Annals of the Institute of Statistical Mathematics (2024) 76:93–110
https://doi.org/10.1007/s10463-023-00877-3

A tuning-free efficient test for marginal linear effects in high-dimensional quantile regression

Kai Xu[1] · **Nan An**[1]

Received: 30 September 2022 / Revised: 26 February 2023 / Accepted: 13 June 2023 /
Published online: 18 July 2023

Abstract

This work is concerned with testing the marginal linear effects of high-dimensional predictors in quantile regression. We introduce a novel test that is constructed using maxima of pairwise quantile correlations, which permit consistent assessment of the marginal linear effects. The proposed testing procedure is computationally efficient with the aid of a simple multiplier bootstrap method and does not involve any need to select tuning parameters, apart from the number of bootstrap replications. Other distinguishing features of the new procedure are that it imposes no structural assumptions on the unknown dependence structures of the predictor vector and allows the dimension of the predictor vector to be exponentially larger than sample size. To broaden the applicability, we further extend the preceding analysis to the censored response case. The effectiveness of our proposed approach in the finite samples is illustrated through simulation studies.

Keywords High dimension · Marginal quantile regression · Multiplier bootstrap · Quantile correlation · Quantile slope · Randomly censored data

1 Introduction

Inference for regression models is of central importance in statistics. As pioneered by Koenker and Bassett (1978), quantile regression models have drawn a great deal of attention in recent years, due to their flexibility for general error distributions and because they provide a more detailed description of the conditional distribution of the response, compared to classical mean regression. More references about quantile regression estimation and interpretation can be found in the seminal book

✉ Kai Xu
tjxxukai@163.com

Nan An
an1985560@163.com

[1] School of Mathematics and Statistics, Anhui Normal University, Wuhu 241002, Anhui, China

of Koenker (2005). For a scalar response variable $Y \in \mathbb{R}$ and a set of predictors $\mathbf{x} = (X_1, \ldots, X_p)^{\mathrm{T}} \in \mathbb{R}^p$, we are interested in exploring whether $\mathbf{x}$ is useful in modeling a certain aspect of the quantiles of Y. This paper is especially concerned with the setting where $p = p(n)$ grows as a function of the sample size n such that p also tends to infinity.

In the context of low and fixed-dimensional predictors, many model specification tests can be directly used to test for significant predictors at the τth conditional quantile of Y. Examples include, but are not limited to, Zheng (1998), He and Zhu (2003), Conde-Amboage et al. (2015) and Xu and Chen (2020). These three types of tests extend, respectively, the mean restriction tests of Zheng (1996), Stute (1997), Escanciano (2006) and Su and Zheng (2017) to the quantile restriction case. However, these omnibus-type tests do not target the case of high-dimensional predictors and are based on fitting a model with all predictors. Therefore, they suffer from the curse of dimensionality and quickly become prohibitive when the predictor dimension p diverges.

How to test the presence of significant predictors that affect the conditional quantile of Y in high dimensions is an urgent challenge. To the best of our knowledge, the recent work by Wang et al. (2018) is perhaps the first to serve the interesting and challenging topic. Their approach is in the spirit of McKeague and Qian (2015), who proposed an adaptive resampling test for detecting significant predictors based on marginal linear mean regression. In other words, Wang et al. (2018) suggested a marginal testing procedure based on fitting the working marginal quantile regression models by regressing Y on X_k, $k = 1, \ldots, p$, for each k separately. Let $I(\cdot)$ denote the indicator function. Here for each k and the quantile loss function $\rho_\tau(u) = \{\tau - I(u \le 0)\}u$, the working marginal quantile regression solves the population minimization problem to obtain

$$(\alpha_{\tau,k}, \theta_{\tau,k}) = \arg\min_{\alpha,\theta} E\{\rho_\tau(Y - \alpha - \theta X_k) - \rho_\tau(Y)\}, \tag{1}$$

for $0 < \tau < 1$. For $1 \le k \le p$, the quantile slope $\theta_{\tau,k}$ is referred as the quantile marginal linear effect of X_k. Instead of assessing the existence of overall predictor effects in high dimensions, Wang et al. (2018) focused on a relatively weaker null hypothesis:

$$H_0 : \theta_{\tau,k} = 0, \text{ for all } 1 \le k \le p. \tag{2}$$

Let $k_{\tau,0} = \arg\min_{k \in \{1,\ldots,p\}} E\{\rho_\tau(Y - \alpha_{\tau,k} - \theta_{\tau,k} X_k) - \rho_\tau(Y)\}$ denote the index of the most informative predictor at the τth quantile. If the values of $\alpha_{\tau,k}$ and $\theta_{\tau,k}$ are unique for $k = 1, \ldots, p$, testing H_0 is equivalent to testing

$$H_0' : \theta_{\tau,k_{\tau,0}} = 0.$$

A rejection of H_0' automatically rejects H_0 and implies that at least one of the p predictors has an effect on the τth conditional quantile of Y. To test H_0 in (2), Wang et al. (2018) used the marginal quantile regression estimator of $\theta_{\tau,k_{\tau,0}}$ as a test statistic. Despite the usefulness of the existing marginal testing approach, its practical performance depends on the selection of tuning parameters including bandwidth, kernel and an adaptive threshold, which affect the inference procedure. By assuming the dimension to be fixed, Wang et al. (2018) suggested a double-bootstrap method for choosing the adaptive threshold, which is too computationally burdensome.

To overcome aforementioned problems, we introduce a novel marginal testing procedure that is constructed using maxima of pairwise quantile correlations (Li et al. 2015) that permit consistent assessment of the marginal linear effects in (2). We summarize the contributions of our work as follows.

- The proposed testing procedure is computationally fast with the aid of a simple multiplier bootstrap method and does not involve any need to select tuning parameters, apart from the number of bootstrap replications.
- It imposes no structural assumptions on the unknown dependence structures of the predictor vector and allows the dimension of the parameter vector of interest to be exponentially larger than sample size.
- It achieves better finite-sample performance by avoiding the slow convergence of maximum-type statistic and are particularly effective when the predictors with nonzero effects are sparse.
- Our methodology is readily applicable to handling censored data, a common problem in survival analysis.

Recently, Zhang et al. (2018) is a closely related work that essentially uses the martingale difference divergence to assess the nonlinear conditional quantile dependence of a response variable on a large number of predictors in a model-free setting. Our proposed test differs from that of Zhang et al. (2018) in two major respects. First, we propose to use a supremum-based test statistic, whereas Zhang et al. (2018) considered a sum-of-squares-based test statistic. We choose the maximum-type test statistic as it has good power against sparse alternatives, and in high-dimensional quantile regression $\{\theta_{\tau,k}\}_{1\le k\le p}$ are more likely to be nonzero in a sparse way. Additionally, our proposed method does not require correlational assumptions, besides some weak conditions on the moments and tail properties of the elements in the predictor vector. In contrast, Zhang et al. (2018) imposed structural assumptions on the unknown dependence structures of the predictor vector.

The rest of the paper is organized as follows. In Sect. 2, we introduce a new testing procedure based on the quantile correlation (Li et al. 2015) and propose a data-driven Gaussian approximation method to provide accurate critical values. Here, we should point out that this approximation relies on an impressive Gaussian approximation (GAR) theory recently developed in Chernozhukov et al. (2013). The application of their GAR theory may be nontrivial in the present paper. This is because our suggested test statistics will be constructed based on the transformed observations $\{\psi_\tau[Y_i - \widehat{Q}_\tau(Y)], \mathbf{x}_i\}_{i=1}^n$, where $\psi_\tau(u) = \tau - I(u \le 0)$ is the derivative of $\rho_\tau(u)$ and $\widehat{Q}_\tau(Y)$ is the sample τth quantile of $\{Y_i\}_{i=1}^n$. Theoretical properties of our test are studied in Sect. 3. To broaden the applicability, Sect. 4 further extends our main result to the case where the response variable is subject to random censoring. Section 5 conducts simulation studies to evaluate the empirical sizes and powers of our proposals. Section 6 concludes the paper. All the technical proofs are gathered in an online Supplementary Material.

2 Methodology

We begin with basic notation and definitions. For a matrix $\mathbf{A} = (a_{kl}) \in \mathbb{R}^{p\times p}$, define $\|\mathbf{A}\|_\infty = \max_{1\le k,l\le p} | a_{kl} |$. Denote by $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$ the smallest and the largest eigenvalues of $\mathbf{A}$, respectively. For a vector $\mathbf{a} = (a_1, \ldots, a_p)^{\mathrm{T}}$ and $q > 0$, denote by $\|\mathbf{a}\|_q = (\sum_{j=1}^p | a_j |^q)^{1/q}$ the ℓ_q norm. For simplicity, set $\|\cdot\| = \|\cdot\|_2$. Write $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. For two sequences $\{a_n\}$ and $\{b_n\}$, write $a_n \asymp b_n$ if there exist positive constants c and C such that $c \le \liminf_n(a_n/b_n) \le \limsup_n(a_n/b_n) \le C$. Write $a_n \lesssim b_n$ if $a_n \le C'b_n$ for some constant $C' > 0$ independent of n. Let $\#(\mathcal{D})$ denote the cardinality of the set $\mathcal{D}$. The notation $N(\boldsymbol{\mu}, \mathbf{A})$ stands for the p-variate normal distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^p$ and covariance matrix $\mathbf{A} \in \mathbb{R}^{p\times p}$.

2.1 A quick review of quantile correlation

For random variables Y and X_k, let $Q_\tau(Y)$ be the τth unconditional quantile of Y and $Q_\tau(Y \mid X_k)$ be the τth quantile of Y conditional on X_k. Note that $Q_\tau(Y \mid X_k)$ is independent of X_k if and only if the random variables $I\{Y - Q_\tau(Y) > 0\}$ and X are independent. Based on this fact, Li et al. (2015) advocated using the quantile correlation

$$\begin{aligned}\mathrm{qcor}_\tau(Y, X_k) &= \frac{\mathrm{qcov}_\tau(Y, X_k)}{\{\mathrm{var}[\psi_\tau\{Y - Q_\tau(Y)\}]\mathrm{var}(X_k)\}^{1/2}}\\ &= \frac{E[\psi_\tau\{Y - Q_\tau(Y)\}\{X_k - E(X_k)\}]}{\{\tau(1-\tau)\sigma_k^2\}^{1/2}},\end{aligned} \tag{3}$$

for measuring dependence between $Q_\tau(Y \mid X_k)$ and X_k, where $\mathrm{qcov}_\tau(Y, X_k) = E[\psi_\tau\{Y - Q_\tau(Y)\}\{X_k - E(X_k)\}]$ is the quantile covariance of Y and X_k, and $\sigma_k^2 = \mathrm{var}(X_k)$ is the variance of X_k.

In the simple linear regression with the quadratic loss function, the slope of the regression line is directly related to the correlation coefficient. Unlike in mean regression, the quantile slope $\theta_{\tau,k}$ in (1) has no explicit form. Lemma 1 enables us to derive a nice relationship between the quantile slope and the quantile correlation.

Lemma 1 (Li et al. 2015) *Suppose that random variables X_k and $\varepsilon_{\tau,k} = Y - \alpha_{\tau,k} - \theta_{\tau,k}X_k$ have a joint density and $E(X_k^2) < \infty$. Then, the quantile correlation* $\mathrm{qcov}_\tau(Y, X_k)$ *increases with the quantile slope $\theta_{\tau,k}$, and* $\mathrm{qcor}_\tau(Y, X_k) = 0$ *if and only if $\theta_{\tau,k} = 0$.*

From the above lemma, the null hypothesis in (2) holds if and only if $\mathrm{qcor}_\tau(Y, X_1) = \ldots = \mathrm{qcor}_\tau(Y, X_p) = 0$. The quantile correlation $\mathrm{qcor}_\tau(Y, X_k)$ lies between − 1 and 1, whereas the range of quantile slope can be not bounded. It is natural to employ the quantile correlation rather than the quantile slope to rank the significance of predictors on the quantile of Y.

2.2 The qcor-based test statistic in high dimension

Suppose that $\{(Y_i, \mathbf{x}_i^{\mathrm{T}}) : i = 1, \dots, n\}$ is a random sample of $(Y, \mathbf{x})$, where $\mathbf{x}_i = (X_{i1}, \dots, X_{ip})^{\mathrm{T}}$. Let $\widehat{Q}_\tau(Y) = \inf\{y : \widehat{F}_Y(y) \ge \tau\}$ be the sample τth quantile of $\{Y_i : i = 1, \dots, n\}$, where $\widehat{F}_Y(y) = n^{-1}\sum_{i=1}^n I(Y_i \le y)$ is the empirical distribution function. Based on Eq. (3), for each k, a natural estimator of the quantile correlation $\mathrm{qcor}_\tau(Y, X_k)$ can be defined as

$$\widehat{\mathrm{qcor}}_\tau(Y, X_k) = \left\{\tau(1-\tau)\widehat{\sigma}_k^2\right\}^{-1/2} n^{-1}\sum_{i=1}^n \psi_\tau\left\{Y_i - \widehat{Q}_\tau(Y)\right\}\left(X_{ik} - \overline{X}_k\right),$$

where $\overline{X}_k = n^{-1}\sum_{i=1}^n X_{ik}$ and $\widehat{\sigma}_k^2 = n^{-1}\sum_{i=1}^n \left(X_{ik} - \overline{X}_k\right)^2$.

By Lemma 1, the null hypothesis in (2) is equivalent to $\max_{1\le k\le p} \mid \mathrm{qcor}_\tau(Y, X_k) \mid = 0$. Moreover, $\max_{1\le k\le p} \mid \mathrm{qcor}_\tau(Y, X_k) \mid$ is nonnegative. It is appropriate to consider a one-sided test in which we aggregate all marginal sample quantile correlations $\{\widehat{\mathrm{qcor}}_\tau(Y, X_k) : k = 1, \dots, p\}$ into the test statistic

$$\widehat{S}_\tau = \max_{1\le k\le p} \mid \widehat{\mathrm{qcor}}_\tau(Y, X_k) \mid .$$

2.3 A new testing procedure

For the nominal significance level α, we propose a new test that rejects (2) when $n^{1/2}\widehat{S}_\tau > c_{\tau,\alpha}$, where $c_{\tau,\alpha}$ is obtained using a simple multiplier bootstrap method. In the following, we describe the bootstrap scheme.

(S1). Independent of $\{(Y_i, \mathbf{x}_i^{\mathrm{T}}) : i = 1, \dots, n\}$, we generate a sequence of independent $N(0, 1)$ random variables $e_1, \dots, e_n$.

(S2). Using the e_i's as multipliers, we calculate the perturbed version of the test statistic

$$\widehat{S}_\tau^* = \max_{1\le k\le p} \mid \widehat{\mathrm{qcor}}_\tau^*(Y, X_k) \mid, \tag{4}$$

where $\widehat{\mathrm{qcor}}_\tau^*(Y, X_k) = \widehat{\mathrm{qcov}}_\tau^*(Y, X_k)/\{\tau(1-\tau)\widehat{\sigma}_k^2\}^{1/2}$ and $\widehat{\mathrm{qcov}}_\tau^*(Y, X_k) = n^{-1}\sum_{i=1}^n \psi_\tau\{Y_i - \widehat{Q}_\tau(Y)\}(X_{ik} - \overline{X}_k)e_i$, for $1 \le k \le p$.

(S3). The critical value $c_{\tau,\alpha}$ is defined as the upper α-quantile of $\widehat{S}_\tau^*$ conditional on $\{(Y_i, \mathbf{x}_i^{\mathrm{T}}) : i = 1, \dots, n\}$. That is, $c_{\tau,\alpha} = \inf\{t : \mathbb{P}^*(n^{1/2}\widehat{S}_\tau^* > t) \le \alpha\}$, where $\mathbb{P}^*$ denotes the probability measure induced by the Gaussian random variables with $\{(Y_i, \mathbf{x}_i^{\mathrm{T}}) : i = 1, \dots, n\}$ being fixed.

This algorithm is in the spirit of Liu and Shao (2013) and Chernozhukov et al. (2017) and combines the ideas of multiplier bootstrap and parametric bootstrap. Unlike the adaptive resampling procedure in Wang et al. (2018), this particular resampling scheme has the practical advantage of not requiring the subjective choice of tuning parameters at each bootstrap replication and is a fast-computing data perturbation procedure. We will prove that our bootstrap-assisted test also resolves three issues at once. First, it achieves better finite-sample performance by avoiding the slow convergence of maximum-type statistic; see Liu et al. (2008) for detailed illustrations in the context of dependence testing. Second, due to the principle of parametric bootstrap in step (S2), the new procedure we present above automatically takes into account correlations among the $\widehat{\text{qcor}}_\tau^*(Y, X_k)$'s and imposes no structural assumptions on the unknown dependence structures of the predictor vector. In addition, our procedure allows the dimension of the predictor vector of interest to be much larger than the sample size.

3 Theoretical properties

We study the asymptotic properties of the proposed test $\Psi_{\hat{S}_\tau,\alpha} = I\left(n^{1/2}\hat{S}_\tau > c_{\tau,\alpha}\right)$ under both null hypothesis (2) and a sequence of local alternatives. Denote $\mathbf{\Sigma} = (\sigma_{kl})_{1\le k,l\le p} \in \mathbb{R}^{p\times p}$ for the covariance matrix of the predictor vector $\mathbf{x}$. For the asymptotic properties, we only require the following relaxed regularity conditions.

(C1). The design matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^{\mathrm{T}}$ has either i.i.d. sub-Gaussian rows (i.e., $\sup_{\|\mathbf{a}\|\le 1} E\{\exp(|\sum_{k=1}^p a_k X_{ik}|^2/C)\} = O(1)$ for some fixed positive constant C) or i.i.d. rows satisfying for some $K_n \ge 1$, $\|\mathbf{X}\|_\infty = O(K_n)$, where K_n is allowed to grow with n. The latter we call the strongly bounded case.

(C2). $\sigma_{11}, \ldots, \sigma_{pp}$ are uniformly bounded away from zero and infinity.

(C3). The cumulative distribution function of the continuous response variable Y, F_Y is continuously differentiable in a small neighborhood of $Q_\tau(Y)$, say $[Q_\tau(Y) - \delta_0, Q_\tau(Y) + \delta_0]$ with $\delta_0 > 0$. Let $G_1(\delta_0) = \inf_{y\in[Q_\tau(Y)-\delta_0, Q_\tau(Y)+\delta_0]} f_Y(y)$ and $G_2(\delta_0) = \sup_{y\in[Q_\tau(Y)-\delta_0, Q_\tau(Y)+\delta_0]} f_Y(y)$, where f_Y is the density function of Y. Assume that $0 < G_1(\delta_0) \le G_2(\delta_0) < \infty$.

(C4). For $1 \le k \le p$, the conditional density $f_{Y|X_k}(\cdot)$ is uniformly integrable on $[Q_\tau(Y) - \delta_0, Q_\tau(Y) + \delta_0]$.

Assumption (C1) is standard for the predictors in high-dimensional regression and has been used in van de Geer et al. (2014), Javanmard and Montanari (2014) and Belloni et al. (2015). Assumption (C2) is much weaker than those such as $C^{-1} \le \lambda_{\min}(\mathbf{\Sigma}) \le \lambda_{\max}(\mathbf{\Sigma}) \le C$ for some constant $C > 0$. No structural assumptions on the unknown covariance matrix $\mathbf{\Sigma}$ are imposed in Assumption (C2). Assumption (C3) was first introduced in Shao and Zhang (2014) and is quite mild. Assumption (C4) is a standard condition in the literature on quantile regression and ensures that the null hypothesis in (2) is equivalent to $\text{qcor}_\tau(Y, X_k) = 0$ for $1 \le k \le p$. It is worth noting that the above assumptions do not impose the existence of any moments of the errors $\varepsilon_{\tau,k}, k = 1, \ldots, p$. Theorem 1 below shows that, under these mild moment and regularity conditions, the proposed test $\Psi_{\hat{S}_\tau,\alpha}$ with $c_{\tau,\alpha}$ defined in Section 2.3 has asymptotical size α.

Theorem 1 *Under H_0 in (2), suppose that Assumptions (C1), (C2), (C3) and (C4) hold. Then, as $n, p \to \infty$, $\mathbb{P}\left(\Psi_{\hat{S}_\tau,\alpha} = 1\right) \to \alpha$ holds if*

$$v^{1/3}\{1 \vee \log(p/v)\}^{2/3} \bigvee \zeta\{1 \vee \log(p/\zeta)\}^{1/2} \bigvee \{n^{-1}K_0^2 \log^7(pn)\}^{1/6} \to 0, \quad (5)$$

where $v \asymp n^{-1}K_n^2 \log(p) \vee n^{-1/2}K_n^2 \log^{1/2}(p)$, $\zeta \asymp n^{-1/4}K_n^3 \log^{3/4}(p)$ and $K_0 = K_n$ in the strongly bounded case, and $v \asymp n^{-1/2}\log(p) \vee n^{-1}\log(p)\log(pn)$, $\zeta \asymp n^{-3/2}\log^2(p)\log^{3/2}(pn)$ and $K_0 = 1$ in the sub-Gaussian case.

Theorem 1 guarantees our testing procedure to maintain significance level asymptotically even when p is much larger than n. The asymptotic validity of the proposed test is also obtained without imposing structural assumptions on $\mathbf{x}$, such as conditions (8), (10) and (11) in Zhang et al. (2018).

Let $\mathcal{T} = \{\tau_1, \dots, \tau_L\}$ be a set of quantile levels of interest. To pool information across quantiles, one may consider the maximum-type test statistic $\hat{S} = \max_{\tau\in\mathcal{T}} \hat{S}_\tau = \max_{1\le l\le L} \max_{1\le k\le p} \mid \widehat{\text{qcor}}_{\tau_l}(Y, X_k) \mid$. Accordingly, define the bootstrap statistic $\hat{S}^* = \max_{\tau\in\mathcal{T}} \hat{S}^*_\tau = \max_{1\le l\le L} \max_{1\le k\le p} \mid \widehat{\text{qcor}}^*_{\tau_l}(Y, X_k) \mid$, the critical value $c_\alpha = \inf\{t : \mathbb{P}^*\left(n^{1/2}\hat{S}^* > t\right) \le \alpha\}$, and the multiple-quantile test $\Psi_{\hat{S},\alpha} = I(n^{1/2}\hat{S} > c_\alpha)$. Suppose that Assumptions (C1) and (C2) are satisfied. Furthermore, (2) and Assumptions (C3)–(C4) hold over $\tau \in \mathcal{T}$. By replacing the dimension p in condition (5) with pL, apply the convergence result in Theorem 1 to establish the consistency of the multiple-quantile test.

In the next, we investigate the asymptotic power of $\Psi_{\hat{S}_\tau,\alpha}$. It is known that statistics of the maximum-type are preferable for detecting relatively sparse signals (Hall and Jin 2010; Arias-Castro et al. 2011). The scenario in which the existence of marginal effects occurs only at a small number of locations is of great interest in high-dimensional quantile regression. From Lemma 1, $\text{qcov}_\tau(Y, X_k)$ is a rescaled version of the quantile slope $\theta_{\tau,k}$ via a nondecreasing function. Note that the test statistic $\hat{S}_\tau$ is defined to be the maximum of marginal sample quantile correlations. The power of the proposed test is in spirit controlled by the maximum of marginal population quantile correlations, i.e., $\max_{1\le k\le p} \mid \text{qcov}_\tau(Y, X_k)/\{\tau(1-\tau)\sigma_{kk}\}^{1/2} \mid$, which can be viewed as a signal-to-noise ratio. Therefore, we focus on the local sparse alternatives characterized by the following class of vector

$$\mathcal{V}_\tau(\gamma) = \{\boldsymbol{\theta}_\tau = (\theta_{\tau,1}, \dots, \theta_{\tau,p})^{\mathrm{T}} : \max_{1\le k\le p} \mid \text{qcov}_\tau(Y, X_k)/\sigma_{kk}^{1/2} \mid \ge \gamma\{\tau(1-\tau)\log(p)/n\}^{1/2}\}.$$

To evaluate powers, we assume the following condition.

(C5). Assume that $\max_{1\le l\le p} \sum_{k=1}^{p} \sigma_{kl}^2 \le c$ for some constant $c > 0$.

Assumption (C5) is slightly weaker than that of $\lambda_{\max}(\boldsymbol{\Sigma}) \le C$ for some constant $C > 0$ and is similar to that of Cai, Liu and Xia (Cai et al. (2014), Lemma 6).

Theorem 2 *Under the assumptions in Theorem* 1 *and Assumption* (*C*5), *then as* $n, p \to \infty$, *we have* $\inf_{\theta_\tau \in \mathcal{V}_\tau(\epsilon_0 + 2^{1/2})} \mathbb{P}\left(\Psi_{\widehat{S}_\tau,\alpha} = 1\right) \to 1$ *for any* $\epsilon_0 > 0$.

Theorem 2, together with Lemma 1, says that the correct rejection of our bootstrap-assisted test can still be triggered even when there exists only one entry of $\boldsymbol{\theta}_\tau$ with a magnitude being larger than $(\epsilon_0 + 2^{1/2})\{\tau(1-\tau)\log(p)/n\}^{1/2}$. An important by-product of the proof of Theorem 2 is that under additional Assumption (C5), the distribution of $n^{1/2}\widehat{S}_\tau$ can be well approximated by $\max_{1 \le k \le p} \mid G_k \mid$ with $(G_1, \ldots, G_p)^{\mathrm{T}} \sim (\boldsymbol{\mu}, \mathbf{D}^{-1/2}\boldsymbol{\Theta}\ \mathbf{D}^{-1/2})$, where $\boldsymbol{\mu} = \{\mathrm{qcor}_\tau(Y, X_1), \ldots, \mathrm{qcor}_\tau(Y, X_p)\}^{\mathrm{T}} \in \mathbb{R}^p$, $\mathbf{D} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp}) \in \mathbb{R}^{p \times p}$ and $\boldsymbol{\Theta} = E[\psi_\tau^2\{Y - Q_\tau(Y)\}\{\mathbf{x} - E(\mathbf{x})\}\{\mathbf{x} - E(\mathbf{x})\}^{\mathrm{T}}] \in \mathbb{R}^{p \times p}$. Therefore, when $\mathbf{D}^{-1/2}\boldsymbol{\Theta}\mathbf{D}^{-1/2}$ is an identity matrix, the constant $2^{1/2}$ turns out to be asymptotically optimal in the minimax sense (see Arias-Castro et al. 2011; Ingster et al. 2010). Under Assumption (C5) and using the boundedness of the function $\psi_\tau(\cdot)$, we further have for any $x \in \mathbb{R}$ and as $p \to \infty$, $\mathbb{P}[n\widehat{S}_\tau^2 - 2\log(p) + \log\{\log(p)\} \le x] \to \exp\{-\pi^{-1/2}\exp(-x/2)\}$ under the null hypothesis. In contrast with our bootstrap-assisted method, the above alternative testing procedure heavily relies on the structural condition on the unknown covariance matrix $\mathbf{D}^{-1/2}\boldsymbol{\Theta}\mathbf{D}^{-1/2}$. In addition, the critical value obtained from the above type I extreme value distribution may not work well in practice since this weak convergence is typically slow.

4 Extension to the censored response case

Our main result can be extended beyond complete responses to a general framework with the response variable being subject to random censoring. Assume that Y_i is subject to random right censoring. Instead of $\{(Y_i, \mathbf{x}_i^{\mathrm{T}}) : i = 1, \ldots, n\}$, we observe the data $\{(\delta_i, Y_i^*, \mathbf{x}_i^{\mathrm{T}}) : i = 1, \ldots, n\}$, consisting of independent copies of $(\delta, Y^*, \mathbf{x}^{\mathrm{T}})$, where

$$\delta_i = I(Y_i \le C_i), \quad Y_i^* = Y_i \wedge C_i,$$

with C_i representing the censoring variable. For ease of exposition, we assume that the censoring distribution is independent of predictors.

To accommodate censoring, we extend (Li et al. 2015, Lemma 1) from the case of complete data to the random censoring case. Let $G(t) = \mathbb{P}(C > t)$ be the survival function of C. The working marginal quantile regression for the censored response case solves the population minimization problem to obtain

$$\begin{aligned}(\alpha_{\tau,k}, \theta_{\tau,k}) &= \arg\min_{\alpha,\theta} E\{\rho_\tau(Y - \alpha - \theta X_k) - \rho_\tau(Y)\} \\ &= \arg\min_{\alpha,\theta} E[\{\delta/G(Y^*)\}\{\rho_\tau(Y^* - \alpha - \theta X_k) - \rho_\tau(Y^*)\}],\end{aligned}$$

for $1 \le k \le p$. Let $F(y) = \mathbb{P}(Y \le y)$ and the weight function

$$w_\tau(F) = \begin{cases} 1 & \text{if } \Delta = 1 \text{ or } F(C) > \tau, \\ \frac{\tau - F(C)}{1 - F(C)} & \text{if } \Delta = 0 \text{ and } F(C) \le \tau, \end{cases}$$

redistributes the masses of censored observations to the right (Portnoy, 2003; Wang and Wang, 2009). Following Portnoy (2003) and Li et al. (2015), it is natural to define the censored quantile correlation

$$\mathrm{cqcor}_\tau(Y, X_k) = \frac{E([\tau - w_\tau(F)I\{Y^* \le Q_\tau(Y)\}]\{X_k - E(X_k)\})}{\{\tau(1-\tau)\sigma_k^2\}^{1/2}}.$$

We then obtain the relationship between $\theta_{\tau,k}$ and $\mathrm{cqcor}_\tau(Y, X_k)$ given below.

Lemma 2 *Suppose that random variables X_k and $\varepsilon_{\tau,k} = Y - \alpha_{\tau,k} - \theta_{\tau,k}X_k$ have a joint density and $E(X_k^2) < \infty$. Then, the censored quantile correlation* $\mathrm{cqcor}_\tau(Y, X_k)$ *increases with the quantile slope $\theta_{\tau,k}$, and* $\mathrm{cqcor}_\tau(Y, X_k) = 0$ *if and only if $\theta_{\tau,k} = 0$.*

Let $1 - \widehat{F}(y)$ be the Kaplan–Meier estimator of Y based on $\{(\delta_i, Y_i^*) : i = 1, \dots, n\}$. The τth sample quantile $\widehat{Q}_\tau(Y)$ is an estimator of $Q_\tau(Y)$ when Y is subject to right censoring. From Lemma 2, a maximal statistic for testing H_0 in the presence of random censoring can be defined as

$$\widehat{T}_\tau = \max_{1 \le k \le p} | \widehat{\mathrm{cqcor}}_\tau(Y, X_k) |,$$

where the sample censored quantile correlation is defined as

$$\widehat{\mathrm{cqcor}}_\tau(Y, X_k) = \{\tau(1-\tau)\widehat{\sigma}_k^2\}^{-1/2} n^{-1} \sum_{i=1}^n \left[\tau - w_{i\tau}(\widehat{F})I\{Y_i^* \le \widehat{Q}_\tau(Y)\}\right]\left(X_{ik} - \overline{X}_k\right),$$

for $1 \le k \le p$, Here

$$w_{i\tau}(\widehat{F}) = \begin{cases} 1 & \text{if } \Delta_i = 1 \text{ or } \widehat{F}(C_i) > \tau, \\ \frac{\tau - \widehat{F}(C_i)}{1 - \widehat{F}(C_i)} & \text{if } \Delta_i = 0 \text{ and } \widehat{F}(C_i) \le \tau. \end{cases}$$

To conduct the testing procedure, we employ the multiplier bootstrap in the following way. Generate a sequence of i.i.d standard normal random variables $\{e_i : i = 1, \dots, n\}$ and define the bootstrap statistic,

$$\widehat{T}_\tau^* = \max_{1 \le k \le p} \{\tau(1-\tau)\widehat{\sigma}_k^2\}^{-1/2} \left| n^{-1} \sum_{i=1}^n \left[\tau - w_{i\tau}(\widehat{F})I\{Y_i^* \le \widehat{Q}_\tau(Y)\}\right](X_{ik} - \overline{X}_k)e_i \right|.$$

The bootstrap critical value is given by $c'_{\tau,\alpha} = \inf\{t : \mathbb{P}^*(n^{1/2}\widehat{T}_\tau^* > t) \le \alpha\}$.

For the random censoring case, in addition to Assumptions (C1)–(C4), we make the following assumption to facilitate the technical proofs.

(C6). $\mathbb{P}(t \leq Y_i \leq C_i) \geq \tau_0 > 0$ for some positive constant τ_0 and any $t \in [0, T]$, where T denotes the maximum follow-up time. Furthermore, $\sup\{t : \mathbb{P}(Y > t) > 0\} \geq \sup\{t : \mathbb{P}(C > t) > 0\}$. The survival function of the censoring variable $G(t)$ has uniformly bounded first derivative.

Assumption (C6) is routinely used in the survival analysis literature (He, et al. 2013) to ensure that the Kaplan–Meier estimator and its inverse function are well behaved. We are now in position to justify the use of the test $\Psi_{\widehat{T}_\tau,\alpha} = I(n^{1/2}\widehat{T}_\tau > c'_{\tau,\alpha})$.

Theorem 3 *Under H_0 in (2), suppose that Assumptions (C1)–(C4) and (C6) hold. Then, as $n, p \to \infty$, $\mathbb{P}\left(\Psi_{\widehat{T}_\tau,\alpha} = 1\right) \to \alpha$ holds under Condition (5).*

5 Finite-sample performance

We conduct simulations to evaluate the finite-sample performance of the proposed testing procedure and compare it with the two existing methods: the adaptive resampling test (Wang, et al. 2018, WMQ for short) and the martingale-difference-divergence-based test (Zhang et al. 2018, ZYS for short). The WMQ test is applied by calling the function QMET.1tau available at http://www.columbia.edu/~im2131/ps/QMET.R. We use the EDMeasure library of R to compute the ZYS test statistic by the mdd function in that package. The calculation of the WMQ test statistic involves a kernel function and a bandwidth parameter. Upon the suggestion of Wang et al. (2018), we use the normal density kernel and choose the bandwidth parameter by following the rule from Hall and Sheather (1988). For the WMQ method and following Wang et al. (2018), we use 200 bootstrap samples, let the threshold $\lambda_n(\tau) = c\{\tau(1-\tau)\log(n)\}^{1/2}$, and choose $c \in (0, 6)$ by a double bootstrap with 100 double-bootstrap samples. The critical value of the ZYS test is calculated based on its asymptotic null distribution. For the proposed testing procedure, 500 independent realizations of in (4) by repeating steps (S1) and (S2) are used. The significance level α is fixed at 0.05, and for each test, the rejection probabilities reported in the simulation are computed based on 1,000 Monte Carlo replications.

5.1 Simulation results for the complete response cases

We consider three data-generating models. They are adapted from Wang et al. (2018) and can be expressed as

$$Y = \varepsilon, \tag{6}$$

$$Y = X_1/3 + \varepsilon, \tag{7}$$

$$Y = \sum_{k=1}^{5} X_k/4 - \sum_{k=6}^{10} 3X_k/20 + \varepsilon, \tag{8}$$

$$Y = \quad X_1^2/3 + \varepsilon, \tag{9}$$

$$Y = \quad X_1^2/3 + X_1/3 + \varepsilon, \tag{10}$$

where the predictor vector $\mathbf{x} = (X_1, \dots, X_{10}, \dots, X_p)^{\mathrm{T}}$ follows the multivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma} = \left(\rho^{|k-l|}\right)_{1\le k,l\le p}$, truncated at − 2 and 2. We take $\rho = 0.1$ and 0.5 for weak and moderately high correlations, respectively. The error term ε independent of $\mathbf{x}$ is generated from the $2^{-1/2}N(0,1)$ or $2^{-1}t(2)$ distribution for models (6)–(8) and from the $2^{-1}N(0,1)$ distribution for non-linear models (9, 10). The error adjustment factors $2^{-1/2}$ and 2^{-1} give less relevant noises, and the powers of tests will have less difficulty distinguishing between deviations. The data contain outliers in the response when $\varepsilon \sim 2^{-1}t(2)$. We consider three quantile levels $\tau = 0.25, 0.5, 0.75$, choose five dimensions p =10, 100, 200, 300, 400, 1000, and set the sample size to 200. Model (6) corresponds to the null hypothesis of no active predictors. Model (7) contains a unique active predictor and is the sparse case. In model (8), there are ten active predictors such that the predictors with nonzero effects are dense when p is moderately large and sparse under the setting where the number of predictors greatly exceeds the sample size. Nonlinear models (9) and (10) are used to assess the performance of the proposed test under the mis-specification of linear models.

Tables 1, 2, 3 and 4 present the empirical sizes and powers of the proposed test, the WMQ test and the ZYS test under models (6)–(8) with two different error distributions and several combinations of (ρ, τ, p) at the 5% significance level. Overall, all tests perform reasonably well under the null, though the WMQ test tends to have unstable size performance. This is probably due to the fact WMQ's size performance is sensitive to the threshold value used. A more time-consuming double-bootstrap procedure will hopefully provide more stable size at the expense of heavy computation. In terms of comparison of power, the overall powers of all tests increase as the correlation ρ increases, and the dependence within the predictors seems to enhance the powers under both dense and sparse alternatives. Since there is less relevant data information available for larger p, the overall powers of all tests decrease under both dense and sparse alternatives as the dimension p increases. Under sparse alternative (7), the proposed test and the WMQ test have quite good power and consistently outperform the ZYS test when the dimension is much larger than 10. As both NEW and WMQ are supremum-based tests, they target for sparse alternatives. Although our suggested test and the WMQ test have satisfactory power performance against the sparse alternative, our proposal gets more power. This is probably due to the fact WMQ's power performance depends on the subjective choice of tuning parameters such as bandwidth, kernel and an adaptive threshold, which are needed to be optimally chosen. Note that in model (8), signals are sparse when $p = 10, 100, 200$, whereas they become dense in situations of $p = 400, 100$. It is not surprised to see that for model (8) with $p \le 200$, the ZYS test is more powerful than the other tests, while our test has very much comparable powers. This is because the ZYS method is quite useful for detecting dense alternatives. Further, we also have reason to observe that the power

Table 1 Empirical sizes and powers for models (6)–(8) with $\rho = 0.1$ and $\varepsilon \sim 2^{-1/2}N(0,1)$ at significance level 5%

Model	p	Results for $\tau = 0.25$			Results for $\tau = 0.5$			Results for $\tau = 0.75$		
		NEW	WMQ	ZYS	NEW	WMQ	ZYS	NEW	WMQ	ZYS
(6)	10	0.047	0.034	0.060	0.057	0.041	0.062	0.060	0.062	0.066
	100	0.046	0.056	0.061	0.054	0.038	0.059	0.043	0.034	0.067
	200	0.042	0.035	0.056	0.049	0.033	0.054	0.048	0.055	0.049
	400	0.053	0.045	0.058	0.046	0.059	0.048	0.052	0.046	0.054
	1000	0.049	0.037	0.054	0.047	0.040	0.052	0.048	0.032	0.051
(7)	10	0.945	0.846	0.870	0.989	0.940	0.939	0.956	0.863	0.874
	100	0.840	0.704	0.353	0.911	0.795	0.401	0.846	0.766	0.285
	200	0.748	0.683	0.182	0.871	0.714	0.275	0.760	0.691	0.236
	400	0.750	0.676	0.144	0.889	0.728	0.216	0.736	0.632	0.134
	1000	0.562	0.556	0.075	0.741	0.577	0.124	0.628	0.609	0.120
(8)	10	0.988	0.811	1.000	0.998	0.906	1.000	0.993	0.791	1.000
	100	0.892	0.599	0.957	0.954	0.677	0.975	0.901	0.644	0.956
	200	0.749	0.501	0.810	0.889	0.537	0.905	0.768	0.614	0.812
	400	0.681	0.550	0.604	0.845	0.499	0.742	0.682	0.586	0.650
	1000	0.524	0.473	0.292	0.697	0.388	0.354	0.537	0.496	0.322

Throughout our numerical studies we refer to our proposed test, and the tests proposed by Wang et al. (2018) and Zhang et al. (2018) as NEW, WMQ and ZYS, respectively

Table 2 Empirical sizes and powers for models (6)–(8) with $\rho = 0.5$ and $\varepsilon \sim 2^{-1/2}N(0,1)$ at significance level 5%

Model	p	Results for $\tau = 0.25$			Results for $\tau = 0.5$			Results for $\tau = 0.75$		
		NEW	WMQ	ZYS	NEW	WMQ	ZYS	NEW	WMQ	ZYS
(6)	10	0.055	0.027	0.065	0.043	0.034	0.064	0.051	0.055	0.067
	100	0.043	0.032	0.052	0.046	0.036	0.059	0.051	0.046	0.065
	200	0.045	0.026	0.063	0.042	0.040	0.058	0.044	0.063	0.055
	400	0.055	0.048	0.054	0.053	0.029	0.047	0.045	0.037	0.046
	1000	0.046	0.042	0.056	0.049	0.056	0.044	0.052	0.043	0.054
(7)	10	0.960	0.861	0.865	0.982	0.928	0.946	0.957	0.866	0.876
	100	0.850	0.747	0.389	0.896	0.764	0.435	0.837	0.721	0.339
	200	0.715	0.697	0.218	0.854	0.715	0.286	0.724	0.672	0.260
	400	0.704	0.690	0.168	0.831	0.665	0.189	0.735	0.670	0.189
	1000	0.633	0.631	0.094	0.842	0.607	0.156	0.654	0.635	0.138
(8)	10	1.000	0.980	1.000	1.000	1.000	1.000	1.000	0.990	1.000
	100	1.000	0.974	1.000	1.000	0.996	1.000	1.000	0.981	1.000
	200	1.000	0.970	0.992	1.000	0.984	1.000	1.000	0.980	1.000
	400	1.000	0.927	0.944	1.000	0.975	0.990	1.000	0.974	0.937
	1000	0.993	0.912	0.687	1.000	0.966	0.753	1.000	0.925	0.694

Refer to the captions in Table 1 for abbreviations

Table 3 Empirical sizes and powers for models (6)–(8) with $\rho = 0.1$ and $\varepsilon \sim 2^{-1}t(2)$ at significance level 5%

Model	p	Results for $\tau = 0.25$			Results for $\tau = 0.5$			Results for $\tau = 0.75$		
		NEW	WMQ	ZYS	NEW	WMQ	ZYS	NEW	WMQ	ZYS
(6)	10	0.048	0.029	0.069	0.045	0.041	0.057	0.049	0.057	0.065
	100	0.040	0.030	0.067	0.057	0.062	0.059	0.052	0.042	0.057
	200	0.049	0.057	0.051	0.041	0.036	0.045	0.048	0.026	0.064
	400	0.042	0.033	0.055	0.045	0.044	0.059	0.047	0.037	0.054
	1000	0.048	0.046	0.051	0.050	0.039	0.049	0.054	0.045	0.048
(7)	10	0.955	0.822	0.878	1.000	0.986	0.985	0.946	0.838	0.872
	100	0.787	0.658	0.290	0.982	0.943	0.539	0.802	0.647	0.346
	200	0.764	0.645	0.211	0.963	0.902	0.406	0.734	0.603	0.218
	400	0.693	0.620	0.141	0.950	0.872	0.264	0.706	0.599	0.177
	1000	0.644	0.609	0.137	0.940	0.847	0.173	0.612	0.581	0.105
(8)	10	0.985	0.749	1.000	0.999	0.909	1.000	0.986	0.782	1.000
	100	0.791	0.536	0.905	0.961	0.685	0.987	0.824	0.542	0.918
	200	0.682	0.497	0.741	0.902	0.603	0.902	0.715	0.500	0.741
	400	0.640	0.481	0.554	0.881	0.535	0.750	0.566	0.473	0.518
	1000	0.472	0.424	0.287	0.734	0.353	0.410	0.486	0.431	0.242

Refer to the captions in Table 1 for abbreviations

Table 4 Empirical sizes and powers for models (6)–(8) with $\rho = 0.5$ and $\varepsilon \sim 2^{-1}t(2)$ at significance level 5%

Model	p	Results for $\tau = 0.25$			Results for $\tau = 0.5$			Results for $\tau = 0.75$		
		NEW	WMQ	ZYS	NEW	WMQ	ZYS	NEW	WMQ	ZYS
(6)	10	0.052	0.035	0.071	0.056	0.028	0.070	0.047	0.044	0.066
	100	0.045	0.058	0.058	0.052	0.034	0.053	0.053	0.063	0.049
	200	0.048	0.030	0.060	0.048	0.059	0.045	0.043	0.037	0.058
	400	0.045	0.026	0.054	0.057	0.040	0.056	0.046	0.041	0.049
	1000	0.053	0.056	0.055	0.049	0.045	0.047	0.052	0.057	0.056
(7)	10	0.946	0.838	0.868	0.999	0.989	0.989	0.944	0.835	0.860
	100	0.817	0.672	0.348	0.981	0.932	0.573	0.819	0.660	0.358
	200	0.772	0.651	0.250	0.974	0.925	0.392	0.756	0.613	0.238
	400	0.650	0.554	0.163	0.953	0.884	0.242	0.749	0.601	0.157
	1000	0.571	0.549	0.077	0.924	0.835	0.138	0.553	0.464	0.087
(8)	10	1.000	0.982	1.000	1.000	0.995	1.000	1.000	0.988	1.000
	100	1.000	0.951	1.000	1.000	0.995	1.000	0.999	0.958	0.998
	200	1.000	0.946	0.984	1.000	0.982	0.996	0.990	0.954	0.955
	400	0.997	0.913	0.906	1.000	0.975	0.970	0.996	0.933	0.914
	1000	0.963	0.902	0.587	1.000	0.960	0.733	0.979	0.914	0.536

Refer to the captions in Table 1 for abbreviations

Table 5 Empirical sizes and powers of the proposed test across three quantiles 0.25, 0.5 and 0.75 for models (6)–(8) with $\rho \in \{0.1, 0.5\}$ and $\varepsilon \in \{2^{-1/2}N(0,1), 2^{-1}t(2)\}$ at significance level 5%

ε	Model	Results for $\rho = 0.1$					Results for $\rho = 0.5$				
		p					p				
		10	100	200	400	1000	10	100	200	400	1000
$2^{-1/2}N(0,1)$	(6)	0.043	0.058	0.053	0.045	0.047	0.041	0.046	0.054	0.056	0.051
	(7)	0.964	0.875	0.797	0.786	0.644	0.989	0.866	0.785	0.750	0.704
	(8)	0.995	0.920	0.811	0.738	0.603	1.000	1.000	1.000	1.000	0.997
$2^{-1}t(2)$	(6)	0.057	0.044	0.052	0.048	0.055	0.049	0.056	0.047	0.054	0.058
	(7)	0.973	0.925	0.882	0.814	0.776	0.972	0.901	0.813	0.774	0.695
	(8)	0.967	0.904	0.851	0.782	0.670	1.000	1.000	1.000	0.998	0.983

Table 6 Empirical powers for nonlinear models (9)–(10) with $(\rho, \tau) \in \{(0.1, 0.5), (0.1, 0.75), (0.5, 0.5), (0.5, 0.75)\}$ and $\varepsilon \sim 2^{-1}N(0,1)$ at significance level 5%

(ρ, τ)	Method	Results for Model (9)					Results for Model (10)				
		p					p				
		10	100	200	400	1000	10	100	200	400	1000
(0.1, 0.5)	NEW	0.085	0.054	0.071	0.065	0.053	0.916	0.807	0.753	0.694	0.535
	WMQ	0.079	0.048	0.045	0.064	0.055	0.887	0.755	0.702	0.661	0.484
	ZYS	0.716	0.326	0.127	0.083	0.062	0.990	0.496	0.359	0.248	0.120
(0.1, 0.75)	NEW	0.052	0.047	0.074	0.070	0.050	0.978	0.966	0.960	0.935	0.893
	WMQ	0.043	0.041	0.048	0.055	0.026	0.952	0.930	0.911	0.890	0.822
	ZYS	0.832	0.449	0.178	0.116	0.086	0.995	0.783	0.537	0.414	0.217
(0.5, 0.5)	NEW	0.045	0.036	0.043	0.039	0.031	0.915	0.808	0.747	0.715	0.589
	WMQ	0.040	0.024	0.032	0.025	0.022	0.871	0.763	0.694	0.680	0.536
	ZYS	0.588	0.322	0.136	0.105	0.064	0.992	0.518	0.376	0.274	0.105
(0.5, 0.75)	NEW	0.091	0.042	0.060	0.044	0.045	0.989	0.971	0.955	0.949	0.914
	WMQ	0.065	0.030	0.046	0.031	0.052	0.972	0.934	0.909	0.901	0.872
	ZYS	0.703	0.329	0.145	0.138	0.073	1.000	0.795	0.603	0.458	0.209

Refer to the captions in Table 1 for abbreviations

performance of the proposed test is better than that of the ZYS test for model (8) with $p \geq 400$. Table 5 summarizes the rejection rates of the proposed test across three quantiles 0.25, 0.5 and 0.75. As suggested by the results in Table 5, the new test across multiple quantiles tends to provide relatively stable size and high power.

In Table 6, we report the rejection rates of the NEW, WMQ and ZYS tests for nonlinear models (9, 10). Note that the NEW and WMQ tests are designed to detect linear covariate effects. As there is no linear relationship between $\mathbf{x}$ and Y at any quantiles in misspecified model (9), the NEW and WMQ tests have difficulty identifying the nonlinear relationship. However, if there is some linear trend in a misspecified model such as in model (10), they have an excellent capability of identifying the covariate effects. By contrast, the ZYS test achieves satisfactory power performance in the two nonlinear situations with $p \le 200$, though it tends to exhibit low power when the dimension is much larger than 200. This is anticipated because the ZYS procedure is built on a sum-of-squares-based statistic and is quite useful for identifying the nonlinear covariate effects under dense alternatives.

5.2 Simulation results for the censored response cases

We consider cases in which the latent response variable Y is generated using the same setup as in models (6)–(8). We take the censoring time C to be $\widetilde{C} \wedge L$, where $\widetilde{C}$ is generated from Un(0,L) with L being the study duration time, which is chosen to yield a censoring rate of 40%. Due to the high censoring rate, the performance of the proposed test procedure is investigated at the median and the 0.25 quantile. Tables 7 and 8 summarize the simulation results of the proposed test for models (6)–(8) with censored responses when $\varepsilon \in \{2^{-1/2}N(0,1), 2^{-1}t(2)\}$, $\rho \in \{0.1, 0.5\}$ and the significance level is 5%. As presented, the sizes are generally precise at 5% level and the powers are comparable to the complete response cases.

6 Discussion

We have proposed a new procedure for detecting marginal effects in quantile regression. The validity of the test is established under a framework where the dimension of the variables can grow nonlinearly with the sample size. This new

Table 7 Empirical sizes and powers of the proposed test for models (6)–(8) with censored responses when $\varepsilon \sim 2^{-1/2}N(0,1)$, $\rho = 0.1, 0.5$ and the significance level is 5%

	p	Model (6)		Model (7)		Model (8)	
		$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.25$	$\tau = 0.5$
$\rho = 0.1$	10	0.046	0.047	0.946	0.977	0.993	1.000
	100	0.041	0.044	0.816	0.948	0.874	0.956
	200	0.043	0.052	0.759	0.881	0.782	0.902
	400	0.037	0.046	0.692	0.846	0.667	0.832
	1000	0.045	0.053	0.610	0.745	0.584	0.752
$\rho = 0.5$	10	0.047	0.048	0.957	0.986	1.000	1.000
	100	0.043	0.054	0.831	0.916	1.000	1.000
	200	0.038	0.051	0.779	0.894	1.000	1.000
	400	0.039	0.045	0.674	0.843	0.999	1.000
	1000	0.042	0.048	0.551	0.697	0.985	0.998

Table 8 Empirical sizes and powers of the proposed test for models (6)–(8) with censored responses when $\varepsilon \sim 2^{-1}t(2)$, $\rho = 0.1, 0.5$ and the significance level is 5%

	p	Model (6)		Model (7)		Model (8)	
		$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.25$	$\tau = 0.5$
$\rho = 0.1$	10	0.044	0.042	0.934	0.996	0.972	1.000
	100	0.046	0.056	0.882	0.984	0.790	0.953
	200	0.049	0.057	0.753	0.963	0.702	0.908
	400	0.049	0.054	0.707	0.955	0.581	0.860
	1000	0.053	0.047	0.598	0.932	0.459	0.784
$\rho = 0.5$	10	0.042	0.050	0.930	0.998	1.000	1.000
	100	0.057	0.062	0.808	0.980	1.000	1.000
	200	0.047	0.059	0.754	0.971	1.000	1.000
	400	0.048	0.045	0.688	0.947	1.000	1.000
	1000	0.046	0.055	0.609	0.925	0.966	1.000

framework can naturally handle censored data arising in survival analysis. The proposed test successfully avoids having to make subjective choices of parameters, such as bandwidths and kernels, and can be computed much faster. Our simulation results demonstrate the good behavior of our test in high dimensions. It is also worth noting that there is no power analysis provided for censored response case. The challenge lies in deriving the limiting distribution of our test statistic from the case of complete data to the random censoring case, which deserves further research.

Appendix: Necessary Lemmas

We start with providing technical lemmas used repeatedly in the online Supplementary Material.

Lemma 3 (Shao and Zhang 2014, Proposition 2) *Suppose the distribution function of Y satisfies Assumption (C3), then there exist $\epsilon_0 > 0$ and $c > 0$ such that for any $\epsilon \in (0, \epsilon_0)$,*

$$\mathbb{P}\left[n^{-1} \sum_{i=1}^{n} \left| \psi_\tau\{Y_i - \hat{Q}_\tau(Y)\} - \psi_\tau\{Y_i - Q_\tau(Y)\} \right| > \epsilon \right] \leq 3\exp\left(-2cn\epsilon^2\right).$$

Lemma 4 (Chernozhukov et al. 2015, Corollary 5.1) *Let $Z, Z_1, \dots, Z_n$ be i.i.d. random variables taking values in a measurable space $(S, \mathcal{S})$, Q denote a probability measure on the measurable space, $\mathcal{F} \subset L^2(Q)$ be a pointwise measurable class of real-valued functions on S with measurable envelope F, and $\mathcal{N}(\mathcal{F}, \|\cdot\|_{Q,2}, \delta)$ denote the δ-covering number for $\mathcal{F}$ with respect to the $L^2(Q)$-seminorm $\|\cdot\|_{Q,2}$.*

Suppose that $\mathcal{F}$ *is VC type, i.e., there exist constants* $A \geq e$ *and* $V \geq 1$ *such that* $\sup_Q \mathcal{N}(\mathcal{F}, \|\cdot\|_{Q,2}, \epsilon\|F\|_{Q,2}) \leq (A/\epsilon)^V$ *where* $\sup_Q$ *is taken over all finitely discrete distributions on S. Furthermore, suppose that* $0 < E\{F^2(Z)\} < \infty$, *and let* $\sigma^2 > 0$ *be any positive constant such that* $\sup_{f\in\mathcal{F}} E\{f^2(Z)\} \leq \sigma^2 \leq E\{F^2(Z)\}$. *Define* $B = E^{1/2}\{\max_{1\leq i\leq n} F^2(Z_i)\}$. *Then*

$$E\left(\|n^{-1/2}\sum_{i=1}^{n}\left[f(Z_i) - E\{f(Z)\}\right]\|_{\mathcal{F}}\right)$$
$$\lesssim \left(V\sigma^2\log\left[AE^{1/2}\{F^2(Z)\}/\sigma\right]\right)^{1/2} + n^{-1/2}VB\log\left[AE^{1/2}\{F^2(Z)\}/\sigma\right],$$

up to a universal constant.

Lemma 5 (Einmahl and Li 2008, Theorem 3.1) *Let* $\mathbf{z}_1, \ldots, \mathbf{z}_n$ *be independent centered random vectors in* $\mathbb{R}^p$ *where* $p \geq 2$. *Write* $\mathbf{z}_i = (Z_{i1}, \ldots, Z_{ip})^{\mathrm{T}}$ *for* $i = 1, \ldots, n$. *If* $E(\max_{1\leq k\leq p} \mid Z_{ik} \mid^r) < \infty$ *for* $i = 1, \cdots, n$, *and some* $r > 2$, *then*

$$\mathbb{P}\left\{\max_{1\leq k\leq p} \mid \sum_{i=1}^{n} Z_{ik} \mid \geq 2E(\max_{1\leq k\leq p} \mid \sum_{i=1}^{n} Z_{ik} \mid) + t\right\}$$
$$\lesssim \exp\{-t^2/(3n\max_{1\leq i\leq n}\max_{1\leq k\leq p} E \mid Z_{ik} \mid^2)\} + t^{-r}\sum_{i=1}^{n} E(\max_{1\leq k\leq p} |Z_{ik}|^r).$$

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10463-023-00877-3.

Acknowledgements This work is supported by National Natural Science Foundation of China (12271005, 11901006), Young Scholars Program of Anhui Province and Natural Science Foundation of Anhui Province (1908085QA06).

References

Arias-Castro, E., Candès, E. J., Plan, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *The Annals of Statistics, 39*, 2533–2556.
Belloni, A., Chernozhukov, V., Kato, K. (2015). Uniform post-selection inference for least absolute deviation regression and other Z-estimation problems. *Biometrika, 102*, 77–94.
Cai, T. T., Liu, W., Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society, Series B, 76*, 349–372.
Chernozhukov, V., Chetverikov, D., Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics, 41*, 2786–2819.
Chernozhukov, V., Chetverikov, D., Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probability Theory and Related Fields, 162*, 47–70.
Chernozhukov, V., Chetverikov, D., Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *The Annals of Probability, 45*, 2309–2352.
Conde-Amboage, M., Sánchez-Sellero, C., González-Manteiga, W. (2015). A lack-of-fit test for quantile regression models with high-dimensional covariates. *Computational Statistics and Data Analysis, 88*, 128–138.

Einmahl, U., Li, D. (2008). Characterization of lil behavior in Banach space. *Transactions of the American Mathematical Society 360*, 6677–6693.

Escanciano, J. C. (2006). A consistent diagnostic test for regression models using projections. *Econometric Theory, 22*, 1030–1051.

Hall, P., Jin, J. (2010). Innovated higher criticism for detecting sparse signals in correlated noise. *The Annals of Statistics, 38*, 1686–1732.

Hall, P., Sheather, S. J. (1988). On the distribution of a studentized quantile. *Journal of the Royal Statistical Society, Series B, 50*, 381–391.

He, X., Zhu, L. (2003). A lack-of-fit test for quantile regression. *Journal of the American Statistical Association, 98*, 1013–1022.

He, X., Wang, L., Hong, H. G. (2013). Quantile-adaptive model-free variable screening for high-dimensional heterogeneous data. *The Annals of Statistics, 41*, 342–369.

Ingster, Y. I., Tsybakov, A. B., Verzelen, N. (2010). Detection boundary in sparse regression. *Electronic Journal of Statistics, 4*, 1476–1526.

Javanmard, A., Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *Journal of Machine Learning Research, 15*, 2869–2909.

Koenker, R. (2005).*Quantile regression*, Cambridge University Press, New York).

Koenker, R., Bassett, G. (1978). Regression quantiles. *Econometrica, 46*, 33–50.

Li, G., Li, Y., Tsai, C. L. (2015). Quantile correlations and quantile autoregressive modeling. *Journal of the American Statistical Association, 110*, 246–261.

Liu, W., Shao, Q. (2013). A Cramér moderate deviation theorem for Hotelling's T^2-statistic with applications to global tests. *The Annals of Statistics, 41*, 296–322.

Liu, W., Lin, Z., Shao, Q. (2008). The asymptotic distribution and Berry-Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. *The Annals of Applied Probability, 18*, 2337–2366.

McKeague, I. W., Qian, M. (2015). An adaptive resampling test for detecting the presence of significant predictors. *Journal of the American Statistical Association, 110*, 1422–1433.

Portnoy, S. (2003). Censored regression quantiles. *Journal of the American Statistical Association, 98*, 1001–1012.

Shao, X., Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. *Journal of the American Statistical Association, 109*, 1302–1318.

Stute, W. (1997). Nonparametric model checks for regression. *The Annals of Statistics, 25*, 613–641.

Su, L., Zheng, X. (2017). A martingale-difference-divergence-based test for specification. *Economics Letters, 156*, 162–167.

van de Geer, S., Bühlmann, P., Ritov, Y., Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *The Annals of Statistics, 42*, 1166–1202.

Wang, H. J., Wang, L. (2009). Locally weighted censored quantile regression. *Journal of the American Statistical Association, 104*, 1117–1128.

Wang, H. J., McKeague, I. W., Qian, M. (2018). Testing for marginal linear effects in quantile regression. *Journal of the Royal Statistical Society: Series B, 80*, 433–452.

Xu, K., Chen, F. (2020). Martingale-difference-divergence-based tests for goodness-of-fit in quantile models. *Journal of Statistical Planning and Inference, 207*, 138–154.

Zheng, J. X. (1996). A consistent test of functional form via nonparametric estimation technique. *Journal of Econometrics, 75*, 263–289.

Zheng, J. X. (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory, 14*, 123–138.

Zhang, X., Yao, S., Shao, X. (2018). Conditional mean and quantile dependence testing in high dimension. *The Annals of Statistics, 46*, 219–246.