Introduction 00	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion O	References

Robust Lasso-Zero for sparse corruption and model selection with missing covariates One World : statistical learning seminars

¹University of Geneva, Switzerland

²Sorbonne University, France

³Ecole Normale Supérieure, France

⁴Ecole Polytechnique, France

⁵INRIA, France

⁶Visiting Researcher Google Brain, France

Introduction ●○	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion 0	References
<u> </u>					

The sparse corruptions problem

For taking into account occasional corruptions:

Sparse corruptions problem

$$y = X\beta^0 + \sqrt{n}\omega^0 + \epsilon$$

- $y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n \times p}$, $\beta^0 \in \mathbb{R}^p$, $\omega^0 \in \mathbb{R}^n$
- $\epsilon \sim \mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbf{I}_n).$
- high-dimension: $p \gg n$, rank(X) = n.
- sparsity: β^0 is s-sparse, ω^0 is k-sparse.

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion	References
•0	0000	000	000000	0	00000

The sparse corruptions problem

For taking into account additional occasional corruptions:

Sparse corruptions problem

$$y = X\beta^0 + \sqrt{n}\omega^0 + \epsilon$$

- $y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n \times p}$, $\beta^0 \in \mathbb{R}^p$, $\omega^0 \in \mathbb{R}^n$
- $\epsilon \sim \mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbf{I}_n).$
- high-dimension: $p \gg n$, rank(X) = n.
- sparsity: β^0 is *s*-sparse, ω^0 is *k*-sparse.

Sparse linear model with an augmented design matrix and sparse vector

$$y = \begin{bmatrix} X & \sqrt{n} I_n \end{bmatrix} \begin{bmatrix} \beta^0 \\ \omega^0 \end{bmatrix} + \epsilon.$$

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion	References
•0	0000	000	000000	0	00000
T I		1.1			

The sparse corruptions problem

Face recognition problem



Figure: Corrupted image, a sparse linear combination of all the training images (middle) plus sparse errors (right) due to corruption. Red (darker) coefficients correspond to training images of the correct individual.

Credit: [Wright et al., 2009].

Introduction ○●	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion 0	References
Existing	works				

With sparse noise ($\omega^0 \neq 0$), without dense noise ($\epsilon = 0$): Justice Pursuit (JP)

$$\min_{\beta \in \mathbb{R}^{p}, \omega \in \mathbb{R}^{n}} \|\beta\|_{1} + \lambda \|\omega\|_{1} \quad \text{s.t.} \quad y = X\beta + \sqrt{n}\omega, \ \lambda > 0$$

	Condition on (y, X)	Recovery of (β^0, ω^0)
Wright, Ma (2010)	y Gaussian	✓ (support)
Laska et al. (2009) Li et al. (2010)	y, X Gaussian	√ (exact)
JP with tuned paramater Li (2013) Nguyen and Tran (2013b)	X sub-orthogonal Gaussian design	√ (exact)

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで

Introduction ○●	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion 0	References
Existing	works				

With sparse noise $(\omega^0 \neq 0)$, without dense noise $(\epsilon = 0)$: Justice Pursuit (JP)

$$\min_{\beta \in \mathbb{R}^{p}, \omega \in \mathbb{R}^{n}} \|\beta\|_{1} + \lambda \|\omega\|_{1} \quad \text{s.t.} \quad y = X\beta + \sqrt{n}\omega, \, \lambda > 0$$

With sparse noise $(\omega^0 \neq 0)$ and dense noise $(\epsilon \neq 0)$: Robust Lasso

$$\min_{\beta \in \mathbb{R}^{p}, \omega \in \mathbb{R}^{n}} \frac{1}{2} \| y - X\beta - \omega \|_{2}^{2} + \lambda_{\beta} \| \beta \|_{1} + \lambda_{\omega} \| \omega \|_{1}.$$

	Condition on (y, X)	Recovery of (β^0, ω^0)
Nguyen and Tran (2013b)	X Gaussian invertible covariance matrix	√ (sign)
ℓ_1 - penalized Huber's <i>M</i> - estimator Dalalyan and Thompson (2019)	X Gaussian invertible covariance matrix	√ (sign)

Introduction 0	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion 0	References
Existing	works				

With sparse noise
$$(\omega^0 \neq 0)$$
, without dense noise $(\epsilon = 0)$:
Justice Pursuit (JP)
$$\min_{\beta \in \mathbb{R}^{p}, \omega \in \mathbb{R}^{n}} \|\beta\|_{1} + \lambda \|\omega\|_{1} \quad \text{s.t.} \quad y = X\beta + \sqrt{n}\omega, \ \lambda > 0$$

With sparse noise
$$(\omega^0 \neq 0)$$
 and dense noise $(\epsilon \neq 0)$:
Robust Lasso
$$\min_{\beta \in \mathbb{R}^p, \omega \in \mathbb{R}^n} \frac{1}{2} \|y - X\beta - \omega\|_2^2 + \lambda_\beta \|\beta\|_1 + \lambda_\omega \|\omega\|_1.$$

• Our proposal: same problem but different strategy, solving Justice Pursuit and thresholding.

Introduction 00	Robust Lasso-Zero ●000	Missing data	Numerical experiments 000000	Conclusion O	References
Our stra Sparse-linear	tegy: "Overf model	it, then tl	hreshold."		

Strategy already introduced for the sparse-linear model $y = X\beta^0 + \epsilon$.

▲□▶ ▲□▶ ▲三▶ ▲三▶ - 三 - のへで

Thresholded Basis Pursuit [Saligrama and Zhao, 2011]

- solving the Basis Pursuit $\min_{\beta \in \mathbb{R}^p} \|\beta\|_1$ s.t. $y = X\beta$.
- setting the small coefficients to zero.

× noise generally overfitted.

Introduction 00	Robust Lasso-Zero ●000	Missing data	Numerical experiments 000000	Conclusion O	References
Our stra Sparse-linea	ategy: "Overf ar model	it, then t	hreshold."		

Strategy already introduced for the sparse-linear model $y = X\beta^0 + \epsilon$.

Thresholded Basis Pursuit [Saligrama and Zhao, 2011]

- solving the Basis Pursuit $\min_{\beta \in \mathbb{R}^p} \|\beta\|_1$ s.t. $y = X\beta$.
- setting the small coefficients to zero.

× noise generally overfitted.

Lasso-Zero [Descloux and Sardy, 2018]

- For $k \in \{0, ..., M\}$
 - use a Gaussian noise dictionary $G^{(k)} \in \mathbb{R}^{n \times q}, q > 0$.
 - solve BP problems with the augmented matrix $[X|G^{(k)}]$.
- Aggregate the obtained estimates $\hat{\beta}^{(1)}, \ldots, \hat{\beta}^{(M)}$ with the component-wise medians.
- Threshold the aggregated estimator at level $\tau > 0$.

Introduction	Robust Lasso-Zero 0●00	Missing data	Numerical experiments	Conclusion 0	References
Our stra	ategy: "Over	fit, then t	hreshold."		
Sparse corru	uption model				

"Thresholded Justice Pursuit"

- solving the Justice Pursuit $\Rightarrow \hat{\beta}_{\lambda}^{\rm JP}, \hat{\omega}_{\lambda}^{\rm JP}$
- hard-thresholding the solution $\hat{\beta}_{(\lambda,\tau)}^{\mathrm{TJP}} = \eta_{\tau}(\hat{\beta}_{\lambda}^{\mathrm{JP}})$ and $\hat{\omega}_{(\lambda,\tau)}^{\mathrm{TJP}} = \eta_{\tau}(\hat{\omega}_{\lambda}^{\mathrm{JP}})$.

Robust Lasso-Zero

- For $k \in \{0, ..., M\}$
 - use a Gaussian noise dictionary $G^{(k)} \in \mathbb{R}^{n \times q}, q > 0$.
 - solve the augmented JP problem

 $\begin{aligned} (\hat{\beta}_{\lambda}^{(k)}, \hat{\omega}_{\lambda}^{(k)}, \hat{\gamma}_{\lambda}^{(k)}) \in & \arg\min_{\beta \in \mathbb{R}^{p}, \ \omega \in \mathbb{R}^{n}, \ \gamma \in \mathbb{R}^{n}} & \|\beta\|_{1} + \lambda \|\omega\|_{1} + \|\gamma\|_{1} \\ \text{s.t.} & y = X\beta + \sqrt{n}\omega + G^{(k)}\gamma. \end{aligned}$

- Aggregate the obtained estimates $\hat{\beta}_{\lambda}^{(1)}, \ldots, \hat{\beta}_{\lambda}^{(M)}$ with the component-wise medians $\Rightarrow \hat{\beta}_{\lambda}^{med}$
- Hard-threshold $\hat{\beta}^{\text{Rlass0}}_{(\lambda,\tau)} := \eta_{\tau}(\hat{\beta}^{\text{med}}_{\lambda}) = \hat{\beta}^{\text{med}}_{\lambda} \mathbf{1}_{(\tau,+\infty)}(|\hat{\beta}^{\text{med}}_{\lambda}|).$

Introduction 00 Robust Lasso-Zero

Missing data

Numerical experiments

Conclusion

References

Theoretical guarantees on Thresholded Justice Pursuit Identifiability as a necessary and sufficient condition for consistent sign recovery

 \Leftrightarrow see the Robust Lasso-Zero as an extension of the Thresholded Justice Pursuit (TJP). \rightarrow theoretical results derived for TJP.

Identifiability for the TJP

= Extension of [Tardivel and Bogdan, 2019] for the TBP.

 $(\beta^0, \omega^0) \in \mathbb{R}^p \times \mathbb{R}^n$ is identifiable with respect to $X \in \mathbb{R}^{n \times p}$ and $\lambda > 0$ if it is the unique solution to JP when $y = X\beta^0 + \sqrt{n}\omega^0$ (noiseless case).

For a fixed matrix $X \in \mathbb{R}^{n \times p}$ and a sequence $\{(\beta^{(r)}, \omega^{(r)})\}_{r \in \mathbb{N}^*}$ assume

- the sign vectors of $\beta^{(r)}$ and $\omega^{(r)}$ are invariant, i.e.

 $\exists \theta \in \{1, -1, 0\}^{\rho} \text{ such that } \operatorname{sign}(\beta^{(r)}) = \theta, \forall r \in \mathbb{N}^{*} \text{ (resp. for } \omega^{(r)}),$

• the nonzero coefficients are large i.e.

$$\lim_{r \to +\infty} \min\{\beta_{\min}^{(r)}, \omega_{\min}^{(r)}\} = +\infty \quad \text{and} \quad \exists q > 0, \frac{\min\{\beta_{\min}^{(r)}, \omega_{\min}^{(r)}\}}{\max\{\|\beta^{(r)}\|_{\infty}, \|\omega^{(r)}\|_{\infty}\}} \ge q,$$

where $\beta_{\min} := \min_{j \in \text{supp}(\beta)} |\beta_j|.$

Introduction

Robust Lasso-Zero

Missing data

Numerical experiments

Conclusion

References

Theoretical guarantees on Thresholded Justice Pursuit Identifiability as a necessary and sufficient condition for consistent sign recovery

 $(\hat{\beta}_{(\lambda,\tau)}^{\mathsf{TJP}(r)},\hat{\omega}_{(\lambda,\tau)}^{\mathsf{TJP}(r)}): \text{ TJP estimates when } y = y^{(r)} := X\beta^{(r)} + \sqrt{n}\omega^{(r)} + \epsilon.$

Theorem 1 (Descloux, Boyer, Josse, S., Sardy, 2020)

Let $\lambda > 0$ and $X \in \mathbb{R}^{n \times n}$ such that for any $y \in \mathbb{R}^n$, the JP solution is unique. Let $\{(\beta^{(r)}, \omega^{(r)})\}_{r \in \mathbb{N}^*}$ be a sequence satisfying assumptions above.

• If the pair of sign vectors $(\theta, \tilde{\theta})$ is identifiable w.r.t. X and λ , then $\exists R, \forall r \ge R$, there is a threshold $\tau = \tau(r) > 0$ for which

$$\operatorname{sign}(\hat{\beta}_{(\lambda,\tau)}^{\mathsf{TJP}(r)}) = \theta \quad \text{and} \quad \operatorname{sign}(\hat{\omega}_{(\lambda,\tau)}^{\mathsf{TJP}(r)}) = \tilde{\theta}. \tag{1}$$

• Conversely, if for some $\epsilon \in \mathbb{R}^n$ and $r \in \mathbb{N}^*$ there is a threshold $\tau > 0$ such that (1) holds, then $(\theta, \tilde{\theta})$ is identifiable w.r.t. X and λ .



 \rightarrow How large the coefficients should scale to be correctly detected? Assume a correlated Gaussian design, i.e. $X_{i.} \in \mathbb{R}^{p} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \Sigma)$ and

- the smallest eigenvalue of the covariance matrix $\boldsymbol{\Sigma}$ is positive,
- the variance of the covariates is equal to one,
- the noise is assumed to be Gaussian.

Theorem 2 (Descloux, Boyer, Josse, S., Sardy, 2020)

Under the correlated Gaussian design above and signal-to-noise ratio high enough, TJP successfully recovers sign(β^0) with high probability, even with a positive fraction of corruptions.



 \rightarrow How large the coefficients should scale to be correctly detected?

Theorem 2 (if Σ is well-conditioned and $p/n \rightarrow \delta > 1$)

• Assume that the eigenvalues of Σ are bounded: $0 < \gamma_1 \leq \lambda_{\min}(\Sigma) \leq \lambda_{\max}(\Sigma) \leq \gamma_2$

• Assume
$$p/n \rightarrow \delta > 1$$
.

TJP achieves sign consistency provided that

$$n = \Omega(s \log p), \ k = \mathcal{O}(n) \ \text{and} \ \beta_{\min}^0 = \Omega(\sqrt{n}).$$

▲日 ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● ● ● ●

(s: sparsity of β^0 , k: sparsity of ω^0)

Introduction 00	Robust Lasso-Zero	Missing data ●00	Numerical experiments	Conclusion 0	References
Missin	g covariates in	high dim	ension		

- X partially known, we observe (y, X^{NA}) instead of (y, X)
- very few works for dealing with missing covariates in the high dimensional setting.

((Liu et al., 2016)	multiple imputation	increasingly complex
(Rosenbaum et al., 2013)	modified Dantzig selector	estimation Σ
(Loh and Wainwright, 2012 Datta and Zou, 2017)	modified LASSO	estimation Σ
(Jiang et al., 2019)	Adaptive Bayesian SLOPE	estimation Σ

- **X** Parametric assumption on the covariates for the estimation of Σ .
- Restrictive assumptions on the missing-data mechanism: missingness completely at random.
- X Not suitable in practice.

Introduction 00	Robust Lasso-Zero	Missing data ○●○	Numerical experiments	Conclusion 0	References
N 4 · ·	1				

Missing data as corruptions

- $\checkmark\,$ What we can do (well): impute "naively" the missing entries (by the mean for example).
- **✗** Bias in the estimates
- ex: Income is missing.



N 4 · · ·	1.				
Introduction	Robust Lasso-Zero	Missing data ○●○	Numerical experiments	Conclusion 0	References

Missing data as corruptions

- $\checkmark\,$ What we can do (well): impute "naively" the missing entries (by the mean for example).
- X Bias in the estimates
- 1. Impute "naively" the missing entries in $X^{\rm NA}$ to get \tilde{X} and then correct the imputation error.
- 2. See the imputation error as a corruption.

$$y = X\beta^0$$

N 4 · ·					
Introduction 00	Robust Lasso-Zero	Missing data ○●○	Numerical experiments	Conclusion 0	References

Missing data as corruptions

- $\checkmark\,$ What we can do (well): impute "naively" the missing entries (by the mean for example).
- ✗ Bias in the estimates
- 1. Impute "naively" the missing entries in $X^{\rm NA}$ to get \tilde{X} and then correct the imputation error.
- 2. See the imputation error as a corruption.

How to solve $y = X\beta^0 + \epsilon$ if we observe (X^{NA}, y) ?

• rewrite the model in the form of the sparse corruption model, where

$$\omega^0 := \frac{1}{\sqrt{n}} (X - \tilde{X}) \beta^0$$

▲日 ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● ● ● ●

is the corruption due to imputations.

Introduction 00	on	Robust 0000	Lasso-Zero		Missing 00●	g data	Numerica	l exp	eriments	5	Conclusion 0	References
			-	c								

Robust Lasso-Zero for dealing with missing data

Robust Lasso-Zero for missing data

- Impute "naively" X^{NA} and rescale the imputed matrix X.
- Run Robust Lasso-Zero algorithm with the design matrix X.
- $\checkmark\,$ without specify a model for the covariates or the missing data mechanism
- $\checkmark\,$ without estimation of the covariates covariance matrix or of the noise variance,

▲日 ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● ● ● ●

 $\checkmark~$ simple method for the user.

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments ●00000	Conclusion 0	References
C' L.'					

Simulation settings

- $X \sim \mathcal{N}(0, \Sigma)$, $\Sigma \in \mathbb{R}^{200 \times 200}$ is a Toeplitz matrix, s.t. $\Sigma_{ij} = \rho^{|i-j|}$.
- noise level $\sigma = 0.5$, coefficient β^0 drawn uniformly from $\{\pm 1\}$.
- Missing values: MCAR (random missingness, a = 0) or MNAR (informative missingness, $a \neq 0$).

$$\mathbb{P}(X_{ij}^{\mathrm{NA}} = \mathtt{NA} \mid X_{ij} = x) = \frac{1}{1 + e^{-a|x|-b}}, \ a \ge 0 \ \mathrm{and} \ b \in \mathbb{R}.$$

Methods

- **Rlass0:** Robust Lasso-Zero using M = 30 noisy dictionaries. The tuning parameters are obtained using $\lambda = 1$ and selecting τ by quantile universal threshold (QUT) at level $\alpha = 0.05$.
- lass0: Lasso-Zero [Descloux and Sardy, 2018]. The automatic tuning is performed by QUT, at level α = 0.05.
- **lasso**: Lasso [Tibshirani, 1996] performed on the mean-imputed matrix where the regularization parameter is tuned by cross-validation.
- NClasso: the nonconvex ℓ_1 estimator of [Loh and Wainwright, 2012].
- ABSLOPE: Adaptive Bayesian SLOPE of [Jiang et al., 2019].

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments ●00000	Conclusion 0	References
C' L.'					

Simulation settings

- $X \sim \mathcal{N}(0, \Sigma)$, $\Sigma \in \mathbb{R}^{200 \times 200}$ is a Toeplitz matrix, s.t. $\Sigma_{ij} = \rho^{|i-j|}$.
- noise level $\sigma = 0.5$, coefficient β^0 drawn uniformly from $\{\pm 1\}$.
- Missing values: MCAR (random missingness, a = 0) or MNAR (informative missingness, $a \neq 0$).

$$\mathbb{P}(X_{ij}^{\mathrm{NA}} = \mathtt{NA} \mid X_{ij} = x) = \frac{1}{1 + e^{-a|x|-b}}, \ a \ge 0 \ \mathrm{and} \ b \in \mathbb{R}.$$

Performance evaluation

- the Probability of Sign Recovery (PSR), $PSR = \mathbb{P}(sign(\hat{\beta}) = sign(\beta^0))$,
- the signed True Positive Rate (sTPR), the proportion of nonzero coefficients whose sign is correctly identified;
- the signed False Discovery Rate (sFDR), the proportion of incorrect signs among all discoveries.

▲ロト ▲周ト ▲ヨト ▲ヨト ヨー のへの

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments 00000	Conclusion 0	References
Results v	vith <i>s</i> -oracle	hyperpara	meter tuning		

Non-correlated case



- 5% NA and high sparsity: similar results.
- 20% NA and high sparsity: Robust Lasso-Zero outperforms other methods for MNAR setting.
- lower sparsity: Robust Lasso-Zero and Lasso-Zero generally give the best results.

Introduction 00	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion 0	References
Results V Correlated ca	with <i>s</i> -oracle	hyperpar	ameter tuning		



- similar results as in the non-correlated case.
- 5% NA and high sparsity: Robust Lasso-Zero for the MNAR setting behaves very well.

Introduction 00	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion 0	References 00000
Results Non-correlat	with automa ted case	tic hyperp	parameter tunin	g	



- 5% NA, high sparsity: except the LASSO, good performances.
- 20% NA, high sparsity: Robust Lasso-Zero has the best PSR.
- lower sparsity: except the LASSO, the methods are comparable in terms of PSR.

• □ ▶ • □ ▶ • □ ▶

• ABSLOPE behaves well in term of sTPR.

IntroductionRobust Lasso-ZeroMissing dataNumerical experimentsConclusionReferencesResults with automatic hyperparameter tuning
Correlated caseCorrelated caseCorrelated caseCorrelated case



< □ > < // >

• ABSLOPE generally behaves well.

IntroductionRobust Lasso-ZeroMissing dataNumerical experimentsConclusionReferencesResultswith automatic hyperparameter tuningsFDR



- Robust Lasso-Zero and Lasso-Zero have better performances than ABSLOPE.
- even with low sparsity and 20 % NA, FDR stability in the MCAR and MNAR settings.

Introduction 00	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion •	References
Conclusi	on				

- Robust Lasso-Zero: overfit by solving the Justice Pursuit and threshold by handling the overfitting with the use of noise dictionaries.
- Theoretical guarantees for Thresholded Justice Pursuit, a simplified version of the Robust Lasso-Zero.
- Applying Robust Lasso-Zero for dealing with missing data, simple method without parametric assumption.

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion	References
00	0000		000000	0	••••
Referenc	es I				



Dalalyan, A. S. and Thompson, P. (2019).

Outlier-robust estimation of a sparse linear model using $\ell_1\text{-penalized}$ huber's m-estimator.

arXiv preprint arXiv:1904.06288.



Datta, A. and Zou, H. (2017).

CoCoLasso for high-dimensional error-in-variables regression. *The Annals of Statistics*, 45(6):2400–2426.



Descloux, P. and Sardy, S. (2018).

Model selection with lasso-zero: adding straw to the haystack to better find needles.

arXiv:1805.05133 [stat]. arXiv: 1805.05133.



Garcia, R. I., Ibrahim, J. G., and Zhu, H. (2010). Variable selection for regression models with missing data. *Statistica Sinica*, 20(1):149.

Introduction 00	Robust Lasso-Zero 0000	Missing data	Numerical experiments	Conclusion 0	References •••••
Referenc	ces II				

Jiang, W., Bogdan, M., Josse, J., Miasojedow, B., Rockova, V., and TraumaBase Group (2019).

Adaptive Bayesian SLOPE – High-dimensional Model Selection with Missing Values.

arXiv e-prints, page arXiv:1909.06631.



Laska, J. N., Davenport, M. A., and Baraniuk, R. G. (2009).

Exact signal recovery from sparsely corrupted measurements through the Pursuit of Justice.

In 2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers, pages 1556–1560.



Li, X. (2013).

Compressed Sensing and Matrix Completion with Constant Proportion of Corruptions.

Constructive Approximation, 37(1):73–99.



Li, Z., Wu, F., and Wright, J. (2010).

On the systematic measurement matrix for compressed sensing in the presence of gross errors.

In 2010 Data Compression Conference, pages 356-365. IEEE.

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion O	References ●●●●●		
References III							
 Liu, Y., Wang, Y., Feng, Y., and Wall, M. M. (2016). Variable selection and prediction with incomplete high-dimensional data. <i>The annals of applied statistics</i>, 10(1):418. 							

High-Dimensional Regression with Noisy and Missing Data: Provable Guarantees with Nonconvexity.

The Annals of Statistics, 40(3):1637–1664.



Nguyen, N. H. and Tran, T. D. (2013a). Exact Recoverability From Dense Corrupted Observations via ℓ_1 -Minimization. *IEEE Transactions on Information Theory*, 59(4):2017–2035.



Nguyen, N. H. and Tran, T. D. (2013b). Robust Lasso With Missing and Grossly Corrupted Observations. *IEEE Transactions on Information Theory*, 59(4):2036–2058.

Rosenbaum, M., Tsybakov, A. B., et al. (2013). Improved matrix uncertainty selector. In From Probability to Statistics and Back: High-Dimensional Models and Processes-A Festschrift in Honor of Jon A. Wellner, pages 276–290. Institute of Mathematical Statistics.

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion	References
00	0000		000000	0	•••••
Referenc	es IV				



Saligrama, V. and Zhao, M. (2011).

Thresholded basis pursuit: LP algorithm for order-wise optimal support recovery for sparse and approximately sparse signals from noisy random measurements. *IEEE Transactions on Information Theory*, 57(3):1567–1586.



Tardivel, P. and Bogdan, M. (2019).

On the sign recovery by lasso, thresholded lasso and thresholded basis pursuit denoising.



Tibshirani, R. (1996).

Regression shrinkage and selection via the lasso.

Journal of the Royal Statistical Society: Series B (Methodological), 58(1):267–288.



Van Buuren, S. (2018). *Flexible imputation of missing data*. Chapman and Hall/CRC.



Wright, J. and Ma, Y. (2010). Dense Error Correction Via ℓ_1 -Minimization. IEEE Transactions on Information Theory, 56(7):3540–3560.

Introduction	Robust Lasso-Zero	Missing data	Numerical experiments	Conclusion	References
00	0000		000000	0	••••
Referenc	es V				

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., and Ma, Y. (2009). Robust Face Recognition via Sparse Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(2):210–227.

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで