisely defines this measure. Defining the Shannon entropy function $H_S(\tilde{x})$ for $\tilde{x} = \tilde{x}(x)$ by

$$H_S(\tilde{x}) = -\sum_{i=1}^{n} \tilde{x}[i] \log \tilde{x}[i], \quad \tilde{x} \geq 0 \,, \tag{13}$$

the differences arise in how one defines $\tilde{x}$ as a function of $x$. These differences affect the properties of $H_S$ as a function of $x$. It is well known that $H_S(\tilde{x})$ defined as a function of $\tilde{x}$ by (13) is Schur-concave [53, 81]. However it is generally *not* the case that $H_S(x)$ is Schur-concave with respect to $x$ [53].

**Shannon Entropy $H_s^{abs}(x)$.** This corresponds to the choice $\tilde{x}[i] = |x[i]|$. It can be readily shown that $H_S^{abs}$ is Schur-concave with respect to $x$ over $\mathcal{Q}_1$. Even more, it can be shown by an application of Theorem 10 that $H_S^{abs}$ is strictly concave on the interior of $\mathcal{Q}_1$. Thus $H_S^{abs}$ is a good measure of diversity and minimizing $H_S^{abs}$ should result in sparse solutions to the best basis problem.

**2-Normalized Shannon Entropy $H_s^{(2)}(x)$.** This corresponds to the choice $\tilde{x}[i] = \frac{x[i]^2}{\|x\|_2^2}$ in (13) [81]. In reference [69] it is argued that the minima of $H_S^{(2)}$ do not correspond to completely sparse solutions to $Ax = b$, and this was demonstrated via simulation. This fact can now be verified utilizing the insights provided by the theory of majorization and Schur-concavity. By a straightforward (but tedious) application of Theorem 2, one can show that the separable function $H_S^{(2)}$ is *not* Schur-concave with respect to $x$ over the positive orthant $\mathcal{Q}_1$, and therefore from Theorem 5 is not concave over the interior of $\mathcal{Q}_1$. The lack of Schur-concavity for $H_S^{(2)}$ is also shown below in the discussion of the Renyi entropy-based sparsity measures. Thus minimizing $H_S^{(2)}$ with respect to $x$ will not generally yield a sparse solution.

This can be seen graphically in Figure 3 where the value of $H_S^{(2)}(x)$ is represented in the positive orthant of $\mathsf{R}^3$ by its height above the positive simplex along the ray from the origin through $x$. Because of the 2-normalization, all other vectors $x$ on this ray have the same value. In Figure 3, it is evident that the minima of $H_S^{(2)}$ occur just shy of the boundaries defined by the coordinate axes (where the sparse solutions reside). This explains why in [69] concentrated, but not completely sparse, solutions were obtained.

**1-Normalized Shannon Entropy $H_s^{(1)}(x)$.** This corresponds to the choice $\tilde{x}[i] = \frac{|x[i]|}{\|x\|_1}$, Applying Theorem 2 it can be shown that $H_S^{(1)}$ is Schur-concave over $\mathcal{Q}_1$. Alternatively, use of Theorem 3 shows that $H_S^{(1)}$ is Schur-concave. Because of the complexity of the Hessian of $H_S^{(1)}$, it is difficult to directly ascertain if $H_S^{(1)}$ is concave. However, below in the discussion of the Renyi entropy-based measures we show that $H_S^{(1)}$ is almost concave (but perhaps not strictly almost concave) over the interior of $\mathcal{Q}_1$. Figure 4 graphically represents $H_S^{(1)}$ in the positive orthant of $\mathsf{R}^3$ by its height above the positive simplex along the ray from the origin through $x$. Because of the 1-normalization, all other vectors $x$ on this ray have the same value.

### 3.2.3 Renyi Entropy

A family of entropies, parameterized by $p$, is described in [70, 71, 45]. These *Renyi entropies*[6] include, as a special case, the Shannon entropy. Given a "probability" $\tilde{x}$, $\tilde{x}[i] \geq 0$, $\sum_i \tilde{x}[i] = 1$, the Renyi entropy is defined for $0 \leq p$ by

$$H_p(\tilde{x}) = \frac{1}{1-p} \log \sum_{i=1}^{n} \tilde{x}[i]^p = \frac{1}{1-p} \log d_p(\tilde{x}) \,, \tag{14}$$

where

$$H_1(\tilde{x}) = \lim_{p \to 1} H_p(\tilde{x}) = -\sum_{i=1}^{n} \tilde{x} \log \tilde{x} = H_S(\tilde{x}) \,.$$

Thus $H_1(\tilde{x})$ is the Shannon entropy of $\tilde{x}$. Because log is monotonic, we see that for purposes of optimization $H_p(\tilde{x})$ is equivalent to the $p$-norm like measure $d_p(\tilde{x})$. Thus, consistent with the discussion given in [27], one can also reasonably refer to the normalized $p$-norm-like measures $\ell_{p \leq 1}$ as entropies.

We now see that the $p$-norm-like, $p \geq 0$, and the Shannon entropy measures are not unrelated, but are different manifestations of the Renyi entropy measures. As in the case of Shannon Entropy, different diversity measures can be obtained depending on how $\tilde{x}$ is defined as a function of $x$.

**1-Normalized Renyi Entropy.** Let $\tilde{x} = x/\|x\|_1$ and $d_p(\tilde{x}) = d_p^{(1)}(x)$ be the 1-normalized $p$-norm-like measure discussed in Section 3.1. The *1-Normalized Renyi entropy*, parameterized by $0 \leq p \leq 1$, is defined by

$$H_p^{(1)}(x) = \frac{1}{1-p} \log d_p^{(1)}(x) \,. \tag{15}$$

One can readily show that $H_p^{(1)}(x)$ for $0 < p < 1$ is almost strictly concave as a consequence of the almost strict concavity of $d_p^{(1)}(x)$ for $0 < p < 1$ (Theorem 14) and the fact that log is an increasing concave function.

In the limit $p \longrightarrow 1$, $H_p^{(1)}(x)$ is the 1-norm Shannon entropy evaluated for $\tilde{x} = x/\|x\|_1$, $H_1^{(1)}(x) = H_S^{(1)}(x)$. If we let $p \longrightarrow 1$ from below, $p < 1$, then $H_p^{(1)}(x)$ is almost concave for all $p$ as we take the limit showing, as claimed earlier, that $H_S^{(1)}(x) = \lim_{p \to 1^-} H_p^{(1)}(x)$ is almost concave.

From (15) and the fact that the measures $d_p^{(1)}(x)$ form a subclass of the measures $d_{\text{sig}}^{(1)}(x)$, the 1-Normalized $\mathcal{S}$-class of functions $d_{\text{sig}}^{(1)}(x)$ can be viewed as a generalization of

[6] The Renyi entropy parameter is commonly denoted by $\alpha$. Thus, this entropy is also referred to as the *$\alpha$-entropy.*

the 1-normalized Renyi entropies $H_p^{(1)}(x)$. With this insight, let us define a *generalized 1-normalized Renyi entropy* function by

$$H_{\text{sig}}^{(1)}(x) = \frac{1}{q - \sum_{j=1}^q p_j} \log d_{\text{sig}}^{(1)}(x) = \frac{1}{q - \sum_{j=1}^q p_j} \log \sum_{j=1}^q \omega_j \; d_{p_j}^{(1)}(x) \,, \tag{16}$$

where $q$, $0 < p_j \leq 1$ and $0 < \omega_j$ are defined in (8) and $\sum_j \omega_j = 1$. Note that (16) includes (15) as a special case and that, as defined, $H_{\text{sig}}^{(1)}(x)$ is less general than $d_{\text{sig}}^{(1)}(x)$ of (8) because of the restriction to positive $p_j$. As a consequence of Theorem 14 and the fact that log is an increasing concave function, it is straightforward to show that $H_{\text{sig}}^{(1)}(x)$ is almost concave on $\mathcal{Q}_1$.

**2-Normalized Renyi Entropy.** Here we take $\tilde{x} = |x|^2/\|x\|_2^2$. Then $d_p(\tilde{x}) = d_p^{(2)}(x)$ is the 2-normalized $p$-norm-like measure discussed in Section 3.1. The *2-normalized Renyi entropy*, parameterized by $0 < p \leq \frac{1}{2}$, is defined by

$$H_p^{(2)}(x) = \frac{1}{1-p} \log d_p^{(2)}(x) \,. \tag{17}$$

It is straightforward to show that $H_p^{(2)}(x)$ is almost strictly concave for $0 \leq p < \frac{1}{2}$ and almost concave for $p = \frac{1}{2}$ as a consequence of the concavity of $d_p^{(2)}(x)$ for $0 \leq p \leq \frac{1}{2}$ (Theorem 16) and the fact that log is an increasing concave function.

Because $d_p^{(2)}(x)$ is not Schur-concave for $p > \frac{1}{2}$, $H_p^{(2)}(x)$ is not Schur-concave (and hence not concave) for $p > \frac{1}{2}$. Note that in the limit $p \to 1$, $H_p^{(2)}(x)$ is the 2-norm Shannon entropy evaluated for $\tilde{x} = |x|^2/\|x\|_2^2$, $H_1^{(2)}(x) = H_S^{(2)}$, showing that $H_S^{(2)}$ is not Schur-concave, and hence not concave, as claimed earlier.

From (17) and the fact that the measures $d_p^{(2)}(x)$ form a subclass of the measures $d_{\text{sig}}^{(2)}(x)$, the 2-normalized $\mathcal{S}$-class of functions $d_{\text{sig}}^{(2)}(x)$ can be viewed as a generalization of the 2-normalized Renyi entropies $H_p^{(2)}(x)$. One can then reasonably define a *generalized 2-normalized Renyi entropy* function by

$$H_{\text{sig}}^{(2)}(x) = \frac{1}{q - \sum_{j=1}^q p_j} \log d_{\text{sig}}^{(2)}(x) = \frac{1}{q - \sum_{j=1}^q p_j} \log \sum_{j=1}^q \omega_j \; d_{p_j}^{(2)}(x) \,, \tag{18}$$

where $q$, $0 < p_j \leq \frac{1}{2}$ and $0 < \omega_j$ are defined in (8) and $\sum_j \omega_j = 1$.

## 3.3 Other Measures

One can proceed to provide a systematic development and analysis of a variety of possible diversity measures and their appropriateness in specific application domains. For example, for a general unnormalized positive vector $x \in \mathcal{Q}_1$, one could explore the possible use of the generalized Renyi entropy [70, 71],

$$H_p(x) = \frac{1}{1-p} \log \left\{ \frac{\sum_{i=1}^{n} x[i]^p}{\sum_{i=1}^{n} x[i]} \right\} , \quad x \geq 0 ,$$

and its connection to the $p$-norm-like diversity measure (7). Other possible entropies that can be analyzed include the Daroczy, Quadratic, and R-norm entropies [45].

More generally, by utilizing the tools and insights provided by majorization theory, one can systematically develop a variety of diversity measures. Toward this end, reference [53] has a wealth of results regarding tests for, and classes of, Schur-concave functions, and an examination of possible diversity measures, and their properties, can be found in references [30, 64, 32, 26]. As one last example of a possible extension, we mention the multinomial-like class of functions generated by the symmetric sums of product terms like $x[i_1]^{p_{j_1}} \cdots x[i_l]^{p_{j_l}}$, $l = 1, \cdots, n$ [53]. The measures presented in Section 3.1 correspond to the case $l = 1$.

# 4 ALGORITHMS FOR SPARSE BASIS SELECTION

Now we discuss some algorithms for minimizing the diversity measures presented in the previous section. The choice of provably well-behaved algorithms is limited because the diversity measures we are considering are concave, and therefore generally have many local minima on their boundaries [73, 63, 42, 41, 12]. A brute force exhaustive search of the extreme points is not practical because of their large number, namely $\binom{n}{m}$, as soon as the dimension of $x$ becomes reasonably large, although branch-and-bound search methods might be applicable in this regard [59, 56, 34].

To minimize the general classes of concave diversity measures developed in this paper, we will extend the affine scaling methodology described in [69] and develop iterative algorithms which converge to a basic or degenerate basic solution of (1). Toward this end, for each possible diversity measure $d(x)$, we will first need to define an associated *scaling matrix* which naturally arises in this formulation, and identify its properties. Then the affine scaling-related optimization methodology of [69] can be used to develop algorithms for minimizing the diversity measure.

Convergence of these algorithms is established under various sets of assumptions, first starting from minimization of general concave diversity measures on $\mathcal{Q}_1$, with relatively severe restrictions on the algorithms, and ending with the minimization of concave diversity measures having positive definitive scaling matrices, which allows minimal restrictions on the algorithm.

## 4.1 Gradient Factorization and The Scaling Matrix $\mathit{\Pi}(x)$

In the algorithms to be developed next, an expression for the gradient of the diversity measure $d(x)$ with respect to $x$ is required. The following factorization of the gradient turns out to be essential for the development of the algorithms of this paper,

$$\nabla d(x) = \alpha(x)\mathit{\Pi}(x)x\,, \tag{19}$$

where $\alpha(x)$ is a positive scalar function, and $\mathit{\Pi}(x)$ is the *Scaling Matrix*, which in this paper is always chosen to be *diagonal*. The scaling matrices for the different diversity measures described in the previous section play an important role and are tabulated in Table 1, along with their properties. The derivation of these scaling matrices is described in the appendix.

An important distinction amongst the diversity measures from an algorithmic point of view is whether their scaling matrix is positive definite or not. For diversity measures with positive definite scaling matrices, we have been able to develop simpler convergent algorithms. An examination of Table 1 shows that this includes the large class of measures provided by the signomial functions, $d_{\mathrm{sig}}(x)$.

Notice, however, that the diversity measures with scale invariance have scaling matrices that are *not* positive definite. For scale invariant diversity measures we have $d(x) = d(\gamma x)\ ,\quad \forall\gamma \in \mathsf{R}$, showing that the projection of the gradient $\nabla d(x)$ along the direction $x$ must be zero, $x^T\nabla d(x) = \alpha(x)x^T\Pi(x)x = 0$, which we summarize as

$$d(x) = d(\gamma x)\ ,\quad \forall\gamma \in \mathsf{R} \quad \Rightarrow \quad x^T\Pi(x)x = 0\,. \tag{20}$$

Under the assumption that $\Pi(x)$ is diagonal, the property (20) implies that the nonzero diagonal elements must change sign and consequently that $\Pi(x)$ is indefinite. Scale invariant measures described in this paper are the 1-normalized and 2-normalized $\mathcal{S}$-functions and Renyi entropies.

## 4.2 A Generalized Affine Scaling Algorithm

**Constrained Optimization.** The affine scaling methodology developed in [69] is readily adapted to address the minimization of the more general diversity measures developed in this paper. This algorithm attempts to solve the optimization problem,

$$\min_x d(x) \quad \text{subject to} \quad Ax = b\,. \tag{21}$$

The standard method of Lagrange multipliers is used, where the Lagrangian is given by

$$L(x,\lambda) = d(x) + \lambda^T(Ax - b)\,, \tag{22}$$

and $\lambda$ is the $m \times 1$ vector of Lagrange multipliers. A necessary condition for a minimizing solution $x^*$ to exist is that $(x^*, \lambda^*)$ be stationary points of the Lagrangian function, i.e.

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) &= \nabla d(x^*) + A^T \lambda^* = 0 \\ \nabla_\lambda L(x^*, \lambda^*) &= \quad Ax^* - b = 0\,. \end{aligned} \tag{23}$$

Using the factored form of the gradient given in (19), after some manipulation the stationary point $x^*$ can be shown to satisfy

$$x^* = \mathit{\Pi}^{-1}(x^*) A^T (A \mathit{\Pi}^{-1}(x^*) A^T)^{-1} b\,. \tag{24}$$

**An Iterative Algorithm.** The necessary condition (24) naturally suggests an iterative algorithm of the form,

$$x_{k+1} = \mathit{\Pi}^{-1}(x_k) A^T (A \mathit{\Pi}^{-1}(x_k) A^T)^{-1} b\,. \tag{25}$$

This algorithm has desirable properties when the scaling matrix $\mathit{\Pi}(x)$ is positive definite. As shown in [69], and discussed below, when $\mathit{\Pi}(x)$ is positive definite it can be used to naturally define an Affine Scaling Transformation (AST) matrix, $W(x) = \mathit{\Pi}^{-\frac{1}{2}}(x) > 0$, and thereby establish a strong connection to affine scaling methods used in optimization theory [24, 29, 60]. Hence the use of the terminology "Affine Scaling" in connection with the algorithms developed here and in [69].

The algorithm developed in [69] has an interesting interpretation as an extension of the standard affine scaling methodology which arises because in general $\mathit{\Pi}(x)$ is *not* always positive definite. In such cases a simple generalization will often still yield a provably convergent algorithm. This is achieved by defining an intermediate variable $x_k^r$ which is given by

$$x_{k+1}^r = \mathit{\Pi}^{-1}(x_k) A^T (A \mathit{\Pi}^{-1}(x_k) A^T)^{-1} b\,. \tag{26}$$

Note that $x_{k+1}^r$ is feasible. Assuming that $x_k$ is also feasible (i.e. $Ax_k = b$), then the increment $x_{k+1}^r - x_k$ is in the nullspace of $A$, and provides a direction along which $d(x)$ can be decreased while maintaining the equality constraints.

The value of $x$ in the next iteration is computed as

$$x_{k+1} = x_k + \mu_k (x_{k+1}^r - x_k) = \mu_k \, x_{k+1}^r + (1 - \mu_k)\, x_k\,, \tag{27}$$

where the step size $\mu_k$ is chosen to ensure that

$$\nabla d(x_k)^T \, (x_{k+1} - x_k) < 0\,, \tag{28}$$

along with a decrease in the diversity measure, $d(x_{k+1}) < d(x_k)$. Note that the choice of $\mu_k = 1$ yields the original, simpler algorithm (25).

**An AST Interpretation.** As discussed in [69], when the diagonal matrix $\mathit{\Pi}(x)$ is positive definite, the algorithm (25) has an interpretation as an Affine Scaling Transformation algorithm [24, 29]. Specifically, if we construct a symmetric affine scaling transformation (AST) $W(x)$ by

$$W(x) \triangleq \mathit{\Pi}^{\frac{1}{2}}(x)\,, \tag{29}$$

we can then form $W_{k+1}$ and, following the AST methodology [24, 29], the scaled quantities $q$ and $A_{k+1}$ by

$$W_{k+1} = W(x_k)\,, \quad x = W_{k+1}q\,, \quad A_{k+1} = AW_{k+1}\,,$$

assuming we have at hand a current estimated feasible solution, $x_k$ (corresponding to the scaled quantity $q_k = W_{k+1}^{-1}x_k$), to the problem (21). We then consider the optimization problem (21) in terms of the scaled variable $q$,

$$\min_q d_{k+1}(q) = d(W_{k+1}q) \quad \text{subject to} \quad A_{k_1}q = b\,.$$

Treating $q_k$ as the current estimate of the solution to this optimization problem we can obtain an updated estimated by projecting the gradient of $d_{k+1}(q_k)$ onto the nullspace of $A_{k+1}$ and moving in this direction an amount proportional to a stepsize given by $1/\alpha(x_k)$. This yields the algorithm

$$q_{k+1} = A_{k+1}^{+}b\,, \quad x_{k+1} = W_{k+1}q_{k+1}\,,$$

with $A_{k+1}^{+}$ the Moore-Penrose pseudoinverse of $A$, which is equivalent to (25). It is common in the affine scaling approach to use the specific AST $W(x)$ given by

$$W(x) = \operatorname{diag}\left(\frac{1}{x[i]^2}\right)\,,$$

which corresponds in (29) to defining $W(x)$ in terms of $\mathit{\Pi}(x) = \mathit{\Pi}_p(x)$, with $\mathit{\Pi}_p(x)$ given in Table 1, for $p = 0$. In contrast, the algorithm (25) corresponds to a "natural" choice of $W(x)$ dictated by the particular choice of the sparsity measure $d(x)$.

**Convergence and Numerical Issues.** The algorithm (26)–(27) can be also interpreted as a Lagrange Multiplier Update method ([50], page 436), as a Gradient Projection method [50, 14], as a Successive Approximate Lagrangian method [35], and as a Frank-Wolfe method[7] [33]. Independently of how it is derived or interpreted, convergence of the algorithm can be shown for specific classes of diversity measures, $d(x)$, using the general convergence theorems of Zangwill, and their variants [82, 8, 50, 55].

A key issue is the choice of the sign and magnitude of the scalar stepsize parameter $\mu_k$ in (27). The stepsize is chosen to obtain the inequality in (28), to ensure that $d(x_{k+1}) < d(x_k)$

[7] Also known as the conditional gradient method [14] or as a linear approximation method [82].

for $x_{k+1} \neq x_k$, and, if necessary, to ensure that $x_{k+1}$ remains in the same orthant as $x_k$ (which, without loss of generality, is taken to be the positive orthant $\mathcal{Q}_1$). The condition $d(x_{k+1}) < d(x_k)$ ensures that the algorithm results in a strict decrease of $d(x_k)$ at each iteration, resulting in convergence to a stationary point, $x^*$, of $d(x)$ in the linear variety $LV(A,b)$ (i.e., to a stationary point of (22)) satisfying condition (24). Such points must be either saddlepoints or local minima of $d(x)$ in $LV(A,b)$ and will be local minima if $d(x)$ is strictly concave, in which case the local minimum $x^*$ must be on the boundary of a quadrant $\mathcal{Q}_l$, ensuring a sparse solution.

Note the requirement in (26) that $\Pi(x_k)$ be invertible for every possible $x_k$. For the diversity measures discussed in this paper, this is either true or generically true (i.e., true except for a set of Lebesgue measure zero). In the latter case, for increased numerical robustness one may chose to numerically solve the system (cf. (23)),

$$\begin{bmatrix} \Pi(x_k) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^r_{k+1} \\ \lambda^r_{k+1} \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} . \tag{30}$$

This system has the solution $x^r_{k+1} = -\Pi^{-1}(x_k)A^T\lambda^r_{k+1}$, $\lambda^r_{k+1} = -(A\Pi^{-1}(x_k)A^T)^{-1}b$, which corresponds to (25). $\lambda^r_{k+1}$ can be interpreted as an estimate for the Lagrange multiplier of the Lagrangian (22).

With this background, we can now proceed to the development of convergent algorithms. In subsection 4.3 we discuss the algorithm (27) applied to permutation invariant concave diversity measures with domain restricted to the positive orthant. Conditions on the stepsize parameter $\mu_k$ are given to ensure convergence of the algorithm. In subsection 4.4 we consider permutation and sign invariant diversity measures. In this case the restriction to the positive orthant can be removed, slightly relaxing the conditions on $\mu_k$. The special case of sign and permutation invariant concave diversity measures with positive definite scaling matrix $\Pi(x)$ is discussed in subsection 4.5. For this case, the simple choice of $\mu_k = 1$ for the stepsize parameter (resulting in the use of the simpler algorithm (25)) results in convergence over $\mathsf{R}^n$.

## 4.3 Concave Diversity Minimization on the Positive Orthant

Assume that the diversity measure $d(x)$ is permutation invariant and concave on the positive orthant $\mathcal{Q}_1 \subset \mathsf{R}^n$, and is otherwise arbitrary. Then from (6) and (19) we have

$$(y - x_k)^T \nabla d(x_k) = \alpha(x_k)(y - x_k)^T \Pi(x_k) x_k \geq d(y) - d(x_k) \tag{31}$$

for all $y, x_k \in \mathcal{Q}_1$, where $\alpha(x_k) > 0$. If $y = x_{k+1} \in \mathcal{Q}_1$ then (31) and condition (28) yield

$$x_{k+1} \neq x_k \Rightarrow d(x_{k+1}) < d(x_k) . \tag{32}$$

Iterating in this manner will result in convergence to a local minimum $d(x^*)$, where $x^* \in \mathcal{Q}_1$ satisfies the first order necessary condition (24).

Suppose that $y = x_{k+1} \in \mathcal{Q}_1$ with $x_{k+1}$ determined by (27). Then (31) and condition (28) yield

$$0 > x_k^T \Pi(x_k)(x_{k+1} - x_k) = \mu_k x_k^T \Pi(x_k)(x_{k+1}^r - x_k) \geq d(x_{k+1}) - d(x_k) \tag{33}$$

when $x_{k+1} \neq x_k$. To ensure that conditions (28) and (33) hold we choose $\mu_k$ such that

$$\operatorname{sgn}(\mu_k) = \operatorname{sgn}[x_k^T \Pi(x_k)(x_k - x_{k+1}^r)] \quad \text{and} \quad x_{k+1} = x_k + \mu_k(x_{k+1}^r - x_k) \in \mathcal{Q}_1 \,. \tag{34}$$

Then the iterates $x_k$ will converge to yield a local minimum of $d(x)$ over the convex set $LV(A,b) \cap \mathcal{Q}_1$. If $d(x)$ is assumed strictly concave, the minimum is a boundary point of this set yielding a sparse solution. The diagonal structure of the scaling matrix allows the sign of $\mu_k$ to be determined using only order $n$ operations.

## 4.4 General Concave Sparsity Minimization

The diversity measures considered in this paper are sign invariant. As a consequence, the requirement of concavity over the positive orthant guarantees concavity over any fixed orthant of $\mathsf{R}^n$. However, generally such a measure is *not* concave across orthants. Therefore, because $x_{k+1}$ generated by (25) may be in a different orthant from $x_k$, we cannot guarantee that at each iteration of (25) we can directly invoke the condition (6) to obtain the crucial inequality (31) used in the proof of the previous subsection. However, we can still apply (6) to $d(\cdot)$ viewed as a function of $|x_k|$ when $d(\cdot)$ is additionally assumed to be sign invariant. This is what we now exploit to generalize the algorithm.

### 4.4.1 Preliminaries

The sign invariance and concavity properties of $d(x)$ on $\mathcal{Q}_1$ are exploited to select the step size in (27) and develop convergent algorithms. In order to do this, we establish a correspondence between $x$ in any quadrant and its corresponding absolute value vector $\bar{x} \triangleq |x| \in \mathcal{Q}_1$. Defining the symmetric *sign matrix*, $S(x)$, by

$$S(x) \triangleq \operatorname{diag}\left(\operatorname{sgn}(x[i])\right) \,, \tag{35}$$

we note that $\bar{x} = |x| = S(x)x$, $x = S(x)\bar{x} = S(x)|x|$, $S^2(x) = I$, $S^{-1}(x) = S(x)$, and $S(x)\Pi(x) = \Pi(x)S(x)$, assuming that $\Pi(x)$ is diagonal.

Now note that

$$\frac{\partial d(x)}{\partial x[i]} = \frac{\partial \bar{x}[i]}{\partial x[i]} \frac{\partial d(\bar{x})}{\partial \bar{x}[i]} = \operatorname{sgn}(x[i]) \, \frac{\partial d(\bar{x})}{\partial \bar{x}[i]} \,,$$

and therefore,

$$\alpha \cdot \frac{1}{\alpha} \cdot \frac{\partial d(x)}{\partial x[i]} \frac{1}{x[i]} = \alpha \cdot \frac{1}{\alpha} \cdot \frac{\partial d(\bar{x})}{\partial \bar{x}[i]} \frac{1}{\bar{x}[i]} \, ,$$

where $\alpha = \alpha(x) = \alpha(\bar{x}) > 0$ is chosen to simplify the quantities

$$\pi_i(x) \triangleq \frac{1}{\alpha} \frac{\partial d(x)}{\partial x[i]} \frac{1}{x[i]} = \frac{1}{\alpha} \frac{\partial d(\bar{x})}{\partial \bar{x}[i]} \frac{1}{\bar{x}[i]} \triangleq \pi_i(\bar{x}) \, . \tag{36}$$

The diagonal scaling matrix of (19) is defined by $\Pi(x) = \text{diag}(\pi_i(x))$. Note that

$$\begin{aligned} \nabla d(x) &= S(x) \nabla_{\bar{x}} d(\bar{x}) = \alpha(x) \Pi(x) x \\ \nabla_{\bar{x}} d(\bar{x}) &= \alpha(\bar{x}) \Pi(\bar{x}) \bar{x} = \alpha(x) \Pi(x) \bar{x} \\ \alpha(x) &= \alpha(\bar{x}) > 0 \\ \Pi(x) &= \text{diag} \left( \frac{1}{\alpha(x)|x[i]|} \frac{\partial d(x)}{\partial |x[i]|} \right) = \text{diag} \left( \frac{1}{\alpha(\bar{x})\bar{x}[i]} \frac{\partial d(\bar{x})}{\partial \bar{x}[i]} \right) = \Pi(\bar{x}) \, . \end{aligned} \tag{37}$$

### 4.4.2 A Convergent Algorithm

Assume that the permutation and sign invariant function $d(x) = d(\bar{x})$, $\bar{x} = |x|$, is concave over $\mathcal{Q}_1$. Then properties (6) and (37) yield the relationship

$$\begin{aligned} \alpha(x_k) \, \bar{x}_k^T \Pi(\bar{x}_k) \, (\bar{x}_{k+1} - \bar{x}_k) &= \nabla_{\bar{x}} g(\bar{x}_k)^T (\bar{x}_{k+1} - \bar{x}_k) \\ &\geq d(\bar{x}_{k+1}) - d(\bar{x}_k) = d(x_{k+1}) - d(x_k) \, , \end{aligned} \tag{38}$$

with $\bar{x} = |x|$ and $x_{k+1}$ given by (27). Note that the condition $\bar{x}_k^T \Pi(\bar{x}_k)(\bar{x}_{k+1} - \bar{x}_k) < 0$ for $\bar{x}_{k+1} \neq \bar{x}_k$ is sufficient to guarantee convergence. To ensure that this occurs, we choose the stepsize $\mu_k$ in (27) as follows:

```
$\mu_k$ := 1;
loop  $j \geq 0$
        If  $\bar{x}_k^T \Pi(\bar{x}_k)(\bar{x}_{k+1} - \bar{x}_k) < 0$,   exit;            (39)
        If $j$ even   $\mu_k := -\mu_k$,   else  $\mu_k := \frac{\mu_k}{2}$;
end;
```

This results in the sequence $\mu_k = 1, -1, -\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{8}, \frac{1}{8}, \cdots$. Note that for $x_k$ and $x_{k+1}$ in the same orthant (which eventually will be true once $|\mu_k|$ becomes small enough), we have

$$\bar{x}_k^T \Pi(\bar{x}_k) \, (\bar{x}_{k+1} - \bar{x}_k) = x_k^T \Pi(x_k) \, (x_{k+1} - x_k) \; .$$

Therefore, the loop (39) is guaranteed to terminate after a finite number of iterations as a consequence of the discussion given above in Subsection 4.3. Note that, unlike the algorithm in Subsection 4.3, we do not have to guarantee that $x_k$ and $x_{k+1}$ are always in the same orthant in order for the test (39) to make sense.

### 4.4.3 Scale Invariant Measures

As discussed earlier, scale invariant measures must satisfy the property (20). Scale invariant diversity measures presented in this paper include the 1- and 2-normalized Renyi Entropies and $\mathcal{S}$-functions. The 1-Norm Renyi Entropy and $\mathcal{S}$-Functions have been shown to be Schur-concave for $p \leq 1$ and almost strictly concave for $0 < p < 1$. They are not Schur-concave (and hence not concave) for $p > 1$. The 2-Norm Renyi Entropy and $\mathcal{S}$-Functions are almost strictly concave (and hence Schur-concave) for $p < \frac{1}{2}$. They are not Schur-concave for $p > \frac{1}{2}$.

As a consequence of property (20), the test (39) can be slightly simplified to the requirement that

$$\bar{x}_k^T \mathit{\Pi}(\bar{x}_k)\bar{x}_{k+1} < 0\,, \tag{40}$$

which will be true if $\bar{x}_{k+1}$ is not parallel to $\bar{x}_k$. If $\bar{x}_{k+1}$ is parallel to $\bar{x}_k$, then because of the assumed scale invariance of $d(x)$ the algorithm has converged. In this case, $\bar{x}_k^T \mathit{\Pi}(\bar{x}_k)\bar{x}_{k+1} = 0$ and the algorithm terminates.

## 4.5 Concave Diversity Minimization, $\mathit{\Pi}(x) > 0$

In this subsection we add the additional constraint that a sign and permutation invariant concave diversity measure $d(x)$ now have a positive definite scaling matrix, $\mathit{\Pi}(x) > 0$ for all $x \in \mathsf{R}^n$. With this condition, we have our strongest results and can show that we can remove the requirement that $x_k$ be constrained to a single orthant *and* the need to compute a value for $\mu_k$. Specifically, it is shown that the simple choice $\mu_k = 1$, corresponding to the algorithm (25), will yield a convergent algorithm.

Interestingly, this constraint does not appear to be overly restrictive and it admits the important class of $\mathcal{S}$-Functions described in Section 3.1. This is a large class of sparsity measures which satisfy the conditions of Theorem 17 and also contains the $p$-norm-like, $p \leq 1$, concave sparsity measures. The general scaling matrix for this class, $\mathit{\Pi}_{\text{sig}}(x)$, is given in Table 1. Note that $\mathit{\Pi}_{\text{sig}}(x)$ is positive definite for all $x$ and has a well defined inverse everywhere. $\mathit{\Pi}_{\text{sig}}^{-1}(x)$ is diagonal with the $i^{th}$ diagonal element given by

$$\frac{|x[i]|^{2-p_j}}{(|p_1|\,\omega_1\,|x[i]|^{p_1-p_j} + \cdots + |p_j|\omega_j + \cdots + |p_q|\,\omega_q\,|x[i]|^{p_q-p_j})} \tag{41}$$

where $p_j = \min(p_1, \cdots, p_q)$.

**Lemma 1** *Given $x_k$, $Ax_k = b$, let $x_{k+1}$ be computed by (25), $\bar{x}_k = S(x_k)x_k = |x_k|$, and $\bar{x}_{k+1} = S(x_{k+1})x_{k+1} = |x_{k+1}|$. If $\mathit{\Pi}(x_k) = \mathit{\Pi}(\bar{x}_k) > 0$, then*

$$(\bar{x}_{k+1} - \bar{x}_k)^T \mathit{\Pi}(\bar{x}_k)\bar{x}_{k+1} \leq 0\,, \tag{42}$$

*with equality if $x_{k+1}$ and $x_k$ are in the same orthant.*

**Proof.** Notice that $x^r_{k+1}$, and hence $x_{k+1}$, is guaranteed to be feasible given feasibility of $x_k$. Furthermore,

$$(x^r_{k+1})^T \Pi(x_k) x^r_{k+1} = (x^r_{k+1})^T \Pi(x_k) x_k \,, \tag{43}$$

which can be readily shown by substituting for $x^r_{k+1}$ from (26) [69]. If $x_{k+1}$ and $x_k$ are in the same orthant, we have $S \overset{\Delta}{=} S(x_{k+1}) = S(x_k)$. Recalling that $S^2 = I$, then $(\bar{x}_{k+1} - \bar{x}_k)^T S^2 \Pi(\bar{x}_k)\bar{x}_{k+1} = (x_{k+1} - x_k)^T \Pi(x_k) x_{k+1} = 0$ from (43). To prove (42), note that $\Pi(x_k) = \mathrm{diag}(\pi_i)$ is positive definite by assumption, so that $\pi_i > 0$. We have

$$\begin{aligned} \bar{x}^T_{k+1} \Pi(\bar{x}_k) \bar{x}_{k+1} &= x^T_{k+1} \Pi(x_k) x_{k+1} = x^T_{k+1} \Pi(x_k) x_k \quad \text{(from (43))} \\ &= \sum_{i=1}^{n} x_{k+1}[i] \pi_i \; x_k[i] \leq \sum_{i=1}^{n} \bar{x}_{k+1}[i] \pi_i \; \bar{x}_k[i] = \bar{x}^T_k \Pi(\bar{x}_k) \bar{x}_{k+1} \,, \end{aligned}$$

from which the result (42) follows. ■

With Lemma 1 at hand we can now prove convergence of the algorithm (25).

**Theorem 17** (Concave Convergence for $\Pi(x) > 0$) *Let $d(x)$ be a sign and permutation invariant function that is strictly concave on the positive orthant $\mathcal{Q}_1$ and for which $\Pi(x) > 0$ for all $x \in \mathsf{R}^n$. Assume that the set $\{x | d(x) \leq d(x_0)\}$ is compact for all $x_0$. Let $x_k$ be generated by the iteration (25) starting with $x_0$ feasible, $Ax_0 = b$. Then for all $\bar{x}_{k+1} \neq \bar{x}_k$, we have $d(x_{k+1}) < d(x_k)$ and the algorithm converges to a local minimum $d(x^*)$, $x_k \to x^*$, where $x^*$ is a boundary point of $\mathcal{Q}_l \cap LV(A, b)$ for some orthant $\mathcal{Q}_l$.*

**Proof.** Let $\bar{x}_{k+1} \neq \bar{x}_k$ and $\Pi = \Pi(x_k) = \Pi(\bar{x}_k) > 0$. Then

$$0 < (\bar{x}_{k+1} - \bar{x}_k)^T \Pi (\bar{x}_{k+1} - \bar{x}_k) = (\bar{x}_{k+1} - \bar{x}_k)^T \Pi \bar{x}_{k+1} - (\bar{x}_{k+1} - \bar{x}_k)^T \Pi \bar{x}_k \tag{44}$$

so that

$$(\bar{x}_{k+1} - \bar{x}_k)^T \Pi \bar{x}_k < (\bar{x}_{k+1} - \bar{x}_k)^T \Pi \bar{x}_{k+1} \leq 0 \,, \tag{45}$$

where the last inequality follows from (42). Then for $\bar{x}_{k+1} \neq \bar{x}_k$, from (42), (45), and (37) we have $0 \geq \alpha(\bar{x})(\bar{x}_{k+1} - \bar{x}_k)^T \Pi \bar{x}_{k+1} > \alpha(\bar{x})(\bar{x}_{k+1} - \bar{x}_k)^T \Pi \bar{x}_k = (\bar{x}_{k+1} - \bar{x}_k)^T \nabla_{\bar{x}} d(\bar{x}_k) \geq d(\bar{x}_{k+1}) - d(\bar{x}_k) = d(x_{k+1}) - d(x_k)$, where the last inequality follows from concavity of $d(\bar{x})$ on the positive orthant, $\bar{x}_k \in \mathcal{Q}_1$. Thus $d(x_{k+1}) < d(x_k)$. It then follows from the general convergence theorem [82, 8, 50] that $x_k$ converges to $x^*$ satisfying the necessary condition (24) for $d(x)$ to have a local minimum. Asymptotically, $x_k$ will eventually remain in one quadrant as it converges to $x^*$. Since $d(x)$ is strictly concave on the interior of any quadrant, $x^*$ is a boundary point, and hence sparse, over some convex set $\mathcal{Q}_l \cap LV(A, b)$ for some quadrant $\mathcal{Q}_l$. ■

## 4.6 Discussion

Other functions can be proved to be minimized by the algorithm (27), with convergence generally being shown on a case-by-case basis. For example, it is proved in [69] that the 2-normalized Shannon entropy–based algorithm is convergent. As discussed earlier, the 2-normalized Shannon entropy diversity measure corresponds to the 2-normalized Renyi entropy for $p = 1$, which is not concave or Schur-concave, and therefore is not expected to result in a minimum associated with complete sparsity; a fact that has been demonstrated in simulation [69]. The lack of concavity requires a convergence proof via different means than the use of (6).

The requirement of the invertibility of $\mathit{\Pi}(x)$ in (25) or (27) appears to give good reason to prefer the non-scale invariant measures provided by the class of (unnormalized) $\mathcal{S}$-functions over the 1-norm and 2-normalized scale invariant $\mathcal{S}$-functions (which effectively include the normalized Renyi entropies). In particular, the very tractable form of the scaling matrices for $\mathcal{S}$-functions allows them to be readily inverted, as shown by (41). Significantly more care is needed when dealing with the generically invertible scaling matrices for the normalized $\mathcal{S}$-functions.

The fact that the some of the diversity measures described in this paper are almost strictly concave, rather than strictly concave does not usually cause any practical difficulty for $n$ reasonably large. Generically, convergence generally proceeds as if the measures are in fact strictly concave. Every step size taken by the algorithm causes a strict descent in the $(n-1)$ directions along which the diversity measure is concave while in all directions Schur-concavity ensures that the Lorentz order is preserved. Example case studies of the effectiveness of the proposed algorithm in obtaining sparse solutions can be found in references [1, 69, 36, 37].

Optimal subset selection is a problem that (in principle) can be solved by exhaustive search. Of course this direct approach quickly (as a function of the number of dictionary vectors) becomes computationally infeasible, and computationally tractable *suboptimal* procedures, such as the methods proposed in this paper, must be utilized [25, 56, 34, 19, 52, 27, 17, 81, 61, 78, 1, 39, 37]. The suboptimal methods generally give acceptably sparse solutions with no guarantee of global optimality. A comparison of some of these methods can be found in [1].

Interestingly, in [1] a simple modification to our sparse-basis selection procedure is given which guarantees with probability approaching 1 that a true global optimum to the sparse basis selection problem will be found *provided* that a sufficiently sparse solution exists and a certain technical condition holds on the dictionary vectors (equivalently, the columns of $A$). This fact leads us to conjecture that simple stochastic, annealing-like modifications to our algorithm should improve the optimality of the solution in more general settings.

# 5 CONCLUSIONS

The concept of *majorization* as a preordering on vectors, one which orders vectors according to their degree of concentration or diversity (a concept also captured by the equivalent concept of the Lorentz order), was presented and we discussed how the majorization-consistent property of *Schur-concavity* is a desirable condition for a functional measure of diversity to have. Such functions are necessarily *permutation invariant* on the domain of definition. We then argued that the subclass of *sign invariant concave functions* are especially good measures of diversity and that their minimization generally results in a good solution to the sparse basis selection problem. Tests for determining the concavity or Schur-concavity of a candidate diversity measure were given. We also discussed a relaxation of the property of concavity of a diversity measure to the newly defined and slightly weaker property of "almost concavity." Almost concave functions are Schur-concave, and therefore respect the Lorentz order, and locally concave in every direction but one.

Candidate diversity measures were analyzed from the perspective of majorization, Schur-concavity, concavity, and almost concavity. We examined the class of $p$-norm-like measures, including the special cases where the vectors are first normalized in the 1-norm and in the 2-norm sense. We also looked at the Gaussian entropy (discussing its equivalence to the $p = 0$ $p$-norm-like measure) and the Shannon entropy as measures of diversity. We then investigated the class of Renyi entropies and showed that this class contains both the $p$-norm-like and Shannon entropy and therefore that all of these measures are intimately related. We also developed and analyzed the large class of signomial diversity measures, and showed that they can reasonably be interpreted as a generalization of the Renyi entropy measures (and hence of the $p$-norm-like, Gaussian entropy, and Shannon Entropy). Finally, we developed an iterative optimization algorithm that is shown to converge to a local minimum of an appropriately chosen diversity measure and thereby provide a sparse solution to the best basis selection problem.

In the development of the basis selection algorithm, it was shown that associated with each of the diversity measures considered in this paper is a corresponding diagonal *scaling matrix*, $\Pi(x)$. The scaling matrix is defined by a particular factorization of the gradient of the diversity measure. When the scaling matrix is positive definite, the iterative algorithm developed in this paper can be interpreted as an Affine Scaling Transformation (AST)-based algorithm, with an AST, $W(x) \triangleq \Pi^{-\frac{1}{2}}(x)$, uniquely defined by the particular choice of the diversity measure $d(x)$. This is in contrast to the general AST methodology where a choice of an AST is taken independently of the objective function to be minimized. More generally, our algorithm is applicable even in the case of a non-positive definite scaling matrix, which is an additional generalization relative to the standard AST methodology. However, the strongest convergence results hold for the case of diversity measures with positive definite scaling matrices. Simulations and case studies showing the convergence behavior of the proposed algorithm can be found in [36, 1, 69, 37].

As discussed in this paper, diversity measures drawn from the class of "$\mathcal{S}$-functions" appear to be particularly well-behaved; they are a large set of separable, concave diversity measures with positive definite, and readily invertible, scaling matrices. Certainly, many questions naturally arise and remain to be answered regarding the properties of this class. For instance, a natural line of inquiry would be to determine the extent to which the parameters defining a diversity function chosen from this class can be optimized for a particular application. It is also interesting to ask whether general separable, concave functions can be approximated arbitrarily well within the class of $\mathcal{S}$-functions, or it extension to the set of symmetric sums of multinomial–like terms [53]—unfortunately, unlike working with general signomials [20], one cannot invoke the Stone-Weierstrass theorem [3] to claim that the subset of $\mathcal{S}$-functions are dense in the space of continuous functions (because the $\mathcal{S}$-functions do not form a subalgebra). Another interesting question (touched upon briefly in Section 4.6) is whether stochastic modifications of the basic algorithm can improve the quality of the solution found, similarly to the provably optimal extension presented in [1]. Much additional research will be required to obtain definitive answers to these and other related questions.

# APPENDIX

In this appendix we derive the forms of the Hessians $\mathcal{H}_p^{(1)}(x)$ and $\mathcal{H}_p^{(2)}(x)$, given by equations (9) and (10), and prove Theorems 14, 15, and 16. We also discuss the derivations of the scaling matrices given in Table 1.

**Derivation of the Hessian $\mathcal{H}_p^{(1)}(x)$, Equation (9).** Note that

$$\tilde{x} = \frac{|x|}{\|x\|_1} \quad \Rightarrow \quad \frac{\partial \phi(\tilde{x})}{\partial x[i]} = \frac{\operatorname{sgn}(x[i])}{\|x\|_1} \left( \frac{\partial \phi(\tilde{x})}{\partial \tilde{x}[i]} - \sum_k \tilde{x}[k] \frac{\partial \phi(\tilde{x})}{\partial \tilde{x}[k]} \right) . \tag{46}$$

This follows from the chain rule,

$$\frac{\partial \phi(\tilde{x})}{\partial x[i]} = \sum_k \frac{\partial \phi(\tilde{x})}{\partial \tilde{x}[k]} \frac{\partial \tilde{x}[k]}{\partial x[i]} ,$$

and the fact that

$$\tilde{x} = \frac{|x|}{\|x\|_1} \quad \Rightarrow \quad \frac{\partial \tilde{x}[k]}{\partial x[i]} = \frac{\operatorname{sgn}(x[i])}{\|x\|_1} \left( \delta_{k,i} - \tilde{x}[k] \right) . \tag{47}$$

Now apply identity (46) to the function

$$d_p^{(1)}(x) = d_p(\tilde{x}) = \operatorname{sgn}(p) \sum_i \tilde{x}[i]^p .$$

This results in

$$\frac{\partial d_p^{(1)}(x)}{\partial x[i]} = \frac{\partial d_p(\tilde{x})}{\partial x[i]} = \frac{|p| \operatorname{sgn}(x[i])}{\|x\|_1} \left( \tilde{x}[i]^{p-1} - |d_p(\tilde{x})| \right) . \tag{48}$$

Differentiating (48) with respect to $x[j]$ then yields the $(i,j)$-th element of the Hessian $\mathcal{H}_p^{(1)}(x)$,

$$\begin{aligned} \frac{\partial^2 d_p^{(1)}(x)}{\partial x[i] \partial x[j]} &= \frac{\partial^2 d_p(\tilde{x})}{\partial x[i] \partial x[j]} \qquad\qquad (49) \\ &= -\frac{|p| \operatorname{sgn}(x[i]x[j])}{\|x\|_1^2} \left\{ p\, \tilde{x}[i]^{p-1} + p\, \tilde{x}[j]^{p-1} + (1-p) \tilde{x}[i]^{p-2} \delta_{i,j} - (1+p)\, |d_p(\tilde{x})| \right\} . \end{aligned}$$

On the *positive orthant*, $\mathcal{Q}_1$, $\operatorname{sgn}(x[i]) = 1$, $i = 1, \cdots n$, in equation (49) and the Hessian can be written as

$$\mathcal{H}_p^{(1)}(x) = -\frac{|p|}{\|x\|_1^2} \left\{ p \sum_{i=1}^n \frac{\mathbf{e}_i \mathbf{1}^T}{\tilde{x}[i]^{1-p}} + p \sum_{j=1}^n \frac{\mathbf{1} \mathbf{e}_j^T}{\tilde{x}[i]^{1-p}} + (1-p) \sum_{i=1}^n \frac{\mathbf{e}_i \mathbf{e}_i^T}{\tilde{x}[i]^{2-p}} - (1+p) \mathbf{1}\mathbf{1}^T \sum_{i=1}^n \tilde{x}[i]^p \right\} , \tag{50}$$

which is just equation (9). Note that the $(i,j)$-th element of the general Hessian (49) differs from the same element of the positive orthant Hessian (50) only by the sign factor $\operatorname{sgn}(x[i]x[j])$. This means that generally, for any $x \in \mathsf{R}^n$, we have

$$\mathcal{H}_p^{(1)}(x) = \operatorname{diag}[\operatorname{sign}(x)] \cdot \mathcal{H}_p^{(1)}(|x|) \cdot \operatorname{diag}[\operatorname{sign}(x)]^T , \tag{51}$$

where $\mathcal{H}_p^{(1)}(|x|)$ can be determined from (50) since $|x| \in \mathcal{Q}_1$. Equation (51) shows that if $\mathcal{H}_p^{(1)}(x)$ is positive definite on the interior of the positive orthant then it is positive definite on the interior of any orthant (and vice versa).

**Derivation of the Hessian $\mathcal{H}_p^{(2)}(x)$, Equation (10).** Note that

$$\tilde{x} = \frac{x^2}{\|x\|_2^2} \quad \Rightarrow \quad \frac{\partial \tilde{x}[k]}{\partial x[i]} = 2\frac{x[i]}{\|x\|_2^2}\left(\delta_{k,i} - \tilde{x}[k]\right) . \tag{52}$$

From (52) and the chain rule, we find

$$\tilde{x} = \frac{x^2}{\|x\|_2^2} \quad \Rightarrow \quad \frac{\partial \phi(\tilde{x})}{\partial x[i]} = 2\frac{x[i]}{\|x\|_2^2}\left(\frac{\partial \phi(\tilde{x})}{\partial \tilde{x}[i]} - \sum_k \tilde{x}[k]\frac{\partial \phi(\tilde{x})}{\partial \tilde{x}[k]}\right) . \tag{53}$$

Thus, with $\tilde{x} = x^2/\|x\|_2^2$, we obtain

$$\frac{\partial d_p^{(2)}(x)}{\partial x[i]} = \frac{\partial d_p(\tilde{x})}{\partial x[i]} = 2\frac{|p|\,x[i]}{\|x\|_2^2}\left(\tilde{x}[i]^{p-1} - |d_p(\tilde{x})|\right) . \tag{54}$$

Differentiating (54) with respect to $x[j]$ yields the $(i,j)$-th element of the Hessian $\mathcal{H}_p^{(2)}(x)$,

$$\begin{aligned}
\frac{\partial^2 d_p^{(2)}(x)}{\partial x[i]\partial x[j]} &= \frac{\partial^2 d_p(\tilde{x})}{\partial x[i]\partial x[j]} \\
&= \frac{2\,|p|}{\|x\|_2^2}\Big\{\delta_{i,j}\left[(2p-1)\tilde{x}[i]^{p-1} - |d_p^{(2)}(x)|\right] \\
&\quad -2(\tilde{x}[i]\,\tilde{x}[j])^{\frac{1}{2}}\left[p\tilde{x}[i]^{p-1} + p\tilde{x}[j]^{p-1} - |d_p^{(2)}(x)|\right]\Big\} \\
&= \frac{2\,|p|}{\|x\|_2^2}\left(\mathrm{A}_{i,j} + \mathrm{B}_{i,j} + \mathrm{C}_{i,j}\right) ,
\end{aligned} \tag{55}$$

where

$$\begin{aligned}
\mathrm{A}_{i,j} &= \delta_{i,j}\left[(2p-1)\tilde{x}[i]^{p-1} - |d_p^{(2)}(x)|\right] ; & (56)\\
\mathrm{B}_{i,j} &= -2(\tilde{x}[i]\,\tilde{x}[j])^{\frac{1}{2}}\left(p\tilde{x}[i]^{p-1} + p\tilde{x}[j]^{p-1}\right) ; & (57)\\
\mathrm{C}_{i,j} &= 2(\tilde{x}[i]\,\tilde{x}[j])^{\frac{1}{2}}|d_p^{(2)}(x)| ; & (58)
\end{aligned}$$

With $\mathrm{A}_{i,j}$, $\mathrm{B}_{i,j}$, and $\mathrm{C}_{i,j}$ the $(i,j)-th$ component of matrices A, B, and C respectively, we have

$$\mathcal{H}_p^{(2)}(x) = \frac{2\,|p|}{\|x\|_2^2}\left(\mathrm{A} + \mathrm{B} + \mathrm{C}\right) , \tag{59}$$

where

$$\begin{aligned}
\mathrm{A} &= -\frac{2\,|p|}{\|x\|_2^2}\sum_{i=1}^{n}\left(\frac{1-2p}{\tilde{x}[i]^{1-p}} + |d_p^{(2)}(x)|\right)\mathbf{e}_i\mathbf{e}_i^T ; & (60)\\
\mathrm{C} &= 2|d_p^{(2)}|\,(1+p)\,\tilde{x}^{\frac{1}{2}}\tilde{x}^{\frac{T}{2}} = 2\sum_{i=1}^{n}\tilde{x}[i]^p\,(1+p)\,\tilde{x}^{\frac{1}{2}}\tilde{x}^{\frac{T}{2}} ; & (61)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{B} &= -2p\sum_{i,j=1}^{n}\left(\tilde{x}[i]^{p-\frac{1}{2}}\mathbf{e}_i\mathbf{e}_j^T\tilde{x}[j]^{\frac{1}{2}}+\tilde{x}[i]^{\frac{1}{2}}\mathbf{e}_i\mathbf{e}_j^T\tilde{x}[j]^{p-\frac{1}{2}}\right) \\
&= -2p\sum_{i=1}^{n}\frac{1}{\tilde{x}[i]^{\frac{1}{2}-p}}\left(\mathbf{e}_i\tilde{x}^{\frac{T}{2}}+\tilde{x}^{\frac{1}{2}}\mathbf{e}_i^T\right)\,. \qquad (62)
\end{aligned}$$

Equations (59)–(62) taken together yield the desired result, Equation (10).

**Proof of Theorem 14.** We can readily show that for each $x$ in the interior of $\mathcal{Q}_1$, $y^T\mathcal{H}_p^{(1)}(x)y \le 0$ for all nonzero $y$ in the $(n-1)$ dimensional subspace of $\mathsf{R}^n$ orthogonal to the direction $\mathbf{1}$, so that the Schur-concave function $d_{\text{sig}}^{(1)}(x)$ is almost concave in the sense of Definition 6. Indeed, for any nonzero $y = z$, $z \perp \mathbf{1}$, we have for $p < 1$,

$$z^T\mathcal{H}_p^{(1)}(x)z = -\frac{|p|(1-p)}{\|x\|_1^2}\sum_{i=1}^{n}\frac{(z_i)^2}{\tilde{x}[i]^{2-p}} < 0\,, \qquad (63)$$

and $z^T\mathcal{H}_p^{(1)}(x)z = 0$ if $p = 1$. The Hessian of $d_{\text{sig}}^{(1)}(x)$ is a sum of Hessians with this property. Therefore $d_{\text{sig}}^{(1)}(x)$ is almost strictly concave for at least one $p_j < 1$ and is otherwise almost concave.

Of course, the most desirable situation would be if $d_{\text{sig}}^{(1)}(x)$ were concave. This is not, however, the case, and the remainder of the proof is concerned with demonstrating this fact.

Let $y = \mathbf{1} \perp z$, then

$$\begin{aligned}
\mathbf{1}^T\mathcal{H}_p^{(1)}(x)\mathbf{1} &= -\frac{|p|}{\|x\|_1^2}\,f(\tilde{x}) \quad \text{where,} \qquad (64) \\
f(\tilde{x}) &= \sum_{i=1}^{n}\left\{2\,n\,p\,\tilde{x}[i]^{p-1}+(1-p)\,\tilde{x}[i]^{p-2}-(1+p)\,n^2\,\tilde{x}[i]^{p}\right\}\,.
\end{aligned}$$

It is straightforward to show that the Hessian of $f(\tilde{x})$, taken with respect to $\tilde{x}$, is positive definite for all $\tilde{x}$ when $0 < p < 1$. In this case, the unique minimum of $f(\tilde{x})$ subject to the constraint that $\sum_i \tilde{x}[i] = 1$ can be shown via the methods of Lagrange multipliers [60] to occur for $\tilde{x}[i] = n^{-1}$, $i = 1, \cdots, n$, and yields the value zero. Thus $\mathbf{1}^T\mathcal{H}_p^{(1)}(x)\mathbf{1} \le 0$ for all $x \in \mathcal{Q}_1$. Furthermore, $\mathbf{1}^T\mathcal{H}_p^{(1)}(x)\mathbf{1} = 0$ only if $x$ is precisely along the line through $\mathbf{1}$, $x \propto \mathbf{1}$, which defines a set of measure zero in $\mathcal{Q}_1$. For $p < 0$, there are values of $\tilde{x}$ for which $f(\tilde{x})$ becomes negative. (E.g., take $p = -\epsilon$, $\epsilon > 0$ small, and $\tilde{x} \to \mathbf{e}_i$ any $i = 1, \cdots, n$.) Obviously we do not have concavity for $p < 0$, but we still do not have a definitive answer for $0 < p < 1$.

Since for $0 < p < 1$, $\mathcal{H}_p^{(1)}(x)$ is concave along the direction $\mathbf{1}$ and on the subspace $\mathbf{1}^\perp$, it can only cease to be concave along a direction oblique to both $\mathbf{1}$ and $\mathbf{1}^\perp$. Thus, most generally we take $y$ to be of the form $y = \mathbf{1} + z$, $z \perp \mathbf{1}$ [8]. Note that $z \to 0$ returns the case $y = \mathbf{1}$, while $\|z\| \to \infty$ yields $y = z$, corresponding to the two cases considered above. Also note that

$$y = \mathbf{1} + z \Rightarrow y_i = 1 + z_i\,, \quad \mathbf{1}^T z = \sum_i z_i = 0 \quad \text{and} \quad \sum_i y_i = n\,. \tag{65}$$

With this choice of $y$ we have

$$\begin{aligned} y^T \mathcal{H}_p^{(1)}(x) y &= -\frac{|p|}{\|x\|_1^2} g(\tilde{x}, y)\,, \\ g(\tilde{x}, y) &= \sum_{i=1}^n \left\{ (1-p)\,\tilde{x}[i]^{p-2} y_i^2 + 2\,n\,p\,\tilde{x}[i]^{p-1} y_i - (1+p)\,n^2\,\tilde{x}[i]^p \right\}\,. \end{aligned} \tag{66}$$

Note that $f(\tilde{x})$ of (64) and $g(\tilde{x}, y)$ of (66) are related by

$$f(\tilde{x}) = g(\tilde{x}, \mathbf{1})\,. \tag{67}$$

Also note that $y^T \mathcal{H}_p^{(1)}(x) y = y^T \mathcal{H}_p^{(1)}(\tilde{x}) y \le 0$ (respectively, $y^T \mathcal{H}_p^{(1)}(x) y < 0$) for all $x$ and $y$ if and only if $g(\tilde{x}, y) \ge 0$ (respectively, $g(\tilde{x}, y) > 0$) for all $\tilde{x}$ and $y$. Thus if it can be shown that $g(\tilde{x}, y) < 0$ for some $y$, then $\mathcal{H}_p^{(1)}(x)$ is not positive definite and $d_{\text{sig}}^{(1)}(x)$ is not concave on $\mathcal{Q}_1$ along some direction that is oblique to both $\mathbf{1}$ and the space orthogonal to $\mathbf{1}$, $\mathbf{1}^\perp$. We will determine the precise situation that holds by computing the minimum admissible value of $g(\tilde{x}, y)$ for any fixed value of $\tilde{x}$ in the interior of $\mathcal{Q}_1$.

To solve the problem,

$$\min_y g(\tilde{x}, y) \quad \text{subject to} \quad \sum_{i=1}^n y_i = n\,, \tag{68}$$

we will apply the method of Lagrange multipliers. It is readily ascertained via the Hessian test that $g(\tilde{x}, y)$ is strictly convex in $y$ for $\tilde{x}$ in the interior of $\mathcal{Q}_1$ and therefore problem (68) has a unique solution. This solution is a stationary point of the Lagrangian,

$$\mathcal{L} = g(\tilde{x}, y) + 2\lambda \left( n - \sum_{i=1}^n y_i \right)\,, \tag{69}$$

where the factor "2" has been added for convenience. The stationarity condition results in the $n+1$ equations

$$(1-p)\tilde{x}[i]^{p-2} y_i + np\tilde{x}[i]^{p-1} - \lambda \;=\; 0\,, \tag{70}$$

$$\text{for} \quad i = 1, \cdots, n \quad \text{and} \quad \sum_{i=1}^n y_i \;=\; n\,. \tag{71}$$

Note that (70) is equivalent to

$$(1-p) y_i + \, n p \tilde{x}[i] - \lambda \tilde{x}[i]^{2-p} = 0\,. \tag{72}$$

[8] The overall scaling of $y$ does not affect the proof.

Summing (72) over $i$ and using (71) and the fact that $\sum_i \tilde{x}[i] = 1$, results in

$$\lambda_{\text{opt}} = \frac{n}{\sum_{i=1}^{n} \tilde{x}[i]^{2-p}} \ . \tag{73}$$

Substitution of (73) into (72) yields the worst case solution,

$$y_{\text{opt},i} = \frac{\lambda_{\text{opt}}\tilde{x}[i]^{2-p} - np\tilde{x}[i]}{(1-p)} = \frac{n}{(1-p)} \left( \frac{\tilde{x}[i]^{2-p}}{\sum_{i=1}^{n} \tilde{x}[i]^{2-p}} - p\,\tilde{x}[i] \right) \ . \tag{74}$$

Rather than determine the minimum value of $g(\tilde{x}, y)$ directly from (74), note that multiplication of (70) by $y_i$ followed by a summation over $i$ gives the necessary conditions,

$$\begin{aligned}
0 &= \sum_{i=1}^{n} \left\{ (1-p)\tilde{x}[i]^{p-2}y_i^2 + n\,p\,\tilde{x}[i]^{p-1}y_i \right\} - n\lambda\,, \\
0 &= \sum_{i=1}^{n} \left\{ (1-p)\tilde{x}[i]^{p-2}y_i^2 + 2np\tilde{x}[i]^{p-1}y_i \right\} - n\,p \sum_{i=1}^{n} \tilde{x}[i]^{p-1}y_i - n\lambda\,, \\
0 &= g(\tilde{x}, y) + (1+p)n^2|d_p^{(1)}(x)| - n\,p \sum_{i=1}^{n} \tilde{x}[i]^{p-1}y_i - n\lambda\,,
\end{aligned}$$

or

$$g(\tilde{x}, y) = n\,\lambda + n\,p \sum_{i=1}^{n} \tilde{x}[i]^{p-1}\,y_i - (1+p)\,n^2\,|d_p^{(1)}(x)|\,. \tag{75}$$

Evaluating (75) at $\lambda_{\text{opt}}$ and $y_{\text{opt}}$, and using the relationship (74) results in,

$$\begin{aligned}
g(\tilde{x}, y_{\text{opt}}) &= \frac{n}{(1-p)} \left( \lambda_{\text{opt}} - n\,|d_p^{(1)}(x)| \right) = \frac{n^2}{(1-p)} \left( \frac{1}{\sum_{i=1}^{n} \tilde{x}[i]^{2-p}} - |d_p^{(1)}(x)| \right) \\
&= \frac{n^2}{(1-p)} \left( \frac{1}{\sum_{i=1}^{n} \tilde{x}[i]^{2-p}} - \sum_{i=1}^{n} \tilde{x}[i]^{p} \right) \,, \\
&= \frac{n^2}{(1-p)} \left( 1 - \sum_{i=1}^{n} \tilde{x}[i]^{p} \cdot \sum_{i=1}^{n} \tilde{x}[i]^{2-p} \right) \cdot \sum_{i=1}^{n} \tilde{x}[i]^{2-p}\,.
\end{aligned} \tag{76}$$

Thus $g(\tilde{x}, y) \geq 0$ for all nonzero $x$ and for all $y$, and hence $\mathcal{H}_p^{(1)}(x)$ is negative semidefinite for all nonzero $x$ if and only if,

$$1 \geq \sum_{i=1}^{n} \tilde{x}[i]^{p} \cdot \sum_{i=1}^{n} \tilde{x}[i]^{2-p}\,. \tag{77}$$

On the other hand, as a simple consequence of the Cauchy-Schwartz inequality we have

$$\sum_{i=1}^{n} \tilde{x}[i]^{p} \cdot \sum_{i=1}^{n} \tilde{x}[i]^{2-p} \geq \left( \sum_{i=1}^{n} \tilde{x}[i]^{\frac{p}{2}} \cdot \tilde{x}[i]^{\frac{2-p}{2}} \right)^2 = \left( \sum_{i=1}^{n} \tilde{x}[i] \right)^2 = 1\,. \tag{78}$$

Equations (77) and (78) indicate that $\mathcal{H}_p^{(1)}(x)$ will be negative semidefinite for all nonzero $x$ if and only if,

$$\sum_{i=1}^{n} \tilde{x}[i]^{p} \cdot \sum_{i=1}^{n} \tilde{x}[i]^{2-p} \equiv 1\,. \tag{79}$$

This condition corresponds to equality in (78), which will occur if and only if the vectors $(\tilde{x}^{\frac{p}{2}}[1], \cdots \tilde{x}^{\frac{p}{2}}[n])$ and $(\tilde{x}^{\frac{2-p}{2}}[1], \cdots \tilde{x}^{\frac{2-p}{2}}[n])$ are parallel [49]. It is readily shown that equality holds whenever $\tilde{x}$ has $m$ zero entries with the remaining $n-m$ entries equal to the constant value $\frac{1}{(n-m)}$ for $m = 0, \cdots n-1$. Generally, however, equality does not hold, and there is one direction, oblique to both the direction $\mathbf{1}$ and the $(n-1)$-dimensional subspace $\mathbf{1}^{\perp}$, for which concavity of $d_p^{(1)}(x)$ is lost. Thus $d_p^{(1)}(x)$ is almost concave, but not concave, on $\mathcal{Q}_1$.

It is a straightforward application of L'hôpital's rule to show that

$$\lim_{p \to 1^-} \frac{1}{(1-p)} \left( 1 - \sum_{i=1}^{n} \tilde{x}[i]^p \cdot \sum_{i=1}^{n} \tilde{x}[i]^{2-p} \right) = -H_S^{(1)}(x)\,. \tag{80}$$

Thus in the limit we have,

$$p \to 1^- \Rightarrow g(\tilde{x}, y_{\text{opt}}) = -n^2 H_S^{(1)}(x) \le 0\,. \tag{81}$$

Note that $H_S^{(1)}(x) = H_S^{(1)}(\tilde{x})$ is equal to zero when $\tilde{x}$ has $m$ zero entries with the remaining $n-m$ entries equal to $\frac{1}{(n-m)}$ for $m = 0, \cdots n-1$. In general, however, we have $-H_S^{(1)}(x) < 0$. ■

**Proof of Theorem 15.** Because $d_{\text{sig}}^{(2)}(x) = \sum_{j=1}^{n} d_{p_j}^{(2)}(x)$, it is enough to show that the theorem is true for the simpler case of $d_{\text{sig}}^{(2)}(x) = d_p^{(2)}(x)$. Earlier in the appendix it was shown that

$$\frac{\partial}{\partial x[i]} d_p^{(2)}(x) = \frac{2|p|}{\|x\|_2^2} x[i] \left( \tilde{x}[i]^{p-1} - |d_p^{(2)}(x)| \right)\,, \quad \tilde{x} = |x|^2 / \|x\|_2^2\,. \tag{82}$$

Set $d(x) = d_p^{(2)}(x)$ for convenience. Let $x[i] > x[j]$, so that $x[i] = x[j] + \Delta$, $\Delta > 0$, and note that $x[i]$ and $x[j]$ must both be positive for $x \in \mathcal{Q}_1$. From (4), $d(x)$ is Schur-concave iff $\frac{\partial d(x)}{\partial x[i]} - \frac{\partial d(x)}{\partial x[j]} \le 0$. Note that

$$\begin{aligned} \frac{\partial d(x)}{\partial x[i]} - \frac{\partial d(x)}{\partial x[j]} &\propto x[i] \left( \tilde{x}[i]^{p-1} - |d(x)| \right) - x[j] \left( \tilde{x}[j]^{p-1} - |d(x)| \right) \\ &= -|d(x)|(x[i] - x[j]) + \left( x[i]\ \tilde{x}[i]^{p-1} - x[j]\ \tilde{x}[j]^{p-1} \right) = T_1 + T_2\,, \end{aligned} \tag{83}$$

where $T_1 = -|d(x)|(x[i] - x[j]) = -|d(x)|\Delta < 0$. The function $d(x)$ is Schur-concave iff $T_1 + T_2 \le 0$. Expanding the term $T_2$, we have

$$\begin{aligned} T_2 &= \frac{x[i]}{\tilde{x}[i]^{p-1}} - \frac{x[j]}{\tilde{x}[j]^{p-1}} = \frac{x[j] + \Delta}{\left( \frac{x[j]+\Delta}{\|x\|_2} \right)^{2(1-p)}} - \frac{x[j]}{\left( \frac{x[j]}{\|x\|_2} \right)^{2(1-p)}} \\ &= \frac{x[j]}{\tilde{x}[j]^{1-p}} \left\{ \left( 1 + \frac{\Delta}{x[j]} \right)^{2p-1} - 1 \right\} = \frac{x[j]}{\tilde{x}[j]^{1-p}} \left\{ \frac{1}{\left( 1 + \frac{\Delta}{x[j]} \right)^{1-2p}} - 1 \right\}\,. \end{aligned}$$

For $p \le \frac{1}{2}$, $T_2 \le 0$, and $d(x)$ is Schur-concave. For $p > \frac{1}{2}$, note that $T_2 > 0$. For $p > \frac{1}{2}$, $d(x) = \sum_{i=1}^{n} \tilde{x}[i]^p = d(\tilde{x})$ can be shown via the Hessian test to be strictly concave and therefore the problem $\max_{\tilde{x}} d(\tilde{x})$ subject to $\sum_i \tilde{x}[i] = 1$ has a unique solution, which can be shown from the method of Lagrange multipliers to be given by $\tilde{x}[i]^{\text{opt}} = n^{-1}$, $i = 1, \cdots, n$, yielding a maximum value of $d(\tilde{x}^{\text{opt}}) = n^{1-p}$. Thus $T_1$ Is bounded from below by $T_1 \ge -n^{1-p}\Delta$. Assume fixed values of $x[i]$ and $x[j]$, and note that $\Delta = x[i] - x[j]$ is then constant. Assume that $n > 2$, then there exists a component $x[k]$ other than $x[i]$ and $x[j]$. By letting $x[k] \to \infty$, we have $\|x\|_2 \to \infty$ so that $\tilde{x}[j]^{p-1} = \|x\|_2^{2(1-p)} x[j]^{2(p-1)} \to \infty$ for $p > \frac{1}{2}$. Thus $T_2$ can be made to dominate $n^{1-p}\Delta$ so that $T_1 + T_2 > 0$, in which case $d(x)$ cannot be Schur-concave.

■

**Proof of Theorem 16.** The proof is very similar to that of Theorem 14. Let $y = z \ne 0$ be any direction orthogonal to $\tilde{x}^{\frac{1}{2}}$, $z \perp \tilde{x}^{\frac{1}{2}}$. Then, $\tilde{x}^{\frac{T}{2}} z = z^T \tilde{x}^{\frac{1}{2}} = 0$ and

$$z^T \mathcal{H}_p^{(2)}(x) z = -\frac{2|p|}{\|x\|_2^2} \sum_{i=1}^{n} \left\{ \frac{1-2p}{\tilde{x}[i]^{1-p}} + |d_p^{(2)}(x)| \right\} z_i^2 < 0\,, \tag{84}$$

for all $p \le \frac{1}{2}$. This shows that $\mathcal{H}_p^{(2)}(x)$ is negative definite and therefore $d_{\text{sig}}^{(2)}(x)$ is locally strictly concave at the point $x$ along any direction perpendicular to $\tilde{x}^{\frac{1}{2}}$ for $p \le \frac{1}{2}$. Thus $d_{\text{sig}}^{(2)}(x)$ is almost strictly concave according to Definition 6 for $p \le \frac{1}{2}$. However, we will show that $d_{\text{sig}}^{(2)}(x)$, $p \le \frac{1}{2}$, is not concave.

Noting that

$$\tilde{x}^{\frac{T}{2}} \tilde{x}^{\frac{1}{2}} = \sum_i \tilde{x}[i] = 1\,, \tag{85}$$

it is readily shown that

$$\tilde{x}^{\frac{T}{2}} \mathcal{H}_p^{(2)}(x) \tilde{x}^{\frac{1}{2}} = 0\;. \tag{86}$$

Thus $\mathcal{H}_p^{(2)}(x)$ is negative semidefinite and $d_{\text{sig}}^{(2)}(x)$ is concave (but not strictly concave) along the direction $y = \tilde{x}^{\frac{1}{2}}$. It is evident that concavity can only be lost on a direction oblique to both the direction $\tilde{x}^{\frac{1}{2}}$ and the subspace $(\tilde{x}^{\frac{1}{2}})^{\perp}$.

Most generally, we let

$$y = \tilde{x}^{\frac{1}{2}} + z \quad \text{where} \quad z^T \tilde{x}^{\frac{1}{2}} = \sum_{i=1}^{n} z_i \tilde{x}^{\frac{1}{2}}[i] = 0\,. \tag{87}$$

Utilizing (86), (84), and (87) we have

$$\begin{aligned} y^T \mathcal{H}_p^{(2)}(x) y &= 2\tilde{x}^{\frac{T}{2}} \mathcal{H}_p^{(2)}(x) z - \frac{2|p|}{\|x\|_2^2} \sum_{i=1}^{n} \left\{ \frac{1-2p}{\tilde{x}[i]^{1-p}} + |d_p^{(2)}(x)| \right\} z_i^2 \\ &= -\frac{4|p|}{\|x\|_2^2} \sum_{i=1}^{n} \left\{ \frac{1-2p}{\tilde{x}[i]^{1-p}} + |d_p^{(2)}(x)| \right\} z_i \tilde{x}^{\frac{1}{2}}[i] - \frac{2|p|}{\|x\|_2^2} \sum_{i=1}^{n} \left\{ \frac{1-2p}{\tilde{x}[i]^{1-p}} + |d_p^{(2)}(x)| \right\} z_i^2 \end{aligned}$$

$$\begin{aligned}
&= -\frac{2|p|}{\|x\|_2^2}\sum_{i=1}^{n}\left[\left((1-2p)\tilde{x}[i]^{p-1}+|d_p^{(2)}(x)|\right)z_i^2+2(1-p)\tilde{x}[i]^{p-1}\tilde{x}^{\frac{1}{2}}z_i\right] \\
&= -\frac{2|p|}{\|x\|_2^2}g(\tilde{x},z)\,, \quad \text{where} \\
g(\tilde{x},z) &= \sum_{i=1}^{n}\left[\left((1-2p)\tilde{x}[i]^{p-1}+|d_p^{(2)}(x)|\right)z_i^2+2(1-p)\tilde{x}[i]^{p-1}\tilde{x}^{\frac{1}{2}}z_i\right]\,. \qquad (88)
\end{aligned}$$

If it can be shown that $g(\tilde{x},z)\geq 0$ for all $z$, then the Hessian is negative definite and $d_{\rm sig}^{(2)}(x)$ is concave on $\mathcal{Q}_1$. On the other hand if it can be shown that $g(\tilde{x},z)<0$ for some $z$, then $\mathcal{H}_p^{(2)}(x)$ is not positive definite and $d_{\rm sig}^{(2)}(x)$ is not concave on $\mathcal{Q}_1$. We will determine the precise situation that holds by computing the minimum admissible value of $g(\tilde{x},z)$ for any fixed value of $\tilde{x}$ in the interior of $\mathcal{Q}_1$.

To solve the problem,

$$\min_z g(\tilde{x},z) \quad \text{subject to} \quad \sum_{i=1}^{n} z_i\tilde{x}^{\frac{1}{2}}[i]=0\,, \qquad (89)$$

we will apply the method of Lagrange multipliers. It is readily ascertained via the Hessian test that $g(\tilde{x},z)$ is strictly convex in $z$ for $\tilde{x}$ in the interior of $\mathcal{Q}_1$ and therefore problem (89) has a unique solution. This solution is a stationary point of the Lagrangian,

$$\mathcal{L}=g(\tilde{x},z)-\lambda\sum_{i=1}^{n} z_i\tilde{x}^{\frac{1}{2}}[i]\,. \qquad (90)$$

The stationarity condition results in the $n+1$ equations

$$\begin{aligned}
2\left((1-2p)\tilde{x}[i]^{p-1}+|d_p^{(2)}(x)|\right)z_i+2(1-p)\tilde{x}[i]^{p-1}\tilde{x}^{\frac{1}{2}}-\lambda\tilde{x}^{\frac{1}{2}}[i] &= 0\,, \qquad (91)\\
\text{for} \quad i=1,\cdots,n \quad \text{and} \quad \sum_{i=1}^{n} z_i\tilde{x}^{\frac{1}{2}}[i] &= 0\,, \qquad (92)
\end{aligned}$$

which can be used to solve for $z\in\mathsf{R}^n$ and $\lambda\in\mathsf{R}$. Rather than do this, however, we can directly solve for the optimal value of $g(\tilde{x},z)$ by multiplying (91) by $z_i$, summing over $i$, and applying condition (92). This yields the condition,

$$\sum_{i=1}^{n}\left[2\left((1-2p)\tilde{x}[i]^{p-1}+|d_p^{(2)}(x)|\right)z_{{\rm opt},i}^2+2(1-p)\tilde{x}[i]^{p-1}\tilde{x}^{\frac{1}{2}}z_{{\rm opt},i}\right]=0\,. \qquad (93)$$

A comparison of (88) and (93) shows that the optimal value, $g(\tilde{x},z_{\rm opt})$, of $g(\tilde{x},z)$ is given by

$$g(\tilde{x},z_{\rm opt})=-\sum_{i=1}^{n}\left((1-2p)\tilde{x}[i]^{p-1}+|d_p^{(2)}(x)|\right)z_{{\rm opt},i}^2\,. \qquad (94)$$

It is readily determined from the conditions (91) that $z_{\rm opt}$ must be nonzero in general, and therefore $g(\tilde{x},z_{\rm opt})<0$. Thus the Hessian is *not* negative semidefinite and $d_{\rm sig}^{(2)}(x)$ is *not* concave on the interior of $\mathcal{Q}_1$.

Note that a relationship between $\lambda_{\text{opt}}$ and $z_{\text{opt}}$ can be found by multiplying (91) by $\tilde{x}^{\frac{1}{2}}[i]$, summing over $i$, and applying conditions (85) and (92). This yields,

$$\lambda_{\text{opt}} = 2(1-2p)\left\{|d_p^{(2)}(x)| + \sum_{i=1}^{n} \tilde{x}[i]^{p-1}\tilde{x}^{\frac{1}{2}} z_{\text{opt},i}\right\} . \tag{95}$$

Finally, note that (91) and (95) can be combined to determine the worst case values $z_{\text{opt},i}$, $i = 1, \cdots, n$. ■

**Derivation of the Scaling Matrices of Table 1.** In Table 1 notice that the expressions for $\mathit{\Pi}_{sig}(x)$, $\mathit{\Pi}^{(1)}_{sig}(x)$, and $\mathit{\Pi}^{(2)}_{sig}(x)$ directly follow from the definitions of $d_{sig(x)}$, $d^{(1)}_{sig}(x)$, and $d^{(2)}_{sig}(x)$ respectively. Also note that the indefiniteness of the scaling matrices $\mathit{\Pi}^{(1)}_p(x)$, $\mathit{\Pi}^{(2)}_p(x)$, $\mathit{\Pi}^{(1)}_{sig}(x)$, $\mathit{\Pi}^{(2)}_{sig}(x)$, $\mathit{\Pi}^{(1)}_S(x)$, and $\mathit{\Pi}^{(2)}_S(x)$ (which can be verified directly) follow from the scale invariance of their corresponding diversity measures, as discussed in Section 4.1. The expressions for the scaling matrices of $H^{(1)}_p$ and $H^{(2)}_p$ follow from the definitions (15) and (17) and the chain rule of differentiation.

The derivations of $\mathit{\Pi}_p(x)$ and $\mathit{\Pi}_G(x)$ from the definition (19) are straightforward. The expression for $\mathit{\Pi}^{(1)}_p(x)$ follows from (19) and the identity (48), while that for $\mathit{\Pi}^{(2)}_p(x)$ follows from (19) and the identity (54). The expressions for $\mathit{\Pi}^{(1)}_S(x)$, and $\mathit{\Pi}^{(2)}_S(x)$ follow from the definition (13), with the appropriate definition for $\tilde{x}$, and the use of the identities (48) and (54) respectively.

**TABLE 1**

| DIVERSITY MEASURE | | | SCALING MATRIX | |
|---|---|---|---|---|
| Type | Concave | Scale Invariant | Expression | Positive Semidefinite |
| $d_p(x)$ | for $p \leq 1$ | No | $\mathit{\Pi}_p(x) = \text{diag}\left(\frac{1}{\lvert x[i]\rvert^{2-p}}\right)$ | Yes |
| $d_p^{(1)}(x)$ | for $p \leq 1$ | Yes | $\mathit{\Pi}_p^{(1)}(x) = \text{diag}\left(\frac{1 - \lvert d_p^{(1)}(x)\rvert \; \tilde{x}[i]^{1-p}}{\tilde{x}[i]^{2-p}}\right) , \quad \tilde{x} = \frac{\lvert x\rvert}{\lVert x\rVert_1}$ | No |
| $d_p^{(2)}(x)$ | for $p \leq \frac{1}{2}$ | Yes | $\mathit{\Pi}_p^{(2)}(x) = \text{diag}\left(\frac{1 - \lvert d_p^{(2)}(x)\rvert \; \tilde{x}[i]^{1-p}}{\tilde{x}[i]^{1-p}}\right) , \quad \tilde{x} = \frac{\lvert x\rvert^2}{\lVert x\rVert_2^2}$ | No |
| $d_{sig}(x)$ | for $p_j \leq 1$ | No | $\mathit{\Pi}_{sig}(x) = \sum_j \lvert p_j\rvert \; \omega_j \; \mathit{\Pi}_{p_j}(x)$ | Yes |
| $d_{sig}^{(1)}(x)$ | for $p_j \leq 1$ | Yes | $\mathit{\Pi}_{sig}^{(1)}(x) = \sum_j \lvert p_j\rvert \; \omega_j \; \mathit{\Pi}_{p_j}^{(1)}(x)$ | No |
| $d_{sig}^{(2)}(x)$ | for $p_j \leq 1$ | Yes | $\mathit{\Pi}_{sig}^{(2)}(x) = \sum_j \lvert p_j\rvert \; \omega_j \; \mathit{\Pi}_{p_j}^{(2)}(x)$ | No |
| $H_G(x)$ | Yes | No | $\mathit{\Pi}_G(x) = \text{diag}\left(\frac{1}{\lvert x[i]\rvert^2}\right)$ | Yes |
| $H_S^{(1)}$ | for $p \leq 1$ | Yes | $\mathit{\Pi}_S^{(1)}(x) = \text{diag}\left\{-\frac{1}{\lvert\tilde{x}[i]\rvert}\left(H_S^{(1)}(x) + \log \tilde{x}[i]\right)\right\}$ | No |
| $H_S^{(2)}$ | for $p \leq \frac{1}{2}$ | Yes | $\mathit{\Pi}_S^{(2)}(x) = \text{diag}\left(-H_S^{(2)} - \log \tilde{x}[i]\right) , \quad \tilde{x} = \frac{\lvert x\rvert^2}{\lVert x\rVert_2^2}$ | No |
| $H_p^{(1)}$ | for $p \leq 1$ | Yes | $\mathit{\Pi}_p^{(1)}(x)$ | No |
| $H_p^{(2)}$ | for $p \leq \frac{1}{2}$ | Yes | $\mathit{\Pi}_p^{(2)}(x)$ | No |

Figure 1: Lorentz Curves. $S[k]$ versus $k$, $1 \leq k \leq n$, for population size $n = 15$ and normalized to $S[n] = 1$. The solid lines are the convex hulls of the plotted points. The minimum diversity (maximum inequality) curve is $\mathcal{L}_I$. Maximum diversity (equality) is given by curve $\mathcal{L}_E$. Curves $\mathcal{L}_X$ and $\mathcal{L}_Y$ correspond to $x \prec y$, i.e., $x$ represents greater diversity (equality) than $y$. Curves that intersect correspond to vectors that *cannot* be ordered via majorization (i.e., via Lorentz curves that are nested one inside the other).

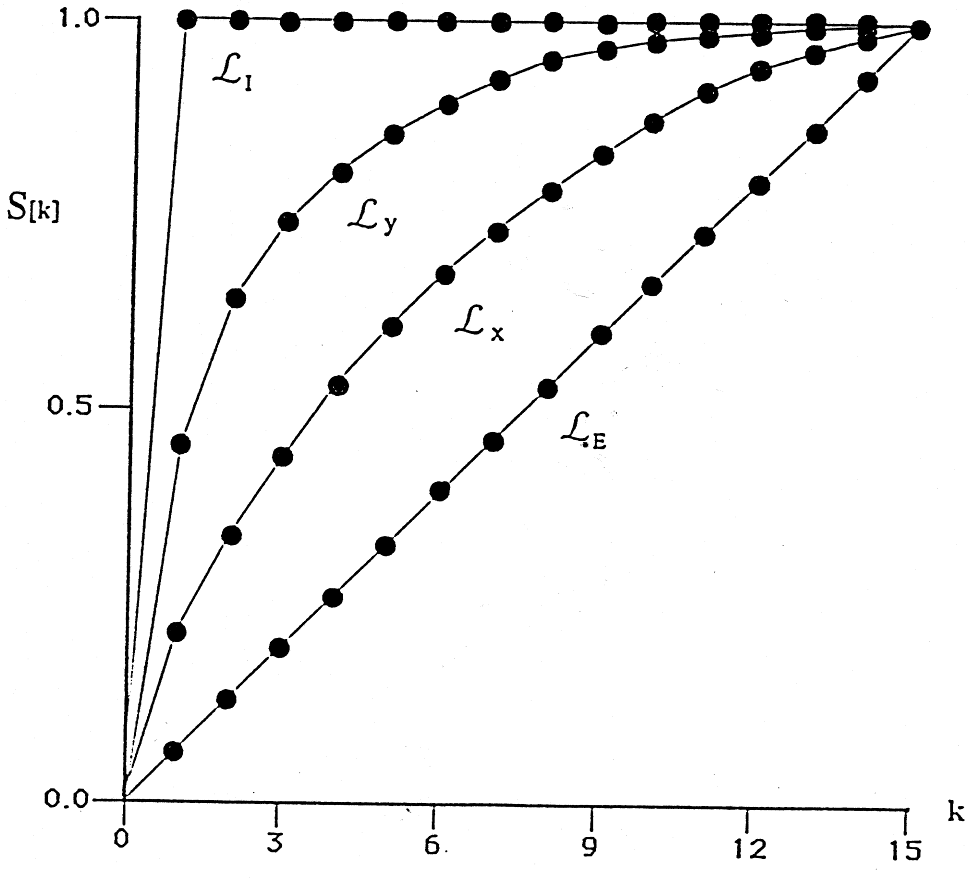


Figure 2: The Diversity Measure $d_p(x) = \mathrm{sgn}(p)\sum_{i=1}^{n} |x[i]|^p$, $p \le 1$. A) 2-D graph of $p = 1.0$; B) 2-D contour plot of $p = 1.0$; C) 2-D graph of $p = 0.1$; D) 2-D contour plot of $p = 0.1$.

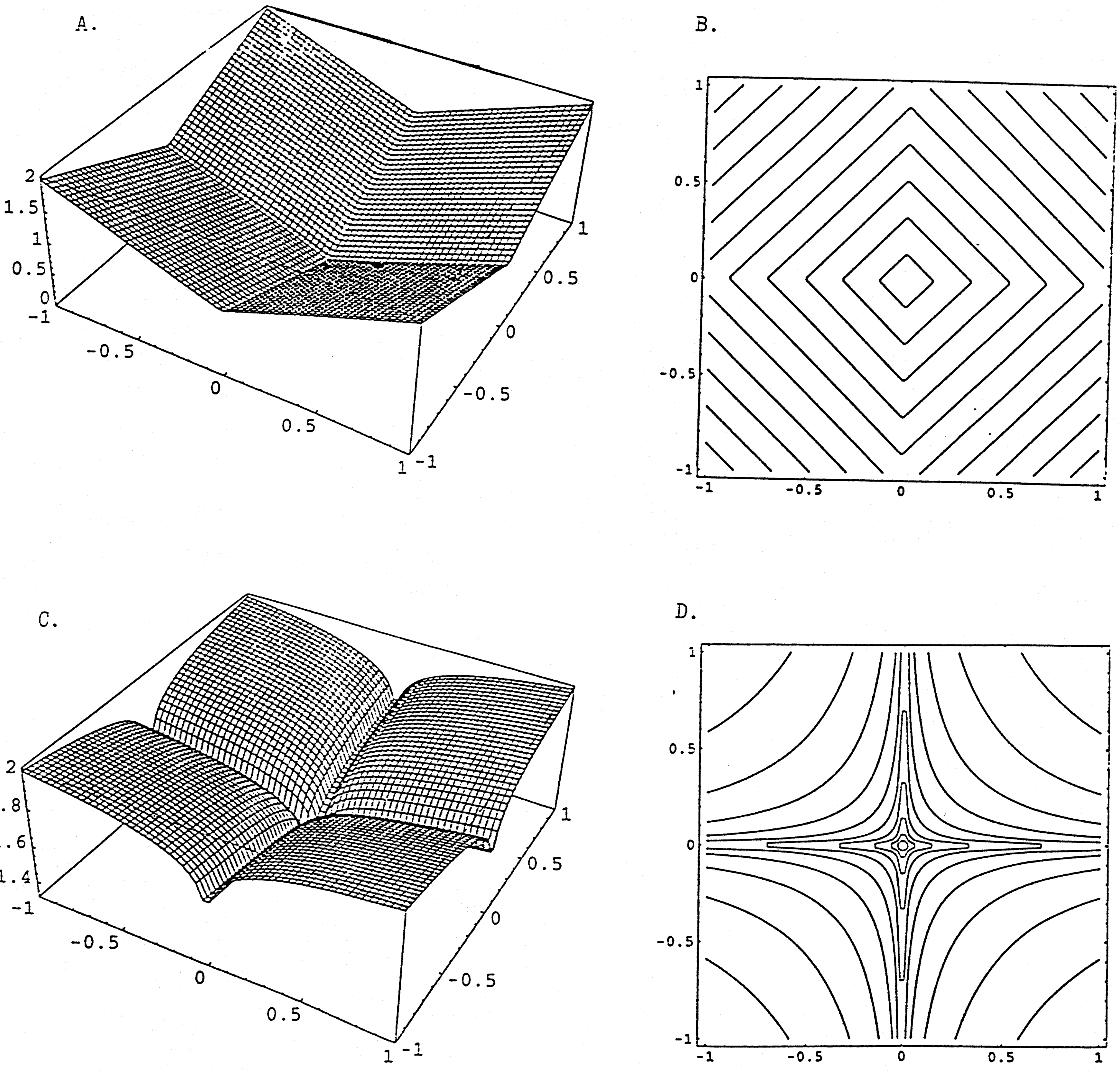


Figure 3: 2-Normalized Shannon Entropy, $H_S^{(2)}(x)$. A) $H_S^{2-\mathrm{norm}}(x)$ evaluated over any single orthant of $\mathrm{R}^3$. The value of the scale- and permutation-invariant measure $H_S^{(2)}(x)$ is given by the height of $H_S^{2-\mathrm{norm}}(x)$ above the unit simplex along the radial direction. Note that $H_S^{(2)}(x)$ is *not* concave over an orthant. B) Sign invariance of $H_S^{(2)}(x)$ with respect to changes in sign of the elements of $x$. The diversity measures considered in this paper are permutation invariant over any single orthant *and* sign invariant with respect to changes of the form $x[i] \rightarrow -x[i]$, $1 \leq i \leq n$. Shown is the graph of $H_S^{(2)}(x)$ over four of the eight orthants of $\mathrm{R}^3$ with values given by the height above the unit simplex (taken along the radial direction) in each orthant. The symmetry with respect to sign changes (i.e., with respect to change of orthant) is evident

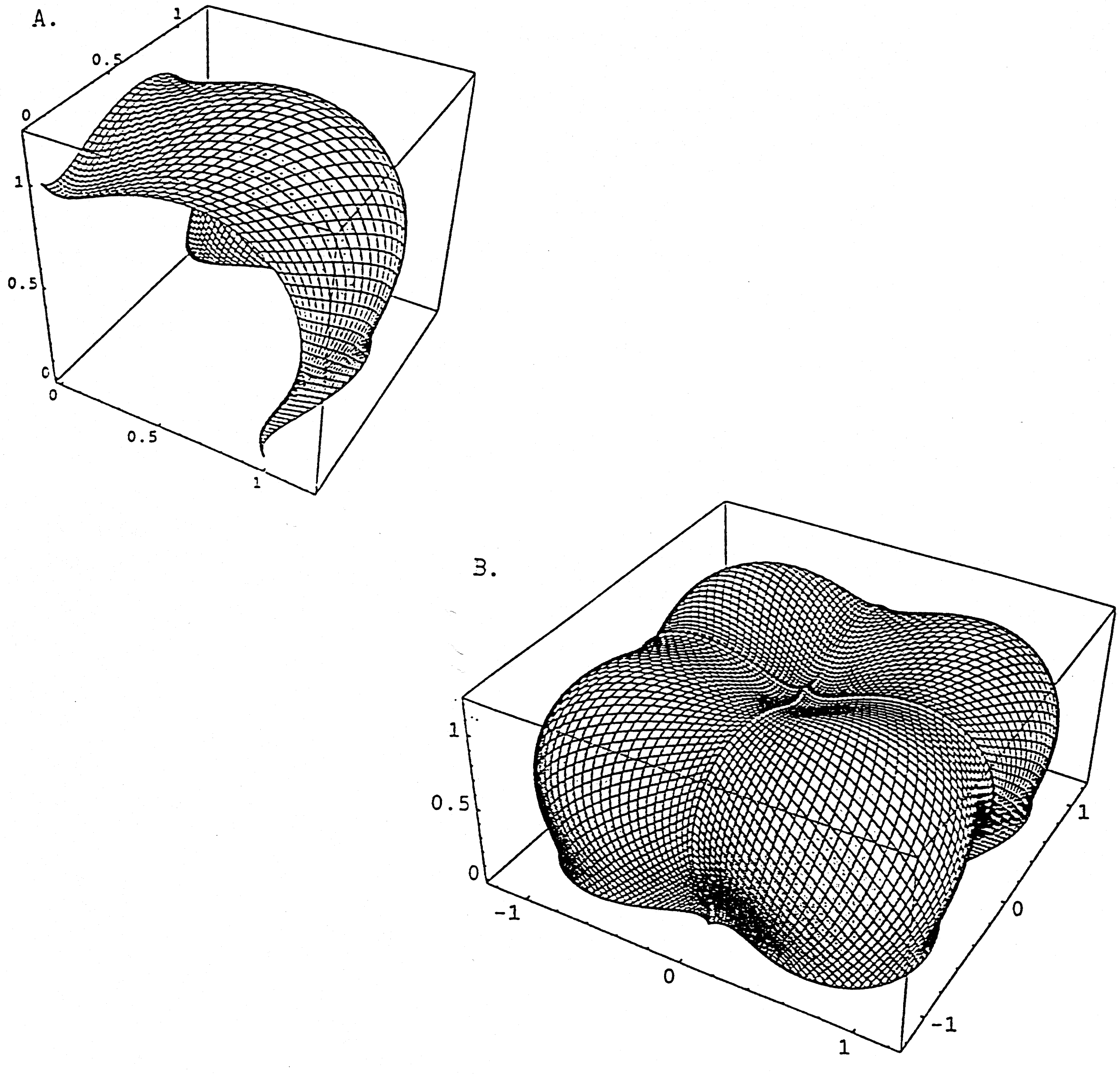
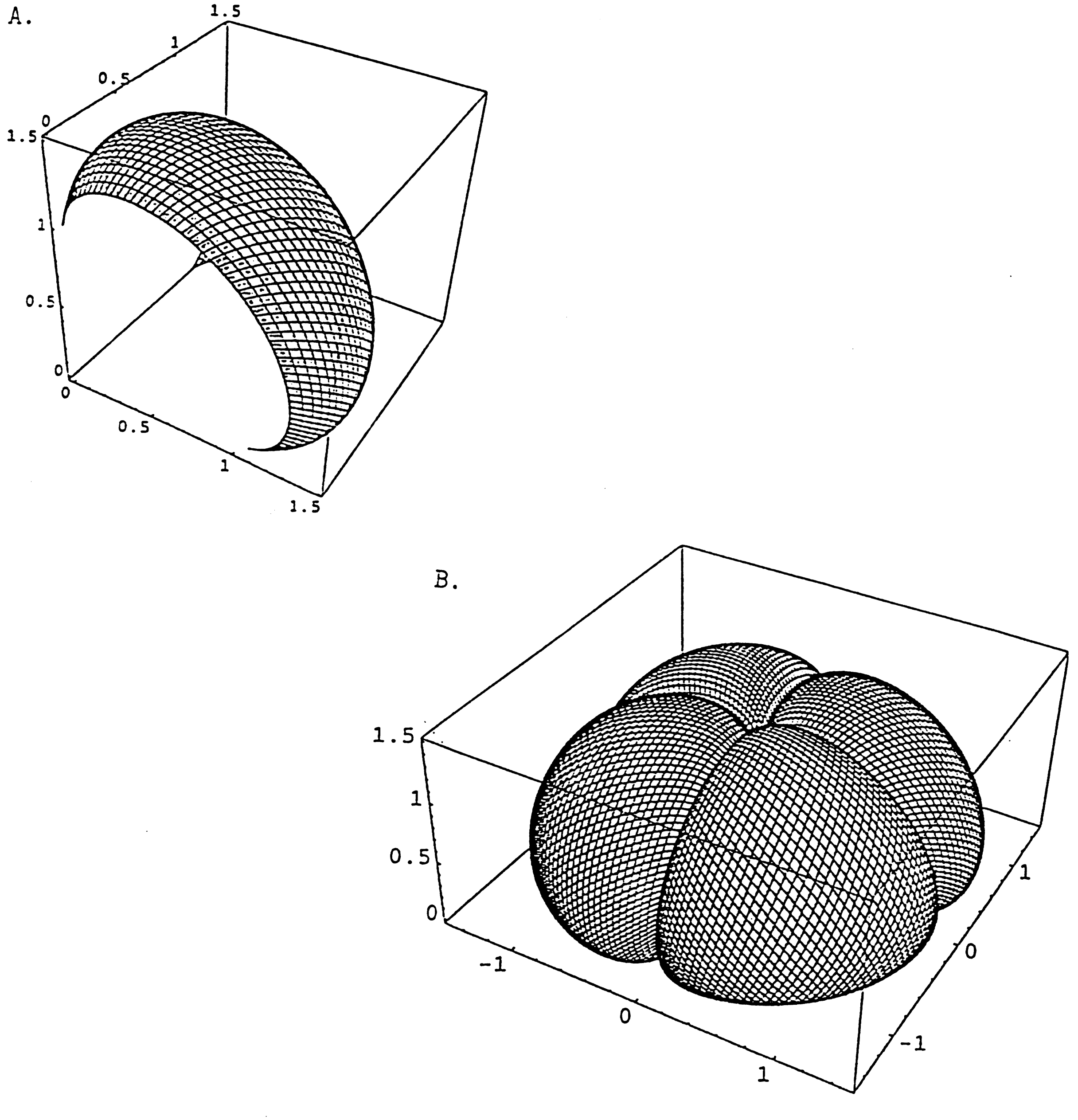


Figure 4: 1-Normalized Shannon Entropy, $H_S^{(1)}(x)$. A) $H_S^{1-\text{norm}}(x)$, evaluated over any single orthant of $\mathsf{R}^3$. The value of the scale- and permutation-invariant measure $H_S^{(1)}(x)$ is given by the height of $H_S^{1-\text{norm}}(x)$ above the unit simplex along the radial direction. Note that $H_S^{(1)}(x)$ is concave over an orthant. B) Sign invariance of $H_S^{(1)}(x)$ with respect to changes in sign of the elements of $x$. Shown is the graph of $H_S^{(1)}(x)$ over four of the eight orthants of $\mathsf{R}^3$ with values given by the height above the unit simplex (taken along the radial direction) in each orthant. The symmetry with respect to sign changes (i.e., with respect to change of orthant) is evident. Even though $H_S^{1-\text{norm}}(x)$ is almost concave over any single orthant, it is not almost concave across orthants.



# References

[1] J.M. Adler, B.D. Rao, and K. Kreutz-Delgado. "Comparison of Basis Selection Methods". In *Proceedings of the* $30^{th}$ *Asilomar Conference on Signals, Systems, and Computers*, November 1996.

[2] P.M. Alberti and A. Uhlmann. *Stochasticity and Partial Order: Doubly Stochastic Maps and Unitary Mixing.* D. Reidel, 1982.

[3] C.D. Aliprantis and O. Burkinshaw. *Principles of Real Analysis.* North Holland, New York, 1981.

[4] T. Ando. "Majorization, Doubly Stochastic Matrices, and Comparison of Eigenvalues". *Linear Algebra and Its Applications*, 118:163–248, 1989.

[5] B.C. Arnold. *Majorization and the Lorentz Order: A Brief Introduction.* Springer-Verlag, 1987.

[6] A.B. Atkinson. "On the Measurement of Inequality". *Journal of Economic Theory*, 2:244–63, 1970.

[7] D.J. Bartholomew. *The Statistical Approach to Social Measurement.* Academic Press, New York, 1996.

[8] M.S. Bazaraa and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms.* Wiley, 1979.

[9] E.F Beckenbach and R. Bellman. *Inequalities.* Springer-Verlag, 1971.

[10] C. Beightler and D.T. Phillips. *Applied Geometric Programming.* Wiley, New York, 1976.

[11] C. Beightler, D.T. Phillips, and D.J. Wilde. *Foundations of Optimization.* Prentice-Hall, Englewood Cliffs, N.J., second edition, 1979.

[12] H.P. Benson. "Deterministic Algorithms for Constrained Concave Minimization: A Unified Critical Survey". *Naval Research Logistics*, 43(6):765–95, September 1996.

[13] C. Berge. *Topological Spaces.* Oliver & Boyd, Edinburgh, 1963. Reprinted by Dover Publications, 1997.

[14] D.P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 1995.

[15] C. Blackorby and D. Donaldson. "Measures of Relative Equality and Their Meaning in Term of Social Welfare". *Journal of Economic Theory*, 18:59–80, 1978.

[16] S.D. Cabrera and T.W Parks. "Extrapolation and Spectral Estimation with Iterative Weighted Norm Modification". *IEEE Trans. on Acoustics, Speech and Signal Processing*, 39(4):842–851, April 1991.

[17] S. Chen and D. Donoho. "Basis Pursuit". In *Proceedings of the* $28^{th}$ *Asilomar Conference on Signals, Systems, and Computers*, volume I, pages 41–44, November 1994.

[18] E.S. Cheng, S. Chen, and B. Mulgrew. "Efficient Computational Schemes for the Orthogonal Least Squares Learning Algorithm". *IEEE Trans. on Signal Processing*, 43(1):373–376, January 1995.

[19] R.R. Coifman and M.V. Wickerhauser. "Entropy-Based Algorithms for Best Basis Selection". *IEEE Transactions on Information Theory*, IT-38(2):713–18, March 1992.

[20] N.E. Cotter. "The Stone Weierstrass Theorem and its Application to Neural Networks". *IEEE Transactions on Neural Networks*, 1(4):290–95, December 1990.

[21] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* Wiley, 1991.

[22] F.A. Cowell and K. Kuga. "Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement". *Mathematical Programming*, 25:131–43, 1981.

[23] P. Dasgupta, A. Sen, and D. Starret. "Notes on the Measurement of Inequality". *Journal of Economic Theory*, 6:180–87, 1973.

[24] D. den Hertog. *Interior Point Approach to Linear, Quadratic and Convex Programming: Algorithms and Complexity.* Kluwer Academic, 1994.

[25] P.A. Devijver and J. Kittler. *Pattern Recognition: A Statistical Approach.* Prentice Hall International, 1982.

[26] S. Dey and P.K. Varshney. "Inquality Measures for Fair Resource Allocation". In *Proceedings of the 1992 IEEE Conference on Systems, Man, and Cybernetics*, volume 2, pages 1539–44, New York, 1992. IEEE.

[27] D. Donoho. "On Minimum Entropy Segmentation". In C.K. Chui, L. Montefusco, and L. Puccio, editors, *Wavelets: Theory, Algorthms, and Applications*, pages 233–269. Academic Press, 1994.

[28] R.J. Duffin, E.L. Peterson, and C. Zener. *Geometric Programming–Theory and Application.* Wiley, New York, 1967.

[29] S.-C. Fang and S. Puthenpura. *Linear Optimization and Extensions: Theory and Algorithms.* Prentice Hall, 1993.

[30] G.S. Fields and J.C.H. Fei. "On Inequality Comparisons". *Econometrica*, 46(2):303–16, 1978.

[31] J.E. Foster. "An Axiomatic Characterization of the Theil Measure of Income Inequality". *Journal of Economic Theory*, 31:105–21, 1983.

[32] J.E. Foster. "inequality measurement". In H.P. Young, editor, *Fair Allocation*, volume 33 of *Proceedings of Symposia in Applied Mathematics*, pages 31–68. American Mathematical Society, 1985.

[33] M. Frank and P. Wolfe. "An Algorithm for Quadratic Programming". *Naval Research Logistics Quarterly*, 3:95–110, March 1956.

[34] K. Fukunaga. *Introduction to Statistical Pattern Recognition.* Academic Press, Boston, second edition, 1990.

[35] P.E. Gill, W. Murray, and M.H. Wright. *Practical Optimization.* Academic Press, 1981.

[36] I.F. Gorodnitsky, J.S. George, and B.D. Rao. "Neuromagnetic Source Imaging with FOCUSS: a Recursive Weighted Minimum Norm Algorithm". *Journal of Electroencephalography and Clinical Neurophysiology*, 95(4):231–251, October 1995.

[37] I.F. Gorodnitsky and B.D. Rao. "Sparse Signal Reconstruction from Limited Data Using FOCUSS: A Re-weighted Minimum Norm Algorithm". *IEEE Trans. Signal Processing*, 45(3):600–616, March 1997.

[38] G.H. Hardy, J.E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, second edition, 1959.

[39] G. Harikumar and Y. Bresler. "A New Algorithm for Computing Sparse Solutions to Linear Inverse Problems". In *Proceedings of ICASSP 96*, volume III, pages 1331–1334, Atlanta, Georgia, May 1996.

[40] P.E. Hart. "Entropy and Other Measures of Concentration". *Journal of the Royal Statistical Society*, 134:73–85, 1971.

[41] R. Horst and P.M. Pardalos. *Introduction to Global Optimization.* Kluwer Academic, 1995.

[42] R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches.* Springer-Verlag, second edition, 1993.

[43] A.A. Ioannides, J.P.R. Bolton, and C.J.S. Clarke. "Continuous Probabilistic Solutions to the Biomagnetic Inverse problem". *Inverse Problems*, pages 523–542, 1990.

[44] B. Jeffs and M. Gunsay. "Restoration of Blurred Star Field Images by Maximally Sparse Optimization". *IEEE Trans. on Image Processing*, 2(2):202–211, 1993.

[45] G. Jumarie. *Relative Information: Theories and Applications.* Springer-Verlag, 1990.

[46] H. Krim. "On the Distributions of Optimized Multiscale Representations". In *Proceedings of the 1997 International Conference on Acoustics and Signal Processing (ICASSP-97)*, April 1997.

[47] *Linear Algebra and its Applications*, March 1994. Special Issue Honoring Ingram Olkin.

[48] M.O. Lorentz. "Methods for Measuring Concentrations of Wealth". *Journal of the American Statistical Association*, 9:209–19, 1905.

[49] D.G. Luenberger. *Optimization by Vector Space Methods.* Wiley, Hayward, CA, 1969.

[50] D.G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley, second edition, 1984.

[51] P. Madden. *Concavity and Optimization in Microeconomics.* Basil Blackwell, 1986.

[52] S.G. Mallat and Z. Zhang. "Matching Pursuits with Time-Frequency Dictionaries". *IEEE Trans. ASSP*, 41(12):3397–415, 1993.

[53] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications.* Academic Press, 1979.

[54] H.M. Merrill. "Failure Diagnosis Using Quadratic Programming". *IEEE Trans. on Reliability*, R-22(4):207–213, October 1973.

[55] R.R. Meyer. "Sufficient Conditions for the Convergence of Monotonic Mathematical Programming Algorithms". *Journal of Conputer and System Sciences*, 12:108–21, 1976.

[56] A.J. Miller. *Subset Selection in Regression.* Chapman and Hall, 1990.

[57] D.S. Mitrinović. *Analytic Inequalities.* Springer-Verlag, 1970.

[58] K. Mosler. "Majorization in Economic Disparity Measures". *Linear Algebra and its Applications*, 199:91–114, March 1994.

[59] P.M. Narendra and K. Fukunaga. "A branch and Bound Algorithm for Feature Subset Selection". *IEEE Transactions on Computers*, C-26(9):917–922, September 1977.

[60] S.G. Nash and A. Sofer. *Linear and Nonlinear Programming.* McGraw-Hill, 1996.

[61] B.K. Natarajan. "Sparse Approximate Solutions to Linear Systems". *SIAM Journal on Computing*, 24(2):227–234, April 1995.

[62] D.M. Nicol, R. Simha, and D. Towsley. "Static Assignment of Stochastic Tasks Using Majorization". *IEEE Transactions on Computers*, 45:730–40, June 1996.

[63] P.M. Pardalos and J.B. Rosen. "Methods for Global Concave Minimization: A Bibliographic Survey". *SIAM Review*, 28(3):367–79, September 1986.

[64] G.P. Patil and C. Taillie. "Diversity as a Concept and its Measurement". *Journal of the American Statistical Association*, 77(379):548–67, 1982.

[65] J.E. Pečarić, F Proschan, and Y.L. Tong. *Convex Functions, Partial Orderings, and Statistical Applications.* Academic Press, 1992.

[66] E.C. Pielou. *Ecological Diversity.* Wiley, New York, 1975.

[67] E.C. Pielou. *Mathematical Ecology.* Wiley Interscience, second edition, 1977.

[68] P. Ram. *"A New Understanding of Greed".* PhD thesis, Department of Computer Science, University of California, Los Angeles, 1993.

[69] B.D. Rao and K. Kreutz-Delgado. *An Affine Scaling Methodology for Best Basis Selection.* Center For Information Engineering Report No. UCSD-CIE-97-1, January 1997. Submitted to "IEEE Trans. on Signal Processing".

[70] A. Renyi. "On Measures of Entropy and Information". In J. Neyman, editor, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 547–61, Berkeley, 1961. University of California Press.

[71] A. Renyi. *Probability Theory.* North-Holland, 1970.

[72] A.W. Roberts and D.E. Varberg. *Convex Functions.* Academic Press, 1973.

[73] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[74] M. Rothschild and J.E. Stiglitz. "Some Further Results on the Measurement of Inequality". *Journal of Economic Theory*, 6:188–204, 1973.

[75] A. Sen. *On Economic Inequality.* Clarendon/Oxford University Press, 1972.

[76] S. Singhal and B.S. Atal. “Amplitude Optimization and Pitch Prediction in Multipulse Coders”. *IEEE Trans. ASSP*, ASSP-37(3):317–327, March 1989.

[77] D.L. Solomon. “A Comparative Approach to Species Diversity”. In et al J.F. Grassle, editor, *Ecological Diversity in Theory and Practice*, pages 29–35. International Co-operative Publishing House, Fairland, Maryland, 1979.

[78] C. Taswell. “Satisficing Search Algorithms for Selecting Near-Best Bases in Adaptive Tree-Structured Wavelet Transforms”. *IEEE Transactions on Signal Processing*, 44(10):2423–38, October 1996.

[79] H. Theil. *Economics and Information Theory*. North-Holland, 1967.

[80] R. Webster. *Convexity*. Oxford University Press, 1994.

[81] M.V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software*. A.K. Peters, Wellesley, MA, 1994.

[82] W.I. Zangwill. *Nonlinear Programming: A Unified Approach*. Prentice-Hall, 1969.