Statistica Sinica: Supplement

An Adaptively Resized Parametric Bootstrap for Inference in High-dimensional Generalized Linear Models

Qian Zhao and Emmanuel J. Candès

University of Massachusetts, Amherst and Stanford University

Supplementary Material

This supplement contains theoretical and simulation results. In Section S1, we hypothesize the theoretical MLE distribution when the covariates are from a multivariate t-distribution and validate our conjecture through simulations. In Section S2, we provide additional detail on the SLOE algorithm we use to estimate the signal strength parameter. Section S3–Section S6.4 includes additional simulation results.

S1 Distribution of the MLE under a MVT covariate model

In this section we conjecture the distribution of a logistic MLE when the covariates follow a multivariate t-distribution (MVT). We begin by describing the model setting (Section S1.1), then describe the conjectured MLE distribution (Section S1.2) and compare with simulation results (Section S1.3). In short, we show that the MLE distribution depends on the degrees of freedom ν and, therefore HDT, does not apply. For simplicity, we consider models without an intercept in this section and we refer readers to (Zhao, Sur and Candès , 2020, Section 7) for the theoretical MLE distribution when the model includes an intercept term.

S1.1 Model assumptions

Suppose we have n i. i. d. observations (X_i, Y_i), $X_i \in \mathbb{R}^p$ and $Y_i \in \{\pm 1\}$ where X_i are from MVT and $Y_i \,|\, X_i$ is sampled from a logistic model. In detail,

$$X_i = \zeta_i Z_i, \quad Z_i \sim \mathcal{N}(0, \Sigma), \quad \zeta_i = \sqrt{\frac{\nu - 2}{\chi^2_\nu}}, \tag{S1.1}$$

Compared to the definition of a MVT, we choose $\nu - 2$ in the numerator because it ensures that each X_j has unit variance. Suppose the model coefficients are $\beta \in \mathbb{R}^p$, then $Y_i \in \{\pm 1\}$ and

$$\mathrm{P}(Y = 1 \,|\, X) = \frac{1}{1 + \exp(-X^\top \beta)}.$$

The maximum likelihood estimator is defined to be the minimizer of the negative log-likelihood

$$\hat{\beta} = \underset{b \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^n f(y_i, x_i^\top b),$$

where $f(y, \eta)$ denotes the loss function. For a logistic regression,

$$f(y, \eta) = \log(1 + \exp(-y\eta)).$$

S1.2 A guess at the MLE distribution

We observe that because (Zhao, Sur and Candès , 2020, Proposition 2.1) and (Zhao, Sur and Candès , 2020, Lemma 2.1) are results of the rotational invariance of a multivariate Gaussian distribution, they continue to hold with minor modification. We rehearse these two results for completeness.

Proposition 1. *For any matrix L obeying $\Sigma = LL^\top$, consider the vectors*

$$\hat{\theta} = L^\top \hat{\beta}, \quad \theta = L^\top \beta. \tag{S1.2}$$

Then $\hat{\theta}$ is the MLE in a logistic model with regression coefficient θ and covariates drawn i.i.d. from ζZ where $Z \sim \mathcal{N}(0, I)$.

Lemma 1. *Let $\hat{\theta}$ denote the MLE in a logistic model with regression vector θ and covariates drawn i.i.d. from $\zeta_i Z_i$ where $Z_i \sim \mathcal{N}(0, I_p)$ ($\zeta_i = \sqrt{(\nu-2)/\chi^2_\nu}$). Define the random variables*

$$\alpha(n) = \frac{\langle \hat{\theta}, \theta \rangle}{\|\theta\|^2}, \quad \sigma(n)^2 = \|P_{\theta^\perp} \hat{\theta}\|^2, \tag{S1.3}$$

where $P_{\theta^\perp}$ is the projection onto $\theta^\perp$, which is the orthogonal complement of θ. Then,

$$\frac{\hat{\theta} - \alpha(n)\theta}{\sigma(n)} \tag{S1.4}$$

is uniformly distributed on a unit sphere lying in $\theta^\perp$.

Next, we conjecture that as $n, p \to \infty$ in a constant ratio $p/n \to \kappa$ and if ν is sufficiently large, then the parameters $\alpha(n)$ and $\sigma(n)$ converges asymptotically to solutions to a system of nonlinear equations (See Eqn. (S1.7)). Combining this conjecture with Lemma 1, we can apply the same argument as in (Zhao, Sur and Candès , 2020, Theorem 3.1) to show that when (κ, γ) is in the region where the MLE exists (see Tang and Ye (2020) for a description of the region where the MLE exists), and if $\sqrt{n}\beta_j\tau_j = O(1)$, then $\hat{\beta}_j$ is asymptotically Gaussian

$$\frac{\sqrt{p}(\hat{\beta}_j - \alpha_\star\beta_j)}{\sigma_\star/\tau_j} \xrightarrow{d} \mathcal{N}(0,1) \tag{S1.5}$$

where $\tau_j^2 = \Theta_{jj}^{-1}$ $(\Theta = \Sigma^{-1})$. In the simulations, we indeed observe that the MLE is approximately Gaussian. The scaling factor $\sqrt{p}$ in the numerator (S1.5) differs from the factor $\sqrt{n}$ in (Zhao, Sur and Candès , 2020, Theorem 3.1) because here we define $\sigma_\star$ to be the limit of $\sigma(n)$, whereas in (Zhao, Sur and Candès , 2020, Lemma 3.1), we have $\sigma(n) \to \sqrt{\kappa}\sigma_\star$.

Conjecture S1.1. *Suppose ν is sufficiently large (for example $\nu \geq 15$), then*

$$\alpha(n) \xrightarrow{P} \alpha_\star, \quad \sigma(n) \xrightarrow{P} \sigma_\star, \tag{S1.6}$$

where the set of parameters $(\alpha_\star, \sigma_\star, \gamma_\star)$ solves a system of three equations:

$$\begin{cases} 0 & = \mathrm{E}\left[f'(Y, \mathrm{Prox}_{\zeta^2\lambda f}(Y, \zeta Z_1\alpha\gamma + \sigma\zeta Z_2)\zeta Z_1\right] \\ \frac{\sigma^2\kappa}{\lambda^2} & = \mathrm{E}\left[f'^2(Y, \mathrm{Prox}_{\zeta^2\lambda f}(Y, \zeta Z_1\alpha\gamma + \sigma\zeta Z_2))\zeta^2\right] \\ 1-\kappa & = \mathrm{E}\left[\frac{1}{1+\lambda\zeta^2 f''(Y, \mathrm{Prox}_{\zeta^2\lambda f}(\zeta Z_1\alpha\gamma + \zeta Z_2\sigma))}\right]. \end{cases} \tag{S1.7}$$

In Eqn.(S1.7)*,* f' *is the derivative of* f *w.r.t. the second argument,* Z_1 *and* Z_2 *are i.i.d. standard Gaussian,* ζ *is defined as* (S1.1)*,* $Y \in \{\pm 1\}$ *and*

$$\mathrm{P}(Y = 1 \,|\, \zeta, Z_1) = \frac{1}{1 + \exp(-\zeta Z_1 \gamma)}.$$

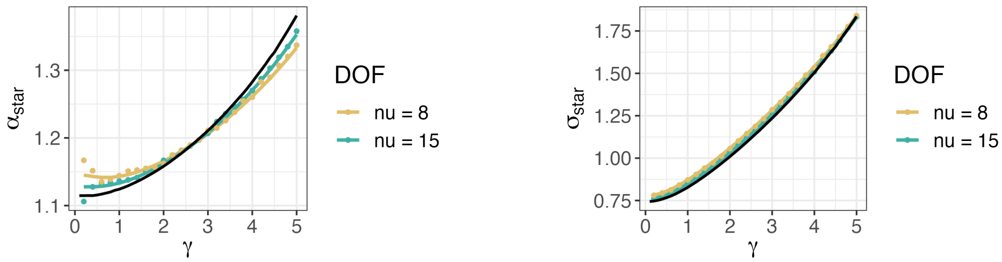
S1.3 Empirical results

We solve the system of equations defined in (S1.7) and compare with empirical values of the two parameters $\alpha_\star$ and $\sigma_\star$. We consider the setting when $\kappa = 0.1$ and $\nu = 8$ or $\nu = 15$ and plot the theoretical parameters $\alpha_\star$ and $\sigma_\star$ as the signal strength γ increases (Figure 1, solid line). We also display the empirical $\alpha_\star$ and $\sigma_\star$ in simulations (Figure 1, points). We simulate $n = 4,000$ observations from a logistic model where the covariates X_i are MVT with $\nu = 8$ or $\nu = 15$ degrees of freedom and $\Sigma = I$ and $Y_i \,|\, X_i$ and regression coefficients obey $\beta_1 = \gamma$ while $\beta_2 = \ldots = \beta_p = 0$. Based on Eqn. (S1.3), we define the observed parameters as

$$\hat{\alpha} = \hat{\beta}_1/\gamma, \quad \hat{\sigma} = \sqrt{\|\hat{\beta}_{-1}\|}, \tag{S1.8}$$

where $\hat{\beta}_{-1} = (\hat{\beta}_2, \ldots, \hat{\beta}_p)$. We observe that the empirical parameters align very well with the theoretical ones.

It is sufficient to consider $\Sigma = I$ and $\theta = (1, 0, \ldots, 0)$ because of Proposition 1 and Lemma 1.



(a) The inflation parameter $\alpha_\star$.

(b) The parameter $\sigma_\star$.

Figure 1: Empirical observations (points) and solutions to the system of equations in Eqn. S1.7 (lines) of the $\alpha_\star$ (Figure 1a) and $\sigma_\star$ (Figure 1b) when $\nu = 8$ (yellow) or $\nu = 15$ (cyan) and $\kappa = 0.1$. The empirical values are the average of observed $\hat{\alpha}$ and $\hat{\sigma}$ in 500 repeated simulations where $n = 4,000$. The black curve shows the theoretical parameters when the covariates are from a multivariate Gaussian distribution.

S2 The SLOE estimator

The Signal Strength Leave-One-Out Estimator (SLOE) provides an analytic expression for estimating $\eta^2 = \lim_{n\to\infty} \mathrm{Var}({X_{\text{new}}}^\top \hat{\beta})$ where $\hat{\beta}$ is the MLE and X_{new} is a new observation Yadlowsky et al. (2021). SLOE was developed to compute the asymptotic distribution of the logistic MLE, which depends on $\kappa = p/n$ and $\gamma^2 = \mathrm{Var}({X_{\text{new}}}^\top \beta)$ and can be reparametrized to depend on κ and η. (Yadlowsky et al. , 2021, Proposition 2) proves that the SLOE estimator consistently estimates η in the high-dimensional setting.

While SLOE was introduced for logistic regression, we generalize the formula to other GLMs; we however do not prove consistency in this broader setting. Define $w_i = x_i^\top \boldsymbol{H}^{-1} x_i$, and $t_i = x_i^\top \hat{\beta}$, $i = 1, \ldots, n$, where $\boldsymbol{H}$ is the Hessian of the negative log-likelihood evaluated at the MLE $\hat{\beta}$. Let

$$S_i = X_i^\top \hat{\beta} + q_i f'_{y_i}(t_i),$$

where

$$q_i = \frac{w_i}{1 - w_i f''_{y_i}(t_i)}.$$

Above, $f_y(t)$ is the negative log-likelihood function when the linear predictor is t and the response is y, In the case of logistic regression, $f_y(t) = \log(1 + e^{-yt})$ for $y \in \{\pm 1\}$.

Then, we define the extended SLOE estimator to be

$$\hat{\eta}^2_{\mathrm{SLOE}} = \frac{1}{n} \sum_i S_i^2 - \left(\frac{1}{n} \sum_i S_i \right)^2. \tag{S2.9}$$

Here, S_i approximates $x_i^\top \hat{\beta}_{(i)}$, where $\hat{\beta}_{(i)}$ is the MLE computed without using the ith observation. Since x_i is independent of $\hat{\beta}_{(i)}$, the variance of S_i approximates $\mathrm{Var}({X_{\mathrm{new}}}^\top \hat{\beta})$.

S3 Additional Simulation Results

Section S3.1 reports the coverage proportion of a null variable when covariates are from a multivariate t-distribution in Section 4. Section S3.2 considers logistic models, where we vary the problem dimension κ, signal strength γ, covariate distribution, and proportion of cases vs. controls in each class. Section S3.3 shows the coverage proportion when covariates are from a modified ARCH model and the responses from a Probit model. Section S3.4 considers a Poisson regression example.

S3.1 Coverage of a null variable

Table 1 reports the coverage proportion of a null variable when the covariates follow a multivariate t-distribution (see Section 4.1 for the simulation design). Coverage using either classical calculations or the standard bootstrap is better than for a non-null, compare with Table 2. This is because we observed that the MLE is unbiased when $\beta_j = 0$.

S3.2 Other logistic regression examples

In this section we consider four examples of logistic regression models. The number of observations in every simulation is $n = 4000$. We summarize the simulation setting in Table 2. In particular, setting (4) considers the situation when the two classes are imbalanced because the intercept is not zero and because the distribution of non-null β_j's is not symmetric around 0. We report

1. The inflation, std.dev. of both a null and a non-null variable.

2. The coverage proportion of both a null and a non-null variable and

Table 1: Coverage proportion of a single *null* variable (q_j in Eqn. (4.10)) with standard deviation between parentheses. This example uses multivariate-t covariates. The numbers closest to the empirical observation are highlighted in bold.

Nominal coverage	Theoretical CI		Standard Bootstrap		Resized Bootstrap			
					Known γ		Estimated γ	
	Classical	High-Dim	Parametric	Pairs	Boot-g	Boot-t	Boot-g	Boot-t
95	93.1	93.6	93.4	**95.1**	95.5	**95.0**	95.6	**95.1**
	(0.3)	(0.2)	(0.9)	(0.7)	(0.6)	(0.7)	(0.7)	(0.7)
90	87.4	88.1	88.7	91.2	**90.0**	90.2	89.7	**90.1**
	(0.3)	(0.3)	(1.1)	(0.9)	(0.9)	(0.9)	(1.0)	(1.0)
80	76.7	77.9	78.0	82.7	78.7	79.4	**80.4**	**80.4**
	(0.4)	(0.4)	(1.4)	(1.1)	(1.2)	(1.2)	(1.4)	(1.4)

proportion of variables covered in a single-shot experiment.

We summarize below the covariate distributions we consider in this section:

- "MVT" is the setting of Section 4.1.
- The "Modified ARCH" model corresponds to the situation where the covariates follow a modified ARCH model $X = \zeta\varepsilon$, where ζ is the inverse of a χ variable, which is distributed as the square root of a chi-squared variable, with $\nu = 8$ degrees of freedom and ε is from

an Autoregressive Conditional Heteroskedasticity (ARCH) model (see (Shumway and Stoffer , 2017, Section 5.4) for a definition of ARCH models). Here, starting with $X_0 \sim \mathcal{N}(0, \alpha_0/(1-\alpha_1))$, we sequentially sample variables so that $X_j = \sigma_j \varepsilon_j$, where $\sigma_j^2 = \alpha_0 + \alpha_1 X_{j-1}^2$ and $\varepsilon_j \sim \mathcal{N}(0,1)$. We work with $\alpha_0 = 0.6$ and $\alpha_1 = 0.4$. Although uncorrelated, the covariates are not independent of each other.

- The "Correlated Pareto" model denotes the setting when $X_i = U_i \Sigma^{1/2}$, where U_i are i.i.d. sampled from a heavy-tailed Pareto distribution, with density $f(x) = \alpha M^\alpha / x^{\alpha+1}$, $x \geq M$, where α is the shape parameter and M is the scale parameter. We set $\alpha = 5$ and $M = 1$. Σ is a circulant matrix as defined in Section 4.1 in the paper.

- The "Gaussian Interaction" is the setting from Section 5 in the paper.

Results

Tables 3 – 10 report the inflation, std.dev. of the logistic MLE of both a null and a non-null variable. We generate $1,000$ bootstrap samples to compute the resized bootstrap estimates. We also report the coverage proportion of a single variable and the proportion of variables in a single shot experiment. The resized bootstrap typically provides the most accurate estimate of the inflation and std.dev. and the relative error of the coverage proportion

Table 2: Simulation design in Section S3.2. The column κ reports the problem dimension $\kappa = p/n$; column γ reports the signal strength $\gamma = \text{Var}(X^\top \beta)^{1/2}$; column β_0 reports the intercept; column "Sparsity" refers to the proportion of non-null variables; column "Covariate" shows the distribution of the covariates.

Setting	κ	γ	β_0	Sparsity	Distribution of non-null β_j	Covariate
(1)	0.1	3.7	0	0.25	$0.5\mathcal{N}(8,2) + 0.5\mathcal{N}(-8,2)$	MVT
(2)	0.3	1.76	0	0.25	$0.5\mathcal{N}(3,2) + 0.5\mathcal{N}(-3,2)$	Modified ARCH
(3)	0.1	4.6	0	0.25	$\mathcal{N}(0,4)$	Gaussian Interaction
(4)	0.2	2.26	1.5	0.1	$0.8\mathcal{N}(6,2) + 0.2\mathcal{N}(-6,2)$	Correlated Pareto

Table 3: Estimated inflation and std. dev. of the logistic MLE in Setting (1). This example is repeated $B = 1,000$ times. We highlight the number closest to the empirical observation in bold.

	inflation				Standard Deviation				
	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical inflation	Classical Theory	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical Std. dev.
$\beta = 0$	-	-	-	-	1.70	1.83	1.86	**1.87**	1.88
$\beta = -8.16$	**1.26**	**1.24**	**1.24**	1.25	1.73	1.83	**1.90**	**1.91**	1.94

compared to target coverage is typically within 3%.

S3.3 A Probit regression example

We now apply the resized bootstrap to construct CIs for model coefficients in a Probit regression models when the covariates follow our modified ARCH model. In all cases, we set $n = 4000$, and $p = 400$. In probit regressions, we sample model coefficients by first picking 50 non-null variables, and sample the corressponding coefficients to be from an equal mixture of $\mathcal{N}(3, 1)$ and $\mathcal{N}(-3, 1)$. The high-dimensional theory also applies to Probit regressions, or a general loss function, with the slight modification of the set of nonlinear equations that the parameters $(\alpha_\star, \sigma_\star, \lambda_\star)$ must obey. The set of equations are the same as in Eqn. (S1.7) except that now we set $\zeta = 1$ and $\mathrm{P}(Y = 1 \,|\, Z_1) = \mathrm{P}(Y = 1 \,|\, X^\top \beta = Z_1)$. As in previous sections, we observe that the resized bootstrap method provides the most accurate estimate of the std. dev. (Table 11) and the corresponding confidence intervals provide most accurate coverage (Table 12).

S3.4 A Poisson regression example

We now consider an example with Poisson regression with log link function, i. e., $Y \,|\, X \sim \text{Poisson}(\mu(X))$ and $\log(\mu(X)) = X^\top \beta$ (Weisberg , 2014, Chapter 12). We use the same simulation design as in Section S3.2. We report the inflation and std. dev. of both a null and a non-null variable in Table 13. We only use the classical theory and the resized bootstrap to estimate the std. dev., since HDT is unavailable for Poisson regression. Table 13 shows that the MLE is nearly unbiased. The estimated std. dev.

using the classical theory and the resized bootstrap are close to each other and slightly underestimates the std. dev. of the MLE. Therefore, we expect that both approaches would produce CI with similar coverage. Indeed, the coverage proportion using both the classical theory and the resized bootstrap method both tend to undercover the true coefficient, and the average coverage proportions are typically within three standard deviations away from the nominal coverage (see Table 14). In sum, both the classical theory and the resized bootstrap yield reasonably accurate CI in case of a Poisson regression.

S4 An example where the sample size is small

We study an example with a small sample size, and set $n = 400$ and $p = 40$. We sample covariates independently from a Pareto distribution (see Section S3.2, here we set $\Sigma = I$) and sample responses from a logistic model where half of the variables are non-nulls and sampled from an equal mixture of $\mathcal{N}(4, 1)$ and $\mathcal{N}(-4, 1)$.

When the covariates are i. i. d., the MLE may be asymptotically Gaussian, however, the normal approximation is inaccurate when n is small. To see this, we can express $\hat{\beta}_j$ as

$$\hat{\beta}_j = \frac{\lambda}{\kappa\sqrt{n}} \sum_{i=1}^{n} x_{ij}(y_i - \rho'(x_{i,-j}^\top \hat{\beta}_{[-j]})) + o_P(1), \tag{S4.10}$$

where $\hat{\beta}_{[-j]}$ refers to MLE computed when leaving out the jth variable and $\rho(t) = \log(1 + e^t)$ (Sur and Candès , 2019, Appendix C). Although Eqn. (S4.10) assumes that the X_{ij} are standard Gaussian, we expect that it holds for other i. i. d. covariates. Since $\hat{\beta}_j$ is approximately a weighted average of the observed $x_{i,j}$, $\hat{\beta}_j$ approaches a Gaussian random variable as $n \to \infty$ by the central limit theorem. Since the rate of convergence depends on the third moment of X_{ij} as a result of the Berry-Esseen theorem, we expect that the distribution of the MLE deviates from a Gaussian distribution when n is small. Indeed, the normal quantile plot of $\hat{\beta}_j$ (Figure 3, Left) confirms that the MLE is skewed and thus not Gaussian. In comparison, the normal quantile plot when $n = 4000$ and $p = 400$ (Figure 2) indicates that the MLE is well-approximated by a Gaussian distribution when n is large.

While the MLE is not Gaussian, a Q-Q plot of the standardized $\hat{\beta}_j^b$ (standardized by estimated inflation and std. dev.) versus the standardized $\hat{\beta}_j$ (standardized by the true inflation and std. dev.) shows that the bootstrap MLE approximates the sampling distribution very well (Figure 3, Right). We thus expect that the resized bootstrap provides correct coverage. This is confirmed in Table 15, which shows that the bootstrap CIs are

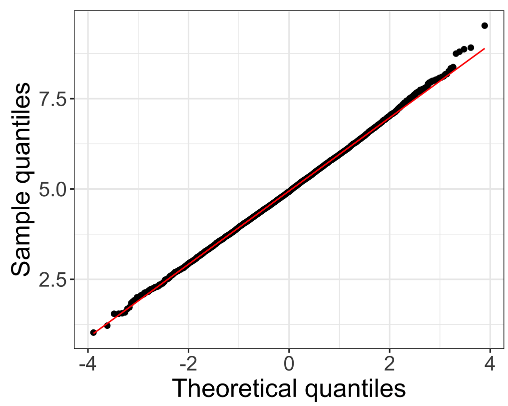


Figure 2: Normal quantile plot of a non-null MLE coordinate. The covariates are i. i. d. from a Pareto distribution. Here, $n = 4000$ and $p = 400$.

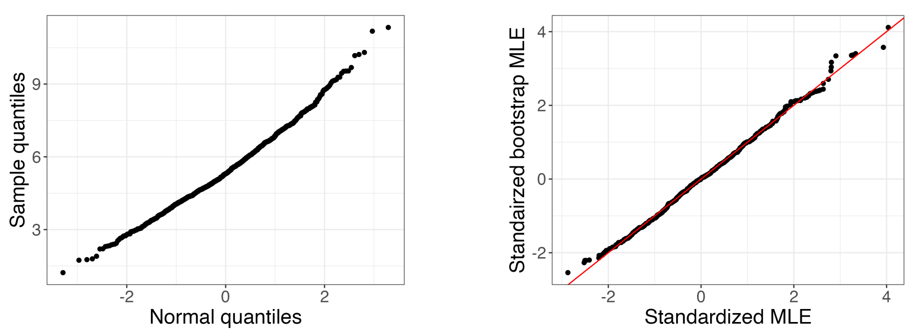


Figure 3: (Left) Normal quantile plot of a single MLE coordinate where $\beta = 4.5$ and $n = 400$, $p = 40$ and the covariates are i. i. d. from a Pareto distribution. (Right) Q-Q plot of the MLE standardized by the true inflation and std. dev. versus the bootstrap MLE standardized by the estimated inflation and std. dev. in a single experiment. The red line is the 45 degree line.

reasonably accurate for both a single variable and a single-shot experiment across all of the confidence levels examined.

Our results in this example suggest that the bootstrap CIs produce reasonable coverage when the sample size is small and the normal approximation is far from valid. This in contrast to methods based on high-dimensional theory since we can see that the corresponding CI's undercover.

S5 A simulated example when the coefficients are sparse

To study the accuracy of our method for large coefficients, we use a simulated example where there are only 10 non-null variables whose coefficients have equal magnitudes, which equals to 10, and ± 1 signs with equal probability. Here, $\tau_j |\beta_j|/\gamma \approx 0.32$ (in this case, $\tau_j^2 = \mathrm{Var}(X_j | X_{-j}) = 1/p = 0.025$).

We first examine the inflation and variance of the MLE (Table 16). We report the average inflation of all of the non-null variables in $N = 10,000$ repeated experiments (Column Empirical) and the estimated inflation using high-dimensional theory (Column High-Dim Theory) and the bootstrap (Column Resized Bootstrap). We observe that the high-dimensional theory slightly under-estimates the inflation, with a relative error of about 1%, whereas the bootstrap estimates are more accurate. We then study the std. dev. of the MLE and report the average std. dev. of all of the null variables (Table 16, Row std. dev. (null)) and the non-null variables (Table 16, Row std. dev. (non-null)). As in Table 1, the variance of the non-null variables are higher than that of the null variables. On the other hand, unlike Table 1, the high-dimensional theory underestimates the variance of both the null and nonnull variables (recall that the covariates are not Gaussian). Lastly, the resized bootstrap method using either a known or estimated γ, also slightly underestimates the variance of the non-null variables. It is however more accurate than HDT with a relative error below 1%. This shows that the bootstrap is reasonably accurate for large coefficients.

We next study the coverage probability of confidence intervals by computing the average coverage proportion of the null and non-null variables. Unsurprisingly, the HDT (Table 17, Column High-Dim Theory) undercovers both for the null and non-null variables, and the coverage proportions are less accurate for non-null variables. The resized bootstrap using the correct γ (Table 17, Column Known γ) slighty undercovers but the coverage is closer to the nominal coverage. Interestingly, the resized bootstrap with an estimated $\hat{\gamma}$ nearly achieves nominal coverage for both null and non-null variables at every considered significance level. In contrast, classical theory CI significantly undercovers the true coefficient because the classical theory does not account for the inflation of the MLE.

Our findings here agree with our hypothesis that the resized bootstrap is less accurate when the coefficient is large. At the same time, our results are reassuring because they suggest that the resized bootstrap is reasonably accurate for relatively large coefficients.

S6 General loss functions and M-estimators

Our discussion thus far concerned the maximum likelihood estimate in a generalized linear model. In this section, we step beyond the MLE and consider M-estimators associated with a *general loss function*, which may not be the negative log-likelihood. As a running example, we consider the situation where we fit a logistic regression when the true likelihood is given by the probit model. In detail, we are given n i.i.d. observations (X_i, Y_i), $i = 1, \dots, n$, where $Y \mid X \sim \text{Binomial}(\mu)$ and $\mu = \Phi(X^\top \beta)$, where Φ is the cumulative distribution of a standard Gaussian. We fit a logistic regression and obtain $\hat{\beta}$. Suppose we repeat the sampling process multiple times, what would the distribution of $\hat{\beta}$ be?

If X follows a multivariate Gaussian distribution, then HDT provides the theoretical distribution of the M-estimator, see Section S1 and Section S3.3 for a discussion. However, when the covariates are not Gaussian, then HDT does not apply and, yet, we can still use the resized bootstrap method. Because $Y \mid X$ follow a probit regression model, we should likewise sample new responses Y from a probit model in Algorithm 2 (Line 6) and Algorithm 1 (Line 3). We note that sticking to the sampling distribution of $Y \mid X$ is important because the estimated distribution will be far from the true one otherwise.

We illustrate HDT and the resized bootstrap through a simulated example where the covariates are either multivariate Gaussian or interaction between these Gaussian variables. An example where the covariates are from a multivariate t-distribution is provided in Section S6.4.

S6.1 Simulation design

We sample covariates as follows: we sample the first half of the covariates from $\mathcal{N}(0, \Sigma_{p/2})$ where Σ is the circulant matrix from Section 4.2; then, we sample the second half of the variables as products between two variables from the first half, i.e., $X_j = X_{j_1} \times X_{j_2}$ $(1 \leq j_1, j_2 \leq p/2)$ if $j = p/2 + 1, \ldots, p$. We standardize X_j to have variance equal to $1/p$. In terms of the model coefficients, we randomly sample 25% of the variables, including interaction terms, to be non-nulls, and sample the non-null coefficients $\beta_j \sim \mathcal{N}(0, 5)$, which leads to a signal strength $\gamma = 2.71$. The number of observations is $n = 4000$ and the number of covariates is $p = 400$ ($\kappa = p/n = 0.1$) in this example.

S6.2 Estimating the signal strength

While the theoretical relationship between $\eta = \mathrm{Var}(X_{\mathrm{new}}^\top \hat{\beta})^{1/2}$ and γ is not available, we empirically observe that η increases monotonically as γ increases (Figure 4). We plot the estimated $\hat{\eta}(\gamma_s)$ for a range of γ_s, which is obtained by re-scaling the MLE (Figure 4, black points), which are centered around the empirical curve shown in orange, indicating that the estimated relationship between η and γ is reasonably accurate. As a result, intersecting the estimated curve, which we do not show because it overlaps with the empirical curve, with the observed $\tilde{\eta} = 6.33$, yields $\hat{\gamma} = 2.56$, which is also close to the true $\gamma = 2.71$.

S6.3 Results

We now report the estimated inflation, std.dev. of the M-estimator and the coverage proportion of the confidence intervals.

Estimated inflation and standard deviation

We plot the averaged estimates versus model coefficients in Figure 5 (Left), where we observe that the inflation is the same for every coordinate. Figure

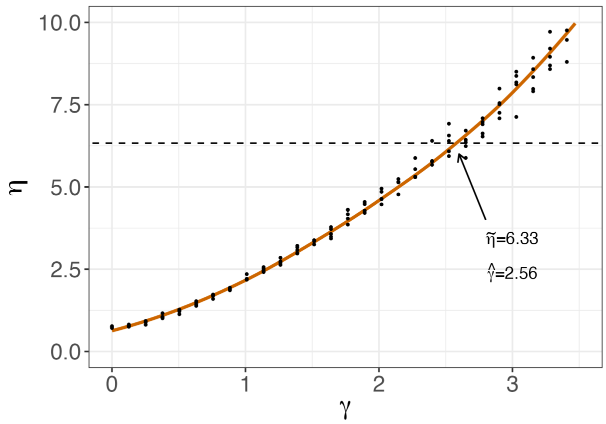


Figure 4: An illustration of using $\eta = \mathrm{sd}({X_{\mathrm{new}}}^\top \hat{\beta})$ to estimate the signal strength $\gamma = \mathrm{sd}({X_{\mathrm{new}}}^\top \beta)$ in case of a general loss function. The orange curve shows the empirical η versus γ (see Figure 3 in the paper for details). The black points show the estimated $\hat{\eta}(\gamma_s)$ for a range of γ_s by using a single dataset. The dashed line shows the estimated $\tilde{\eta}$ using the SLOE estimator evaluated at the MLE. The estimated $\hat{\gamma} = 2.56$ is close to $\gamma = 2.71$.

5 (Right) shows the average bootstrap estimates versus the resized bootstrap coefficient. Because the points align, the M-estimates in the bootstrap samples are also inflated with the same factor, which also happens to be close to the empirical inflation. We report the estimated inflation and std.dev. of both a null and non-null variable using HDT and the resized bootstrap in Table 18. In this example, the HDT estimate of the std.dev. is the most accurate, and the estimated inflation using the resized bootstrap is the most accurate. While the resized bootstrap underestimates the

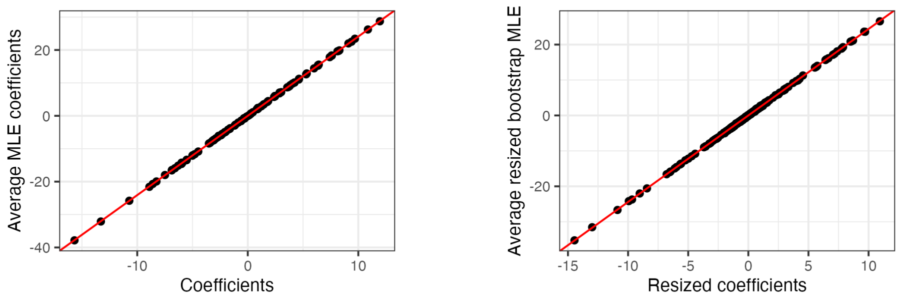


Figure 5: (Left) The average of the M-estimators versus the true model coefficient for $\beta_j \neq 0$. The red line has intercept 0 and slop equal to 2.41. (Right) The average of the M-estimators in the bootstrap samples using one data set versus the resized coefficients $\beta_\star$. The red line has intercept 0 and slop equal to 2.44. In this example, the M-estimators are the logistic regression estimates while the true model is probit.

std.dev., the relative error is within 15%.

Coverage proportion

We report both the proportion of times a single variable is covered (q_j) and the proportion of variables covered in a single shot experiment $\bar{q}$ (Table 19). In this example, we sample $1,000$ bootstrap samples to compute both confidence intervals using the Gaussian approximation (Boot-g) and using the bootstrap MLE (Boot-t). As we have observed before, the classical theory significantly undercovers the true coefficients, while both the HDT and the resized bootstrap provide reasonably correct coverage probability of the model coefficients. For example, while the resized bootstrap undercovers model coefficients on average, the relative error is within 3%.

S6.4 An example of MVT covariates

In this section we report simulation results when the covariates are sampled from a multivariate t-distribution as in Section 4.1; again, we fit $\hat{\beta}$ via a logistic regression while $Y \,|\, X$ is sampled from a probit model. In this example, we set $n = 4,000$, $p = 800$ and we randomly sample 25% of the variables to be non-nulls and their magnitudes are i.i.d. from an equal mixture of $\mathcal{N}(3,1)$ and $\mathcal{N}(-3,1)$. We report the inflation and std.dev. of the M-estimators (Table 20) and the coverage proportion (Table 21). The high-dimensional theory is more accurate compared to the resized bootstrap method, and in both cases the relative error is within 3%.

References

Shumway, R. H., and Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples* 4th Edition. Springer.

Sur, P., andès & J., E. (2019). A Modern Maximum-Likelihood Theory for High-Dimensional Logistic Regression. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 14516–14525.

Tang, W. & Ye, Y. (2020). The Existence of Maximum Likelihood Estimate in High-Dimensional Binary Response Generalized Linear Models. *Electron. J. Stat.* **14**, 4028 – 4053.

Weisberg, S. (2014). *Applied Linear Regression* 4th Edition. John Wiley & Sons.

Yadlowsky, S., Yun, T., McLean, C. & D'Amour, A. (2021). SLOE: A Faster Method for Statistical Inference in High-Dimensional Logistic Regression. In: *NeurIPS 2021*.

Zhao, Q., Sur, P. & Candès, E. (2020). The Asymptotic Distribution of the MLE in High-Dimensional Logistic Models: Arbitrary Covariance. *Bernoulli* **28**, 1835 – 1861.

Table 4: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between the parentheses. Here we compute the logistic MLE from Setting (1). We use both Gaussian approximation (Column Boot-g) and distribution of the bootstrap MLE (Column Boot-t) to construct the CI. We highlight the number closest to the target coverage in bold.

		Theoretical CI		Resized Bootstrap			
	Nominal			Known γ		Estimated γ	
	coverage	Classical	High-Dim	Boot-g	Boot-t	Boot-g	Boot-t
Single Null	95	92.1	94.5	**94.9**	94.5	**95.2**	**94.9**
		(0.9)	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)
	90	85.4	88.1	**88.9**	88.5	**89.2**	**89.2**
		(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)
	80	74.4	77.2	78.0	77.8	**78.7**	78.5
		(1.3)	(1.3)	(1.3)	(1.3)	(1.3)	(1.3)
Single Non-null	95	76.2	93.2	**94.2**	**94.3**	**94.3**	93.8
		(1.4)	(0.8)	(0.7)	(0.7)	(0.7)	(0.7)
	90	65.4	87.6	88.7	89.0	**89.2**	**89.2**
		(1.5)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)
	80	51.9	77.5	**79.8**	78.9	**79.9**	79.3
		(1.6)	(1.3)	(1.3)	(1.3)	(1.3)	(1.3)
Single Experiment	95	88.7	94.2	**94.8**	94.6	**94.9**	**94.8**
		(0.08)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)
	90	81.7	88.9	89.6	89.4	**89.9**	89.6
		(0.1)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)
	80	70.2	78.5	79.5	79.0	**79.7**	79.3
		(0.2)	(0.09)	(0.1)	(0.1)	(0.1)	(0.1)

Table 5: Estimated inflation and std. dev. of the logistic MLE in Setting (2). This example is repeated $B = 832$ times.

	inflation				Standard Deviation				
	High-dim Theory	Resized Bootstrap Known-γ	Estimated-γ	Empirical inflation	Classical Theory	High-dim Theory	Resized Bootstrap Known-γ	Estimated-γ	Empirical Std. dev.
$\beta = 0$	-	-	-	-	2.80	3.78	3.44	**3.82**	3.94
$\beta = -3.56$	**2.11**	1.92	2.02	2.11	2.81	**3.78**	3.46	3.84	3.77

Table 6: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between parentheses. Here we compute the logistic MLE from Setting (2).

		Theoretical CI		Resized Bootstrap			
	Nominal			Known γ		Estimated γ	
	coverage	Classical	High-Dim	Boot-g	Boot-t	Boot-g	Boot-t
Single Null	95	84.4	93.4	91.2	91.2	**93.9**	93.6
		(1.2)	(0.9)	(1.0)	(1.1)	(0.8)	(1.0)
	90	76.7	88.8	86.5	85.0	89.2	**89.5**
		(1.5)	(1.1)	(1.2)	(1.4)	(1.1)	(1.2)
	80	64.3	78.9	74.9	74.3	**79.5**	79.2
		(1.7)	(1.4)	(1.5)	(1.7)	(1.4)	(1.6)
Single Non-null	95	67.2	**95.0**	92.6	92.1	95.4	96.0
		(1.6)	(0.8)	(0.9)	(1.0)	(0.7)	(0.8)
	90	56.9	**90.0**	85.9	85.2	90.5	90.3
		(1.7)	(1.0)	(1.2)	(1.4)	(1.0)	(1.1)
	80	44.5	**80.2**	75.0	74.2	80.3	**80.5**
		(1.7)	(1.4)	(1.5)	(1.7)	(1.3)	(1.5)
Single Experiment	95	80.0	94.0	91.5	91.0	**94.6**	94.4
		(0.08)	(0.07)	(0.09)	(0.1)	(0.03)	(0.04)
	90	71.9	88.8	85.2	84.5	**89.4**	89.0
		(0.09)	(0.10)	(0.10)	(0.20)	(0.04)	(0.06)
	80	60.0	78.5	74.1	73.2	**79.1**	78.4
		(0.09)	(0.13)	(0.14)	(0.2)	(0.06)	(0.08)

Table 7: Estimated inflation and std. dev. of the logistic MLE in Setting (3). This example is repeated $B = 800$ times.

	inflation				Standard Deviation				
	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical inflation	Classical Theory	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical Std. dev.
$\beta = 0$	-	-	-	-	1.00	1.13	**1.08**	**1.08**	1.07
$\beta = 6.76$	**1.17**	**1.17**	**1.17**	1.18	1.10	1.05	**1.13**	**1.13**	1.14

Table 8: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between the parentheses. Here we compute the logistic MLE from *Setting (3)*.

	Nominal coverage	Theoretical CI: Classical	Theoretical CI: High-Dim	Resized Bootstrap, Known γ: Boot-g	Resized Bootstrap, Known γ: Boot-t	Resized Bootstrap, Estimated γ: Boot-g	Resized Bootstrap, Estimated γ: Boot-t
Single Null	95	93.5	96.3	**95.3**	**95.3**	95.4	95.5
		(0.9)	(0.7)	(0.8)	(0.8)	(0.7)	(0.7)
	90	87.1	91.9	**89.9**	**90.0**	90.5	90.6
		(1.2)	(1.0)	(1.1)	(1.1)	(1.0)	(1.0)
	80	76.8	82.1	79.5	79.1	**79.8**	79.1
		(1.5)	(1.4)	(1.4)	(1.4)	(1.4)	(1.4)
Single Non-null	95	77.0	94.5	94.6	**94.8**	94.6	94.6
		(1.5)	(0.8)	(0.8)	(0.8)	(0.8)	(0.8)
	90	67.1	88.8	**90.0**	89.4	91.0	90.5
		(1.7)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)
	80	52.6	77.6	78.6	77.8	**79.0**	78.9
		(1.8)	(1.5)	(1.5)	(1.5)	(1.4)	(1.4)
Single Experiment	95	91.6	95.3	94.8	94.7	**94.9**	94.8
		(0.06)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)
	90	85.4	90.6	89.8	89.6	**89.9**	89.8
		(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)
	80	74.4	80.7	79.7	79.5	**79.9**	79.7
		(0.10)	(0.09)	(0.09)	(0.09)	(0.08)	(0.08)

Table 9: Estimated inflation and std. dev. of the logistic MLE in Setting (4). This example is repeated $B = 1,300$ times.

	inflation				Standard Deviation				
	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical inflation	Classical Theory	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical Std. dev.
$\beta = 0$	-	-	-	-	3.36	2.49	2.86	**3.03**	3.06
$\beta = 9.35$	1.66	1.52	**1.58**	1.60	3.36	2.88	3.19	**3.41**	3.62

Table 10: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between the parentheses. Here we compute the logistic MLE from *Setting (4)*.

	Nominal coverage	Theoretical CI		Resized Bootstrap			
				Known γ		Estimated γ	
		Classical	High-Dim	Boot-g	Boot-t	Boot-g	Boot-t
Single Null	95	88.5	96.5	92.6	91.2	**94.4**	93.6
		(0.9)	(0.5)	(0.7)	(1.1)	(0.6)	(1.0)
	90	81.8	92.5	86.5	85.0	88.9	**89.5**
		(1.1)	(0.7)	(0.9)	(1.4)	(0.9)	(1.2)
	80	70.8	83.2	77	74.4	**79.9**	79.2
		(1.3)	(1.0)	(1.2)	(1.7)	(1.1)	(1.6)
Single Non-null	95	49.4	92.4	91.6	92.1	93.8	**96.0**
		(1.4)	(0.7)	(0.8)	(1.0)	(0.7)	(0.8)
	90	38.3	87.5	85.2	85.3	88.8	**90.3**
		(1.4)	(0.9)	(1.0)	(1.4)	(0.9)	(1.2)
	80	27.3	77.2	73.3	74.2	77.7	**80.5**
		(1.2)	(1.2)	(1.1)	(1.7)	(1.2)	(1.5)
Single Experiment	95	85.8	95.6	92.2	91.0	94.1	**94.4**
		(0.06)	(0.04)	(0.07)	(0.14)	(0.03)	(0.03)
	90	78.3	91.1	86.2	84.5	88.7	**89.0**
		(0.06)	(0.06)	(0.09)	(0.2)	(0.04)	(0.06)
	80	66.6	81.7	75.3	73.2	**78.4**	**78.4**
		(0.07)	(0.09)	(0.11)	(0.2)	(0.08)	(0.08)

Table 11: Estimated inflation and std. dev. of the MLE from a Probit regression when the covariates are from a modified ARCH model. This example is repeated $B = 798$ times.

	inflation				Standard Deviation				
	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical inflation	Classical Theory	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical Std. dev.
$\beta = 0$	-	-	-	-	0.548	0.564	0.590	**0.592**	0.594
$\beta = -2.76$	**1.135**	1.146	1.147	1.139	0.556	0.564	**0.597**	**0.598**	0.614

Table 12: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between the parentheses. This example uses modified ARCH covariates and a Probit model.

	Nominal coverage	Theoretical CI		Resized bootstrap			
				Known γ		Estimated γ	
		Classical	High-Dim	Boot-g	Boot-t	Boot-g	Boot-t
Single Null	95	93.5	94.0	**95.1**	94.9	94.7	94.5
		(0.9)	(0.9)	(0.8)	(0.8)	(0.8)	(0.8)
	90	86.8	87.8	**90.1**	**89.0**	90.4	89.7
		(1.2)	(1.2)	(1.1)	(1.1)	(1.1)	(1.1)
	80	76.2	77.2	79.7	79.2	**80.0**	78.8
		(1.5)	(1.5)	(1.4)	(1.4)	(1.4)	(1.5)
Single Non-null	95	86.7	93.1	94.0	93.6	94.0	**94.5**
		(1.2)	(0.9)	(0.8)	(0.9)	(0.8)	(0.8)
	90	79.5	86.8	88.9	89.6	89.2	**89.4**
		(1.43)	(1.20)	(1.10)	(0.06)	(1.1)	(1.1)
	80	65.8	77.6	**79.2**	78.7	**79.2**	78.8
		(1.7)	(1.5)	(1.4)	(1.5)	(1.4)	(1.5)
Single Experiment	95	92.2	93.7	94.9	94.7	**95.0**	94.8
		(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)
	90	86.0	88.0	89.8	89.7	**89.9**	89.7
		(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)
	80	75.1	77.6	79.9	79.6	**80.0**	79.8
		(0.09)	(0.08)	(0.09)	(0.09)	(0.76)	(0.08)

Table 13: Estimated inflation and std. dev. of the MLE from a Poisson regression when the covariates are from a modified ARCH model. This example is repeated $B = 800$ times.

	inflation			Standard Deviation			
	Resized Bootstrap		Empirical	Classical	Resized Bootstrap		Empirical
	Known-γ	Estimated-γ	inflation	Theory	Known-γ	Estimated-γ	Std. dev.
$\beta = 0$	-	-	-	0.269	0.268	0.268	0.278
$\beta = 5.15$	0.990	0.990	0.994	0.266	0.265	0.265	0.273

Table 14: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between the parentheses. This example uses modified ARCH covariates and a Poisson regression.

	Nominal coverage	Classical Theory	Resized Bootstrap: Known γ Boot-g	Known γ Boot-t	Estimated γ Boot-g	Estimated γ Boot-t
Single Null	95	94.0	93.9	93.9	94.1	**94.5**
		(0.8)	(0.9)	(0.9)	(0.8)	(0.8)
	90	**89.1**	89.0	88.3	**89.1**	88.4
		(1.1)	(1.1)	(1.1)	(1.1)	(1.1)
	80	77.6	77.4	77.4	**77.8**	77.3
		(1.2)	(1.5)	(1.5)	(1.5)	(1.5)
Single Non-null	95	92.9	**93.4**	93.1	**93.4**	92.9
		(0.9)	(0.90)	(0.90)	(0.90)	(0.90)
	90	**88.4**	**88.4**	87.8	88.0	88.0
		(1.1)	(1.1)	(1.2)	(1.2)	(1.2)
	80	**80.6**	79.0	79.3	78.8	79.0
		(1.4)	(1.4)	(1.4)	(1.5)	(1.4)
Single Experiment	95	95.2	**95.0**	94.8	**95.0**	94.8
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
	90	90.3	**90.1**	89.8	**90.0**	89.8
		(0.06)	(0.06)	(0.06)	(0.05)	(0.06)
	80	80.2	**80.0**	79.7	**80.0**	79.7
		(0.08)	(0.07)	(0.08)	(0.08)	(0.07)

Table 15: Coverage proportion of a single non-null variable (Column I) and of all of the variables (Column II) in 1,000 repeated experiments with $n = 400$ and $p = 40$. We use the resized bootstrap method with known or estimated parameters. The standard deviations are reported between parentheses.

	I. Single variable			II. Single Experiment		
Nominal	High-dim	Resized Bootstrap		High-dim	Resized Bootstrap	
Coverage	Theory	boot-g	boot-t	Theory	boot-g	boot-t
95	87.7 (1.0)	96.5 (0.6)	**95.8** (0.6)	92.0 (0.2)	**95.7** (0.1)	**95.7** (0.1)
90	81.2 (1.2)	**92.2** (0.9)	**92.1** (0.9)	86.5 (0.2)	91.0 (0.2)	**90.8** (0.2)
80	67.2 (1.5)	81.0 (1.2)	**79.9** (1.3)	76.4 (0.3)	81.2 (0.2)	**80.5** (0.2)

Table 16: The average inflation and variance of the MLE when the coefficients are large ($\tau_j \beta_j / \gamma \approx 0.32$). The second and third row report average std. dev. for a single null or non-null variable. In this simulation, the covariates are from a modified ARCH and the responses are from a logistic regression. The resized bootstrap estimates are averaged in $N = 1,000$ simulations.

		Classical	High-Dim	Resized Bootstrap	
	Empirical	Theory	Theory	Known γ	Estimated γ
inflation	1.15	-	1.14	1.15	1.15
Std. dev. (null)	0.98	0.92	0.93	0.98	0.98
Std. dev. (non-null)	1.05	0.97	0.93	1.02	1.03

Table 17: Average coverage proportion of a single null or nonnull variable. The std. dev. is reported inside the parentheses. We use bootstrap-t confidence intervals in this example.

	Nominal	Classical	High-Dim	Resized Bootstrap	
Variable	Coverage	Theory	Theory	Known γ	Unknown γ
Null	95	93.4	93.5	95.2	95.6
		(0.2)	(0.3)	(0.6)	(0.6)
	90	87.9	88.0	88.9	89.6
		(0.3)	(0.3)	(0.9)	(0.9)
	80	77.1	77.3	78.8	78.8
		(0.4)	(0.4)	(1.2)	(1.2)
Non-null	95	64.0	91.4	94.4	94.7
		(0.5)	(0.3)	(0.7)	(0.6)
	90	52.0	84.9	88.8	89.0
		(0.5)	(0.4)	(0.9)	(0.9)
	80	38.0	73.8	78.6	78.8
		(0.5)	(0.4)	(1.2)	(1.2)

Table 18: Estimated inflation and std.dev. of the M-estimator of logistic regression when the true model is probit. The example is repeated $B = 800$ times. The resized bootstrap estimates are computed by using an estimated signal strength γ. The values closest to the empirical observations are highlighted in bold.

	Inflation			Standard Deviation			
	High-dim Theory	Resized Bootstrap	Empirical Bias	Classical Theory	High-dim Theory	Resized Bootstrap	Empirical Std.dev.
$\beta = 0$	-	-	-	1.53	**1.85**	1.75	1.85
$\beta = 3.66$	2.46	**2.40**	2.39	1.96	**2.39**	2.25	2.60

Table 19: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between the parentheses. Here, we fit a logistic regression when the true model is probit. The values closest to the target coverage are highlighted in bold.

	Nominal coverage	Theoretical CI		Resized Bootstrap			
				Known γ		Estimated γ	
		Classical	High-Dim	Boot-g	Boot-t	Boot-g	Boot-t
Single Null	95	89.8	**95.9**	93.5	93.6	93.6	93.6
		(1.1)	(0.7)	(0.9)	(0.9)	(0.9)	(0.9)
	90	84.1	**89.5**	88.0	87.8	88.3	88.5
		(1.3)	(1.1)	(1.2)	(1.2)	(1.1)	(1.1)
	80	72.5	81.9	**79.9**	79.1	79.6	79.9
		(1.6)	(1.4)	(1.4)	(1.4)	(1.4)	(1.4)
Single Non-null	95	28.8	**93.0**	91.9	92.1	92.4	92.6
		(1.6)	(0.9)	(1.0)	(1.0)	(0.9)	(0.9)
	90	18.1	**89.4**	88.0	88.3	88.1	88.13
		(1.4)	(1.1)	(1.2)	(1.1)	(1.1)	(1.1)
	80	11.5	**81.1**	78.0	78.3	78.5	78.4
		(1.1)	(1.4)	(1.5)	(1.5)	(1.5)	(1.5)
Single Experiment	95	77.4	95.4	94.2	94.3	94.7	**94.7**
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
	90	71.1	90.8	89.0	89.1	89.6	**89.7**
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
	80	61.0	81.4	78.8	78.7	79.4	**79.4**
		(0.1)	(0.1)	(0.2)	(0.2)	(0.1)	(0.1)

Table 20: Estimated inflation and std. dev. of the logistic MLE when the true model is probit and when covariates are from a MVT distribution. This example is repeated $B = 1,100$ times.

	Inflation				Standard Deviation				
	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical inflation	Classical Theory	High-dim Theory	Resized Bootstrap Known-γ	Resized Bootstrap Estimated-γ	Empirical Std. dev.
$\beta = 0$	-	-	-	-	2.79	**3.49**	3.26	3.40	3.45
$\beta = 4.52$	2.97	2.74	**2.80**	2.87	2.83	**3.49**	3.31	3.46	3.54

Table 21: Coverage proportion of a single *null* variable, single *non-null* variable, and in a *single-shot* experiment with standard deviation between parentheses. Here we compute the logistic MLE when the true model is Probit and when covariates are from a MVT distribution.

		Theoretical CI		Resized Bootstrap			
	Nominal			Known γ		Estimated γ	
	coverage	Classical	High-Dim	Boot-g	Boot-t	Boot-g	Boot-t
Single Null	95	88.6	95.5	93.8	93.99	95.3	**95.2**
		(1.0)	(0.6)	(0.7)	(0.7)	(0.6)	(0.7)
	90	80.6	90.9	87.7	87.4	89.4	**89.9**
		(1.2)	(0.9)	(1.0)	(1.0)	(0.9)	(0.9)
	80	68.7	**78.8**	75.8	74.9	77.9	77.6
		(1.4)	(1.2)	(1.3)	(1.3)	(1.3)	(1.3)
Single Non-null	95	19.0	94.5	92.4	92.8	94.4	**94.6**
		(1.2)	(0.7)	(0.8)	(0.8)	(0.7)	(0.7)
	90	12.6	**90.1**	87.4	87.4	89.2	89.0
		(1.0)	(0.9)	(1.0)	(1.0)	(0.9)	(0.9)
	80	7.18	**80.0**	77.9	77.8	81.0	80.6
		(0.8)	(1.2)	(1.3)	(1.3)	(1.2)	(1.2)
Single Experiment	95	78.27	**94.8**	93.2	93.1	94.5	94.4
		(0.06)	(0.05)	(0.07)	(0.07)	(0.03)	(0.03)
	90	70.7	**89.7**	87.4	87.2	89.2	88.9
		(0.07)	(0.07)	(0.09)	(0.09)	(0.05)	(0.05)
	80	59.3	**79.7**	76.7	76.4	78.8	78.4
		(0.07)	(0.10)	(0.11)	(0.11)	(0.06)	(0.06)