Ultra-high Dimensional Multiple Output Learning With Simultaneous Orthogonal Matching Pursuit: Screening Approach

Supplemental Material

Mladen Kolar **Eric P. Xing**
Machine Learning Department
Carnegie Mellon University

1 Some Technical Results

In what follows, C_1, C_2 will denote arbitrary positive constants.

The following argument is quite standard (e.g. Zhou et al. (2009); Wang (2009))

Lemma 1. *Let* $\mathbf{x} \sim \mathcal{N}(0, \mathbf{\Sigma})$ *and* $\hat{\mathbf{\Sigma}} = n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i'$ *be the empirical estimate from* n *independent realizations of* $\mathbf{x}$*. Denote* $\mathbf{\Sigma} = [\sigma_{ab}]$ *and* $\hat{\mathbf{\Sigma}} = [\hat{\sigma}_{ab}]$*. Assume* $\phi_{min} \leq \Lambda_{\min}(\mathbf{\Sigma}) \leq \Lambda_{\max}(\mathbf{\Sigma}) \leq \phi_{\max}$*. Then*

$$\mathbb{P}[\max_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \Lambda_{\max}(\hat{\mathbf{\Sigma}}_{\mathcal{M}}) \geq 2\phi_{\max}] \leq \frac{2\sqrt{2}\phi_{\max}}{\sqrt{n}} \exp(-\frac{n\phi_{max}}{16s^2} + (s+3)\log p) \tag{1}$$

and

$$\mathbb{P}[\min_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \Lambda_{\min}(\hat{\mathbf{\Sigma}}_{\mathcal{M}}) \leq \phi_{\min}/2] \leq \frac{2\phi_{\min}}{\sqrt{2n}} \exp(-\frac{n\phi_{min}^2}{16\phi_{max}s^2} + (s+3)\log p). \tag{2}$$

Proof. Under the assumptions, it is enough to prove that

$$\min_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \Lambda_{\max}(\hat{\mathbf{\Sigma}}_{\mathcal{M}} - \mathbf{\Sigma}_{\mathcal{M}}) \leq \epsilon \tag{3}$$

with probability tending to 1. Using the union bound, we have

$$\begin{aligned}
&\mathbb{P}[\max_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \Lambda_{\max}(\hat{\mathbf{\Sigma}}_{\mathcal{M}} - \mathbf{\Sigma}_{\mathcal{M}}) \geq \epsilon] \\
&\leq \sum_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \mathbb{P}[\sup_{z \in \mathbb{R}^{|\mathcal{M}|}, ||z||_2 = 1} |z'(\hat{\mathbf{\Sigma}}_{\mathcal{M}} - \mathbf{\Sigma}_{\mathcal{M}})z| \geq \epsilon] \\
&\leq \sum_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \mathbb{P}[\max_{a,b \in \mathcal{M}} |\hat{\sigma}_{ab} - \sigma_{ab}| \sup_{z \in \mathbb{R}^{|\mathcal{M}|}, ||z||_2 = 1} \sum_{i_1, i_2 \in \mathcal{M}} |z_{i_1}||z_{i_2}| \geq \epsilon] \\
&\leq \sum_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \mathbb{P}[\max_{a,b \in \mathcal{M}} |\hat{\sigma}_{ab} - \sigma_{ab}| \sup_{z \in \mathbb{R}^{|\mathcal{M}|}, ||z||_2 = 1} (\sum_{i_1 \in \mathcal{M}} |z_{i_1}|)^2 \geq \epsilon] \\
&\leq \sum_{\mathcal{M} \subseteq [p], |\mathcal{M}| < s} \mathbb{P}[\max_{a,b \in \mathcal{M}} |\hat{\sigma}_{ab} - \sigma_{ab}| \geq \epsilon / |\mathcal{M}|] \\
&\leq p^{s+1} \mathbb{P}[\max_{a,b \in [p]} |\hat{\sigma}_{ab} - \sigma_{ab}| \geq \epsilon / s] \\
&\leq p^{s+1} p^2 \max_{a,b \in [p]} \mathbb{P}[|\hat{\sigma}_{ab} - \sigma_{ab}| \geq \epsilon / s] \\
&\leq \frac{2\sqrt{2}\phi_{\max}}{\sqrt{n}} \exp(-\frac{n\epsilon^2}{4\phi_{max}s^2} + (s+3)\log p).
\end{aligned}$$

Setting $\epsilon = \phi_{\max}$ gives Eq. (1). Similarly, we can prove Eq. (2). □

The following result is a modification of Lemma A.3 in Bickel and Levina (2008) with explicit constants.

Lemma 2. *Let* $\mathbf{x} \sim \mathcal{N}(0, \mathbf{\Sigma})$ *and* $\hat{\mathbf{\Sigma}} = n^{-1}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i'$ *be the empirical estimate from* n *independent realizations of* $\mathbf{x}$*. Denote* $\mathbf{\Sigma} = [\sigma_{ab}]$ *and* $\hat{\mathbf{\Sigma}} = [\hat{\sigma}_{ab}]$*. If* $\Lambda_{\max}(\mathbf{\Sigma}) \le \phi_{\max} < \infty$ *then*

$$\mathbb{P}[|\hat{\sigma}_{ab} - \sigma_{ab}| \ge \epsilon] \le \frac{2\sqrt{2}\phi_{\max}}{\sqrt{n}} \exp(-\frac{n\epsilon^2}{4\phi_{\max}}). \tag{4}$$

Proof. Let $\rho_{ab} = \sigma_{ab}(\sigma_{aa}\sigma bb)^{-1/2}$, $x_{ia}^* = x_{ia}/\sqrt{\sigma_{aa}}$ and $x_{ib}^* = x_{ib}/\sqrt{\sigma_{bb}}$. Write

$$\mathbb{P}[|\sum_{i=1}^n (x_{ia}x_{ib} - \sigma_{ab})| \ge n\epsilon] = \mathbb{P}[|\sum_{i=1}^n (x_{ia}^*x_{ib}^* - \rho_{ab})| \ge n\epsilon(\sigma_{aa}\sigma_{bb})^{-1/2}]. \tag{5}$$

Simple algebra shows that

$$\sum_{i=1}^n (x_{ia}^*x_{ib}^* - \rho_{ab}) = \frac{1}{4}[\sum_{i=1}^n [(x_{ia} + x_{ib})^2 - 2(1+\rho_{ab})] - \sum_{i=1}^n [(x_{ia} - x_{ib})^2 - 2(1-\rho_{ab})]], \tag{6}$$

so

$$\begin{aligned} \mathbb{P}[|\sum_{i=1}^n (x_{ia}x_{ib} - \sigma_{ab})| \ge n\epsilon] &\le \\ &\le \mathbb{P}[|\sum_{i=1}^n [(x_{ia} + x_{ib})^2 - 2(1+\rho_{ab})]| \ge 2n\epsilon(\sigma_{aa}\sigma_{bb})^{-1/2}] + \\ &\quad \mathbb{P}[|\sum_{i=1}^n [(x_{ia} - x_{ib})^2 - 2(1-\rho_{ab})]| \ge 2n\epsilon(\sigma_{aa}\sigma_{bb})^{-1/2}] \\ &\le 2\mathbb{P}[\chi_n^2 \ge n + n\epsilon\phi_{\max}^{-1}] \\ &\le \frac{2\sqrt{2}\phi_{\max}}{\sqrt{n}} \exp(-\frac{n\epsilon^2}{4\phi_{\max}}), \end{aligned} \tag{7}$$

since $\phi_{\max} \ge |1 - \rho_{ab}|(\sigma_{aa}\sigma_{bb})^{1/2}$. □

Exponential inequality for chi-squared distribution Laurent and Massart (2000).

Lemma 3. *Let* χ_n^2 *be a central chi-squared r.v. with* n *degrees of freedom. For any positive* ϵ*,*

$$\mathbb{P}[\chi_n^2 \ge n + 2\sqrt{n\epsilon} + 2\epsilon] \le \exp(-\epsilon) \tag{8}$$

$$\mathbb{P}[\chi_n^2 \le \epsilon - 2\sqrt{n\epsilon}] \le \exp(-\epsilon). \tag{9}$$

From Obozinski et al. (2009) we have

Lemma 4. *Let $X_1, \dots, X_m$ be i.i.d. central chi-squared r.v. with n degrees of freedom. Then for any $\epsilon > n$,*

$$\mathbb{P}[\max_{i\in[m]} X_i \geq 2\epsilon] \leq m\exp(-\epsilon(1-2\sqrt{\frac{n}{\epsilon}})). \tag{10}$$

2 Proofs

2.1 Proof of Theorem 4

Under the assumptions of the theorem, the number of relevant variables s is relatively small compared to the sample size n. The proof strategy can be outlined as follows: i) we are going to show that, with high probability, at least one relevant variable is going to be identified within the following m^*_{one} steps, conditioning on the already selected variables $\mathcal{M}^{(k)}$ and this holds uniformly for all k; ii) we can conclude that all the relevant variables are going to be selected within $m^*_{\max} = sm^*_{\text{one}}$ steps. Exact values for m^*_{one} and $m^*_{\max}$ are given below. Without loss of generality, we analyze the first step of the algorithm, i.e., we show that the first relevant variable is going to be selected within the first m^*_{one} steps.

Assume that in the first $m^*_{\text{one}} - 1$ steps, there were no relevant variables selected. Assuming that the m^*_{one}-th selected variable is still an irrelevant one, we will arrive to a contradiction, which shows that at least one relevant variable has been selected in the first m^*_{one} steps. For any step k, the squared error reduction is given as

$$\Delta(k) := \text{RSS}(k-1) - \text{RSS}(k) = \sum_t ||\mathbf{H}^{(k)}_{t,\hat{j}_k}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\mathbf{y}_t||_2^2 \tag{11}$$

with $\mathbf{H}^{(k)}_{t,j} = \mathbf{X}^{(k)}_{t,j}\mathbf{X}^{(k)\prime}_{t,j}||\mathbf{X}^{(k)}_{t,j}||^{-2}$ and $\mathbf{X}^{(k)}_{t,j} = (\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\mathbf{X}_{t,j}$. We are interested in the quantity $\sum_{k=1}^{m^*_{\text{one}}} \Delta(k)$, when all the selected variables $\hat{j}_k$ belong to $[p]\backslash\mathcal{M}_*$.

In what follows, we will derive a lower bound for $\Delta(k)$. We perform our analysis on the event

$$\begin{aligned}\mathcal{E} = \{&\min_{t\in[T]} \min_{\mathcal{M}\subseteq[p],|\mathcal{M}|\leq m^*_{\max}} \Lambda_{\min}(\hat{\mathbf{\Sigma}}_{\mathcal{M}}) \geq \phi_{\min}/2\} \\ &\bigcap\{\max_{t\in[T]} \max_{\mathcal{M}\subseteq[p],|\mathcal{M}|\leq m^*_{\max}} \Lambda_{\max}(\hat{\mathbf{\Sigma}}_{\mathcal{M}}) \leq 2\phi_{\max}\}.\end{aligned} \tag{12}$$

From the definition of $\hat{j}_k$, we have

$$\begin{aligned}
\Delta(k) &\geq \max_{j\in\mathcal{M}_*} \sum_t ||\mathbf{H}_{t,j}^{(k)}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\mathbf{y}_t||_2^2 \\
&\geq \max_{j\in\mathcal{M}_*} \Big(\sum_t ||\mathbf{H}_{t,j}^{(k)}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}||_2^2 \\
&\qquad - \sum_t ||\mathbf{H}_{t,j}^{(k)}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\boldsymbol{\epsilon}_t||_2^2 \Big) \\
&\geq \max_{j\in\mathcal{M}_*} \sum_t ||\mathbf{H}_{t,j}^{(k)}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}||_2^2 \\
&\qquad - \max_{j\in\mathcal{M}_*} \sum_t ||\mathbf{H}_{t,j}^{(k)}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\boldsymbol{\epsilon}_t||_2^2 \\
&= (I) - (II).
\end{aligned} \tag{13}$$

We deal with these two terms separately. Let $\mathbf{H}_{t,\mathcal{M}}^{\perp} = \mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}}$ denote the projection matrix. With this notation, the first term (I) is lower bounded by

$$\begin{aligned}
&\max_{j\in\mathcal{M}_*} \sum_t ||\mathbf{H}_{t,j}^{(k)}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}||_2^2 \\
&\quad = \max_{j\in\mathcal{M}_*} \sum_t ||\mathbf{X}_{t,j}^{(k)}||_2^{-2}|\mathbf{X}_{t,j}^{(k)'}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}|^2 \\
&\quad \geq \min_{t\in[T],j\in\mathcal{M}_*}\{||\mathbf{X}_{t,j}^{(k)}||_2^{-2}\} \max_{j\in\mathcal{M}_*} \sum_t |\mathbf{X}_{t,j}^{(k)'}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}|^2 \\
&\quad \geq \{\max_{t\in[T],j\in\mathcal{M}_*} ||\mathbf{X}_{t,j}||_2^2\}^{-1} \max_{j\in\mathcal{M}_*} \sum_t |\mathbf{X}_{t,j}'\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}|^2,
\end{aligned} \tag{14}$$

where the last inequality follows form noting that $||\mathbf{X}_{t,j}||_2 \geq ||\mathbf{X}_{t,j}^{(k)}||_2$ and $\mathbf{X}_{t,j}^{(k)'}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp} = \mathbf{X}_{t,j}'\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}$. A simple calculation shows that

$$\begin{aligned}
&\sum_t ||\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}||_2^2 \\
&\quad = \sum_t \sum_{j\in\mathcal{M}_*} \beta_{t,j}\mathbf{X}_{t,j}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*} \\
&\quad \leq \sum_{j\in\mathcal{M}_*} \sqrt{\sum_t \beta_{t,j}^2}\sqrt{\sum_t (\mathbf{X}_{t,j}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*})^2} \\
&\quad \leq ||\boldsymbol{\beta}||_{2,1} \max_{j\in\mathcal{M}_*} \sqrt{\sum_t (\mathbf{X}_{t,j}\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*})^2}.
\end{aligned} \tag{15}$$

Plugging (15) back into (14), the following lower bound is achieved

$$(I) \geq \{\max_{t\in[T],j\in\mathcal{M}_*} ||\mathbf{X}_{t,j}||_2^2\}^{-1} \frac{(\sum_t ||\mathbf{H}_{t,\mathcal{M}^{(k)}}^{\perp}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}||_2^2)^2}{||\boldsymbol{\beta}||_{2,1}^2}. \tag{16}$$

On the event $\mathcal{E}$, $\max_{t\in[T],j\in\mathcal{M}_*}||\mathbf{X}_{t,j}||_2^2 \leq 2n\phi_{\max}$. Since we have assumed that no additional relevant predictors have been selected by the procedure, it holds that $\mathcal{M}_* \not\subseteq \mathcal{M}^{(k)}$. This leads to

$$\sum_t ||\mathbf{H}^{\perp}_{t,\mathcal{M}^{(k)}}\mathbf{X}_{t,\mathcal{M}_*}\boldsymbol{\beta}_{t,\mathcal{M}_*}||_2^2 \geq 2^{-1}n\phi_{\min}||\mathbf{B}_{\min}||_2^2, \tag{17}$$

on the event $\mathcal{E}$. Using the Cauchy-Schwarz inequality, $||\boldsymbol{\beta}||_{2,1}^{-2} \geq s^{-1}T^{-1}C_{\boldsymbol{\beta}}^{-2}$. Plugging back into (16), we have that

$$\begin{aligned}(I) &\geq 2^{-3}\phi_{\min}^2\phi_{\max}^{-1}C_{\boldsymbol{\beta}}^{-2}ns^{-1}T^{-1}||\mathbf{B}_{\min}||_2^4 \\ &\geq 2^{-3}\phi_{\min}^2\phi_{\max}^{-1}C_{\boldsymbol{\beta}}^{-2}C_s^{-1}n^{1-\delta_s}T^{-1}||\mathbf{B}_{\min}||_2^4\end{aligned} \tag{18}$$

Next, we deal with the second term in (13). Recall that $\mathbf{X}^{(k)}_{t,j} = \mathbf{H}^{\perp}_{t,\mathcal{M}^{(k)}}\mathbf{X}_{t,j}$, so that $||\mathbf{X}^{(k)}_{t,j}||_2^2 \geq 2^{-1}n\phi_{\min}$, on the event $\mathcal{E}$. We have

$$\begin{aligned}&\sum_t ||\mathbf{H}^{(k)}_{t,j}(\mathbf{I}_{n\times n} - \mathbf{H}_{t,\mathcal{M}^{(k)}})\boldsymbol{\epsilon}_t||_2^2 \\ &\quad= \sum_t ||\mathbf{X}^{(k)}_{t,j}||^{-2}(\mathbf{X}'_{t,j}\mathbf{H}^{\perp}_{t,\mathcal{M}^{(k)}}\boldsymbol{\epsilon}_t)^2 \\ &\quad\leq 2\phi_{\min}^{-1}n^{-1}\max_{j\in\mathcal{M}_*}\max_{|\mathcal{M}|\leq m^*_{\max}}\sum_t(\mathbf{X}'_{t,j}\mathbf{H}^{\perp}_{t,\mathcal{M}}\boldsymbol{\epsilon}_t)^2.\end{aligned} \tag{19}$$

Under the conditions of the theorem, $\mathbf{X}'_{t,j}\mathbf{H}^{\perp}_{t,\mathcal{M}}\boldsymbol{\epsilon}_t$ is normally distributed with mean 0 and variance $||\mathbf{H}^{\perp}_{t,\mathcal{M}}\mathbf{X}_{t,j}||_2^2$. Furthermore,

$$\max_{j\in\mathcal{M}_*}\max_{|\mathcal{M}|\leq m^*_{\max}}\max_{t\in[T]}||\mathbf{H}^{\perp}_{t,\mathcal{M}}\mathbf{X}_{t,j}||_2^2 \leq 2n\phi_{\max}. \tag{20}$$

Plugging back in (19), we have

$$(II) \leq 2^2\phi_{\min}^{-1}\phi_{\max}\max_{j\in\mathcal{M}_*}\max_{|\mathcal{M}|\leq m^*_{\max}}\chi^2_T. \tag{21}$$

The total number of possibilities for $j \in \mathcal{M}_*$ and $|\mathcal{M}| \leq m^*_{\max}$ is bounded by $p^{m^*_{\max}+2}$. Using Lemma 4, with $\epsilon = 2T(m^*_{\max}+2)\log(p)$, we obtain

$$\begin{aligned}(II) &\leq 2^3\phi_{\min}^{-1}\phi_{\max}T(m^*_{\max}+2)\log p \\ &\leq 9\phi_{\min}^{-1}\phi_{\max}C_pn^{\delta_p}Tm^*_{\max}\end{aligned} \tag{22}$$

with probability at least

$$1 - p^{m^*_{\max}+2}\exp\left(-2T(m^*_{\max}+2)\log(p)\left(1 - 2\sqrt{\frac{1}{2(m^*_{\max}+2)\log(p)}}\right)\right). \tag{23}$$

Going back to Eq. (13), we have the following

$$\begin{aligned}
n^{-1}T^{-1}\Delta(k) \geq\ & 2^{-3}\phi_{\min}^{2}\phi_{\max}^{-1}C_{\boldsymbol{\beta}}^{-2}C_{s}^{-1}n^{-\delta_s}T^{-2}||\mathbf{B}_{\min}||_2^4 \\
& - 9\phi_{\min}^{-1}\phi_{\max}C_p n^{\delta_p-1}m_{\max}^{*} \\
\geq\ & 2^{-3}\phi_{\min}^{2}\phi_{\max}^{-1}C_{\boldsymbol{\beta}}^{-2}C_{s}^{-1}c_{\boldsymbol{\beta}}^{2}n^{-\delta_s-2\delta_{\min}} \\
& - 9\phi_{\min}^{-1}\phi_{\max}C_p n^{\delta_p-1}m_{\max}^{*} \\
\geq\ & 2^{-3}\phi_{\min}^{2}\phi_{\max}^{-1}C_{\boldsymbol{\beta}}^{-2}C_{s}^{-1}c_{\boldsymbol{\beta}}^{2}n^{-\delta_s-2\delta_{\min}} \\
& \times (1 - 72\phi_{\min}^{-3}\phi_{\max}^{2}C_{\boldsymbol{\beta}}^{2}C_pC_sc_{\boldsymbol{\beta}}^{-2}n^{\delta_s+2\delta_{\min}+\delta_p-1}m_{\max}^{*}).
\end{aligned} \tag{24}$$

Since the bound in Eq. (24) holds uniformly for $k \in \{1, \ldots, m_{\text{one}}^{*}\}$, we have that $n^{-1}T^{-1}\sum_{t\in[T]}||\mathbf{y}_t||_2^2 \geq n^{-1}T^{-1}\sum_{k=1}^{m_{\text{one}}^{*}}\Delta(k)$. Setting

$$m_{\text{one}}^{*} = \lfloor 2^4\phi_{\min}^{-2}\phi_{\max}C_{\boldsymbol{\beta}}^{2}C_sc_{\boldsymbol{\beta}}^{-2}n^{\delta_s+2\delta_{\min}} \rfloor \tag{25}$$

and recalling that $m_{\max}^{*} = sm_{\text{one}}^{*}$, the lower bound becomes

$$n^{-1}T^{-1}\sum_{t\in[T]}||\mathbf{y}_t||_2^2 \geq 2(1 - Cn^{3\delta_s+4\delta_{\min}+\delta_p-1}), \tag{26}$$

for a positive constant C independent of p, n, s and T. Under the conditions of the theorem, the right side of (26) is bounded below by 2. We have arrived to a contradiction, since under the assumptions $\text{Var}(y_{t,i}) = 1$ and by the weak law of large numbers, $n^{-1}T^{-1}\sum_{t\in[T]}||\mathbf{y}_t||_2^2 \to 1$ in probability. Therefore, at least one relevant variable will be selected in m_{one}^{*} steps.

To complete the proof, we lower bound the probability in Eq. (22) and the probability of the event $\mathcal{E}$. Plugging in the value for $m_{\max}^{*}$, the probability in (22) can be lower bounded by $1 - \exp(-C(2T-1)n^{2\delta_s+2\delta_{\min}+\delta_p})$ for some positive constant C. The probability of the event $\mathcal{E}$ is lower bounded, using Lemma 1 together with the union bound, as $1 - C_1\exp(-C_2\frac{n^{1-6\delta_s-6\delta_{\min}}}{\max\{\log p, \log T\}})$, for some positive constants C_1 and C_2. Both of these probabilities converge to 1 under the conditions of the theorem.

2.2 Proof of Theorem 5

To prove the theorem, we use the same strategy as in Wang (2009). From Theorem 4 we have that $\mathbb{P}[\exists k \in \{0, \ldots, n-1\} : \mathcal{M}_* \subseteq \mathcal{M}^{(k)}] \to 1$, so $k_{\min} := \min_{k\in\{0,\ldots,n-1\}}\{k : \mathcal{M}_* \subseteq \mathcal{M}^{(k)}\}$ is well defined and $k_{\min} \leq m_{\max}^{*}$, for $m_{\max}^{*}$ defined in Theorem 4. We show that

$$\mathbb{P}[\min_{k\in\{0,\ldots,k_{\min}-1\}}(\text{BIC}(\mathcal{M}^{(k)}) - \text{BIC}(\mathcal{M}^{(k+1)})) > 0] \to 1, \tag{27}$$

so that $\mathbb{P}[\hat{s} < k_{\min}] \to 0$ as $n \to \infty$. We proceed by lower bounding the difference in the BIC scores as

$$\begin{aligned}\mathrm{BIC}(\mathcal{M}^{(k)}) - \mathrm{BIC}(\mathcal{M}^{(k+1)}) &= \log\left(\frac{\mathrm{RSS}(\mathcal{M}^{(k)})}{\mathrm{RSS}(\mathcal{M}^{(k+1)})}\right) - \frac{\log(n) + 2\log(p)}{n} \\ &\geq \log\left(1 + \frac{\mathrm{RSS}(\mathcal{M}^{(k)}) - \mathrm{RSS}(\mathcal{M}^{(k+1)})}{\mathrm{RSS}(\mathcal{M}^{(k+1)})}\right) - 3n^{-1}\log(p),\end{aligned} \tag{28}$$

where we have assumed $p > n$. Define the event $\mathcal{A} := \{n^{-1}T^{-1}\sum_{t\in[T]} ||\mathbf{y}_t||_2^2 \leq 2\}$. Note that $\mathrm{RSS}(\mathcal{M}^{(k+1)}) \leq \sum_{t\in[T]} ||\mathbf{y}_t||_2^2$, so on the event $\mathcal{A}$ the difference in the BIC scores is lower bounded as

$$\log(1 + 2n^{-1}T^{-1}\Delta(k)) - 3n^{-1}\log(p), \tag{29}$$

where $\Delta(k)$ is defined in (11). Using the fact that $\log(1+x) \geq \min(\log(2), 2^{-1}x)$ and the lower bound from Eq. (24), we have

$$\mathrm{BIC}(\mathcal{M}^{(k)}) - \mathrm{BIC}(\mathcal{M}^{(k+1)}) \geq \min(\log 2, Cn^{-\delta_s - 2\delta_{\min}}) - 3n^{-1}\log p, \tag{30}$$

for some positive constant C. It is easy to check that $\log 2 - 3n^{-1}\log p > 0$ and $Cn^{-\delta_s-2\delta_{\min}} - 3n^{-1}\log p > 0$ under the conditions of the theorem. The lower bound in (30) is uniform for $k \in \{0, \ldots, k_{\min}\}$, so the proof is complete if we show that $\mathbb{P}[\mathcal{A}] \to 1$. But this easily follows from the tail bounds on the central chi-squared random variable given in Lemma 3.

3 Extended Simulation Studies

We conduct a number of numerical studies to evaluate the finite sample performance of the S-OMP. We consider three procedures that perform estimation on individuals outputs: Sure Independence Screening (SIS) and Iterative SIS (ISIS) (Fan and Lv, 2008), and the OMP, for comparison purposes. SIS and ISIS are used to select a subset of variables and then the ALasso is used to further refine the selection. We denote this combination as SIS-ALasso and ISIS-ALasso. The size of the model selected by SIS is fixed as $n-1$, while the ISIS selects $\lfloor n/\log(n)\rfloor$ variables in each of the $\lfloor \log(n)-1\rfloor$ iterations. From the screened variables, the final model is selected using the ALasso, together with the BIC criterion to select the penalty parameter λ. We use the OMP without further refinement using the ALasso, since it was observed from the numerical studies in Wang (2009) that the combination does not gain much improvement. The S-OMP is used to reduce the dimensionality below the sample size jointly using the regression outputs. Next, the ALasso is used on each of the outputs to further perform the estimation. This combination is denoted SOMP-ALasso.

Let $\hat{\mathbf{B}} = [\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_T] \in \mathbb{R}^{p\times T}$ be an estimate obtained by one of the estimation procedures. We evaluate the performance averaged over 200 simulation runs. Let $\hat{\mathbb{E}}_n$ denote the empirical average over the simulation runs. We measure the size of the union support $\hat{S} = S(\hat{\mathbf{B}}) := \{j \in [p] : ||\hat{\mathbf{B}}_j||_2^2 > 0\}$. Next, we estimate the probability that the screening property is satisfied $\hat{\mathbb{E}}_n[\mathbb{I}\{\mathcal{M}_* \subseteq S(\hat{\mathbf{B}})\}]$, which we call coverage probability. For the union support, we define fraction of correct zeros $(p-s)^{-1}\hat{\mathbb{E}}_n[|S(\hat{\mathbf{B}})^C \cap \mathcal{M}_*^C|]$, fraction of incorrect zeros $s^{-1}\hat{\mathbb{E}}_n[|S(\hat{\mathbf{B}})^C \cap \mathcal{M}_*|]$ and fraction of correctly fitted $\hat{\mathbb{E}}_n[\mathbb{I}\{\mathcal{M}_* = S(\hat{\mathbf{B}})\}]$. Similar quantities are defined for the exact support recovery. In addition, we measure the estimation error $\hat{\mathbb{E}}_n[||\mathbf{B} - \hat{\mathbf{B}}||_2^2]$ and the prediction performance on the test set. On the test data $\{\mathbf{x}_i^*, \mathbf{y}_i^*\}_{i\in[n]}$, we compute

$$R^2 = 1 - \frac{\sum_{i\in[n]}\sum_{t\in[T]}(y_{t,i}^* - (\mathbf{x}_{t,i}^*)'\hat{\boldsymbol{\beta}}_t)^2}{\sum_{i\in[n]}\sum_{t\in[T]}(y_{t,i}^* - \bar{y_t^*})^2}, \tag{31}$$

where $\bar{y_t^*} = n^{-1}\sum_{i\in[n]} y_{t,i}$.

The following simulation studies are used to comparatively access the numerical performance of the procedures.

Simulation 1: The following toy model is based on the simulation I in Fan and Lv (2008) with $(n, p, s, T) = (400, 20000, 18, 500)$. Each $\mathbf{x}_i$ is drawn independently from a standard multivariate normal distribution, so that the variables are mutually independent. For $j \in [s]$ and $t \in [T]$, the non-zero coefficients of $\mathbf{B}$ are given as $\beta_{t,j} = (-1)^u(4n^{-1/2}\log n + |z|)$, where $u \sim \text{Bernoulli}(0.4)$ and $z \sim \mathcal{N}(0,1)$. The number of non-zero elements in $\mathbf{B}_j$ is given as a parameter $T_{\text{non-zero}} \in \{500, 300, 100\}$. The positions of non-zero elements are chosen uniformly at random from $[T]$. The noise is Gaussian with the standard deviation σ set to control the signal-to-noise ratio (SNR). SNR is defined as $\text{Var}(\mathbf{x}\boldsymbol{\beta})/\,\text{Var}(\boldsymbol{\epsilon})$ and we vary SNR $\in \{15, 10, 5, 1\}$.

Simulation 2: The following scenario is used to evaluate the performance of the methods as the number of non-zero elements in a row of $\mathbf{B}$ varies. We set $(n, p, s) = (100, 500, 10)$ and vary the number of outputs $T \in \{500, 750, 1000\}$. For each number of outputs T, we vary $T_{\text{non-zero}} \in \{0.8T, 0.5T, 0.2T\}$. $\mathbf{x}_i$ and $\mathbf{B}$ are given as in Simulation 1, i.e., $\mathbf{x}_i$ is drawn from a multivariate standard normal distribution and the non-zero coefficients $\mathbf{B}$ are given as $\beta_{t,j} = (-1)^u(4n^{-1/2}\log n + |z|)$, where $u \sim \text{Bernoulli}(0.4)$ and $z \sim \mathcal{N}(0,1)$. The noise is Gaussian, with the standard deviation defined thorough the SNR, which varies in $\{10, 5, 1\}$.

Simulation 3: The following model is borrowed from Wang (2009). We assume a correlation structure between variables given as $\text{Var}(\mathbf{X}_{j_1}, \mathbf{X}_{j_2}) = \rho^{|j_1 - j_2|}$, where $\rho \in \{0.2, 0.5, 0.7\}$. This correlation structure appears naturally among ordered variables. We set $(n, p, s, T) = (100, 5000, 3, 150)$ and $T_{\text{non-zero}} = 80$. The relevant variables are at positions $(1, 4, 7)$ and non-zero coefficients are given as $3, 1.5$ and 2 respectively. The SNR varies in $\{10, 5, 1\}$.

Simulation 4: The following model assumes a block compound correlation structure. For a parameter ρ, the correlation between two variables $\mathbf{X}_{j_1}$ and $\mathbf{X}_{j_2}$ is given as ρ, ρ^2 or ρ^3 when $|j_1 - j_2| \le 10$, $|j_1 - j_2| \in (10, 20]$ or $|j_1 - j_2| \in (20, 30]$ and it is set to 0 otherwise. We set $(n, p, s, T) = (150, 4000, 8, 150)$, $T_{\text{non-zero}} = 80$ and the parameter $\rho \in \{0.2, 0.5\}$. The relevant variables are located at positions 1, 11, 21, 31, 41, 51, 61, 71 and 81, so that each block of highly correlated variables has exactly one relevant variable. The values of relevant coefficients is given as in Simulation 1. The noise is Gaussian and the SNR varies in $\{10, 5, 1\}$.

Simulation 5: This model represents a difficult setting. It is modified from Wang (2009). We set $(n, p, s, T) = (200, 10000, 5, 500)$. The number of non-zero elements in each row varies as $T_{\text{non-zero}} \in \{400, 250, 100\}$. For $j \in [s]$ and $t \in [T]$, the non-zero elements equal $\beta_{t,j} = 2j$. Each row of $\mathbf{X}$ is generated as follows. Draw independently $\mathbf{z}_i$ and $\mathbf{z}'_i$ from a p-dimensional standard multivariate normal distribution. Now, $x_{ij} = (z_{ij} + z'_{ij})/\sqrt{(2)}$ for $j \in [s]$ and $x_{ij} = (z_{ij} + \sum_{j' \in [s]} z_{ij'})/2$ for $j \in [p]\backslash[s]$. Now, $\text{Corr}(x_{i,1}, y_{t,i})$ is much smaller then $\text{Corr}(x_{i,j}, y_{t,i})$ for $j \in [p]\backslash[s]$, so that it becomes difficult to select variable 1. The noise is Gaussian with the standard deviation $\sigma \in \{1.5, 2.5, 4.5\}$.

Simulation 1: $(n, p, s, T) = (500, 20000, 18, 500)$, $T_{\text{non-zero}} = 500$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 15				
Union Support	SIS-ALASSO	100.0	100.0	0.0	10.0	20.2	-	-
	ISIS-ALASSO	100.0	100.0	0.0	18.0	19.6	-	-
	OMP	100.0	100.0	0.0	0.0	23.9	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	0.7	0.0	8940.5	0.97	0.93
	ISIS-ALASSO	100.0	100.0	0.0	18.0	9001.6	0.33	0.93
	OMP	100.0	100.0	0.0	0.0	9005.9	0.20	0.93
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	9000.0	0.20	0.93
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	0.0	25.3	-	-
	ISIS-ALASSO	100.0	100.0	0.0	0.0	25.7	-	-
	OMP	100.0	100.0	0.0	0.0	23.9	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	1.6	0.0	8861.0	2.06	0.89
	ISIS-ALASSO	100.0	100.0	0.0	0.0	9007.7	0.65	0.90
	OMP	100.0	100.0	0.0	0.0	9005.9	0.31	0.91
	S-OMP-ALASSO	65.0	100.0	0.1	65.0	8987.4	0.41	0.90
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	64.0	18.4	-	-
	ISIS-ALASSO	100.0	100.0	0.0	57.0	18.6	-	-
	OMP	100.0	100.0	0.0	0.0	24.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	92.8	0.0	645.8	74.61	0.06
	ISIS-ALASSO	0.0	100.0	90.9	0.0	822.2	73.06	0.07
	OMP	100.0	100.0	0.0	0.0	9006.0	0.61	0.83
	S-OMP-ALASSO	0.0	100.0	70.3	0.0	2668.9	56.65	0.24
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	99.9	0.0	0.0	-	-
	ISIS-ALASSO	0.0	100.0	100.0	0.0	0.0	-	-
	OMP	100.0	100.0	0.0	0.0	25.9	-	-
	S-OMP	0.0	100.0	94.4	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	99.0	0.0	0.2	-	-
Exact Support	SIS-ALASSO	0.0	100.0	100.0	0.0	0.0	80.27	-0.00
	ISIS-ALASSO	0.0	100.0	100.0	0.0	0.0	80.27	-0.00
	OMP	0.0	100.0	86.5	0.0	1222.8	71.40	0.05
	S-OMP-ALASSO	0.0	100.0	100.0	0.0	0.2	80.27	-0.00

Simulation 1: $(n, p, s, T) = (500, 20000, 18, 500)$, $T_{\text{non-zero}} = 300$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 15				
Union Support	SIS-ALASSO	100.0	100.0	0.0	97.0	18.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	98.0	18.0	-	-
	OMP	100.0	100.0	0.0	0.0	23.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	55.0	100.0	0.0	53.0	5399.3	0.10	0.93
	ISIS-ALASSO	100.0	100.0	0.0	98.0	5400.0	0.09	0.93
	OMP	100.0	100.0	0.0	0.0	5405.0	0.07	0.93
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5400.0	0.07	0.93
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	82.0	18.2	-	-
	ISIS-ALASSO	100.0	100.0	0.0	91.0	18.1	-	-
	OMP	100.0	100.0	0.0	0.0	23.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	42.0	100.0	0.0	33.0	5399.2	0.18	0.90
	ISIS-ALASSO	100.0	100.0	0.0	91.0	5400.1	0.16	0.90
	OMP	100.0	100.0	0.0	0.0	5405.0	0.11	0.90
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5400.0	0.11	0.90
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	3.0	21.1	-	-
	ISIS-ALASSO	100.0	100.0	0.0	6.0	20.8	-	-
	OMP	100.0	100.0	0.0	0.0	23.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	24.0	100.0	0.0	1.0	5400.9	0.61	0.82
	ISIS-ALASSO	99.0	100.0	0.0	6.0	5402.8	0.52	0.82
	OMP	100.0	100.0	0.0	0.0	5405.0	0.22	0.82
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5400.0	0.23	0.82
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	97.9	0.0	0.4	-	-
	ISIS-ALASSO	0.0	100.0	97.9	0.0	0.4	-	-
	OMP	100.0	100.0	0.0	0.0	25.9	-	-
	S-OMP	0.0	100.0	94.4	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	94.4	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	100.0	0.0	0.4	48.16	-0.00
	ISIS-ALASSO	0.0	100.0	100.0	0.0	0.4	48.16	-0.00
	OMP	0.0	100.0	10.2	0.0	4858.1	5.76	0.43
	S-OMP-ALASSO	0.0	100.0	99.9	0.0	6.1	48.12	-0.00

Simulation 1: $(n, p, s, T) = (500, 20000, 18, 500)$, $T_{\text{non-zero}} = 100$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 15				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
	OMP	100.0	99.9	0.0	0.0	28.8	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.01	0.91
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.01	0.91
	OMP	100.0	100.0	0.0	0.0	1810.8	0.01	0.91
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.01	0.91
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
	OMP	100.0	99.9	0.0	0.0	28.8	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.01	0.88
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.01	0.88
	OMP	100.0	100.0	0.0	0.0	1810.8	0.01	0.88
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.01	0.88
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
	OMP	100.0	99.9	0.0	0.0	28.8	-	-
	S-OMP	100.0	100.0	0.0	100.0	18.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	18.0	-	-
Exact Support	SIS-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.04	0.79
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.03	0.79
	OMP	100.0	100.0	0.0	0.0	1810.8	0.03	0.79
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1800.0	0.02	0.79
				SNR = 1				
Union Support	SIS-ALASSO	100.0	100.0	0.0	19.0	19.6	-	-
	ISIS-ALASSO	100.0	100.0	0.0	35.0	19.0	-	-
	OMP	100.0	99.9	0.0	0.0	28.8	-	-
	S-OMP	0.0	100.0	94.4	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	94.4	0.0	1.0	-	-
Exact Support	SIS-ALASSO	59.0	100.0	0.0	10.0	1800.9	0.74	0.45
	ISIS-ALASSO	89.0	100.0	0.0	32.0	1800.8	0.63	0.45
	OMP	100.0	100.0	0.0	0.0	1810.8	0.13	0.47
	S-OMP-ALASSO	0.0	100.0	95.3	0.0	84.6	15.31	0.02

Simulation 2.a: $(n, p, s, T) = (200, 5000, 10, 500)$, $T_{\text{non-zero}} = 400$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	39.0	10.9	-	-
	ISIS-ALASSO	100.0	100.0	0.0	12.0	12.2	-	-
	OMP	100.0	99.8	0.0	0.0	21.6	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	6.4	0.0	3746.6	3.58	0.85
	ISIS-ALASSO	41.0	100.0	0.2	3.0	3992.8	0.53	0.90
	OMP	100.0	100.0	0.0	0.0	4011.7	0.22	0.90
	S-OMP-ALASSO	99.0	100.0	0.0	98.0	3999.9	0.22	0.90
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	45.0	11.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	37.0	10.9	-	-
	OMP	100.0	99.8	0.0	0.0	22.2	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	65.9	0.0	1363.5	33.32	0.30
	ISIS-ALASSO	0.0	100.0	63.1	0.0	1477.0	31.89	0.33
	OMP	100.0	100.0	0.0	0.0	4012.2	0.45	0.82
	S-OMP-ALASSO	0.0	100.0	48.0	0.0	2081.5	24.19	0.46
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	98.2	0.0	0.2	-	-
	ISIS-ALASSO	0.0	100.0	98.7	0.0	0.1	-	-
	OMP	100.0	99.5	0.0	0.0	35.2	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	95.4	0.0	0.5	-	-
Exact Support	SIS-ALASSO	0.0	100.0	100.0	0.0	0.2	49.94	-0.00
	ISIS-ALASSO	0.0	100.0	100.0	0.0	0.1	49.94	-0.00
	OMP	0.0	100.0	76.5	0.0	964.4	40.05	0.09
	S-OMP-ALASSO	0.0	100.0	100.0	0.0	0.8	49.94	-0.00

Simulation 2.a: $(n, p, s, T) = (200, 5000, 10, 500)$, $T_{\text{non-zero}} = 250$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	99.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	98.0	10.0	-	-
	OMP	100.0	99.8	0.0	0.0	19.9	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	22.0	100.0	0.2	22.0	2495.4	0.19	0.89
	ISIS-ALASSO	100.0	100.0	0.0	98.0	2500.0	0.12	0.89
	OMP	100.0	100.0	0.0	0.0	2509.9	0.09	0.90
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	2500.0	0.08	0.90
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	44.0	10.8	-	-
	ISIS-ALASSO	100.0	100.0	0.0	46.0	10.8	-	-
	OMP	100.0	99.8	0.0	0.0	19.9	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	12.0	100.0	0.9	6.0	2479.5	0.69	0.80
	ISIS-ALASSO	62.0	100.0	0.2	29.0	2496.7	0.43	0.81
	OMP	100.0	100.0	0.0	0.0	2509.9	0.18	0.81
	S-OMP-ALASSO	95.0	100.0	0.0	95.0	2499.6	0.18	0.81
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	65.3	0.0	3.5	-	-
	ISIS-ALASSO	0.0	100.0	61.3	0.0	3.9	-	-
	OMP	100.0	99.7	0.0	0.0	24.7	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	90.0	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	99.8	0.0	4.6	31.16	-0.00
	ISIS-ALASSO	0.0	100.0	99.8	0.0	5.2	31.15	-0.00
	OMP	0.0	100.0	17.2	0.0	2083.7	6.09	0.39
	S-OMP-ALASSO	0.0	100.0	99.6	0.0	10.4	31.11	-0.00

Simulation 2.a: $(n, p, s, T) = (200, 5000, 10, 500)$, $T_{\text{non-zero}} = 100$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	OMP	100.0	98.8	0.0	0.0	69.8	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	98.0	100.0	0.0	98.0	1000.0	0.02	0.80
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1000.0	0.01	0.80
	OMP	100.0	100.0	0.0	0.0	1060.2	0.02	0.79
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1000.0	0.01	0.80
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	OMP	100.0	98.8	0.0	0.0	69.8	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	98.0	100.0	0.0	98.0	1000.0	0.04	0.73
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1000.0	0.04	0.73
	OMP	100.0	100.0	0.0	0.0	1060.2	0.05	0.72
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1000.0	0.03	0.73
				SNR = 1				
Union Support	SIS-ALASSO	100.0	100.0	0.0	61.0	10.6	-	-
	ISIS-ALASSO	100.0	100.0	0.0	60.0	10.5	-	-
	OMP	100.0	98.8	0.0	0.0	69.8	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	90.0	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	12.7	0.0	873.9	2.23	0.37
	ISIS-ALASSO	0.0	100.0	9.8	0.0	902.8	1.79	0.38
	OMP	100.0	100.0	0.0	0.0	1060.2	0.25	0.42
	S-OMP-ALASSO	0.0	100.0	93.3	0.0	67.4	11.66	0.03

Simulation 2.b: $(n, p, s, T) = (200, 5000, 10, 750)$, $T_{\text{non-zero}} = 600$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	25.0	11.3	-	-
	ISIS-ALASSO	100.0	99.9	0.0	5.0	13.3	-	-
	OMP	100.0	99.7	0.0	0.0	26.6	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	6.9	0.0	5585.0	3.87	0.84
	ISIS-ALASSO	29.0	100.0	0.3	4.0	5986.6	0.56	0.90
	OMP	100.0	100.0	0.0	0.0	6016.7	0.22	0.90
	S-OMP-ALASSO	91.0	100.0	0.0	91.0	5999.1	0.23	0.90
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	27.0	11.4	-	-
	ISIS-ALASSO	100.0	100.0	0.0	28.0	11.3	-	-
	OMP	100.0	99.7	0.0	0.0	27.3	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	66.5	0.0	2011.9	33.60	0.30
	ISIS-ALASSO	0.0	100.0	63.6	0.0	2185.7	32.14	0.32
	OMP	100.0	100.0	0.0	0.0	6017.5	0.45	0.82
	S-OMP-ALASSO	0.0	100.0	48.3	0.0	3104.4	24.34	0.45
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	97.8	0.0	0.2	-	-
	ISIS-ALASSO	0.0	100.0	98.2	0.0	0.2	-	-
	OMP	100.0	99.2	0.0	0.0	47.6	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	94.7	0.0	0.5	-	-
Exact Support	SIS-ALASSO	0.0	100.0	100.0	0.0	0.2	49.94	-0.01
	ISIS-ALASSO	0.0	100.0	100.0	0.0	0.2	49.94	-0.01
	OMP	0.0	100.0	76.7	0.0	1436.7	40.13	0.09
	S-OMP-ALASSO	0.0	100.0	100.0	0.0	1.0	49.94	-0.01

Simulation 2.b: $(n, p, s, T) = (200, 5000, 10, 750)$, $T_{\text{non-zero}} = 375$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	99.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	93.0	10.1	-	-
	OMP	100.0	99.7	0.0	0.0	24.7	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	16.0	100.0	0.2	16.0	3741.3	0.21	0.89
	ISIS-ALASSO	100.0	100.0	0.0	93.0	3750.1	0.12	0.89
	OMP	100.0	100.0	0.0	0.0	3764.8	0.09	0.89
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3750.0	0.09	0.89
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	41.0	10.9	-	-
	ISIS-ALASSO	100.0	100.0	0.0	25.0	11.4	-	-
	OMP	100.0	99.7	0.0	0.0	24.7	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	6.0	100.0	1.0	3.0	3713.5	0.73	0.80
	ISIS-ALASSO	53.0	100.0	0.2	13.0	3744.9	0.43	0.80
	OMP	100.0	100.0	0.0	0.0	3764.8	0.18	0.81
	S-OMP-ALASSO	91.0	100.0	0.0	91.0	3749.0	0.19	0.81
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	55.8	0.0	4.4	-	-
	ISIS-ALASSO	1.0	100.0	52.8	1.0	4.7	-	-
	OMP	100.0	99.6	0.0	0.0	32.0	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	90.0	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	99.8	0.0	6.6	31.16	-0.00
	ISIS-ALASSO	0.0	100.0	99.8	0.0	7.3	31.16	-0.00
	OMP	0.0	100.0	17.6	0.0	3111.8	6.21	0.39
	S-OMP-ALASSO	0.0	100.0	99.6	0.0	15.1	31.11	-0.00

Simulation 2.b: $(n, p, s, T) = (200, 5000, 10, 750)$, $T_{\text{non-zero}} = 150$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	OMP	100.0	98.0	0.0	0.0	108.5	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	98.0	100.0	0.0	98.0	1500.0	0.02	0.79
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1500.0	0.02	0.79
	OMP	100.0	100.0	0.0	0.0	1599.5	0.03	0.78
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1500.0	0.01	0.79
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	OMP	100.0	98.0	0.0	0.0	108.5	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	98.0	100.0	0.0	98.0	1500.0	0.04	0.72
	ISIS-ALASSO	100.0	100.0	0.0	100.0	1500.0	0.04	0.72
	OMP	100.0	100.0	0.0	0.0	1599.5	0.05	0.71
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1500.0	0.03	0.72
				SNR = 1				
Union Support	SIS-ALASSO	100.0	100.0	0.0	46.0	10.8	-	-
	ISIS-ALASSO	100.0	100.0	0.0	42.0	10.8	-	-
	OMP	100.0	98.0	0.0	0.0	108.5	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	90.0	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	12.1	0.0	1318.9	2.16	0.37
	ISIS-ALASSO	0.0	100.0	9.4	0.0	1360.3	1.74	0.38
	OMP	100.0	100.0	0.0	0.0	1599.5	0.26	0.42
	S-OMP-ALASSO	0.0	100.0	93.4	0.0	98.9	11.68	0.03

Simulation 2.c: $(n, p, s, T) = (200, 5000, 10, 1000)$, $T_{\text{non-zero}} = 800$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\vert\hat{S}\vert$	$\vert\vert\mathbf{B} - \hat{\mathbf{B}}\vert\vert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	21.0	11.7	-	-
	ISIS-ALASSO	100.0	99.9	0.0	5.0	14.4	-	-
	OMP	100.0	99.6	0.0	0.0	32.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	7.7	0.0	7382.7	4.26	0.84
	ISIS-ALASSO	17.0	100.0	0.4	1.0	7976.0	0.60	0.90
	OMP	100.0	100.0	0.0	0.0	8022.1	0.22	0.90
	S-OMP-ALASSO	86.0	100.0	0.0	86.0	7998.3	0.23	0.90
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	14.0	11.9	-	-
	ISIS-ALASSO	100.0	100.0	0.0	17.0	11.7	-	-
	OMP	100.0	99.5	0.0	0.0	33.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	65.5	0.0	2759.0	33.13	0.31
	ISIS-ALASSO	0.0	100.0	62.7	0.0	2984.0	31.71	0.33
	OMP	100.0	100.0	0.0	0.0	8023.1	0.45	0.82
	S-OMP-ALASSO	0.0	100.0	48.1	0.0	4152.9	24.25	0.46
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	97.6	0.0	0.2	-	-
	ISIS-ALASSO	0.0	100.0	97.3	0.0	0.3	-	-
	OMP	100.0	99.0	0.0	0.0	59.5	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	93.0	0.0	0.7	-	-
Exact Support	SIS-ALASSO	0.0	100.0	100.0	0.0	0.3	49.94	-0.01
	ISIS-ALASSO	0.0	100.0	100.0	0.0	0.3	49.94	-0.01
	OMP	0.0	100.0	76.4	0.0	1942.8	39.98	0.10
	S-OMP-ALASSO	0.0	100.0	100.0	0.0	1.8	49.94	-0.01

Simulation 2.c: $(n, p, s, T) = (200, 5000, 10, 1000)$, $T_{\text{non-zero}} = 500$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	89.0	10.1	-	-
	ISIS-ALASSO	100.0	100.0	0.0	95.0	10.1	-	-
	OMP	100.0	99.6	0.0	0.0	29.1	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	15.0	100.0	0.2	13.0	4990.2	0.19	0.89
	ISIS-ALASSO	100.0	100.0	0.0	95.0	5000.1	0.12	0.89
	OMP	100.0	100.0	0.0	0.0	5019.2	0.09	0.89
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5000.0	0.09	0.89
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	27.0	11.4	-	-
	ISIS-ALASSO	100.0	100.0	0.0	14.0	11.6	-	-
	OMP	100.0	99.6	0.0	0.0	29.1	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	1.0	100.0	0.8	0.0	4958.9	0.69	0.80
	ISIS-ALASSO	39.0	100.0	0.2	10.0	4991.9	0.44	0.81
	OMP	100.0	100.0	0.0	0.0	5019.2	0.18	0.81
	S-OMP-ALASSO	88.0	100.0	0.0	87.0	4998.8	0.19	0.81
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	46.3	0.0	5.4	-	-
	ISIS-ALASSO	1.0	100.0	42.8	1.0	5.7	-	-
	OMP	100.0	99.4	0.0	0.0	38.9	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	90.0	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	99.8	0.0	8.6	31.16	-0.00
	ISIS-ALASSO	0.0	100.0	99.8	0.0	9.6	31.16	-0.00
	OMP	0.0	100.0	17.5	0.0	4155.6	6.16	0.39
	S-OMP-ALASSO	0.0	100.0	99.6	0.0	20.1	31.11	-0.00

Simulation 2.c: $(n, p, s, T) = (200, 5000, 10, 1000)$, $T_{\text{non-zero}} = 200$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
SNR = 10								
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	OMP	100.0	97.4	0.0	0.0	139.6	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.02	0.79
	ISIS-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.02	0.79
	OMP	100.0	100.0	0.0	0.0	2131.6	0.03	0.78
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.01	0.79
SNR = 5								
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
	OMP	100.0	97.4	0.0	0.0	139.6	-	-
	S-OMP	100.0	100.0	0.0	100.0	10.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	10.0	-	-
Exact Support	SIS-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.04	0.72
	ISIS-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.04	0.72
	OMP	100.0	100.0	0.0	0.0	2131.6	0.05	0.71
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.03	0.72
SNR = 1								
Union Support	SIS-ALASSO	100.0	100.0	0.0	37.0	11.1	-	-
	ISIS-ALASSO	100.0	100.0	0.0	44.0	10.8	-	-
	OMP	100.0	97.4	0.0	0.0	139.6	-	-
	S-OMP	0.0	100.0	90.0	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	90.0	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	12.0	0.0	1761.3	2.15	0.37
	ISIS-ALASSO	0.0	100.0	9.1	0.0	1819.3	1.71	0.38
	OMP	99.0	100.0	0.0	0.0	2131.6	0.26	0.42
	S-OMP-ALASSO	0.0	100.0	93.2	0.0	136.0	11.65	0.03

Simulation 3: $(n, p, s, T) = (100, 5000, 3, 150)$, $T_{\text{non-zero}} = 80$, $\rho = 0.2$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
	OMP	100.0	99.8	0.0	0.0	20.0	-	-
	S-OMP	100.0	100.0	0.0	100.0	3.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
Exact Support	SIS-ALASSO	96.0	100.0	0.0	96.0	239.9	0.02	0.73
	ISIS-ALASSO	100.0	100.0	0.0	100.0	240.0	0.02	0.73
	OMP	100.0	100.0	0.0	0.0	257.1	0.03	0.72
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	240.0	0.01	0.73
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
	OMP	100.0	99.8	0.0	0.0	19.6	-	-
	S-OMP	100.0	100.0	0.0	100.0	3.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
Exact Support	SIS-ALASSO	100.0	100.0	0.0	100.0	240.0	0.02	0.72
	ISIS-ALASSO	100.0	100.0	0.0	100.0	240.0	0.02	0.72
	OMP	100.0	100.0	0.0	0.0	256.6	0.03	0.72
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	240.0	0.01	0.72
				SNR = 1				
Union Support	SIS-ALASSO	100.0	100.0	0.0	92.0	3.1	-	-
	ISIS-ALASSO	100.0	100.0	0.0	94.0	3.1	-	-
	OMP	100.0	99.8	0.0	0.0	20.3	-	-
	S-OMP	100.0	100.0	0.0	100.0	3.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
Exact Support	SIS-ALASSO	99.0	100.0	0.0	92.0	240.1	0.04	0.70
	ISIS-ALASSO	100.0	100.0	0.0	94.0	240.1	0.03	0.70
	OMP	100.0	100.0	0.0	0.0	257.3	0.04	0.69
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	240.0	0.02	0.70

Simulation 3: $(n, p, s, T) = (100, 5000, 3, 150)$, $T_{\text{non-zero}} = 80$, $\rho = 0.5$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lVert\mathbf{B} - \hat{\mathbf{B}}\rVert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	98.0	3.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
	OMP	100.0	99.8	0.0	0.0	20.1	-	-
	S-OMP	100.0	100.0	0.0	100.0	3.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
Exact Support	SIS-ALASSO	87.0	100.0	0.2	85.0	239.5	0.08	0.62
	ISIS-ALASSO	88.0	100.0	0.1	88.0	239.8	0.07	0.62
	OMP	100.0	100.0	0.0	0.0	257.1	0.06	0.62
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	240.0	0.03	0.63
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	97.0	3.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	96.0	3.0	-	-
	OMP	100.0	99.8	0.0	0.0	19.6	-	-
	S-OMP	100.0	100.0	0.0	100.0	3.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
Exact Support	SIS-ALASSO	60.0	100.0	0.2	57.0	239.5	0.10	0.61
	ISIS-ALASSO	84.0	100.0	0.1	80.0	239.8	0.08	0.61
	OMP	100.0	100.0	0.0	0.0	256.6	0.06	0.61
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	240.0	0.03	0.62
				SNR = 1				
Union Support	SIS-ALASSO	100.0	100.0	0.0	56.0	3.5	-	-
	ISIS-ALASSO	100.0	100.0	0.0	70.0	3.4	-	-
	OMP	100.0	99.8	0.0	0.0	19.9	-	-
	S-OMP	100.0	100.0	0.0	100.0	3.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	3.0	-	-
Exact Support	SIS-ALASSO	1.0	100.0	2.3	1.0	235.1	0.21	0.58
	ISIS-ALASSO	5.0	100.0	1.5	3.0	236.8	0.16	0.58
	OMP	96.0	100.0	0.0	0.0	256.9	0.08	0.58
	S-OMP-ALASSO	67.0	100.0	0.2	67.0	239.5	0.05	0.59

Simulation 3: $(n, p, s, T) = (100, 5000, 3, 150)$, $T_{\text{non-zero}} = 80$, $\rho = 0.7$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	80.0	100.0	6.7	80.0	2.8	-	-
	ISIS-ALASSO	85.0	100.0	5.0	85.0	2.9	-	-
	OMP	100.0	99.8	0.0	0.0	22.0	-	-
	S-OMP	0.0	100.0	51.0	0.0	1.5	-	-
	S-OMP-ALASSO	0.0	100.0	51.0	0.0	1.5	-	-
Exact Support	SIS-ALASSO	0.0	100.0	63.3	0.0	88.1	3.93	0.15
	ISIS-ALASSO	0.0	100.0	61.0	0.0	93.6	3.70	0.16
	OMP	0.0	100.0	12.0	0.0	230.2	0.73	0.28
	S-OMP-ALASSO	0.0	100.0	57.6	0.0	101.8	2.89	0.19
				SNR = 5				
Union Support	SIS-ALASSO	79.0	100.0	7.0	79.0	2.8	-	-
	ISIS-ALASSO	85.0	100.0	5.0	83.0	2.9	-	-
	OMP	100.0	99.8	0.0	0.0	22.5	-	-
	S-OMP	0.0	100.0	56.7	0.0	1.3	-	-
	S-OMP-ALASSO	0.0	100.0	56.7	0.0	1.3	-	-
Exact Support	SIS-ALASSO	0.0	100.0	66.0	0.0	81.6	4.15	0.14
	ISIS-ALASSO	0.0	100.0	64.2	0.0	85.9	3.95	0.15
	OMP	0.0	100.0	16.5	0.0	219.8	0.96	0.26
	S-OMP-ALASSO	0.0	100.0	61.2	0.0	93.0	3.16	0.18
				SNR = 1				
Union Support	SIS-ALASSO	89.0	100.0	3.7	45.0	3.5	-	-
	ISIS-ALASSO	92.0	100.0	2.7	49.0	3.5	-	-
	OMP	100.0	99.8	0.0	0.0	27.7	-	-
	S-OMP	0.0	100.0	60.3	0.0	1.2	-	-
	S-OMP-ALASSO	0.0	100.0	60.3	0.0	1.2	-	-
Exact Support	SIS-ALASSO	0.0	100.0	71.4	0.0	69.4	4.76	0.11
	ISIS-ALASSO	0.0	100.0	68.9	0.0	75.3	4.46	0.12
	OMP	0.0	100.0	29.3	0.0	196.8	1.96	0.23
	S-OMP-ALASSO	0.0	100.0	64.6	0.0	85.0	3.53	0.16

Simulation 4: $(n, p, s, T) = (150, 4000, 8, 150)$, $T_{\text{non-zero}} = 80$, $\rho = 0.2$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	100.0	8.0	-	-
	ISIS-ALASSO	100.0	100.0	0.0	97.0	8.0	-	-
	OMP	100.0	99.9	0.0	2.0	11.7	-	-
	S-OMP	100.0	100.0	0.0	100.0	8.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	8.0	-	-
Exact Support	SIS-ALASSO	35.0	100.0	1.4	35.0	631.3	0.55	0.88
	ISIS-ALASSO	100.0	100.0	0.0	97.0	640.0	0.14	0.89
	OMP	100.0	100.0	0.0	2.0	643.7	0.10	0.89
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	640.0	0.09	0.89
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	85.0	8.2	-	-
	ISIS-ALASSO	100.0	100.0	0.0	78.0	8.3	-	-
	OMP	100.0	99.9	0.0	2.0	11.7	-	-
	S-OMP	100.0	100.0	0.0	100.0	8.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	8.0	-	-
Exact Support	SIS-ALASSO	2.0	100.0	4.5	2.0	611.7	1.78	0.77
	ISIS-ALASSO	7.0	100.0	2.9	6.0	621.5	1.29	0.78
	OMP	100.0	100.0	0.0	2.0	643.7	0.20	0.80
	S-OMP-ALASSO	39.0	100.0	1.0	39.0	633.8	0.48	0.80
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	90.5	0.0	0.8	-	-
	ISIS-ALASSO	0.0	100.0	87.6	0.0	1.0	-	-
	OMP	100.0	99.8	0.0	0.0	14.9	-	-
	S-OMP	0.0	100.0	87.5	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	88.5	0.0	0.9	-	-
Exact Support	SIS-ALASSO	0.0	100.0	99.9	0.0	0.8	29.62	-0.01
	ISIS-ALASSO	0.0	100.0	99.8	0.0	1.1	29.61	-0.01
	OMP	0.0	100.0	31.1	0.0	447.7	10.11	0.32
	S-OMP-ALASSO	0.0	100.0	99.6	0.0	2.7	29.56	-0.00

Simulation 4: $(n, p, s, T) = (150, 4000, 8, 150)$, $T_{\text{non-zero}} = 80$, $\rho = 0.5$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				SNR = 10				
Union Support	SIS-ALASSO	100.0	100.0	0.0	80.0	8.2	-	-
	ISIS-ALASSO	100.0	100.0	0.0	89.0	8.1	-	-
	OMP	100.0	99.9	0.0	2.0	11.9	-	-
	S-OMP	100.0	100.0	0.0	100.0	8.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	8.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	13.1	0.0	556.5	4.24	0.80
	ISIS-ALASSO	80.0	100.0	0.2	70.0	638.9	0.23	0.89
	OMP	100.0	100.0	0.0	2.0	643.9	0.11	0.89
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	640.0	0.10	0.89
				SNR = 5				
Union Support	SIS-ALASSO	100.0	100.0	0.0	69.0	8.4	-	-
	ISIS-ALASSO	100.0	100.0	0.0	47.0	8.9	-	-
	OMP	100.0	99.9	0.0	2.0	12.3	-	-
	S-OMP	100.0	100.0	0.0	100.0	8.0	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	8.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	23.8	0.0	487.8	7.53	0.65
	ISIS-ALASSO	0.0	100.0	7.6	0.0	592.5	2.75	0.75
	OMP	99.0	100.0	0.0	2.0	644.4	0.22	0.80
	S-OMP-ALASSO	7.0	100.0	2.8	7.0	622.2	1.04	0.79
				SNR = 1				
Union Support	SIS-ALASSO	0.0	100.0	60.6	0.0	3.2	-	-
	ISIS-ALASSO	1.0	100.0	56.8	1.0	3.5	-	-
	OMP	100.0	99.6	0.0	0.0	23.5	-	-
	S-OMP	0.0	100.0	87.5	0.0	1.0	-	-
	S-OMP-ALASSO	0.0	100.0	87.5	0.0	1.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	99.3	0.0	4.7	29.45	-0.00
	ISIS-ALASSO	0.0	100.0	99.2	0.0	5.1	29.43	-0.00
	OMP	0.0	100.0	44.9	0.0	369.3	15.05	0.28
	S-OMP-ALASSO	0.0	100.0	98.5	0.0	9.9	29.39	0.01

Simulation 5: $(n, p, s, T) = (200, 10000, 5, 500)$, $T_{\text{non-zero}} = 400$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lVert\mathbf{B} - \hat{\mathbf{B}}\rVert_2^2$	R^2
				$\sigma = 1.5$				
Union Support	SIS-ALASSO	53.0	99.6	9.4	0.0	41.1	-	-
	ISIS-ALASSO	100.0	99.8	0.0	0.0	28.1	-	-
	OMP	100.0	99.9	0.0	12.0	10.0	-	-
	S-OMP	100.0	100.0	0.0	44.0	5.6	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	68.9	0.0	936.0	84.66	0.66
	ISIS-ALASSO	0.0	100.0	16.2	0.0	1791.9	5.80	0.96
	OMP	100.0	100.0	0.0	12.0	2090.3	0.06	0.99
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	2000.0	0.05	0.99
				$\sigma = 2.5$				
Union Support	SIS-ALASSO	53.0	99.4	9.4	0.0	61.4	-	-
	ISIS-ALASSO	100.0	99.3	0.0	0.0	77.7	-	-
	OMP	100.0	99.9	0.0	10.0	13.2	-	-
	S-OMP	100.0	100.0	0.0	44.0	5.6	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	69.2	0.0	910.2	85.82	0.64
	ISIS-ALASSO	0.0	100.0	17.5	0.0	1834.1	7.23	0.93
	OMP	100.0	100.0	0.0	10.0	2093.3	0.16	0.96
	S-OMP-ALASSO	93.0	100.0	0.0	93.0	1999.9	0.13	0.96
				$\sigma = 4.5$				
Union Support	SIS-ALASSO	40.0	99.1	12.0	0.0	92.5	-	-
	ISIS-ALASSO	100.0	97.8	0.0	0.0	226.8	-	-
	OMP	100.0	99.8	0.0	1.0	25.7	-	-
	S-OMP	92.0	100.0	1.6	46.0	5.5	-	-
	S-OMP-ALASSO	92.0	100.0	1.6	92.0	5.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	70.0	0.0	850.2	88.65	0.56
	ISIS-ALASSO	0.0	100.0	27.4	0.0	1847.2	15.79	0.83
	OMP	0.0	100.0	3.2	0.0	2040.9	1.15	0.88
	S-OMP-ALASSO	0.0	100.0	10.2	0.0	1795.3	2.38	0.87

	Simulation 5: $(n, p, s, T) = (200, 10000, 5, 500)$, $T_{\text{non-zero}} = 250$							
	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				$\sigma = 1.5$				
Union Support	SIS-ALASSO	100.0	99.7	0.0	0.0	31.5	-	-
	ISIS-ALASSO	100.0	99.9	0.0	1.0	14.3	-	-
	OMP	100.0	99.7	0.0	0.0	30.8	-	-
	S-OMP	100.0	100.0	0.0	20.0	5.8	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	45.9	0.0	768.9	25.98	0.79
	ISIS-ALASSO	0.0	100.0	5.3	0.0	1200.7	1.00	0.92
	OMP	100.0	100.0	0.0	0.0	1287.6	0.05	0.92
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	1250.0	0.03	0.92
				$\sigma = 2.5$				
Union Support	SIS-ALASSO	100.0	99.6	0.0	0.0	40.5	-	-
	ISIS-ALASSO	100.0	99.6	0.0	0.0	44.3	-	-
	OMP	100.0	99.7	0.0	0.0	32.0	-	-
	S-OMP	100.0	100.0	0.0	23.0	5.8	-	-
	S-OMP-ALASSO	100.0	100.0	0.0	100.0	5.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	46.2	0.0	757.5	26.30	0.74
	ISIS-ALASSO	0.0	100.0	7.5	0.0	1205.2	1.55	0.86
	OMP	100.0	100.0	0.0	0.0	1288.6	0.14	0.87
	S-OMP-ALASSO	92.0	100.0	0.0	92.0	1249.9	0.08	0.87
				$\sigma = 4.5$				
Union Support	SIS-ALASSO	98.0	99.6	0.4	0.0	48.0	-	-
	ISIS-ALASSO	100.0	99.0	0.0	0.0	104.0	-	-
	OMP	100.0	99.7	0.0	0.0	36.1	-	-
	S-OMP	1.0	100.0	19.8	1.0	4.7	-	-
	S-OMP-ALASSO	1.0	100.0	19.8	1.0	4.2	-	-
Exact Support	SIS-ALASSO	0.0	100.0	48.4	0.0	713.1	27.64	0.62
	ISIS-ALASSO	0.0	100.0	22.8	0.0	1080.7	5.57	0.71
	OMP	0.0	100.0	2.3	0.0	1264.0	0.70	0.75
	S-OMP-ALASSO	0.0	100.0	19.9	0.0	1002.0	2.26	0.73

Simulation 5: $(n, p, s, T) = (200, 10000, 5, 500)$, $T_{\text{non-zero}} = 100$

	Method name	$\mathcal{M}_* \subseteq \hat{S}$	Correct zeros	Incorrect zeros	$\mathcal{M}_* = \hat{S}$	$\lvert\hat{S}\rvert$	$\lvert\lvert\mathbf{B} - \hat{\mathbf{B}}\rvert\rvert_2^2$	R^2
				$\sigma = 1.5$				
Union Support	SIS-ALASSO	100.0	99.9	0.0	1.0	10.9	-	-
	ISIS-ALASSO	100.0	100.0	0.0	56.0	5.7	-	-
	OMP	100.0	98.0	0.0	0.0	205.8	-	-
	S-OMP	99.0	100.0	0.2	4.0	6.0	-	-
	S-OMP-ALASSO	99.0	100.0	0.2	99.0	5.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	19.4	0.0	411.0	2.86	0.60
	ISIS-ALASSO	17.0	100.0	0.5	16.0	498.0	0.06	0.62
	OMP	100.0	100.0	0.0	0.0	726.4	0.19	0.60
	S-OMP-ALASSO	99.0	100.0	0.2	99.0	499.0	0.02	0.62
				$\sigma = 2.5$				
Union Support	SIS-ALASSO	100.0	99.9	0.0	1.0	11.0	-	-
	ISIS-ALASSO	100.0	99.9	0.0	0.0	12.4	-	-
	OMP	100.0	98.0	0.0	0.0	205.8	-	-
	S-OMP	0.0	100.0	20.0	0.0	4.9	-	-
	S-OMP-ALASSO	0.0	100.0	20.0	0.0	4.0	-	-
Exact Support	SIS-ALASSO	0.0	100.0	19.6	0.0	408.8	2.92	0.54
	ISIS-ALASSO	0.0	100.0	2.5	0.0	495.2	0.21	0.56
	OMP	100.0	100.0	0.0	0.0	726.4	0.54	0.53
	S-OMP-ALASSO	0.0	100.0	20.0	0.0	400.0	0.83	0.52
				$\sigma = 4.5$				
Union Support	SIS-ALASSO	98.0	100.0	0.4	1.0	9.8	-	-
	ISIS-ALASSO	97.0	99.9	0.6	0.0	17.4	-	-
	OMP	100.0	98.0	0.0	0.0	206.4	-	-
	S-OMP	0.0	100.0	41.2	0.0	3.6	-	-
	S-OMP-ALASSO	0.0	100.0	41.2	0.0	3.4	-	-
Exact Support	SIS-ALASSO	0.0	100.0	27.6	0.0	367.3	3.48	0.41
	ISIS-ALASSO	0.0	100.0	19.9	0.0	413.1	1.33	0.42
	OMP	4.0	100.0	1.4	0.0	720.0	1.79	0.41
	S-OMP-ALASSO	0.0	100.0	41.2	0.0	295.9	4.66	0.35

References

Peter J. Bickel and Elizaveta Levina. Regularized estimation of large covariance matrices. *Annals of Statistics*, 36(1):199–227, 2008.

Jianqing Fan and Jinchi Lv. Sure independence screening for ultrahigh dimensional feature space. *Journal Of The Royal Statistical Society Series B*, 70 (5):849–911, 2008.

Béatrice Laurent and Pascal Massart. Adaptive estimation of a quadratic functional by model selection. *Annals of Statistics*, 28(5):1302–1338, 2000.

Guillaume Obozinski, Martin Wainwright, and Michael Jordan. High-dimensional support union recovery in multivariate regression. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1217–1224. 2009.

Hansheng Wang. Forward Regression for Ultra-High Dimensional Variable Screening. *SSRN eLibrary*, 2009.

Shuheng Zhou, Sara van de Geer, and Peter Buhlmann. Adaptive lasso for high dimensional regression and gaussian graphical modeling, 2009.