Variable selection in saturated and supersaturated designs via ℓ_p-ℓ_q minimization

Alessandro Buccini* Omar De la Cruz Cabrera†
Christos Koukouvinos‡ Marilena Mitrouli§ Lothar Reichel¶

July 31, 2021

Abstract

In many real world problems it is of interest to ascertain which factors are most relevant for determining a given outcome. This is the so-called variable selection problem. The present paper proposes a new regression model for its solution. We show that the proposed model satisfies continuity, sparsity, and unbiasedness properties. A generalized Krylov subspace method for the practical solution of the minimization problem involved is described. This method can be used for the solution of both small-scale and large-scale problems. Several computed examples illustrate the good performance of the proposed model. We place special focus on screening studies using saturated and supersaturated experimental designs.

Keywords: variable selection, ℓ_p-ℓ_q minimization, nonconvex minimization, saturated design, supersaturated design, screening experiments.

*Department of Mathematics and Computer Science, University of Cagliari, Via Ospedale 72, 09124 Cagliari, Italy

†Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

‡Department of Mathematics, National Technical University of Athens, Zografou 15773, Athens, Greece

§Department of Mathematics, National and Kapodistrian University of Athens, University Campus, Zografou 15784, Athens, Greece

¶Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

1 Introduction

A crucial aspect in the design and analysis of experiments is the process of selection of important experimental factors, called *active factors*, and elimination of factors that have little to no effect on the response variable. This falls within the field of variable selection, which has been a topic of great and increasing interest in Statistics. In the experimental setting, this can take the form of a two-step process: a screening experiment, in which many factors are considered and only a few are selected for the second step, a confirmatory experiment. In the screening regime, the conclusions are not regarded as final, and it is expected that any purported discoveries will be further scrutinized in follow-up studies.

Thus, selection procedures are very important in screening experiments, where a large number of factors need to be considered but only a few are expected to influence the outcome. This is common in industrial experiments, but has become popular in the sciences as well. Typically, each experimental run is relatively expensive, and the goal is to produce a low-cost identification of a few dominating factors.

Experimental Design Terminology. We will now quickly define a handful of basic terms in the field of experimental design and analysis; readers familiar with the field can skip this part. A full development can be found in, e.g., [29].

Formally, a *factor* is a qualitative variable with 2 or more possible values, called *levels*. In practice, factors can be truly qualitative, but many times they are the result of picking a finite set of values (levels) out of a quantitative variable; often, there are 2 levels, called "low" and "high." For regression models, qualitative variables are usually encoded using 0-1 dummy variables; in experimental design, "low" often is encoded by -1 and "high" by 1. A *treatment* is a particular combination of levels, specifying one level for each of the factors considered. A *run* is a particular instance of the given experimental setup, using one treatment and resulting in one measurement of the response. If there are multiple runs using the same treatment, these are called *replicates*.

Then, an *experimental design* is a choice of treatments and (possibly) of replicates. A full *factorial design* uses all possible treatments; if there are replicate runs, it is called a *replicated factorial design*. Since the number of possible treatments grows exponentially with the number of factors considered, it is often useful to consider *fractions* of a factorial design, that is, using only a subset of all possible treatments; how to create good *fractional factorial designs* is a large and important field of statistical experimental design.

Once an experiment has been performed, i.e., a value of the response has been recorded for each run (together with the treatment used), several types of models can be fitted to the data. In this article we will concern ourselves only with linear models, in which $E(y) = X\beta$, where y is the vector of responses, X is the *design matrix* (or *model matrix*), whose columns are dummy variables for the factors, and rows correspond to run treatments, and β is the vector of coefficients to be estimated. We will use the terms *effect* and *coefficient* interchangeably. The coefficients corresponding to levels of one factor are called *main effects*, but *interactions* also can be included in the model: these are synthetic factors, derived from the original ones, which capture the non-additive part of the combined influence of two or more factors.

Saturated and Supersaturated Models. A model is said to be *saturated* (with respect to an experimental design) if it has as many coefficients to estimate as there are runs in the design; in this case X is square. A model is *supersaturated* if it has more coefficients than runs, i.e., X has more columns than rows. If there are k factors, each with 2 levels, the full factorial model will have 2^k runs, and the linear model with all interactions up to the k-way interaction has 2^k coefficients, resulting in a saturated model. On the other hand, the linear model without interactions has only $k+1$ coefficients (including the intercept); a suitable fractional factorial design can be used, and that can make the model saturated, or even supersaturated.

A well-known criterion for evaluating the efficiency of a supersaturated design (SSD) is the $E(s^2)$ criterion proposed in [4]. This criterion measures the average correlation between the factors; minimization of this value can be used as a criterion for designing good SSDs. The correlation between the factors and the departure from orthogonality are of high importance; they have significant influence on the detection of the true active factors. For more details on the construction and analysis of SSDs; see [19]. A comprehensive review on the construction and analysis of SSDs also is presented in [22].

The Regression Problem. We consider the problem

$$y = X\beta + \varepsilon, \qquad \varepsilon \sim N_n(0_n, \sigma^2 I_n), \tag{1}$$

where $X \in \mathbb{R}^{n\times(m+1)}$, $\beta \in \mathbb{R}^{m+1}$, $\sigma \in \mathbb{R}$, and $y, \varepsilon \in \mathbb{R}^n$. Further, 0_n and I_n denote the zero and identity matrices of order n, respectively. Since only few of the involved factors are expected to be important, we may assume β to be sparse. This work considers two-level supersaturated designs (SSDs), which are coded as -1 and 1 to denote the low and high levels, respectively. We assume the SSDs to be balanced, i.e., in each column of the matrix X, there are the same number of '-1's and '1's.

Standard regression techniques for fitting a linear prediction function using all candidate variables fail in the supersaturated design setting, because the normal equations cannot be solved uniquely. Specialized fitting and model selection techniques are required. Fan and Li [17] introduce a penalized likelihood approach, which employs symmetric nonconcave penalty functions on $(0,\infty)$ with singularities at the origin to produce sparse solutions, Candès and Tao [10] describe a method for selecting variables based on linear programming, and Georgiou [18] identifies the active factors by using the singular value decomposition of X. This paper considers the regression model

$$\beta^*_\lambda = \arg\min_\beta \left\{ \frac{1}{p} \|X\beta - y\|_p^p + \lambda \frac{1}{q} \|\beta\|_q^q \right\}, \tag{2}$$

with $0 < p, q \leq 2$, for the determination of the sparse active factors, where $z \mapsto \|z\|_p^p$ for $z = [z_1, z_2, \ldots, z_n]^T \in \mathbb{R}^n$ is defined as $\sum_{i=1}^n |z_i|^p$. Here and throughout this paper the superscript ^{T} denotes transposition. Note that $\|\cdot\|_p$ is a norm for $p \geq 1$, while it is not a norm for $0 < p < 1$ since the triangle inequality is not satisfied; see, e.g., [20]. We will refer to the mapping $z \mapsto \|z\|_p$ as a quasinorm when $0 \leq p < 1$. The model (2) appears in many applications in different areas, including numerical linear algebra [3, 39],

compressive sensing [11, 12, 15, 16, 28], and image restoration [24, 34]. The first term in the right-hand side of (2) is commonly referred to as the *fidelity term* and the second term as the *regularization term*.

We briefly discuss the choice of q. Since we assume β to be sparse, we would like to choose q to promote sparsity in the solution of (2). We may be tempted to consider using the ℓ_0-quasinorm $\|\cdot\|_0$ instead of the quasinorm $\|\cdot\|_q$ for some $0 < q < 1$, where the ℓ_0-quasinorm $\|\beta\|_0$ counts the number of nonzero entries of the vector β. However, minimizing the ℓ_0-quasinorm of a vector is very difficult. A common approach to overcome this complication is to approximate the ℓ_0-quasinorm by the ℓ_1-norm. The main advantage of this approximation is the convexity of the ℓ_1-norm, which makes the computation of a solution of (2) easier. However, ℓ_q-quasinorms with $0 < q < 1$ are better approximations of the ℓ_0-quasinorm. In particular, the smaller q, the better the approximation. On the other hand, using $0 < q < 1$ results in a nonconvex minimization problem; see Lanza et al. [23] for a recent discussion on the choice of q in the context of image restoration.

We turn to the choice of p. Since the p-norm measures the residual error, it follows that p should depend on the type of noise in the data y. For white Gaussian noise, it is common and appropriate to let $p = 2$. However, for other types of noise, $p = 2$ usually produces solutions of poor quality. It has been shown, see, e.g., [7, 21, 23], that in the context of image restoration, letting $0 < p < 1$ in (2) gives accurate restorations in the case of salt-and-pepper and other impulse noise.

Finally, we discuss the choice of $\lambda > 0$. This parameter determines the balance between the fidelity and regularization terms in (2). A small value of $\lambda > 0$ typically gives a solution with many non-vanishing coefficients, low bias (but high variance) and a closer fit to the data y (which in fact may be an overfit); meanwhile, when $0 < q < 1$ larger values of λ tend to yield a sparser solutions, biased toward zero (shrunken) with lower variance, and worse fit to the data y. Algorithms for the selection of this parameter are described in [7–9], but are not always appropriate for supersaturated designs, as there are too few observations for cross-validation. We will see that for the application considered in this paper, the solution of (2) is not very sensitive to the choice of λ, as wide ranges of values result in essentially equivalent solutions.

In this paper, we use the optimization problem (2) to fit the regression model (1). In other words, from the minimizer $\hat{\beta}$ of (2) we will select as active variables the ones that correspond to nonzero entries of $\hat{\beta}$. We first show, following [17], that this model satisfies continuity, sparsity, and unbiasedness properties. Then we review the fast and accurate numerical method for the solution of (2) proposed in [21]. This method employs generalized Krylov subspaces (GKS) to speed up the computations. The use of these subspaces allows us to analyze large data sets. To the best of our knowledge, this is the first time that GKS methods have been used in the context of variable selection for SSDs. Finally, we illustrate the performance of our approach when applied to some selected examples.

The present paper continues the investigation of the application of the minimization problem (2) to regression problems begun in [5]. We show new theoretical results and exploit properties of SDDs, which has not been done before. In particular, we show continuity of the solution.

This paper is structured as follows: Section 2 shows the continuity, sparsity, and unbi-

asedness properties of the regression model (2). In Section 3, we describe the algorithm used for the solution of the minimization problem (2). Section 4 reports numerical comparisons between our approach and other methods in the literature when applied to both synthetic and real data. Section 5 contains another example using real data, involving saturated and supersaturated models with interaction terms. Finally, we draw some conclusions and outline future work in Section 6.

2 Analysis of the Model

We discuss the theoretical properties of the minimization problem (2). In particular, we address the sparsity, stability, and unbiasedness of the model; see [17] for related discussions.

Sparsity and Unbiasedness. The first issues that we would like to consider are the sparsity inducing property and the unbiasedness of the method. These properties already have been discussed in [5]. We report here the main points of that discussion for the convenience of the reader.

We can derive the regression model (2) with a Bayesian argument as a maximum-a-posteriori estimation. Consider the generalized error distribution (GED) [35],

$$\xi_{\nu,\lambda,\sigma}(x) = c_{\nu,\sigma} \exp\left(-\frac{|x-\lambda|^\nu}{\nu\sigma^\nu}\right),$$

with $c_{\nu,\sigma} > 0$ chosen so that $\xi_{\nu,\lambda,\sigma}(x)$ is a probability distribution. We may construct a GED-based Bayesian version of the regression model (1). Let $\beta = [\beta_0, \beta_1, \ldots, \beta_m]^T$ be a random vector with prior distribution

$$\beta_j \overset{\text{i.i.d}}{\sim} f_\beta = f_{q,0,\sigma_\beta}, \qquad j = 0, 1, \ldots, m.$$

Assume that the $x_1, \ldots, x_n$ are either fixed or independent of β, and that the conditional density for y is $f_{y|\beta,x} = \xi_{p,x^T\beta,\sigma_y}$. After observing independent samples $y_1, \ldots, y_n$ with corresponding covariate vectors $x_1, x_2, \ldots, x_n$, the prior density for β is updated to the posterior as follows:

$$\begin{aligned} f_{\beta|y,X} &\propto \exp\left(-\frac{1}{p\sigma_y^p}\sum_{i=1}^n |y_i - x_i^T\beta|^p\right) \exp\left(-\frac{1}{q\sigma_\beta^q}\sum_{j=1}^m |\beta_j|^q\right) \\ &= \exp\left(-\frac{1}{p\sigma_y^p}||y - X\beta||_p^p - \frac{1}{q\sigma_\beta^q}||\beta||_q^q\right). \end{aligned}$$

Then the maximum-a-posteriori estimate for β is obtained by minimizing

$$\frac{1}{p}||y - X\beta||_p^p + \frac{\sigma_y^p}{\sigma_\beta^q}\frac{1}{q}||\beta||_q^q,$$

which is the optimization problem (2) with $\lambda = \sigma_y^p/\sigma_\beta^q$.

Summarizing, we have shown that the regression (2) can be obtained by assuming that the entries of β are realizations of a random variable with GED distribution. It is easy to see that, if $0 < q < 1$, this probability distribution will favor values that are close to 0, thus, leading to sparse solutions.

In [5] the authors furnish a deeper analysis in the one-dimensional case, showing both that the ℓ_p-ℓ_q method can be interpreted as a thresholding rule and that, if $|y|$ is large enough, then the method is unbiased. In more detail, let $m + 1 = n = 1$ and $X = 1$. Then (1) reduces to

$$y = \beta + \varepsilon,$$

with $\beta \sim \xi_{q,0,\lambda}$ and $\varepsilon \sim \xi_{p,0,1}$, independently. Then the following holds:

Proposition 1 ([5]). *Assume $0 < q < 1 < p \leq 2$, and consider the function $U : y \mapsto \hat{\beta}_{\text{lplq}}$ that associates to y the minimizer of $\mathcal{J}(\beta) = \frac{1}{p}|y - \beta|^p + \lambda\frac{1}{q}|\beta|^q$.*

1. *There exists $t > 0$ (which depends on p, q, and λ), such that U is constantly zero for $|y| < t$.*
2. *U has jump discontinuities at $-t$ and t.*
3. *For $|y| > t$, $U(y)$ has the same sign as y, and $|U(y)| < |y|$.*
4. *A lower bound for t is $\lambda^{1/(p-q)}$.*
5. *An upper bound for t is $(\lambda p/q)^{1/(p-q)}$.*
6. *As $|y| \to \infty$, $|y - U(y)| = |y| - |U(y)| \to 0$.*

In particular, point 1 shows that ℓ_p-ℓ_q regression is a thresholding method and point 6 shows that the method is unbiased. The following discussion sheds further light on the properties of the minimizer(s) of (2) and extends the analysis presented in [5].

Continuity. We turn to the continuity of the solution of the minimization problem (2), namely, that if the data y is slightly perturbed, then the minimizer(s) of (2) do(es) not change much. We establish this by showing that the minima of (2) are stable under perturbation. First, we observe the following:

Lemma 2. *Let $\{\beta_j\}_{j\in\mathbb{N}}$ be a sequence in $\mathbb{R}^{m+1}$ and let $q > 0$. If for all $j \in \mathbb{N}$, it holds that $\|\beta_j\|_q^q < c$ for a given $c > 0$, independent of β_j, then there exists a constant $\tilde{c} > 0$, depending only on c and q, such that $\|\beta_j\|_2^2 < \tilde{c}$ for all $j \in \mathbb{N}$.*

We omit the proof of this lemma since it follows immediately by the definition of the q-(quasi)norm. We are now in a position to show our main result.

Theorem 3. *Let $\{\varepsilon_j\}_{j\in\mathbb{N}}$ be a sequence of vectors in $\mathbb{R}^n$ such that $\varepsilon_j \to \varepsilon$ as $j \to \infty$. Let $\{y_j\}_{j\in\mathbb{N}}$ denote the sequence of vectors given by $y_j = X\beta + \varepsilon_j$, with X and β defined in* (1).

Let $y = X\beta + \varepsilon$. Then $y_j \to y$ as $j \to \infty$. Introduce the functionals

$$\begin{aligned} \mathcal{J}_j(\beta) &= \frac{1}{p}\,\|X\beta - y_j\|_p^p + \frac{\lambda}{q}\,\|\beta\|_q^q\,, \\ \mathcal{J}(\beta) &= \frac{1}{p}\,\|X\beta - y\|_p^p + \frac{\lambda}{q}\,\|\beta\|_q^q\,. \end{aligned} \tag{3}$$

Let $\{\beta_j\}_{j\in\mathbb{N}}$ be a sequence such that $\beta_j \in \arg\min_\beta \mathcal{J}_j(\beta)$ for all $j \in \mathbb{N}$, i.e., β_j belongs to the set of minimizers of $\mathcal{J}_j$. Then there exists a subsequence $\{\beta_{j_k}\}_{j_k\in\mathbb{N}}$ such that

$$\beta_{j_k} \to \bar{\beta} \quad as \quad k \to \infty \quad and \quad \bar{\beta} \in \overline{\arg\min_\beta \mathcal{J}(\beta)},$$

where $\overline{S}$ denotes the closure of the set S.

Proof. Since β_j is a minimizer of $\mathcal{J}_j$, it holds that

$$\mathcal{J}_j(\beta_j) \le \mathcal{J}_j(\beta) \quad \forall \beta \in \mathbb{R}^{m+1}. \tag{4}$$

For any fixed $\beta \in \mathbb{R}^{m+1}$, we have due to the continuity of $\mathcal{J}_j$ as a function of y_j, that $\mathcal{J}_j(\beta) \to \mathcal{J}(\beta)$ as $j \to \infty$. It follows that there is a constant c_β (which depends on β) and an index j_0, such that $\mathcal{J}_j(\beta) \le c_\beta$ for all $j \ge j_0$.

Fix $\beta \in \mathbb{R}^{m+1}$. Then for all $j \ge j_0$, we have by (4) that

$$c_\beta \ge \mathcal{J}_j(\beta) \ge \mathcal{J}_j(\beta_j) \ge \frac{\lambda}{q}\,\|\beta_j\|_q^q\,.$$

It follows from Lemma 2 that the sequence $\|\beta_j\|_2^2$, $k = 1, 2, \ldots$, is uniformly bounded and, therefore, has a convergent subsequence $\{\beta_{j_k}\}_{k=1}^\infty$. Let $\beta_{j_k} \to \bar{\beta}$ as $k \to \infty$. Then, for all $\beta \in \mathbb{R}^{m+1}$, it holds

$$\mathcal{J}(\bar{\beta}) \le \liminf_j \mathcal{J}_j(\beta_j) \le \limsup_j \mathcal{J}_j(\beta) = \lim_j \mathcal{J}_j(\beta) = \mathcal{J}(\beta),$$

where the second inequality follows from (4). In particular, let $\beta^* \in \arg\min_\beta \mathcal{J}(\beta)$. Then, by the inequality above, we have $\mathcal{J}(\bar{\beta}) \le \mathcal{J}(\beta^*)$. However, since β^* is a minimizer of $\mathcal{J}$, we also have that $\mathcal{J}(\bar{\beta}) = \mathcal{J}(\beta^*)$. Thus, $\bar{\beta}$ is a minimizer of $\mathcal{J}$. □

3 The Numerical Method

We outline the computational method for the solution of the minimization problem (2) that is used for the computed examples of Section 4. Further details can be found in [21]. First note that the functional $\mathcal{J}(\beta)$ defined in (3) is convex and smooth when $1 < p, q \le 2$, but nonconvex and nonsmooth when $0 < p < 1$ or $0 < q < 1$. The method of this section is shown in [21] to determine a minimizer of $\mathcal{J}(\beta)$ when $\mathcal{J}$ is convex, and a stationary point otherwise.

The iteratively reweighted norm (IRN) algorithm [33], also known as the iteratively reweighted least-squares (IRLS) algorithm [37], is among the most popular and effective methods for the solution of minimization problems of the form (2), when

$$0 < p, q \leq 2, \quad \text{with} \quad pq < 4. \tag{5}$$

The IRN algorithm is an iterative procedure in which the functional (3) is approximated by a sequence of weighted least-squares problems. At each iteration the diagonal weighting matrix is updated and each least-square problem is solved by the conjugate gradient (CG) algorithm; see [33]. We remark that the references [33,37] are concerned with $1 \leq p, q \leq 2$, but the method also can be used for p and q values that satisfy (5). The case $p = q = 2$ is excluded, because it is of no interest to us, and because the solution can be computed by a simpler method than the IRN algorithm.

The authors in [21,23] observed that the weighting matrices in the IRN method change fairly slowly with the iterations. This suggested that the solution of the sequence of least-squares problems be computed by a generalized Krylov subspace (GKS) method; see [21, 23]. Instead of solving each least-square problem separately, the solutions are determined in nested generalized Krylov subspaces of increasing dimension. Computed examples in [23] illustrate that this approach may require significantly fewer matrix-vector product evaluations than the IRN scheme [33], thereby making it cheaper. The use of generalized Krylov subspaces is well suited for large-scale problems. The computed examples of this paper illustrate that these spaces also can be applied when solving small problems.

The GKS methods proposed in [21, 23] are majorization-minimization (MM) iterative methods. At each iteration the minimization problem (3) is replaced by a convex quadratic problem. Specifically, the functional $\mathcal{J}$ is approximated by a quadratic functional that majorizes $\mathcal{J}$ and is tangent to $\mathcal{J}$ at the current iterate. The quadratic majorant is minimized to obtain the new iterate. Two quadratic majorization techniques are described in [21], one of them is the "adaptive aperture" method, which we use in the computations of the present paper.

We remark that other solution methods for the problem (2) have been described in the literature; see, e.g., [12,31,38,39]. It is outside the scope of the present paper to compare these methods. Here we only note that the method used allows flexibility in the choice of p and q, which makes it attractive for the problems considered in the present paper.

We outline the calculation of the quadratic tangent majorants used in the computations. Let $\mathcal{G}(\beta) : \mathbb{R}^{m+1} \to \mathbb{R}$ be a continuously differentiable functional. Then the functional $\mathcal{Q}(\beta, \gamma) : \mathbb{R}^{m+1} \times \mathbb{R}^{m+1} \to \mathbb{R}$ is said to be a quadratic tangent majorant for $\mathcal{G}(\beta)$ if and only if for any $\gamma \in \mathbb{R}^{m+1}$ the following conditions hold:

c1) $\mathcal{Q}(\beta, \gamma)$ is quadratic in β,

c2) $\mathcal{Q}(\gamma, \gamma) = \mathcal{G}(\gamma)$,

c3) $\nabla_\beta \mathcal{Q}(\gamma, \gamma) = \nabla_\beta \mathcal{G}(\gamma)$,

c4) $\mathcal{Q}(\beta, \gamma) \geq \mathcal{G}(\beta)$ for all $\beta \in \mathbb{R}^{m+1}$,

where ∇_β denotes the gradient with respect to the variable β.

Quadratic tangent majorants only can be determined for continuously differentiable functionals, cf. condition c3) above. We therefore smooth the functional $\mathcal{J}$ in (3) when either p or q are in the interval $]\,0,1\,]$. Introduce the function

$$\phi_{z,\theta}(t) = \left(\sqrt{t^2+\theta^2}\right)^z \quad \text{with} \quad \begin{cases} \theta > 0 \ \text{ for } \ 0 < z \leq 1, \\ \theta = 0 \ \text{ for } \ 1 < z \leq 2, \end{cases}$$

with smoothing parameter $\theta > 0$. Define the smoothed version of the minimization problem (2) by

$$\min_{\beta\in\mathbb{R}^{m+1}} \mathcal{J}_\theta(\beta), \qquad \mathcal{J}_\theta(\beta) = \frac{1}{p}\sum_{i=1}^{n} \phi_{p,\theta}\big((X\beta - y)_i\big) \ + \ \frac{\lambda}{q}\sum_{j=1}^{m+1} \phi_{q,\theta}\big(\beta_j\big)\,. \tag{6}$$

Let $\beta^{(k)}$ be the current approximation of the minimizer of (6). One can show that

$$\mathcal{Q}(\beta,\beta^{(k)}) = \frac{1}{2}\left\|(W_{\text{fid}}^{(k)})^{1/2}(X\beta - y)\right\|_2^2 + \frac{\lambda}{2}\left\|(W_{\text{reg}}^{(k)})^{1/2}\beta\right\|_2^2 + c, \tag{7}$$

is a quadratic tangent majorant for (6) at $\beta = \beta^{(k)}$ for a suitable constant c, which is independent of β. The matrices $W_{\text{fid}}^{(k)}$ and $W_{\text{reg}}^{(k)}$ are defined by

$$\begin{aligned} W_{\text{fid}}^{(k)} &= \text{diag}(w_{\text{fid}}^{(k)}), \quad w_{\text{fid}}^{(k)} = \left(\left(X\beta^{(k)} - y\right)^2 + \theta^2\right)^{p/2-1}, \\ W_{\text{reg}}^{(k)} &= \text{diag}(w_{\text{reg}}^{(k)}), \quad w_{\text{reg}}^{(k)} = \left(\left(\beta^{(k)}\right)^2 + \theta^2\right)^{q/2-1}, \end{aligned}$$

where all the operations are meant element-wise; see [21] for details. The minimizer $\beta^{(k+1)}$ of (7) is the next approximate solution of (6).

The normal equations associated with the minimization of (7) are given by

$$(X^T W_{\text{fid}}^{(k)} X \ + \ \lambda\, W_{\text{reg}}^{(k)})\,\beta \ = \ X^T \left(W_{\text{fid}}^{(k)}\right)^{1/2} y, \tag{8}$$

It follows that the minimizer $\beta^{(k+1)}$ of (7) or, equivalently, the solution of (8), is uniquely defined for all $\lambda > 0$ if

$$\text{Ker}\left(X^T W_{\text{fid}}^{(k)} X\right) \cap \text{Ker}\left(W_{\text{reg}}^{(k)}\right) \ = \ \{0\},$$

where $\text{Ker}(M)$ denotes the null space of the matrix M. This condition typically is satisfied. In particular it was satisfied in all our experiments.

To speed up the computations for large-scale problems, we determine approximations of the minimizers $\beta^{(k)}$, $k = 1, 2, \ldots$, in generalized Krylov subspaces. Let the matrix $V_k \in \mathbb{R}^{(m+1)\times d_k}$, with $1 \leq d_k \ll m$, have orthonormal columns. We seek to determine an approximation of the minimizer of (7) in the subspace determined by the columns of V_k. We also denote this minimizer by $\beta^{(k+1)}$. It is computed by evaluation

$$z^{(k+1)} = \arg\min_{z\in\mathbb{R}^{d_k}} \left\| \begin{bmatrix} \left(W_{\text{fid}}^{(k)}\right)^{1/2} X V_k \\ \lambda^{1/2}\left(W_{\text{reg}}^{(k)}\right)^{1/2} V_k \end{bmatrix} z - \begin{bmatrix} \left(W_{\text{fid}}^{(k)}\right)^{1/2} y \\ 0 \end{bmatrix} \right\|_2^2 \tag{9}$$

and letting

$$\beta^{(k+1)} = V_k z^{(k+1)}. \tag{10}$$

The vector $z^{(k+1)}$ is computed by substituting the QR factorizations

$$\begin{aligned} \left(W_{\text{fid}}^{(k)}\right)^{1/2} XV_k &= Q_1 R_1 \quad \text{with} \quad Q_1 \in \mathbb{R}^{n\times d_k}, \quad R_1 \in \mathbb{R}^{d_k\times_k}, \\ \left(W_{\text{reg}}^{(k)}\right)^{1/2} V_k &= Q_2 R_2 \quad \text{with} \quad Q_2 \in \mathbb{R}^{(m+1)\times d_k}, \quad R_2 \in \mathbb{R}^{d_k\times d_k} \end{aligned} \tag{11}$$

into (9). We recall that the matrices Q_1 and Q_2 have orthonormal columns, and the matrices R_1 and R_2 are upper triangular. Combining (11) and (9) yields

$$z^{(k+1)} = \arg\min_{z\in\mathbb{R}^{d_k}} \left\| \begin{bmatrix} R_1 \\ \lambda^{1/2} R_2 \end{bmatrix} z - \begin{bmatrix} Q_1^T y \\ 0 \end{bmatrix} \right\|_2^2.$$

Substituting (10) into (8) gives the residual vector

$$r = X^T \left(W_{\text{fid}}^{(k)}\right)^{1/2} \left(\left(W_{\text{fid}}^{(k)}\right)^{1/2} XV_k z^{(k+1)} - y \right) + \lambda W_{\text{reg}}^{(k)} V_k z^{(k+1)}.$$

We expand V_k by adding the scaled residual vector $v_{\text{new}} = r/\|r\|_2$ as a new column. Thus, $V_{k+1} = [V_k, v_{\text{new}}] \in \mathbb{R}^{(m+1)\times d_{k+1}}$, $d_{k+1} = d_k + 1$. Note that the vector v_{new} is orthogonal to the columns of the matrix V_k. The columns of V_{k+1} form an orthonormal basis for the expanded solution subspace. The so determined solution subspaces are referred to as *generalized Krylov subspaces*. We summarize the described procedure in Algorithm 1.

Algorithm 1 (MM-GKS)**.** *Let* $0 < p, q \leq 2$ *and* $\lambda > 0$. *Consider* $X \in \mathbb{R}^{(m+1)\times n}$. *Fix* $\theta > 0$.

$\beta^{(0)} = X^T y$;
$V_0 = \beta^{(0)} / \left\|\beta^{(0)}\right\|_2$;
for $k = 0, 1, \ldots$ **do**
  $w_{\text{fid}}^{(k)} = \left(\left(X\beta^{(k)} - y\right)^2 + \theta^2 \right)^{p/2-1}$;
  $W_{\text{fid}}^{(k)} = \text{diag}(w_{\text{fid}}^{(k)})$;
  $w_{\text{reg}}^{(k)} = \left(\left(\beta^{(k)}\right)^2 + \theta^2 \right)^{q/2-1}$;
  $W_{\text{reg}}^{(k)} = \text{diag}(w_{\text{reg}}^{(k)})$;
  Compute the QR factorizations $W_{\text{fig}}^{(k)} XV_k = Q_1 R_1$ *and* $W_{\text{reg}}^{(k)} V_k = Q_2 R_2$;
  $z^{(k+1)} = (R_1^T R_1 + \lambda R_2^T R_2^T)^{-1} (R_1^T Q_1 y)$;
  $r = X^T \left(W_{\text{fid}}^{(k)}\right)^{1/2} \left(\left(W_{\text{fid}}^{(k)}\right)^{1/2} XV_k z^{(k+1)} - y \right) + \lambda W_{\text{reg}}^{(k)} V_k z^{(k+1)}$;
  $v_{\text{new}} = r / \|r\|_2$;
  $V_{k+1} = [V_k, v_{\text{new}}]$;
  $\beta^{(k+1)} = V_k z^{(k+1)}$;
end

To improve the computational performance of Algorithm 1, we store the matrix XV_k at each iteration, so that the matrices $W_{\text{fig}}^{(k)} XV_k$ and $W_{\text{reg}}^{(k)} V_k$ can be computed very cheaply. Thus, the main cost of Algorithm 1 is the matrix-vector product evaluations in each iteration with the matrices X and X^T. Here we assume that $d_k \ll m$. Then the dominating computational work in each iteration with Algorithm 1 requires $O(nm)$ arithmetic floating point operations (flops). We observe that for large-scale problems this approach is much cheaper than the one proposed in [18], which requires the computation of a singular value decomposition of the matrix X at a cost of $O(m^2 n)$ flops.

4 Numerical Comparison

In this section we assess the performance of the proposed ℓ_p-ℓ_q method as a variable selection procedure in terms of its ability to identify the true active factors of an experiment. In the following examples all simulations are carried out using MATLAB.

4.1 An Illustrative Example

We first consider the example in [18, Section 2.1]. Here $X = [x_0, x_1, \ldots, x_{10}] \in \mathbb{R}^{6\times 11}$. Let

$$y_{\text{exact}} = 5x_0 + 4x_2 + 3x_5, \qquad y = 5x_0 + 4x_2 + 3x_5 + \varepsilon, \tag{12}$$

where ε is a "noise vector", whose components model white Gaussian noise. The desired solution, β_{exact}, is a sparse vector that satisfies the noise-free system of equations $X\beta_{\text{exact}} = y_{\text{exact}}$. The available data vector is

$$y = [-1.54, 12.02, 6.82, 12.44, 4.62, -1.21]^T.$$

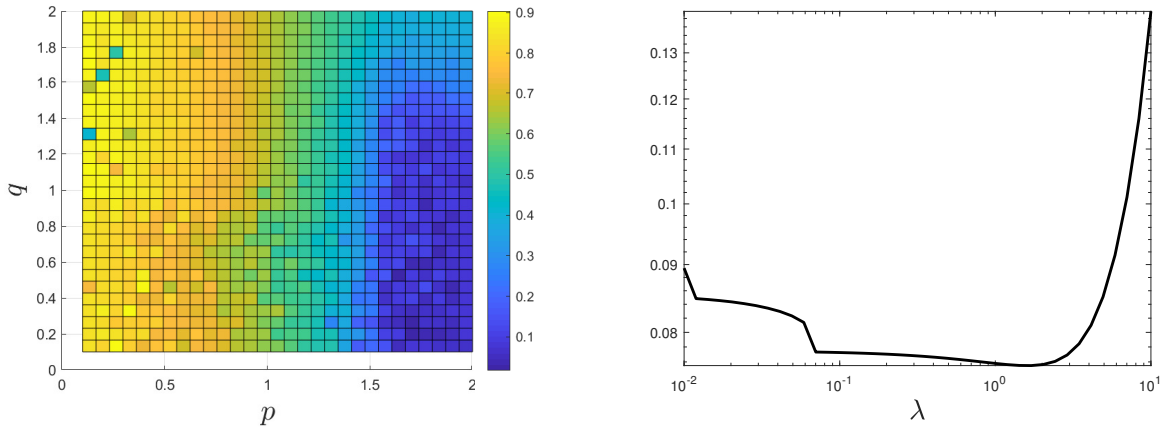
We first would like to confirm that the optimal choice for p is 2, and that the smaller the value of $q > 0$ is, the more accurate is the computed approximate solution. To this end we solve (2) for several values of p and q. For each pair of p- and q-values, we determine the optimal value of λ, i.e., the value that minimizes $\|\beta_\lambda^* - \beta_{\text{exact}}\|_2$, where β_λ^* denotes the solution of (2) and β_{exact} is the desired solution of the noise-free problem associated with (12). For each computed approximate solution β_λ^*, we evaluate the relative restoration error (RRE) defined by

$$\text{RRE}(\beta_\lambda^*) = \frac{\|\beta_\lambda^* - \beta_{\text{exact}}\|_2}{\|\beta_{\text{exact}}\|_2}.$$

Figure 1(a) displays the RRE for the different choices of p and q. It is evident by visual inspection of this image that the best choice is $p = 2$. Moreover, for fixed p, we can observe that the RRE decreases with q. This is in agreement with the theory developed in [5].

Let us now fix $p = 2$ and $q = 0.1$, and analyze how the computed restorations change with λ. We determine approximate solutions β_λ^* for several values of λ and compute the associated RREs. Figure 1(b) shows the RRE as a function of λ. We can observe that the error is stable with respect to different choices of λ, especially for small λ-values.

Variable selection seeks to determine which columns of X are the ones that contribute to the response y. Figure 2 shows reconstructions obtained for several values of λ. We



(a) (b)

Figure 1: Illustrative example in [18, Section 2.1]: (a) RRE for different choices of p and q. For each parameter pair (p, q), the parameter λ is chosen so that the RRE is minimal, (b) RRE of the computed solutions versus λ for $p = 2$ and $q = 0.1$.

observe that all the reconstructions, regardless of the choice of λ, correctly identify the columns x_0, x_2, and x_5 of X as active. For very small values of $\lambda > 0$, also some other columns are identified as active, however, the values of the associated coefficients β_j are very small.

Finally, we compare our reconstructed solutions to the one reported in [18]. To do so, for each of the three active columns, we compare the relative error of the single entries of β^*_λ for several choices of λ. Figure 3 shows the values of

$$\frac{|(\beta^*_\lambda)_j - (\beta_{\text{exact}})_j|}{|(\beta_{\text{exact}})_j|}, \quad j \in \{0, 2, 5\},$$

for different values of λ and compares them to the ones obtained in [18]. We can observe that, for most choices of λ our method is able to provide a better estimate of the coefficients β_j.

4.2 Random Examples

We would like to illustrate that the results shown in Section 4.1 are typical. To this end, we consider a set of 20 test cases randomly generated. Thus, we generate at random 20 matrices $X_i \in \mathbb{R}^{6\times 11}$ such that the first column is made of only ones and the other columns have 3 entries that are equal to -1 and 3 entries that are equal to 1. Moreover, we make sure that all the columns of each X_i are different. For all $i = 1, \ldots, 20$, we generated a vector $\beta^*_i \in \mathbb{R}^{11}$ as follows: Each vector β^*_i has 3 non-vanishing entries, the first one, and two picked at random between the other 10 entries. The nonzero entries of β^*_i are random integers between 1 and 10. We computed the right-hand sides y_i by

$$y_i = X_i \beta^*_i + \varepsilon_i,$$

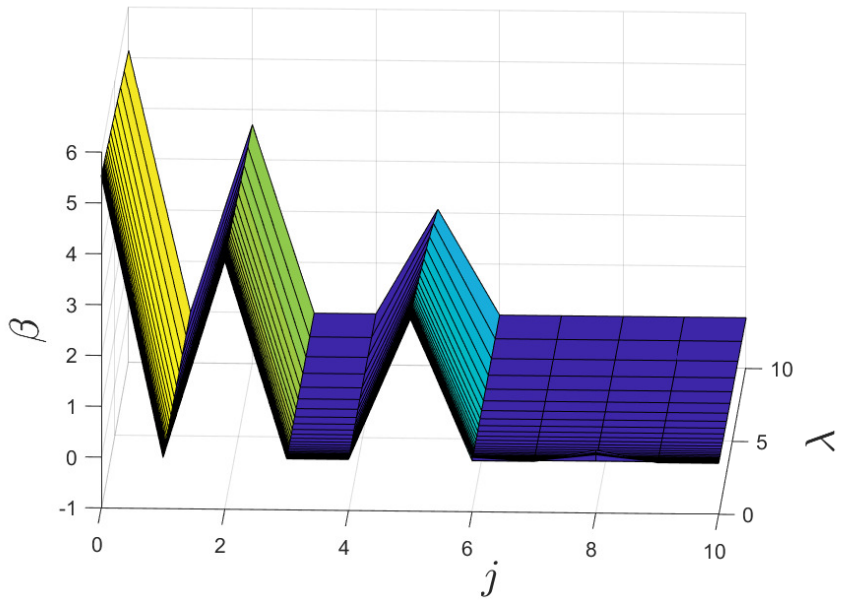
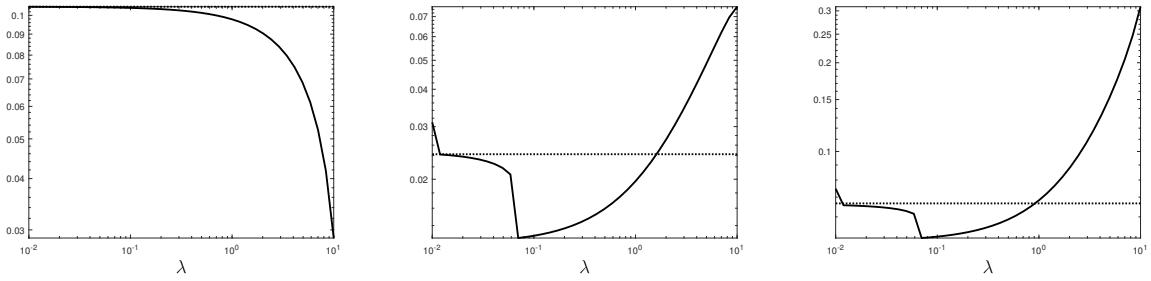
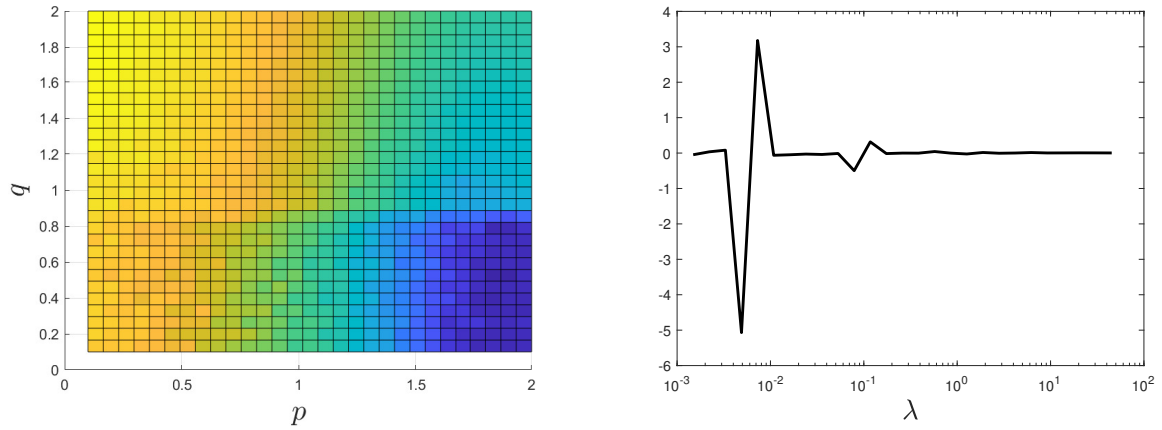


Figure 2: Illustrative example in [18, Section 2.1]: Reconstructions obtained with different λs with $p = 2$ and $q = 0.1$.



(a) (b) (c)

Figure 3: Illustrative example in [18, Section 2.1]: The solid curve represents the relative error of the single component obtained for different values of λ, the dotted curve is the relative error of the same component in the reconstruction computed in [18]. Panel (a) displays the 0th component, (b) the 2nd component, and (c) the 5th component.



(a) (b)

Figure 4: Aggregated results for the randomly generated examples of Section 4.2: (a) Averaged RREs for different choices of p and q, for each parameter pair (p, q), the parameter λ is chosen so that the RRE is minimal, (b) Averaged ∂RRE for $p = 2$ and $q = 0.1$.

where ε_i represents Gaussian noise, i.e., $\varepsilon_i \sim N_n(0_n, \sigma_i^2 I_n)$, where σ_i is chosen such that $\|\varepsilon_i\|_2 = 0.05 \|X_i \beta_i^*\|_2$.

For each example, we apply Algorithm 1 for several values of p, q, and λ. We first illustrate that, as already mentioned, the value of p should be determined by the statistical properties of the noise, and the choice of q should be informed by the desired sparsity of β_i^*. For each i and each p and q, we select the λ-value that produces the smallest RRE. We average the RREs obtained for each choice of p and q, and display the results obtained in Figure 4(a). We can clearly see that the best results are achieved when $p = 2$ (or very close to it) and q is as small as possible. This confirms our theoretical analysis of the roles of p and q.

We turn to the analysis of the behavior of the RRE with respect to the choice of λ. Our aim is to show the stability of our approach to the selection of the regularization parameter. Therefore, let $\mathrm{RRE}(\lambda)$ denote the RRE achieved for the parameter value λ. We would like to illustrate that $\mathrm{RRE}(\lambda)$ does not vary quickly with small changes of λ. To do this we approximate the derivative of $\lambda \to \mathrm{RRE}(\lambda)$ by a first order finite difference. Thus, let $\{\lambda_1, \ldots, \lambda_\ell, \lambda_{\ell+1}\}$ define a grid of logarithmically equispaced λ-values and define the finite differences

$$\partial\mathrm{RRE}(\lambda_j) = \frac{\mathrm{RRE}(\lambda_{j+1}) - \mathrm{RRE}(\lambda_j)}{\lambda_{j+1} - \lambda_j}, \quad j = 1, \ldots, \ell.$$

The analysis above suggest that we fix $p = 2$ and $q = 0.1$. We average all the ∂RRE-values in the considered examples and report the results in Figure 4(b). Note that $\partial\mathrm{RRE}(\lambda)$ is fairly small over large intervals of λ-values. This implies that our approach is fairly stable with respect of the choice of λ.

4.3 The William Data

Synthetic Cases. We now turn our attention to William's data [18,27]. This data set is composed of 10 synthetic models and a single real one and the matrix $X \in \mathbb{R}^{14\times 24}$. This data set has been studied in almost every paper that deals with the analysis of supersaturated designs and thus it acts as a base of comparison among the various proposed methods.

Since the results of the analysis of the effects of the parameters p, q, and λ are the same as for the previous example, we do not report them here. We first consider the synthetic cases and report them in Table 1. The only model that differs in a significant way from the other ones is the first one. In this model, since $y \sim 15x_1 + 8x_5 - 6x_9 + 3x_5x_9 + \varepsilon$, we see that it contains the interaction of factors x_5 and x_9 which is unbalanced. The product x_5x_9 is meant element-wise. To handle this case, we add the column x_5x_9 to the matrix X and add an unknown to the problem.

Our analysis above suggests that we may fix $p = 2$ and $q = 0.1$. Moreover, we have seen that the results obtained are fairly stable with respect to changes in λ. For the sake of simplicity, we therefore fix $\lambda = 10$. Small changes in λ do not affect the overall results that we are going to discuss.

To acquire more data, we simulate 1000 data vectors y with different realizations of ε, and compute both the average of Type I and Type II errors. The Type I error is the ratio of the number of zero entries of β_{exact} that are incorrectly set to a nonzero value by the method and the correct number of zero entries, while the Type II error is the ratio of the number of nonzero entries of β_{exact} that are incorrectly set to zero by the method and the correct number of nonzero entries. Thus,

$$\text{Error Type I}(\beta) = \frac{|\{j : \beta_j \neq 0, (\beta_{\text{exact}})_j = 0\}|}{|\{j : (\beta_{\text{exact}})_j = 0\}|},$$
$$\text{Error Type II}(\beta) = \frac{|\{j : \beta_j = 0, (\beta_{\text{exact}})_j \neq 0\}|}{|\{j : (\beta_{\text{exact}})_j \neq 0\}|}.$$

Table 2 reports the results obtained with our method and with six other methods described in the literature. We observe that, except for the very first model, we obtain excellent results with near-zero errors in all the considered cases. However, the first model turns out to be challenging for our method. This may be due to the fact that the exact solution β_{exact} is close to the null space of the matrix X. To see this, we compute

$$\left\|(I - X^\dagger X)\beta_{\text{exact}}\right\|_2 / \left\|\beta_{\text{exact}}\right\|_2, \tag{13}$$

where $X^\dagger$ denotes the Moore-Penrose pseudoinverse of X. Thus, $(I - X^\dagger X)$ is the orthogonal projector onto the null space of X. The quantity (13) is reported in Table 1 for all models. Observe that for the first model, this quantity is closer to 1 (its maximum value) than for the other models. We remark that models that include interactions of the factors usually are not considered in the analysis of supersaturated designs.

A Real Data Model. We now consider a real data model, which has been studied in many papers that deal with the analysis of SSDs. Obviously, we do not have access to β_{exact},

Table 1: William's data synthetic models. We note that the product x_5x_9 in Model 1 is meant element-wise. In all models $\varepsilon \sim N(0_{14}, I_{14})$.

Model no.	Model	$\frac{\left\|(I - X^\dagger X)\beta_{\text{exact}}\right\|_2}{\left\|\beta_{\text{exact}}\right\|_2}$
1	$y \sim 15x_1 + 8x_5 - 6x_9 + 3x_5x_9 + \varepsilon$	0.7923
2	$y \sim 8x_1 + 5x_{12} + \varepsilon$	0.5811
3	$y \sim 20x_1 + \varepsilon$	0.7372
4	$y \sim 20x_2 + 20x_7 + \varepsilon$	0.6921
5	$y \sim 14x_2 + 20x_7 + 20x_{16} + \varepsilon$	0.5427
6	$y \sim 5x_1 + 5x_2 + \varepsilon$	0.5145
7	$y \sim 5x_1 + 5x_2 + 5x_3 + \varepsilon$	0.5700
8	$y \sim 5x_1 + 5x_2 + 5x_3 + 5x_4 + \varepsilon$	0.4048
9	$y \sim 5x_1 + 5x_2 + 5x_3 + 5x_4 - 5x_5 + \varepsilon$	0.4954
10	$y \sim 10x_1 + 9x_2 + 2x_3 + \varepsilon$	0.5181

Table 2: William's data synthetic models. Type I and II errors obtained with our approach and with the methods described in References [1,2,18,26,27,36]. In [18] the method is tuned so that the Type I and II errors are the same. Therefore, we report them together. When we write 0.00, we mean that the first two decimals equal 0, but the number is nonzero, while when we write 0, we mean that we have an exact 0.

Model no.	Error for $\ell_p - \ell_q$		Error in References		Error in [18] for an opt. signif. level $\alpha = 4\%$
	Type I	Type II	Type I	Type II	Type I and II
1	0.16	0.59	0.20	0.12	0.01
2	0	0	0.06	0.07	0.04
3	0.00	0	0.08	0.14	0.00
4	0	0	0.01	0.01	0.00
5	0	0	0.00	0.00	0.00
6	0.00	0	0.09	0.18	0.02
7	0.00	0	0.02	0.11	0.01
8	0	0	0.08	0.10	0.00
9	0	0	0.11	0.17	0.00
10	0.00	0.33	0.12	0.22	0.26

however, other studies suggest that very few entries of β_{exact} should be non-vanishing. In particular, the 15th entry should be nonzero. Observing the data vector y suggests that all its entries are integers. This immediately implies that the noise that corrupts the data y is not Gaussian. Therefore, we cannot expect the choice $p = 2$ to give satisfactory results. We therefore set $p = 0.8$ in the computations. Similarly as above, we let $q = 0.1$ to promote sparsity of the computed solution. We minimize (2) for several values of λ and analyze the results. The solution β is expected to be sparser for larger values of λ. Moreover, we expect the most important components of β to be non-vanishing also for large values of λ. Figure 5 shows the computed solutions β for several values of λ, as well as the absolute values of the components of the computed β (excluding the first component). We observe that the first displayed component is larger than all the other ones and is nonzero for all choices of λ. From the entries shown in Figure 5(b), we can see that the 15th entry of β appears in most of the computed solutions. This entry is seen to be significantly larger than the other displayed entries. This suggests that the three most important entries of β are the first, the 15th and the 17th ones. This result is in agreement with the Dantzig selector method [10], which identifies as important factors 15 and 17. Also in [18] the factor 15 is identified as the most influential one.

Finally, we determine the 3 entries of β that are nonvanishing for most of the considered values of λ. We consider ℓ values of λ, denoted by $\lambda_1, \ldots, \lambda_\ell$. For each λ_j, the minimization of (2) produces an approximate solution $\beta^{(j)} \in \mathbb{R}^{24}$. For each index $i = 1, \ldots, 24$, we compute

$$\hat{\iota} = |\{j : \beta_i^{(j)} \neq 0\}|,$$

i.e., the number of times that the ith component is nonzero in the reconstructions $\beta^{(j)}$. We then determine the indices with maximum index $\hat{\iota}$. We report in Table 3 the ratio $\hat{\iota}/\ell$ and the averaged value assumed by each entry (considering only nonzero entries), i.e.,

$$\bar{\beta}_i = \frac{\sum_{j=1}^{\ell} (\beta_j)_i}{\hat{\iota}}.$$

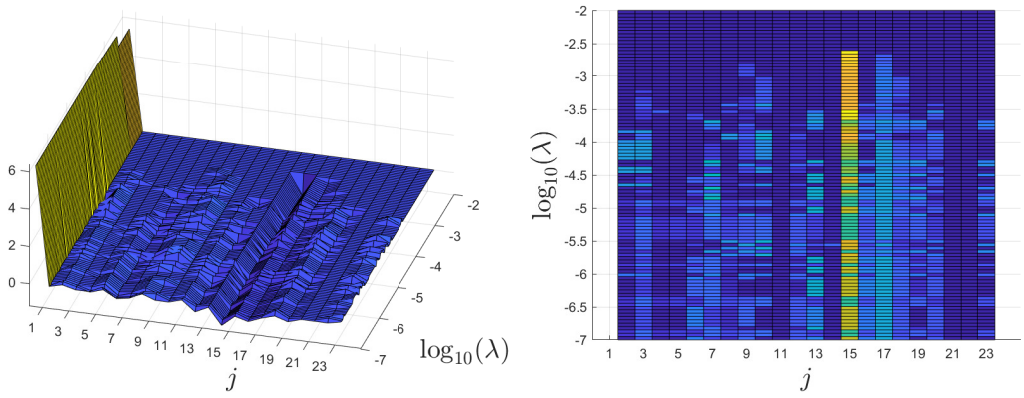
In our experiments, we set $\ell = 100$ and let the λ_j be logarithmically spaced between 10^{-7} and 10^{-2}.

We can observe that the first entry is always nonvanishing and that the 15th entry is nonzero in 87% of the computed reconstructions. In fact, from Figure 5(b), we can observe that the 15th entry vanishes only for very large values of λ, i.e., when we require the solution to be extremely sparse. Our analysis shows that the 17th entry is relevant as it is nonvanishing for 86% of the cases, even though, the magnitude of β_{17} is smaller than the magnitude of β_{15}.

Additional examples in [27]. In [27] the authors proposed three additional examples with the William matrix. In particular they consider three cases:

Case I: One active factor, $\beta_1 = 10$, and the other entries are equal to zero;

Case II: Three active factors, $\beta_1 = -15$, $\beta_5 = 8$, $\beta_9 = -2$, and all the other entries vanish;



(a) (b)

Figure 5: William's test data, real data case: (a) computed β for different values of λ, (b) absolute values of the computed β_j, for $j = 2, 3, \ldots, m+1$, for different values of λ.

Table 3: William's test data, real data case: 3 most frequent nonvanishing entries in the computed reconstructions. For each component we show the averaged values assumed (counting only the non zero instances) and the ratio $\hat{\imath}/\ell$, where $\hat{\imath}$ is the number of times that the coordinate i is nonvanishing and ℓ is the number of values of λ tested.

	Intercept	β_{15}	β_{17}
Averaged value	6.2153	-0.8520	-0.4000
$\hat{\imath}/\ell$	1	0.87	0.86

Table 4: William's data additional examples in [27]. Type I and II errors obtained with our approach. When we write 0.00, we mean that the first two decimals equal 0, but the number is nonzero, while when we write 0, we mean that we have an exact 0. The results are obtained averaging 1000 realizations of the noise.

Case	Error Type I	Error Type 2
I	0	0
II	0.00	0.33
III	0.00	0.20

Case III: Five active factors, $\beta_1 = -15$, $\beta_5 = 12$, $\beta_9 = -8$, $\beta_{13} = 6$, $\beta_{17} = -2$, and other components of β equal 0.

Our method is applied with $p = 2$, $q = 0.1$, and $\lambda = 10$. We generate 1000 realizations of white Gaussian noise with variance 1, and recover β for each realization. The Errors of type I and II are computed as well as their averages. Table 4 reports the results obtained. We observe that our method never identifies a vanishing entry of β as nonvanishing, however, it fails to identify the vanishing entries β in the cases II and III. In both cases, our method fails to identify the smallest nonvanishing entry (in absolute value) of β and sets it to 0. This behavior is consistent with the one described in [18].

5 Pollutant Absorption Example: Screening for Interaction Effects with Saturated and Supersaturated Designs

Here we use data from a study on pollutant absorption by moss by Cesa et al. [14]. Bags filled with moss are used in biomonitoring of water pollution, by submerging them at sites of interest for a fixed interval of time, and then measuring the amount of pollutant absorbed; knowing the uptake rate for each element, and how the presence of other pollutants can increase or decrease such rates, is crucial for the successful application of this monitoring method. The authors of [14] selected 11 chemical elements: Al, As, Cd, Cr, Cu, Fe, Hg, Mn, Ni, Pb, and Zn. Their goal was to assess element uptake based on the concentration in the water of the element itself, as well as interaction effects caused by the other elements, that is, antagonism/competition (absorption of the element is reduced by high concentrations of another element) or synergism (absorption is improved). They fitted 11 different linear models, each one having the uptake of a different element as the response, but all with the same design matrix, which we describe now.

Regarding each chemical element as a factor with two levels (high and low), the researchers used a Rechtschaffner experiment design [32]: one run with all factors at the low level; 11 runs with all but one factor at the high level; and 55 runs with all but two factors at the low level. With this design, the linear model with first order interactions (which has

67 parameters: intercept, 11 main effects, and 55 interaction effects) is a saturated model; the corresponding design matrix X is of size 67×67.

The coefficient estimates in [14] are given by $X^{-1}y$, where y is the response (element uptake measurements), leaving zero residual degrees of freedom. This is a common approach; the authors applied Lenth's method [25] to classify each one of the 66 factors (i.e., leaving out the intercept) as *inactive*, *possibly active*, and *probably active*. (Lenth's method is based on comparing the magnitude of each fitted coefficient with some particular multiples of robust estimates of spread of the 66 fitted coefficients.) They concluded, for example, that aluminum (Al) uptake has "probably active" factors Al and Al:Cu, with positive effects, "probably active" factors Al:Cr, Al:Fe, Fe, and Cr, with negative effects, and "possibly active" factors Al:Pb and Al:Cd, with negative effects. We will see below that there is reason to doubt that Al:Cu is an active factor.

We have two goals in this section: first, to analyze the moss data from [14] using ℓ_p-ℓ_q regression, and to compare the results with those obtained with existing methods used in [14]. Second, to use ℓ_p-ℓ_q regression on subsets of the moss data, determined by supersaturated fractions of the full Rechtschaffner design. We used three approximately optimal fractions identified by Cela et al. [13] that use only 40, 33, and 24 of the 67 Rechtschaffner runs, respectively; for comparison we also use three random fractions of the same sizes.

5.1 Methods

For this section we used a version of Algorithm 1 coded in *R* [30]. We considered all combinations of p and q with $p \in \{0.8, 1, 1.1, 1.5, 1.9, 2\}$ and $q \in \{0.1, 0.5, 0.9, 1, 1.5, 2\}$. For each such combination, we produced a solution path by solving the ℓ_p-ℓ_q minimization problem for 50 different values of λ, equally spaced in log scale from 10^{-2} to 10^4. In each case, we started from the largest value of λ, which produces zeros for all fitted values, and at each subsequent step used the previously fitted values as warm start for the fit with the next smaller value of λ.

This was repeated 11 times, each time having as response the uptake for one of the chemical elements. Furthermore, the process was repeated 7 times: once with the full saturated design matrix, once for each of the 3 approximately optimal supersaturated subdesigns from [13], and once for each of 3 (fixed) random subdesigns. This resulted in 138,600 sets of fitted values; to speed up the computations, parallel processing was used.

5.2 Results

Examination of the solution paths obtained yields the following results.

Regularization is Needed. Some effects can be dramatically different at different levels of regularization. As an example, consider the solution path in Figure 6. The response is aluminum uptake (Al), using $p = q = 2$ (ridge regression), with the full saturated design. Very small values of λ result in a solution that is essentially equal to $X^{-1}y$. The coefficient for the interaction term Al:Cu is 54.6, which Lenth's method classifies as "probably active,"

as mentioned above. However, as λ increases, it's magnitude decreases faster than for other factors, and for $\lambda \geq 62.5$ it actually becomes negative. This strongly suggests that the effect of Al:Cu is not really active. Thus, what seemed like a large effect at zero regularization, might not be truly interesting. Conversely, other factors whose coefficients become larger (relatively to the others) at higher levels of regularization might be truly interesting, yet are missed altogether following the traditional approach.

Strong evidence that some amount of regularization is needed arises when the same analysis is performed on the 40-, 33-, and 24-run fragments of the design. Stronger correlation between the fitted coefficients is obtained at higher values of λ, indicating that the regularized estimates are more stable and reproducible. Figure 7 shows the correlation between the coefficients estimated from the full saturated design and the coefficients estimated using each of the three supersaturated fragments from [13], for different values of λ. High concordance can be obtained between the estimated coefficients resulting from supersaturated fragments and the full saturated design, *as long as regularization is used*; see Figure 8. This shows that supersaturated designs, which are considerably more economical, can be used with some success in screening for active factors. Of course, in general, the larger the subdesign, the better the concordance, but even the small 24-run fragment can consistently capture the largest effects.

Properties of the Solution Paths. The choices of p and q affect the overall properties of the solution path. Here are some important properties:

Having $0 < q \leq 1$ gives a sparse coefficient vector for larger intervals of λ than when $q = 1$; $q > 1$ does not give a sparse coefficient vector. However, very small values of q (like $q = 0.1$) tend not to be useful for obtaining sparse solutions: Typically, there is a threshold λ_0 such that the coefficients are all nonzero when $0 < \lambda < \lambda_0$, and all zero when $\lambda \geq \lambda_0$. Better results are obtained for q close to 1 (like $q = 0.9$), which gives several intervals for λ with different levels of sparsity. The Lasso approach, that is, $p = 2$ and $q = 1$, succeeds in producing sparse solutions, yet the coefficients are often highly shrunken.

Having $0 < p < 2$ results in paths that are nearly piece-wise constant with respect to λ, that is, the coefficients are nearly unchanged on intervals of values of λ, with abrupt changes in between. There is some evidence as well that the correlation between coefficient estimates from the saturated and supersaturated models is improved as well for $0 < p < 2$; see Figure 9. This is likely due to the increased ability to deal with outliers for these p-values.

Random Subsets. Compared with the supersaturated fractions from [13], random fractions of the same size produced different solution paths. The estimated coefficients tend to be less correlated with the full data estimates, and the curves tend to have worse sparsity properties, dropping suddenly from low sparsity to all zeros. This agrees with the observation in [4] that random fractions of saturated designs are suboptimal.

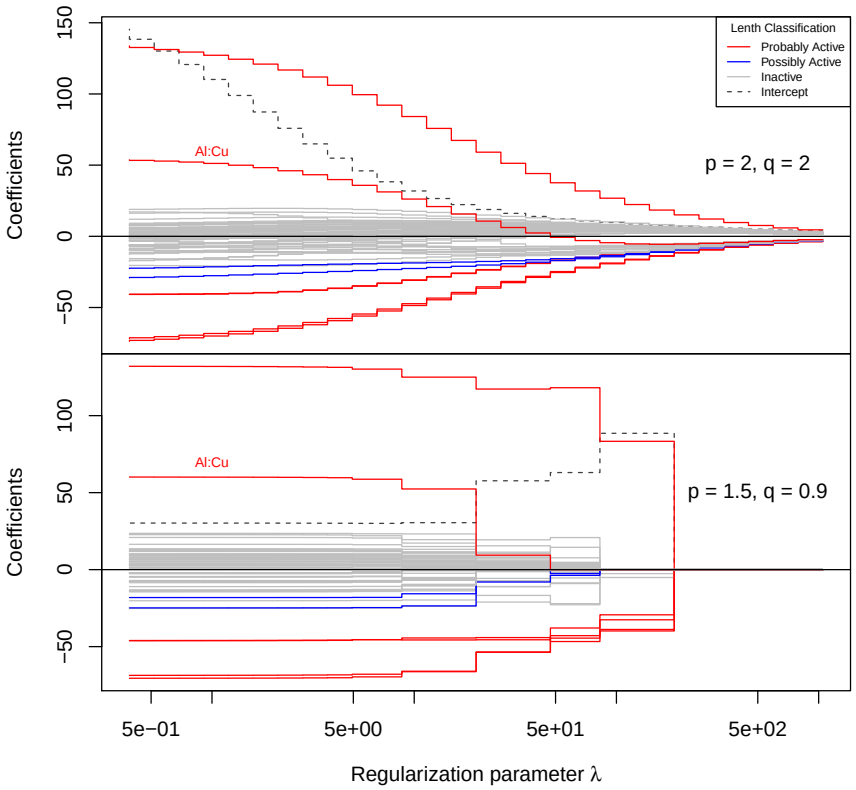


Figure 6: *Solution paths for two ℓ_p-ℓ_q models on the same data.* (Based on the experiment data from [14], using aluminum uptake as response.) The top panel uses $p = 2$ and $q = 2$ (i.e., ridge regression). The bottom panel uses $p = 1.5$ and $q = 0.9$. Colors correspond to the factor classification from [14], using Lenth's method [25]. Notice that the coefficient for factor Al:Cu starts out large (and positive), for small values of the regularization parameter λ; in the top panel, it becomes smaller and eventually negative, while in the bottom panel it becomes zero at about $\lambda = 50$. The second model offers several ranges of values of λ with different levels of sparsity, with essentially constant coefficient values.

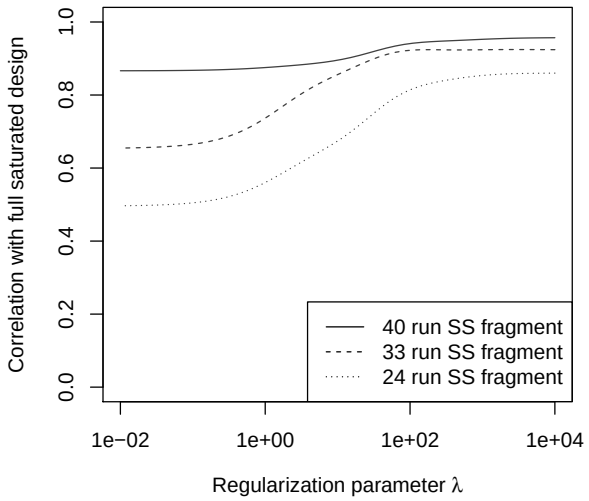


Figure 7: *Coefficient correlation between full saturated design and supersaturated fragments.* (Based on the experiment data from [14], using aluminum uptake as response.) Each curve describes how the correlation between the estimated coefficients from an ℓ_2-ℓ_2 model fitted to the full saturated design and to each of three supersaturated fractions described in [13]. Substantially better concordance is obtained for larger values of the regularization parameter λ. This is evidence both that regularization is needed, and that even small supersaturated fractions can capture information about the likely active factors.

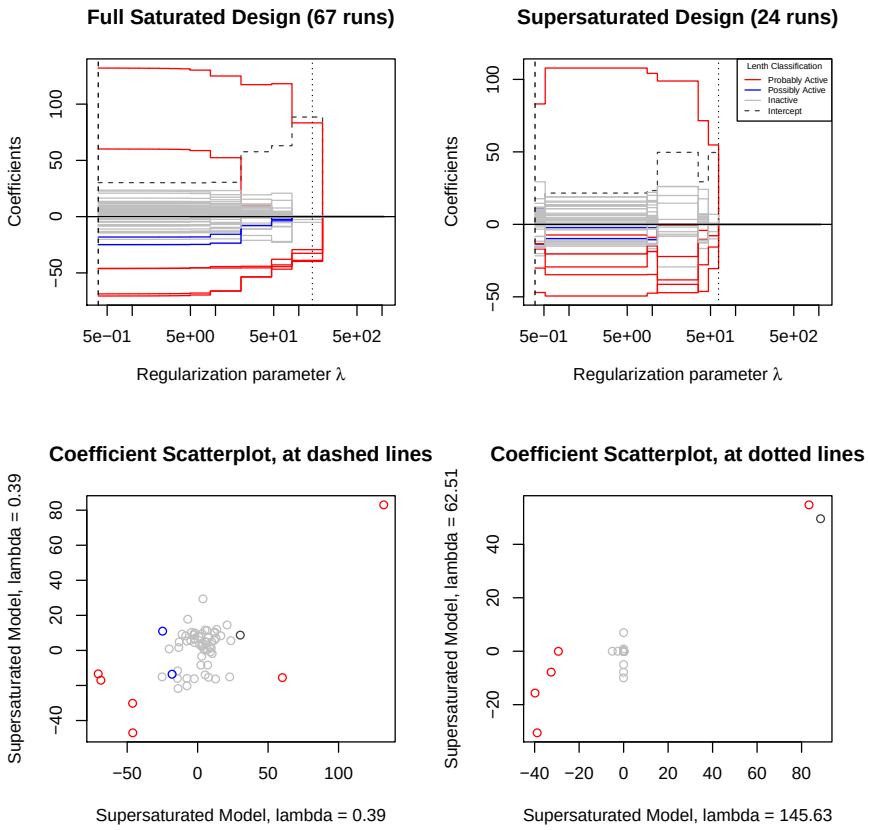


Figure 8: *Scatterplots for ℓ_p-ℓ_q coefficients from saturated and supersaturated designs.* (Based on the experiment data from [14], using aluminum uptake as response.) The top panels use, $p = 1.5$ and $q = 0.9$, one on the full saturated design and the other in the 24-run fragment from [13]. Colors correspond to the factor classification from [14], using Lenth's method [25]. The bottom panels show scatterplots of fitted coefficients; the bottom left shows coefficients estimated using small values of λ, while the bottom right shows the coefficients at large values of λ. This is evidence both that regularization is needed, and that even small supersaturated fractions can detect the most of the likely active factors.

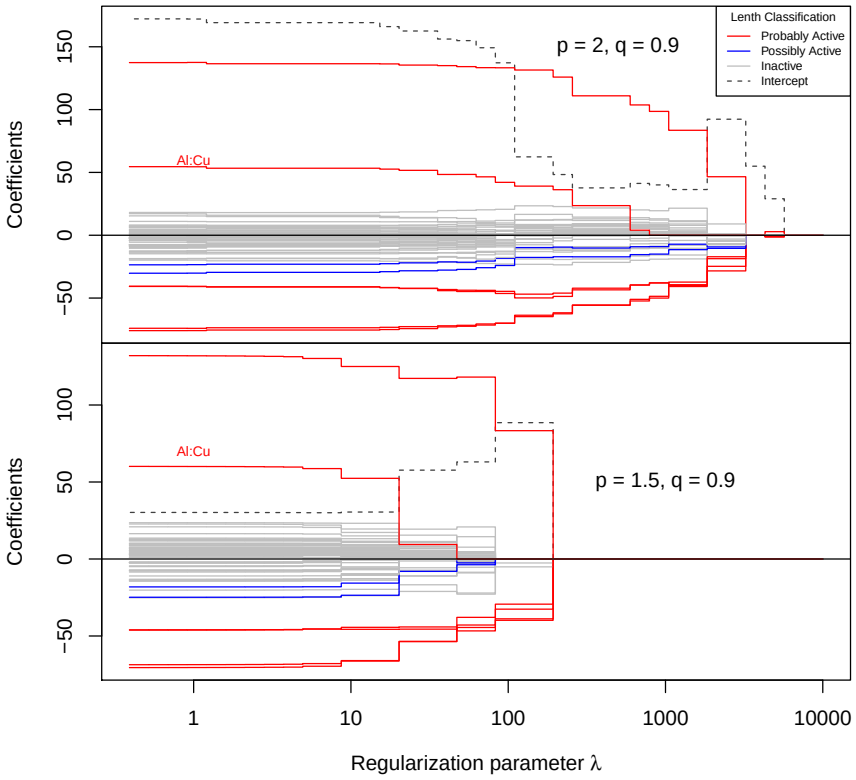


Figure 9: *Solution paths for two ℓ_p-ℓ_q models on the same data.* (Based on the experiment data from [14], using aluminum uptake as response.) The top panel uses $p = 2$ and $q = 0.9$, while the bottom panel uses $p = 1.5$ and $q = 0.9$. Colors correspond to the factor classification from [14], using Lenth's method [25]. Notice that changing p from 2 to 1.5 results in a path that is more strongly piece-wise constant with respect to λ, that is, the coefficients are nearly unchanged on intervals of values of λ, with abrupt changes in between.

6 Conclusions

This paper describes a method for the solution of the variable selection problem. We considered a model that is well studied in the literature. A new theoretical property of our model, namely continuity, is shown. Our model is able to outperform state-of-the-art methods and provides accurate estimations. Thanks to the usage of generalized Krylov subspaces, the method described may be applied to the solution of large-scale problems. Following the ideas in [6], the introduction of constraints in the model will be subject of future research.

Our work here is focused mostly on data from designed experiments, and in particular for saturated and supersaturated designs. Supersaturated designs have been proposed over the decades as a way to perform screening experiments in an economical way. They have not gained much popularity, partly due to the impossibility of performing classical regression inference. However, they seem promising as a way to perform more economical screening studies, when coupled with regularization methods like the one described in this paper.

Choice of λ. The usual methods for choosing the regularization parameter λ are based on either an *a priori* estimate of the magnitude of the error term, or on using a cross-validation scheme. In principle, cross-validation can be applied to saturated designs, but it is usually not applicable to supersaturated designs, because there are few observations to start with. Furthermore, the designs are carefully chosen for balance and to minimize aliasing of effects [19]; these properties are lost when choosing random subsets of observations for cross-validation. Besides, cross-validation seeks to minimize prediction error, while the goal in our case is variable selection.

Based on the fact that screening studies are not meant to provide formal inference on the significance of particular effects, but rather provide a reduced list of factors for follow-up studies, it seems reasonable to choose λ based on the expected or desired level of sparsity. By choosing $0 < q < 1$ close to 1, and $1 < p < 2$, we tend to obtain solution paths that are approximately piece-wise constant (i.e., relatively unchanged on intervals of λ, with abrupt changes in between), and with staggered levels of sparsity. This often will provide an interval of values for λ with the appropriate level of sparsity. That is, the step-wise behavior makes it less crucial to identify a single correct value of λ.

Identifying Likely Active Factors. We propose a criterion for deciding on interesting coefficients, or *likely active factors*, based on sparse fitted coefficients. This is an alternative to Lenth's method [25]. It starts by fitting a solution path for the ℓ_p-ℓ_q model with $q < 1$ but close to 1 (say, $q = 0.9$) and $1 < p < 2$ (say, $p = 1.5$). In the solution path, identify an interval of values for λ for which only a small proportion (say, approximately 10%) of the coefficients are nonzero. The factors with nonzero coefficients are classified as likely active.

Acknowledgments

The authors thank the referees for comments that lead to an improved presentation. A.B. is a member of the GNCS-INdAM group that partially supported this work with the Young Researchers Project (Progetto Giovani Ricercatori) "Variational methods for the approximation of sparse data". Moreover, A.B. research is partially supported by the Regione Autonoma della Sardegna research project "Algorithms and Models for Imaging Science [AMIS]" (RASSR57257, intervento finanziato con risorse FSC 2014-2020 - Patto per lo Sviluppo della Regione Sardegna). The work of O.D.C and L.R. is supported in part by NSF grant DMS-1720259.

References

[1] Abraham, B., Chipman, H., Vijayan, K.: Some risks in the construction and analysis of supersaturated designs. Technometrics **41**, 135–141 (1999)

[2] Beattie, S., Fong, D., Lin, D.: A two-stage bayesian model selection strategy for supersaturated designs. Technometrics **44**, 55–63 (2002)

[3] Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences **2**, 183–202 (2009)

[4] Booth, K.H., Cox, D.R.: Some systematic supersaturated designs. Technometrics **4**(4), 489–495 (1962)

[5] Buccini, A., De la Cruz Cabrera, O., Donatelli, M., Martinelli, A., Reichel, L.: Large-scale regression with non-convex loss and penalty. Applied Numerical Mathematics **157**, 590–601 (2020)

[6] Buccini, A., Pasha, M., Reichel, L.: Modulus-based iterative methods for constrained ℓ_p-ℓ_q minimization. Inverse Problems **36**(8), 084,001 (2020)

[7] Buccini, A., Reichel, L.: An ℓ^2-ℓ^q regularization method for large discrete ill-posed problems. Journal of Scientific Computing **78**, 1526–1549 (2019)

[8] Buccini, A., Reichel, L.: An ℓ_p-ℓ_q minimization method with cross-validation for the restoration of impulse noise contaminated images. Journal of Computational and Applied Mathematics **375**, 112,824 (2020)

[9] Buccini, A., Reichel, L.: Generalized cross validation for ℓ^p-ℓ^q minimization. Numerical Algorithms **In press**, 1–22 (2021). DOI 10.1007/s11075-021-01087-9

[10] Candès, E., Tao, T.: The Dantzig selector: Statistical estimation when p is much larger than n. The Annals of Statistics **35**, 2313–2351 (2007)

[11] Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory **52**(2), 489–509 (2006)

[12] Candès, E.J., Wakin, M.B., Boyd, S.P.: Enhancing sparsity by reweighted ℓ_1 minimization. Journal of Fourier Analysis and Applications **14**(5-6), 877–905 (2008)

[13] Cela, R., Phan-Tan-Luu, R., Claeys-Bruno, M., Sergent, M.: Fractions of Rechtschaffner matrices as supersaturated designs in screening experiments aimed at evaluating main and two-factor interaction effects. Analytica Chimica Acta **721**, 44–54 (2012)

[14] Cesa, M., Campisi, B., Bizzotto, A., Ferraro, C., Fumagalli, F., Nimis, P.: A factor influence study of trace element bioaccumulation in moss bags. Archives of Environmental Contamination and Toxicology **55**(3), 386–396 (2008)

[15] Chan, R.H., Liang, H.X.: Half-quadratic algorithm for ℓ_p-ℓ_q problems with applications to TV-ℓ_1 image restoration and compressive sensing. In: Efficient algorithms for global optimization methods in computer vision, Lecture Notes in Computer Science, Vol. 8293, pp. 78–103. Springer, Berlin (2014)

[16] Donoho, D.L.: Compressed sensing. IEEE Transactions on Information Theory **52**, 1289–1306 (2006)

[17] Fan, J., Li, R.: Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association **96**(456), 1348–1360 (2001)

[18] Georgiou, S.D.: Modelling by supersaturated designs. Computational Statistics & Data Analysis **53**, 428–435 (2008)

[19] Georgiou, S.D.: Supersaturated designs: A review of their construction and analysis. Journal of Statistical Planning and Inference **144**, 92–109 (2014)

[20] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)

[21] Huang, G., Lanza, A., Morigi, S., Reichel, L., Sgallari, F.: Majorization-minimization generalized Krylov subspace methods for ℓ_p-ℓ_q optimization applied to image restoration. BIT Numerical Mathematics **57**, 351–378 (2017)

[22] Huang, H., Yang J.Y.and Liu, M.: Functionally induced priors for componentwise gibbs sampler in the analysis of supersaturated designs. Computational Statistics & Data Analysis **72**, 1–12 (2014)

[23] Lanza, A., Morigi, S., Reichel, L., Sgallari, F.: A generalized krylov subspace method for ℓ_p-ℓ_q minimization. SIAM Journal on Scientific Computing **37**(5), S30–S50 (2015)

[24] Lanza, A., Morigi, S., Sgallari, F.: Constrained TV$_p$-ℓ_2 model for image restoration. Journal of Scientific Computing **68**, 64–91 (2016)

[25] Lenth, R.V.: Quick and easy analysis of unreplicated factorials. Technometrics **31**(4), 469–473 (1989)

[26] Li, R., Lin, D.: Data analysis in supersaturated designs. Statistics and Probability Letters **59**, 135–144 (2002)

[27] Li, R., Lin, D.K.: Analysis methods for supersaturated design: some comparisons. Journal of Data Science **1**(3), 249–260 (2003)

[28] Liu, Z., Wei, Z., Sun, W.: An iteratively approximated gradient projection algorithm for sparse signal reconstruction. Applied Mathematics and Computation **228**, 454–462 (2014)

[29] Mead, R.: The Design of Experiments: Statistical Principles for Practical Applications. Cambridge University Press, Cambridge (1990)

[30] R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2019). URL `https://www.R-project.org/`

[31] Ramlau, R., Zarzer, C.A.: On the minimization of a tikhonov functional with a non-convex sparsity constraint. Electron. Trans. Numer. Anal. **39**, 476–507 (2012)

[32] Rechtschaffner, R.L.: Saturated fractions of 2^n and 3^n factorial designs. Technometrics **9**(4), 569–575 (1967)

[33] Rodríguez, P., Wohlberg, B.: Efficient minimization method for a generalized total variation functional. IEEE Transactions on Image Processing **18**(2), 322–332 (2008)

[34] Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)

[35] Subbotin, M.T.: On the law of frequency of error. Matematicheskii Sbornik **31**(2), 296–301 (1923)

[36] Westfall, P., Young, S., Lin, D.: Forward selection error control in the analysis of supersaturated designs. Statistica Sinica **8**, 101–117 (1998)

[37] Wolke, R., Schwetlick, H.: Iteratively reweighted least squares: algorithms, convergence analysis, and numerical comparisons. SIAM Journal on Scientific and Statistical Computing **9**, 907–921 (1988)

[38] Yuille, A.L., Rangarajan, A.: The concave-convex procedure. Neural Computation **15**(4), 915–936 (2003)

[39] Zhao, Y.B., Li, D.: Reweighted ℓ_1-minimization for sparse solutions to underdetermined linear systems. SIAM Journal on Optimization **22**(3), 1065–1088 (2012)