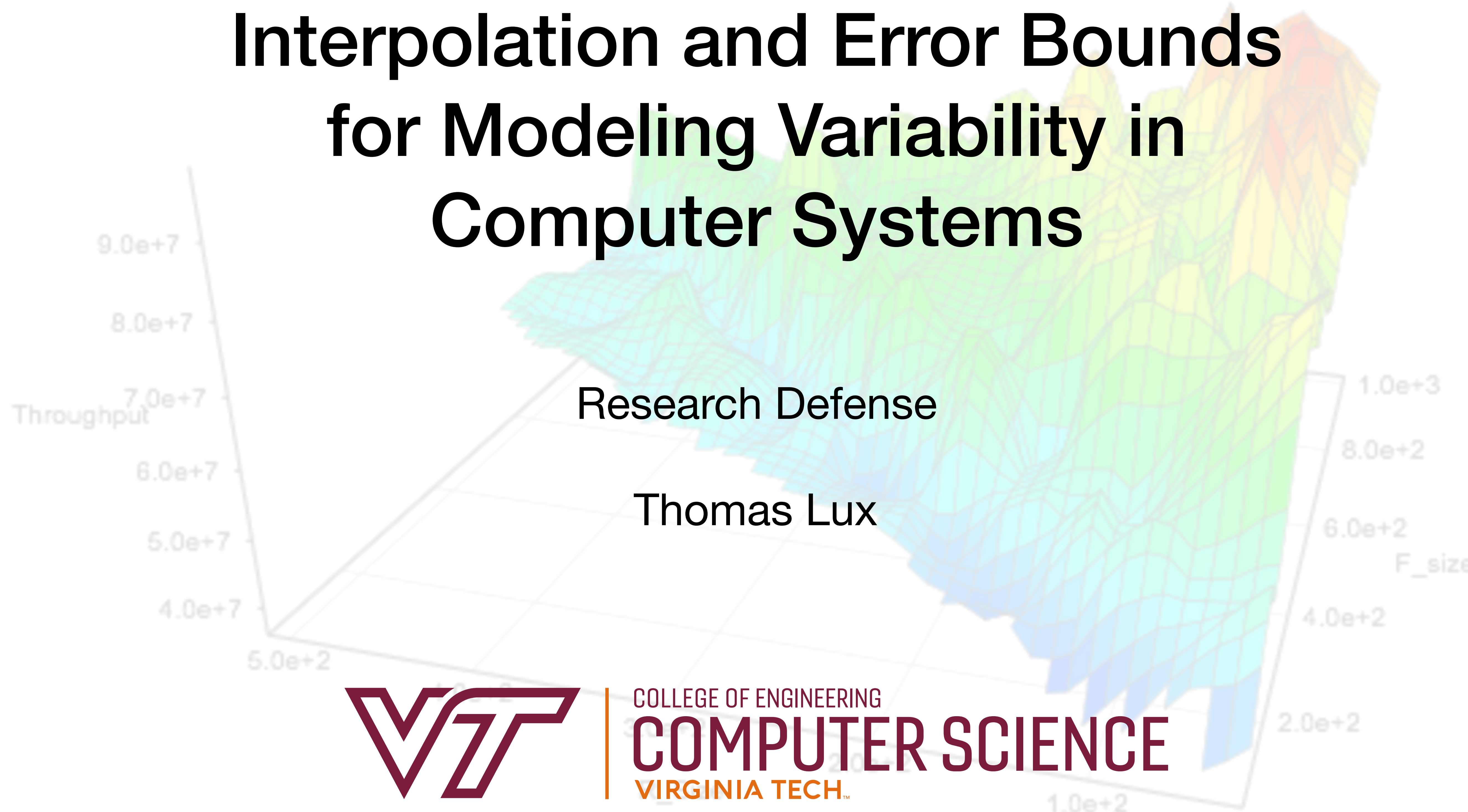
Interpolation and Error Bounds for Modeling Variability in Computer Systems



Research Defense

Thomas Lux

COLLEGE OF ENGINEERING
COMPUTER SCIENCE
VIRGINIA TECH

A Reminder, Chapters of Dissertation

1. The Importance and Applications of Variability
   define variability, why is it important?
2. Algorithms for Constructing Approximations
   approximation, regression and interpolation techniques
3. Naive Approximations of Variability
   mean, variance, and standard deviation prediction with IOzone
4. Box-Splines: Uses, Constructions, and Applications
   spline overview, box splines, meshes, fitting, and data sets
5. Stronger Approximations of Variability
   predicting distributions, measuring error, and tuning
6. An Error Bound for Piecewise Linear Interpolation
   theoretical results, importance, and empirical results
7. Improving Variability estimates
   empirical distributions, monotone splines, and goals

Goals

Incorporate additional comparison between regressors and interpolants with analytic function to test theoretical bound.

Produce a mathematical software for the computation of monotone quintic spline interpolants.

Apply this technique as an improved CDF estimator for the VarSys distribution approximation test problems.

Extend empirical VarSys test, write a journal paper, and submit as a mathematical software package to TOMS.

Goals

Incorporate additional comparison between regressors and interpolants with analytic function to test theoretical bound.

Produce a mathematical software for the computation of monotone quintic spline interpolants.

Apply this technique as an improved CDF estimator for the VarSys distribution approximation test problems.

Extend empirical VarSys test, write a journal paper, and submit as a mathematical software package to TOMS.

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

COMPUTER SCIENCE
VIRGINIA TECH.

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

times the distance to the nearest known point squared

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d} \|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

times the distance to the nearest known point squared

plus the square root of the dimension times the max change in slope

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

times the distance to the nearest known point squared

plus the square root of the dimension times the max change in slope

times the longest edge length between points defining the linear interpolant squared

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

times the distance to the nearest known point squared

plus the square root of the dimension times the max change in slope

times the longest edge length between points defining the linear interpolant squared

divided by how close the interpolated points are to being planar.

COMPUTER SCIENCE
VIRGINIA TECH

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

times the distance to the nearest known point squared

plus the square root of the dimension times the max change in slope

times the longest edge length between points defining the linear interpolant squared

divided by how close the interpolated points are to being planar.

COMPUTER SCIENCE
VIRGINIA TECH.

The Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma\|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

The absolute error of a linear interpolant is tightly upper bounded by

the max change in slope of the function

times the distance to the nearest known point squared

plus the square root of the dimension times the max change in slope

times the longest edge length between points defining the linear interpolant squared

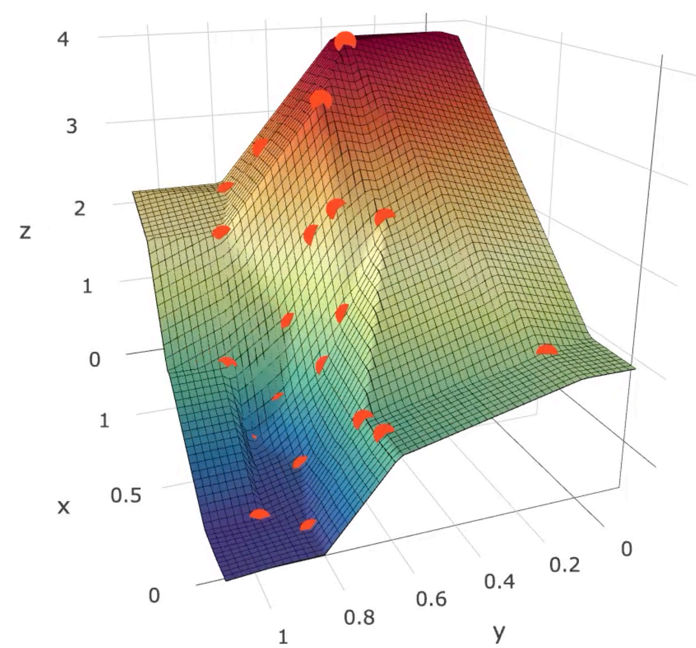
divided by how close the interpolated points are to being planar.

Piecewise Linear Approximations

Delaunay
(interpolant)

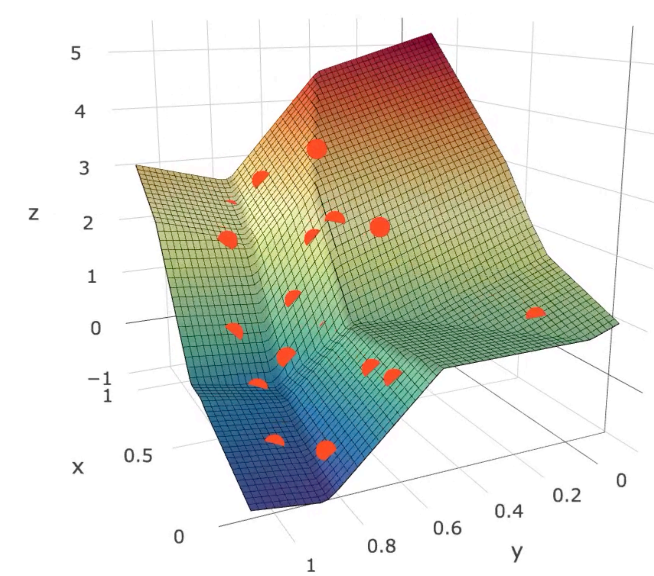
$$y = \sum_{i=0}^{d} w_i x^{(i)}, \quad \sum_{i=0}^{d} w_i = 1, \quad w_i \geq 0, \quad i = 0, \dots, d$$

$$\hat{f}(y) = \sum_{i=0}^{d} w_i f\left(x^{(i)}\right)$$



Multilayer Perceptron
(regressor)

$$l(u) = \left(u^t W_l + c_l\right)_+$$

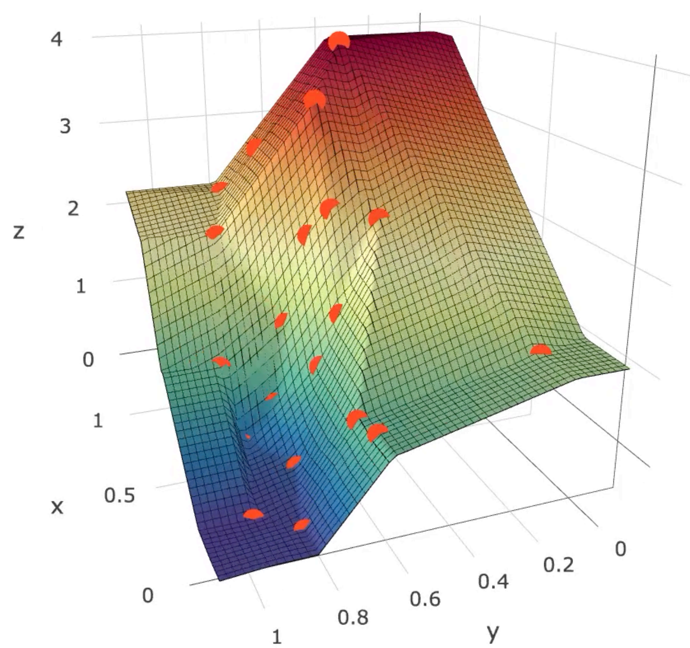


Piecewise Linear Approximations

Delaunay
(interpolant)

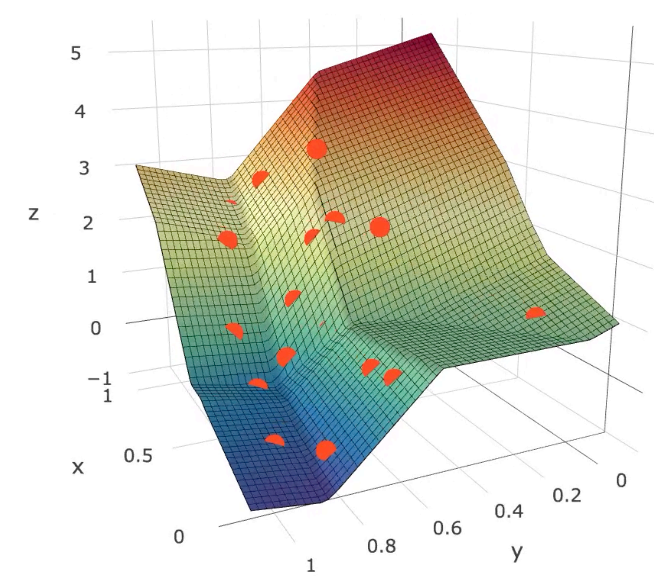
$$y = \sum_{i=0}^{d} w_i x^{(i)}, \quad \sum_{i=0}^{d} w_i = 1, \quad w_i \geq 0, \quad i = 0, \ldots, d$$

$$\hat{f}(y) = \sum_{i=0}^{d} w_i f\left(x^{(i)}\right)$$



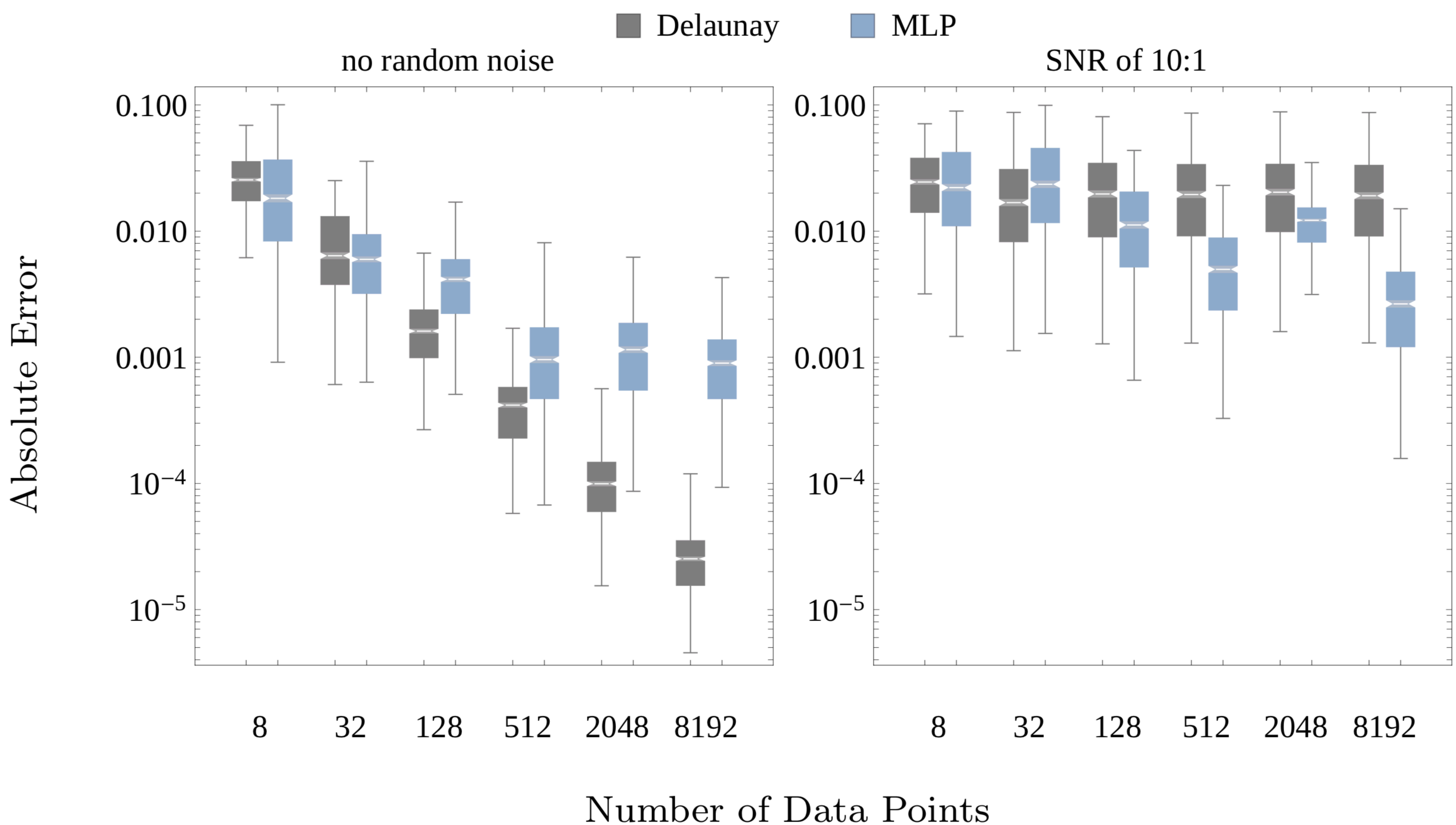
Multilayer Perceptron
(regressor)

$$l(u) = \left(u^t W_l + c_l\right)_+$$



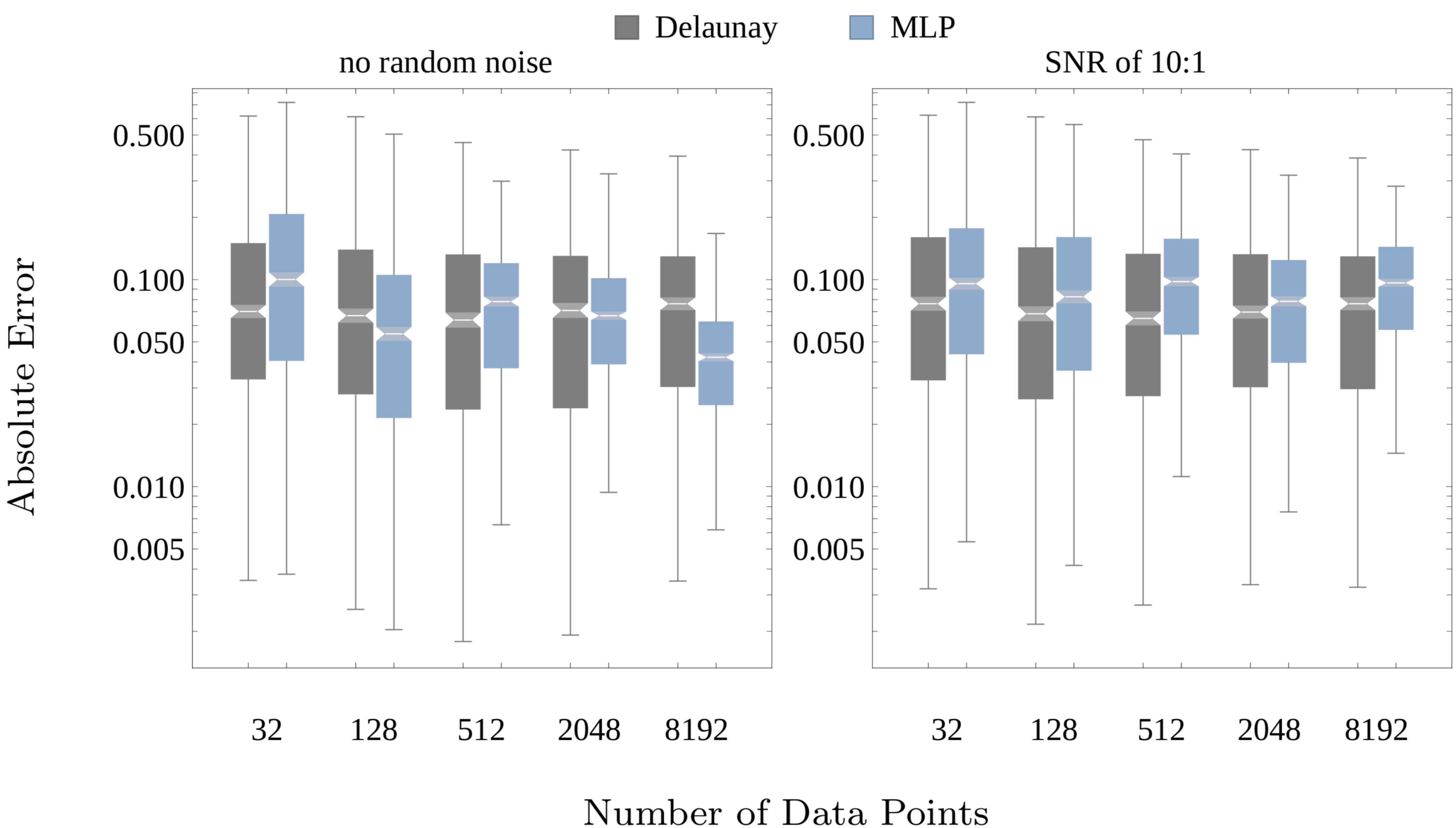
Approximating $f(x) = \cos(\|x\|_2)$

In 2 dimensions, we get expected results.
Delaunay is better at interpolation, MLP better at regression.

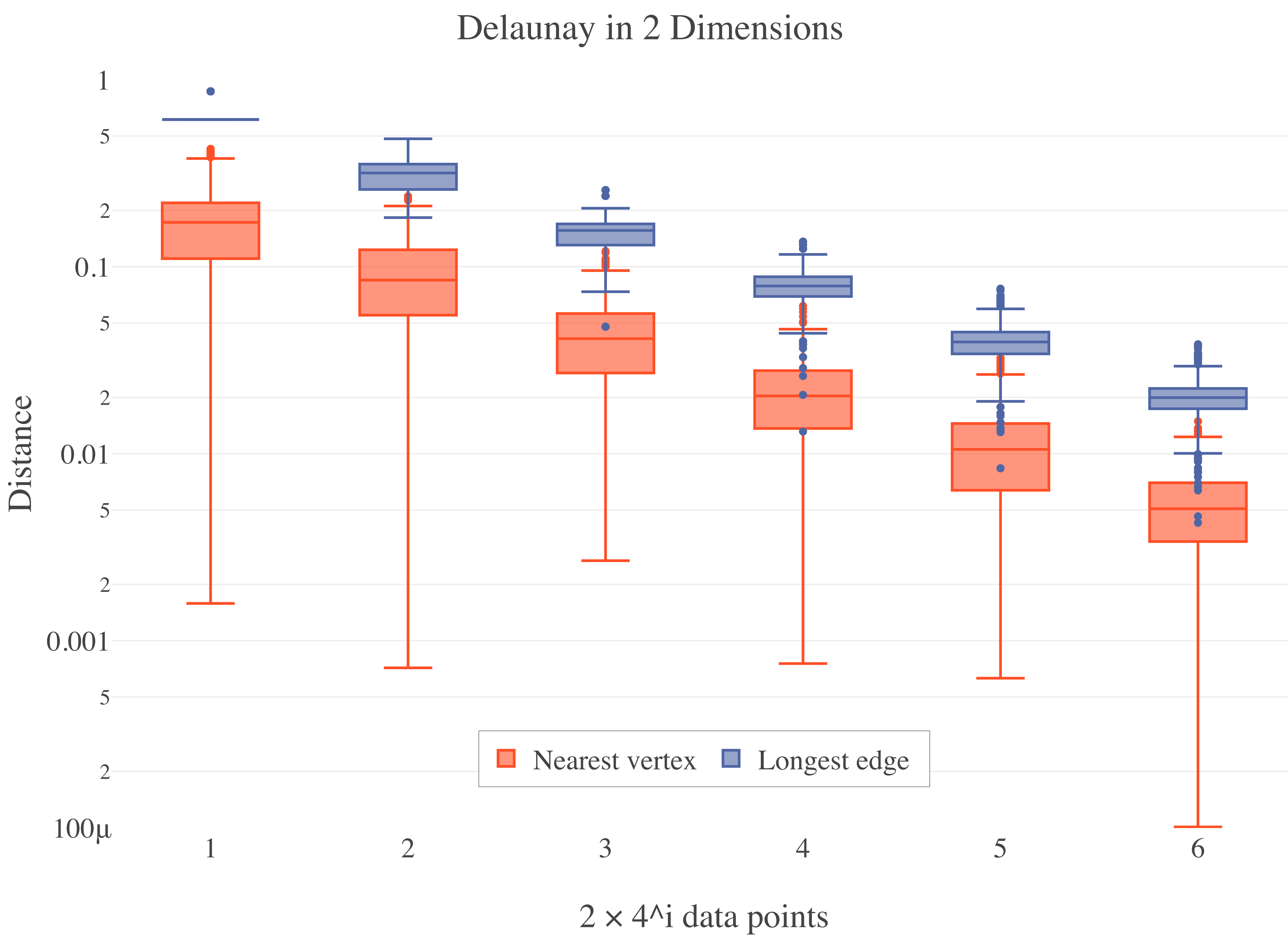


Approximating $f(x) = \cos(\|x\|_2)$

In 20 dimensions, the *intuitive* trend disappears!
Delaunay and MLP look the same.

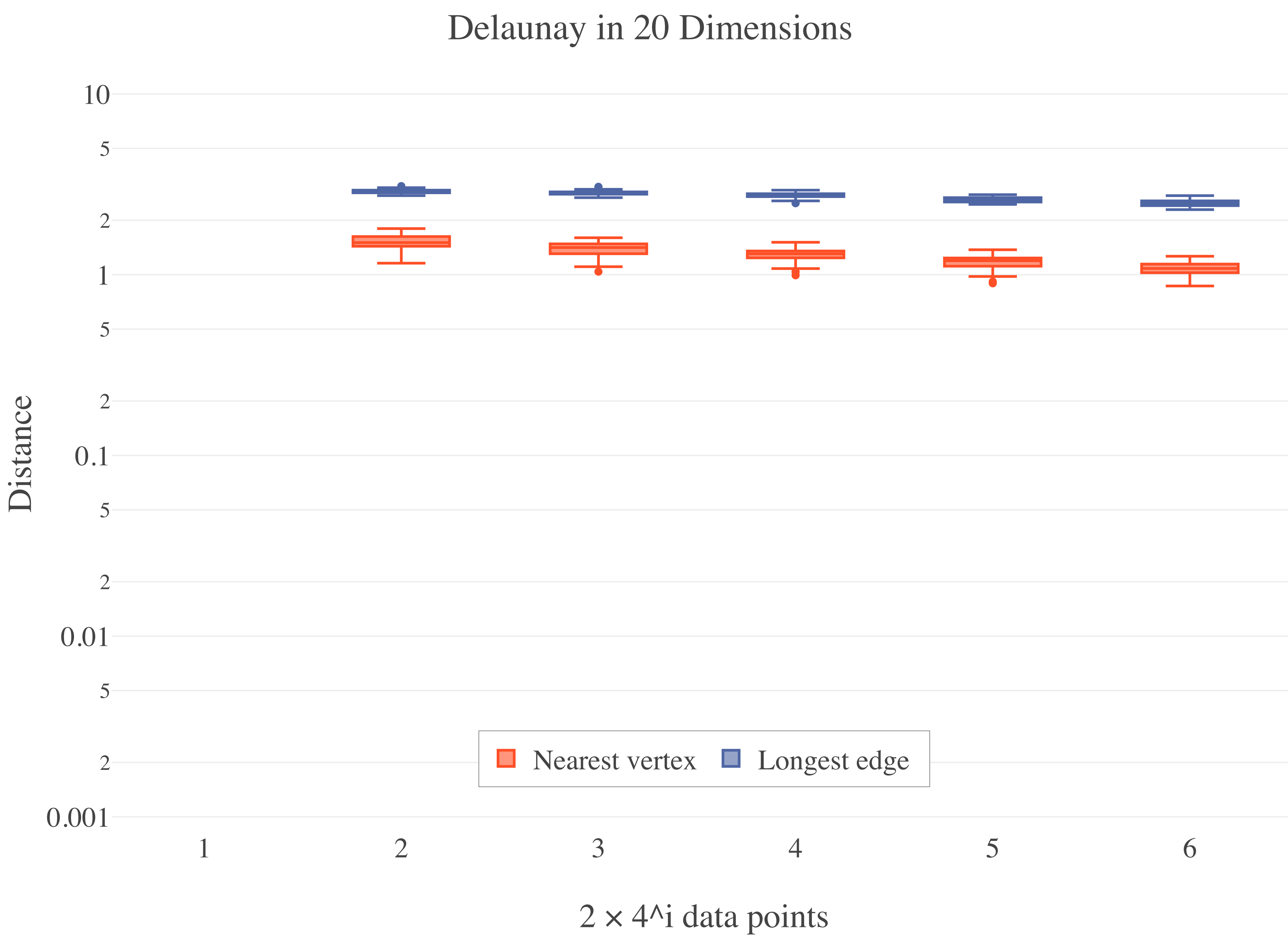


Why does it converge in 2D?



COMPUTER SCIENCE
VIRGINIA TECH.

Why doesn't it converge in 20D?



COMPUTER SCIENCE
VIRGINIA TECH.

Connecting Back to the Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d} \|z - x_0\|_2$$

Connecting Back to the Theory

$$\left|f(z)-\hat{f}(z)\right| \le \frac{\gamma\|z-x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z-x_0\|_2$$

the average closest
point doesn't get
much closer and …

Connecting Back to the Theory

$$\left|f(z) - \hat{f}(z)\right| \leq \frac{\gamma\|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d}\|z - x_0\|_2$$

the average closest point doesn't get much closer and …

the longest edge does not meaningfully shrink in 20 dimensions with thousands of points

Connecting Back to the Theory

$$\left| f(z) - \hat{f}(z) \right| \leq \frac{\gamma \|z - x_0\|_2^2}{2} + \frac{\sqrt{d}\,\gamma k^2}{2\sigma_d} \|z - x_0\|_2$$

the average closest
point doesn't get
much closer and …

the longest edge does
not meaningfully shrink
in 20 dimensions with
thousands of points

Goals

Incorporate additional comparison between regressors and interpolants with analytic function to test theoretical bound.

Produce a mathematical software for the computation of monotone quintic spline interpolants.

Apply this technique as an improved CDF estimator for the VarSys distribution approximation test problems.

Extend empirical VarSys test, write a journal paper, and submit as a mathematical software package to TOMS.

Monotone Piecewise Quintic Splines

Theory exists to create splines that are composed of quintics,
but it has not been done in 30 years … **because it's difficult to do.**

G. Ulrich and L. Watson. Positivity conditions for quartic polynomials. SIAM Journal on Scientific Computing, 15(3):528–544, 1994. doi: 10.1137/0915035.

Walter Hess and Jochen W Schmidt. Positive quartic, monotone quintic c2-spline interpolation in one and two dimensions. Journal of Computational and Applied Mathematics, 55(1): 51–67, 1994. doi: 10.1016/0377-0427(94)90184-8.

Dougherty, Randall L., Alan S. Edelman, and James M. Hyman. Nonnegativity-, monotonicity-, or convexity-preserving cubic and quintic Hermite interpolation. Mathematics of Computation 52.186 (1989): 471-494.

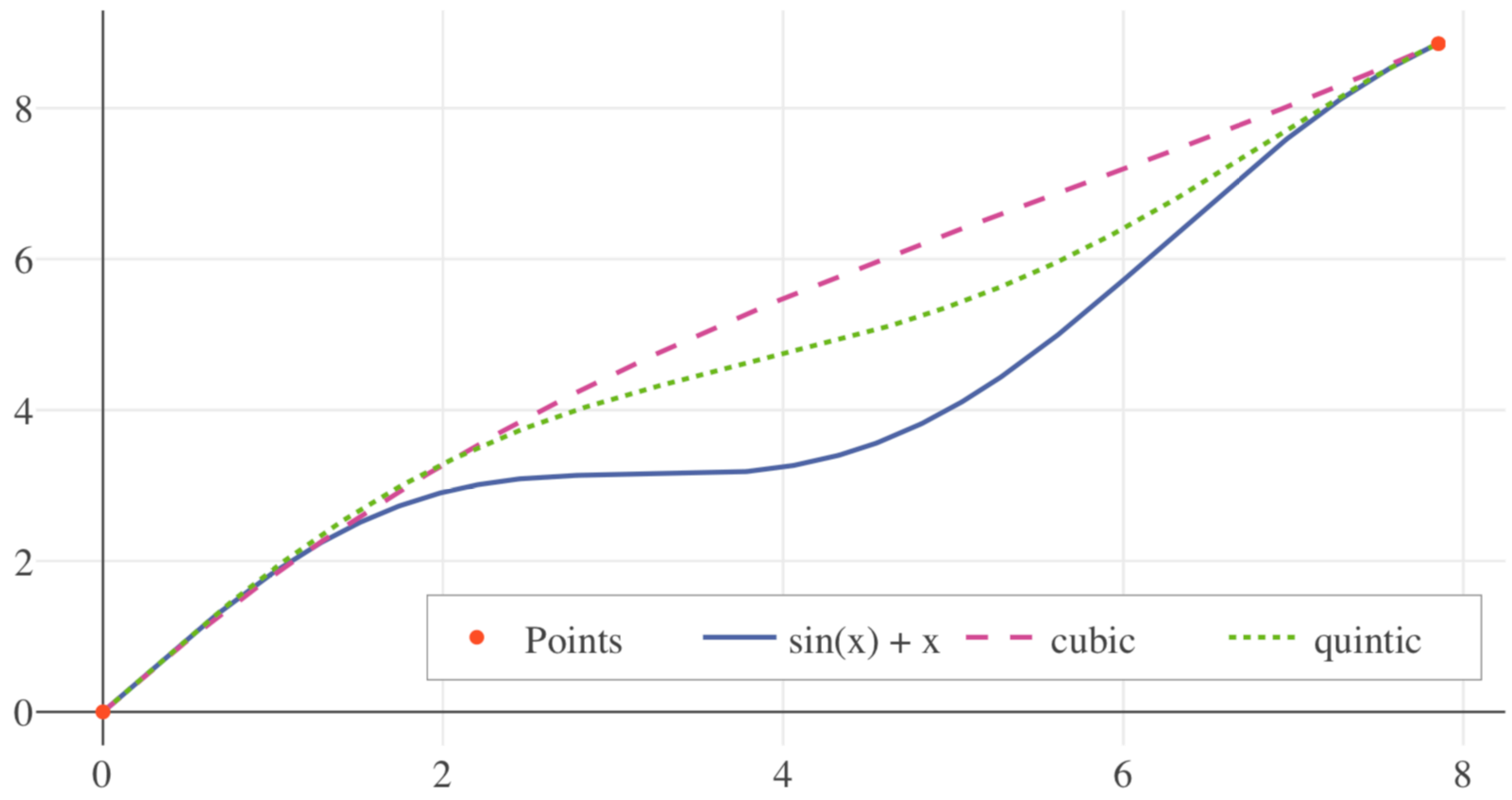
The Algorithms Presented

Algorithm 1: `is_monotone`(x_0, x_1, f)
where x_0, $x_1 \in \mathbb{R}$, $x_0 < x_1$, and f is an order six polynomial defined by $f(x_0)$, $Df(x_0)$, $D^2 f(x_0)$, $f(x_1)$, $Df(x_1)$, $D^2 f(x_1)$. Returns `TRUE` if f is monotone increasing on $[x_0, x_1]$.

Algorithm 2: `make_monotone`(x_0, x_1, f)
where x_0, $x_1 \in \mathbb{R}$, $x_0 < x_1$, and f is an order six polynomial defined by $f(x_0)$, $Df(x_0)$, $D^2 f(x_0)$, $f(x_1)$, $Df(x_1)$, $D^2 f(x_1)$. Returns f monotone on $[x_0, x_1]$.

Algorithm 3: `monotone_spline`$((k_1, \ldots, k_n), f, s)$
where $(k_1, \ldots, k_n)$ is an increasing sequence of real numbers, f is an order six piecewise polynomial with breakpoints $k_1, \ldots, k_n$ defined by the data $\{f(k_i)\}_{i=1}^n$, $\{Df(k_i)\}_{i=1}^n$, $\{D^2 f(k_i)\}_{i=1}^n$, and $s > 1$ is an integer shrink factor.

A Test Approximation

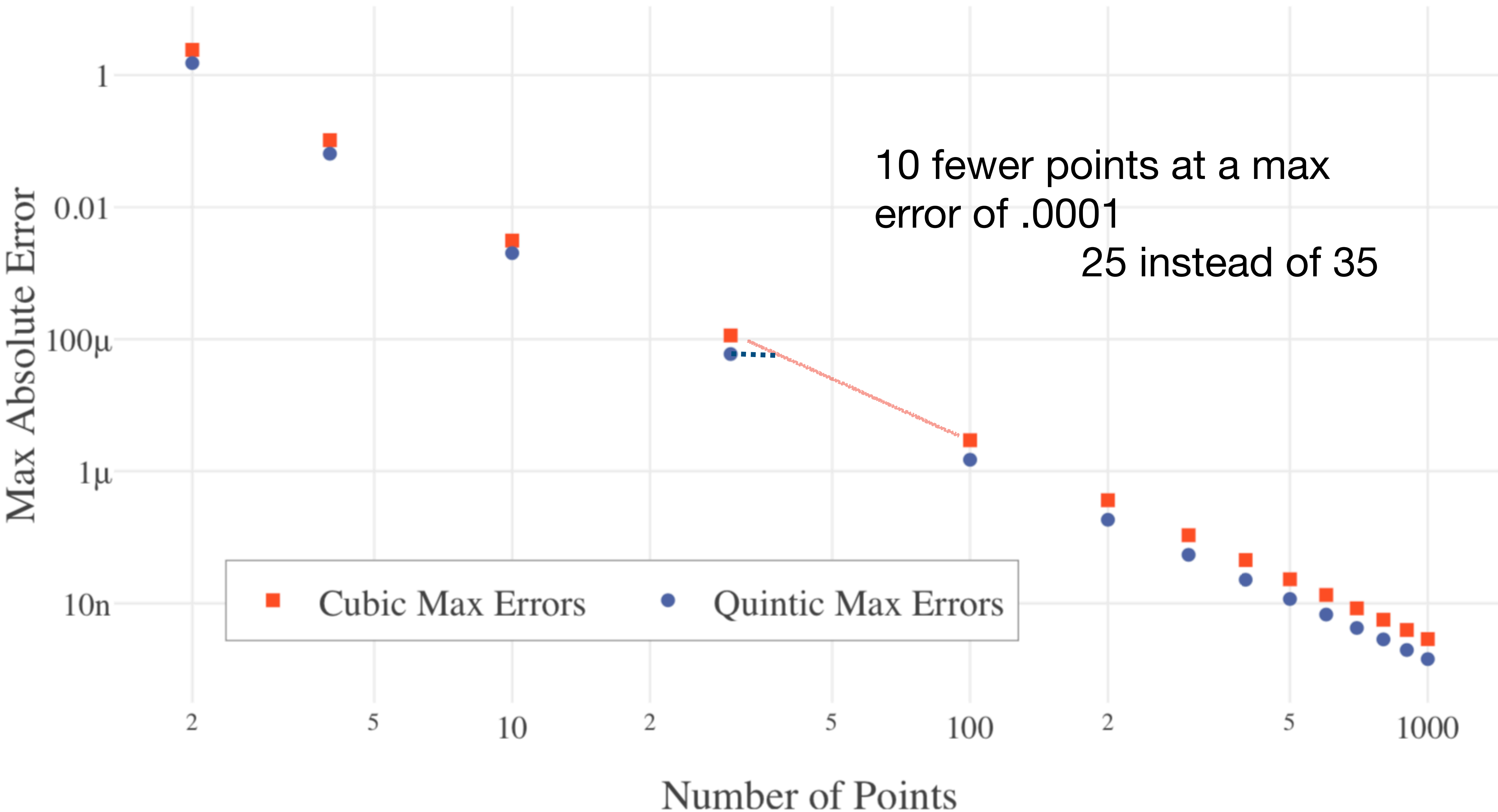


COMPUTER SCIENCE
VIRGINIA TECH.

The Maximum Error with More Data

COMPUTER SCIENCE
VIRGINIA TECH

The Maximum Error with More Data



COMPUTER SCIENCE
VIRGINIA TECH.

The Maximum Error with More Data

30 fewer points at
a max error of 10^{-6}
100 instead of 130

COMPUTER SCIENCE
VIRGINIA TECH

The Maximum Error with More Data

200 fewer points at a max error of 10^{-8}

500 instead of 700

COMPUTER SCIENCE
VIRGINIA TECH.

Goals

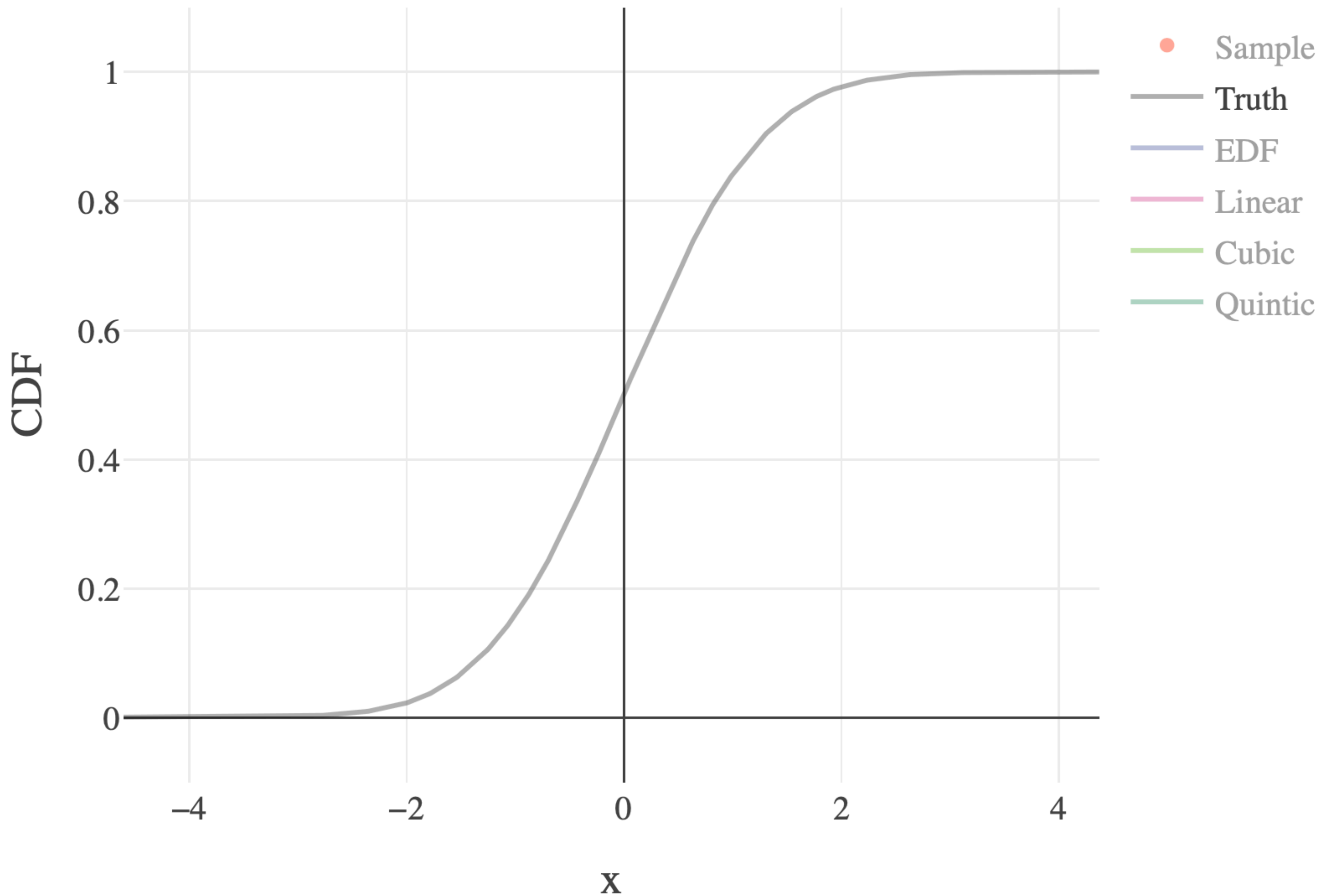
Incorporate additional comparison between regressors and interpolants with analytic function to test theoretical bound.

Produce a mathematical software for the computation of monotone quintic spline interpolants.

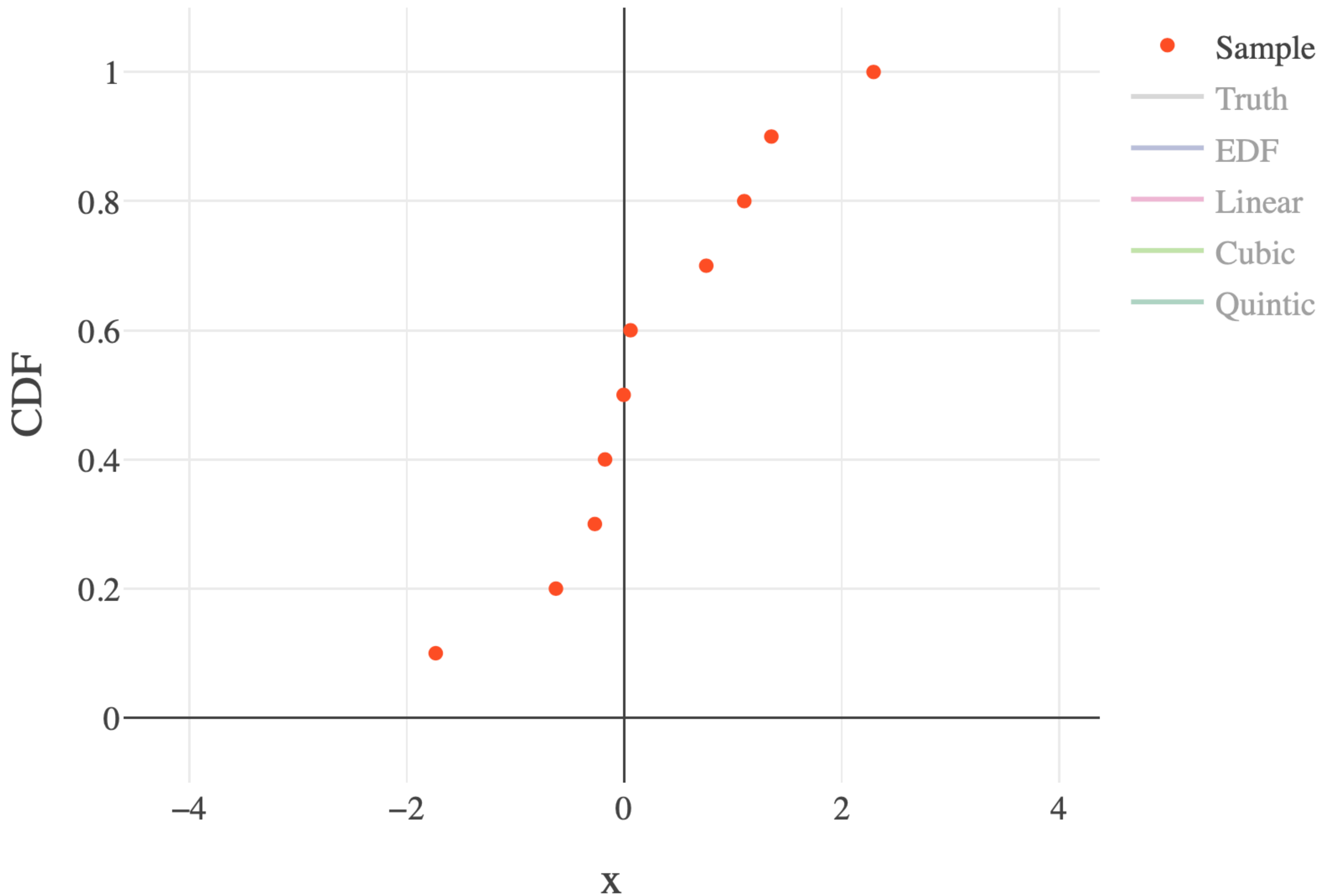
Apply this technique as an improved CDF estimator for the VarSys distribution approximation test problems.

Extend empirical VarSys test, write a journal paper, and submit as a mathematical software package to TOMS.

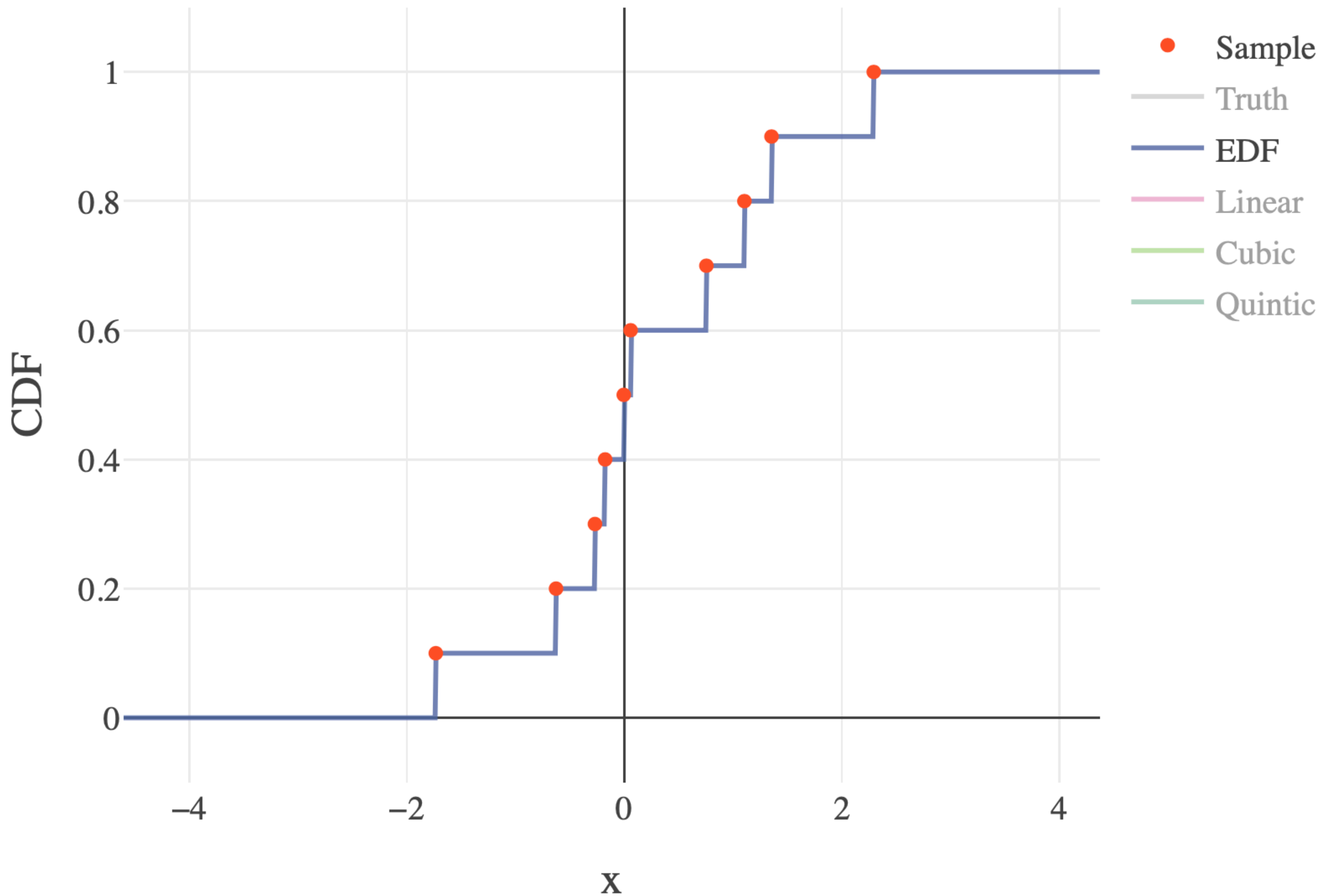
Empirical Distribution Approximations



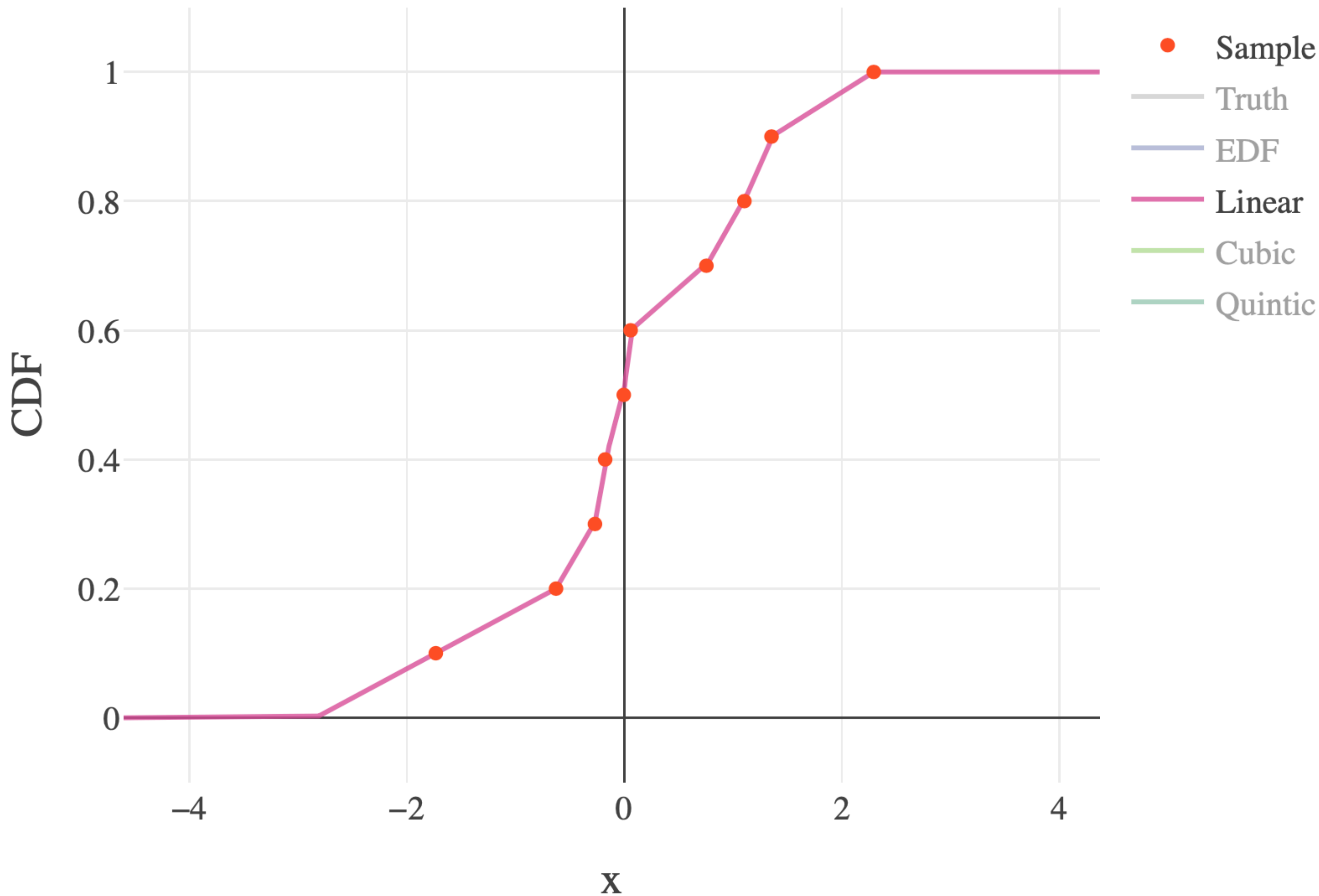
Empirical Distribution Approximations



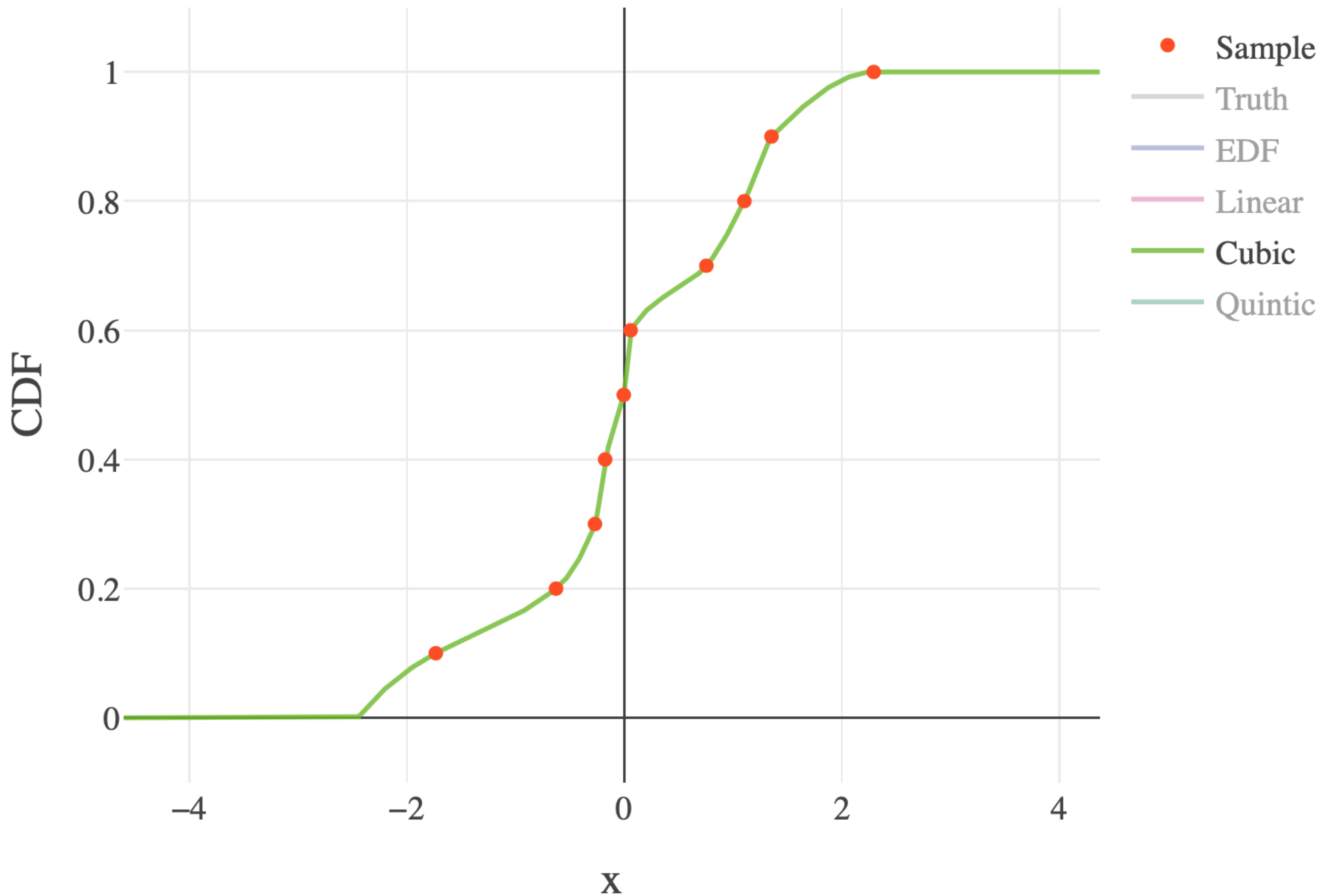
Empirical Distribution Approximations



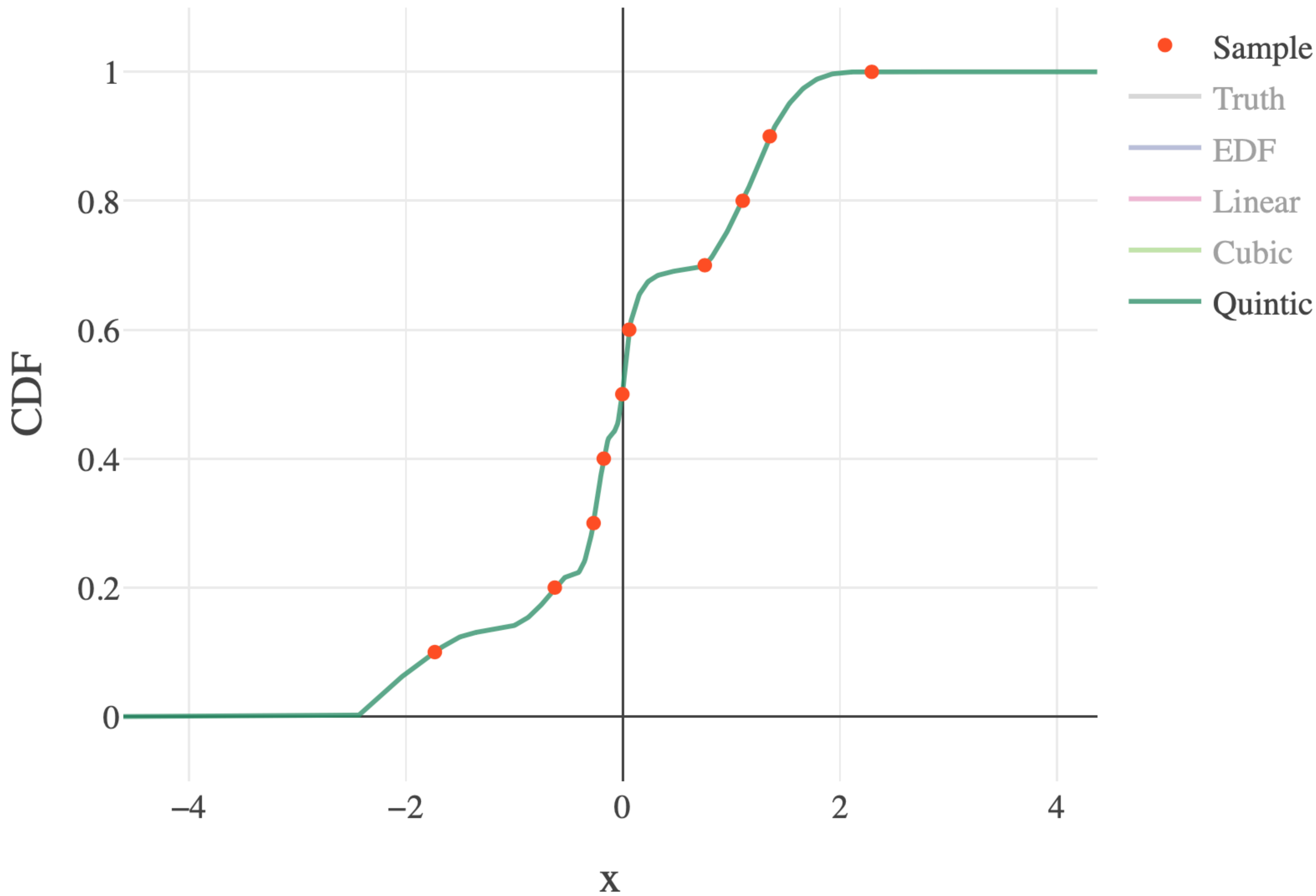
Empirical Distribution Approximations



Empirical Distribution Approximations

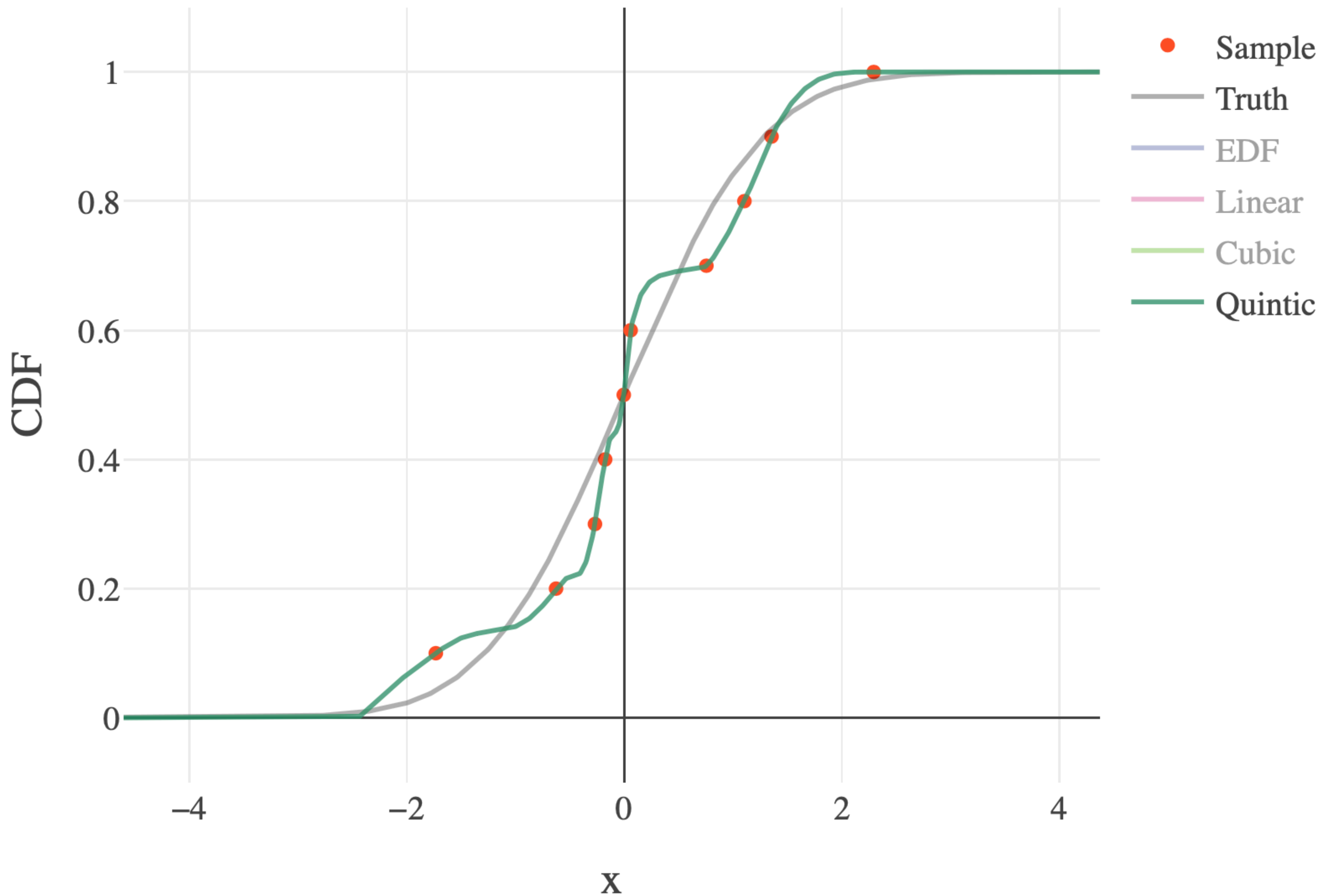


Empirical Distribution Approximations



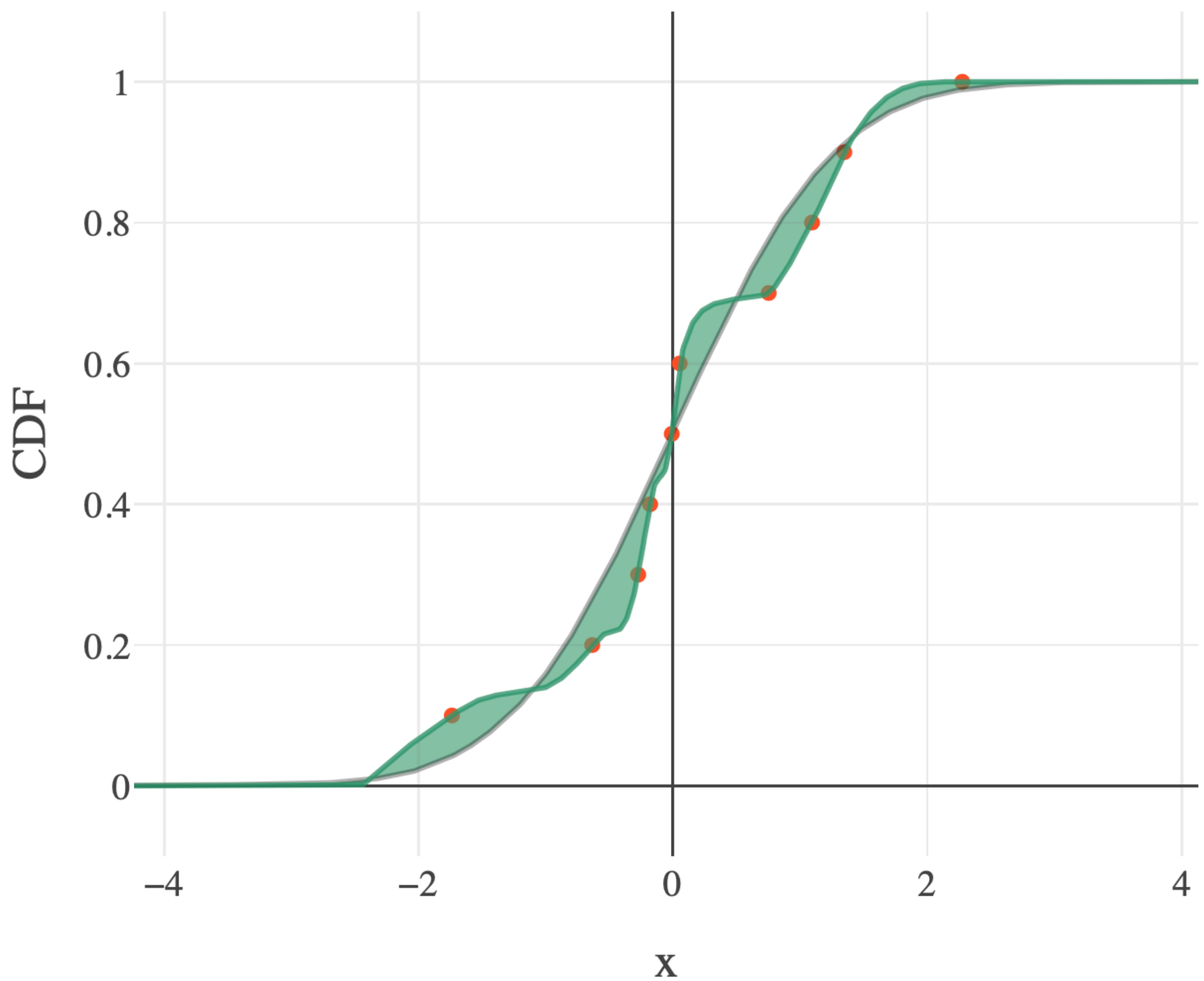
COMPUTER SCIENCE
VIRGINIA TECH.

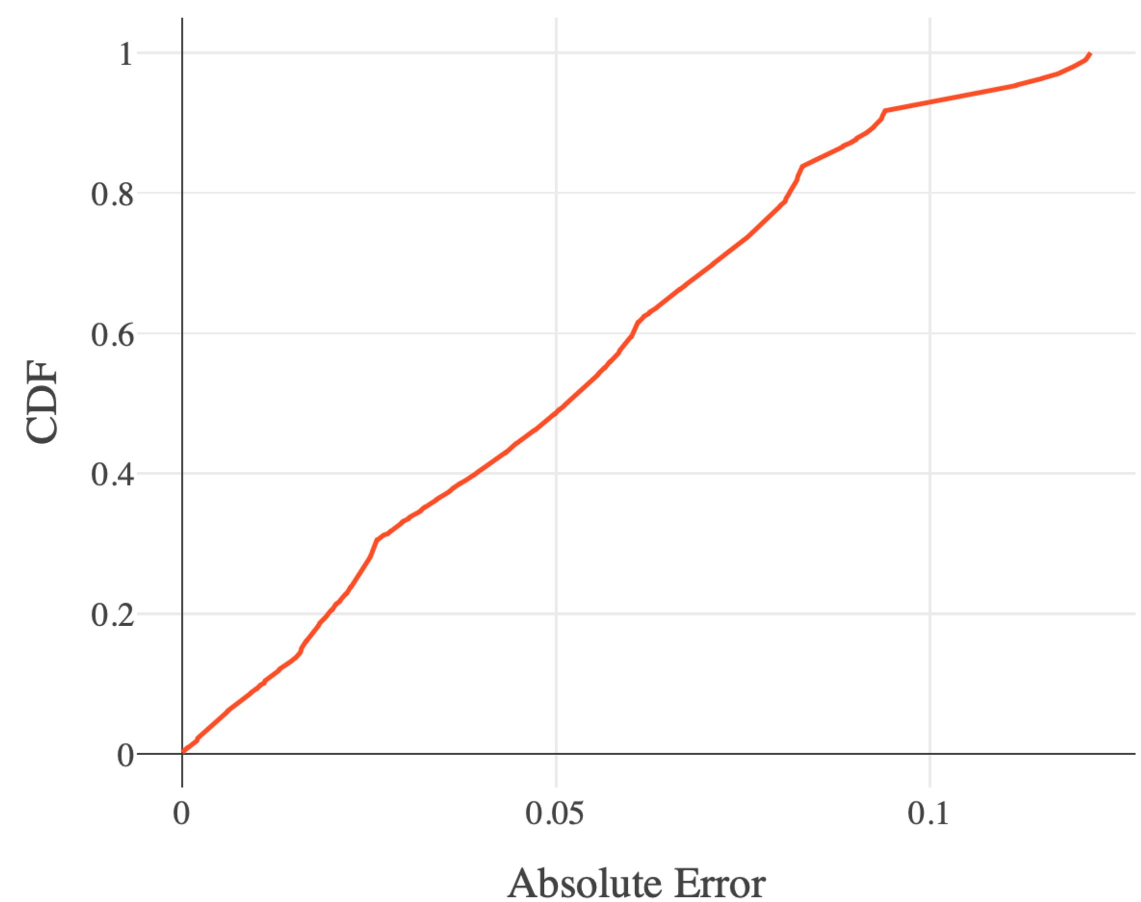
Empirical Distribution Approximations



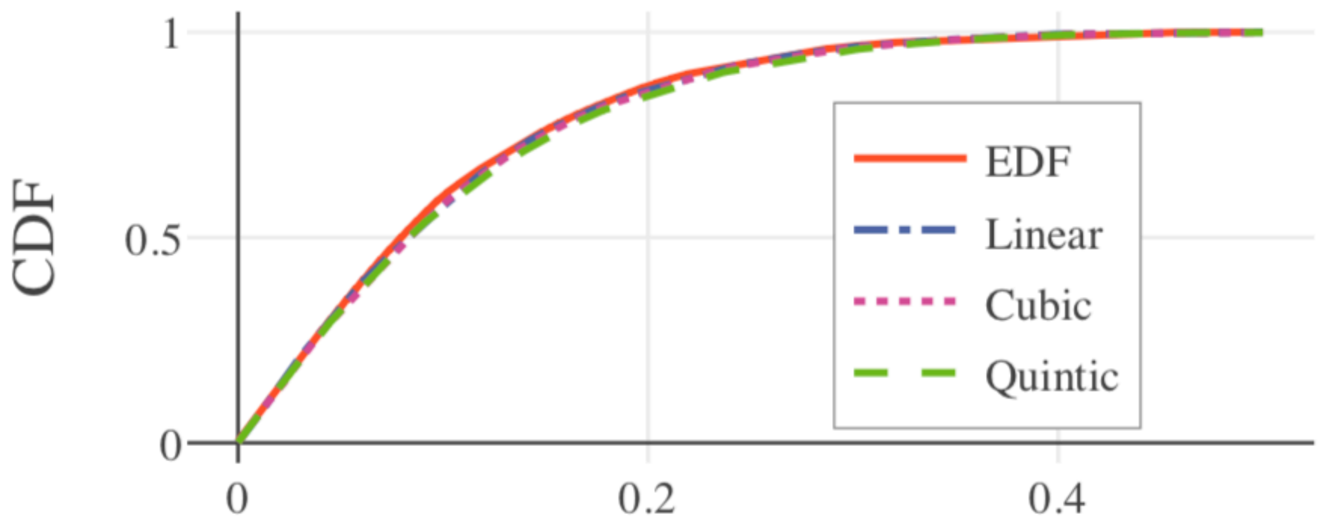
COMPUTER SCIENCE
VIRGINIA TECH.

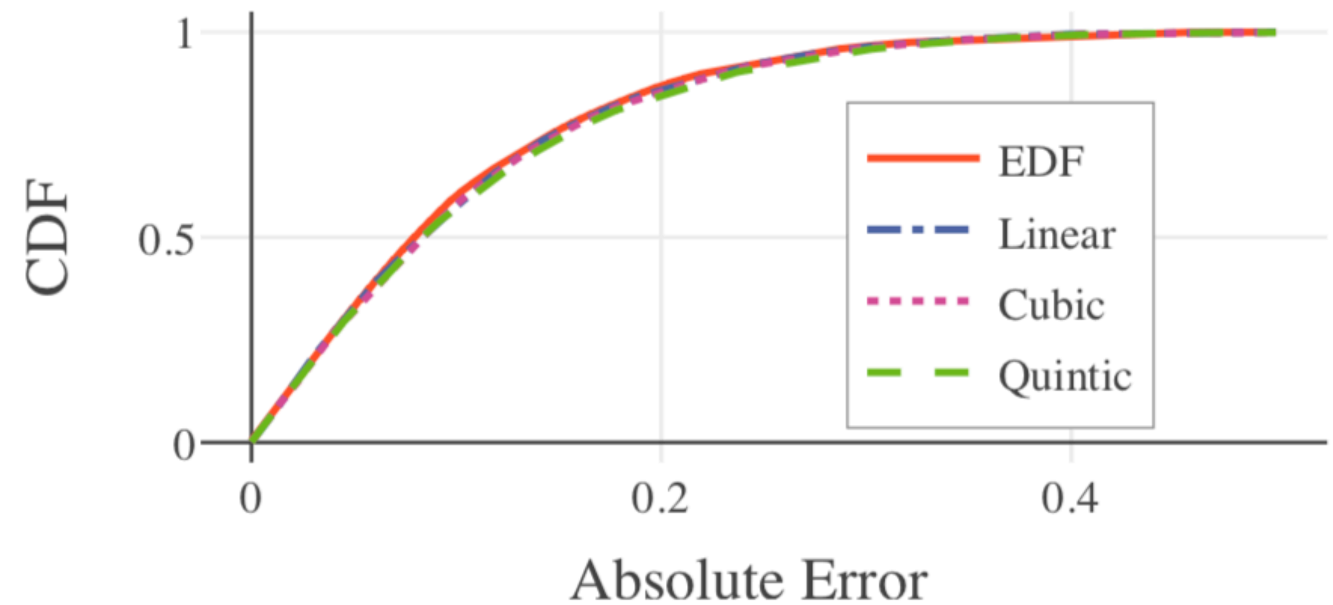
Error of a Distribution Approximation

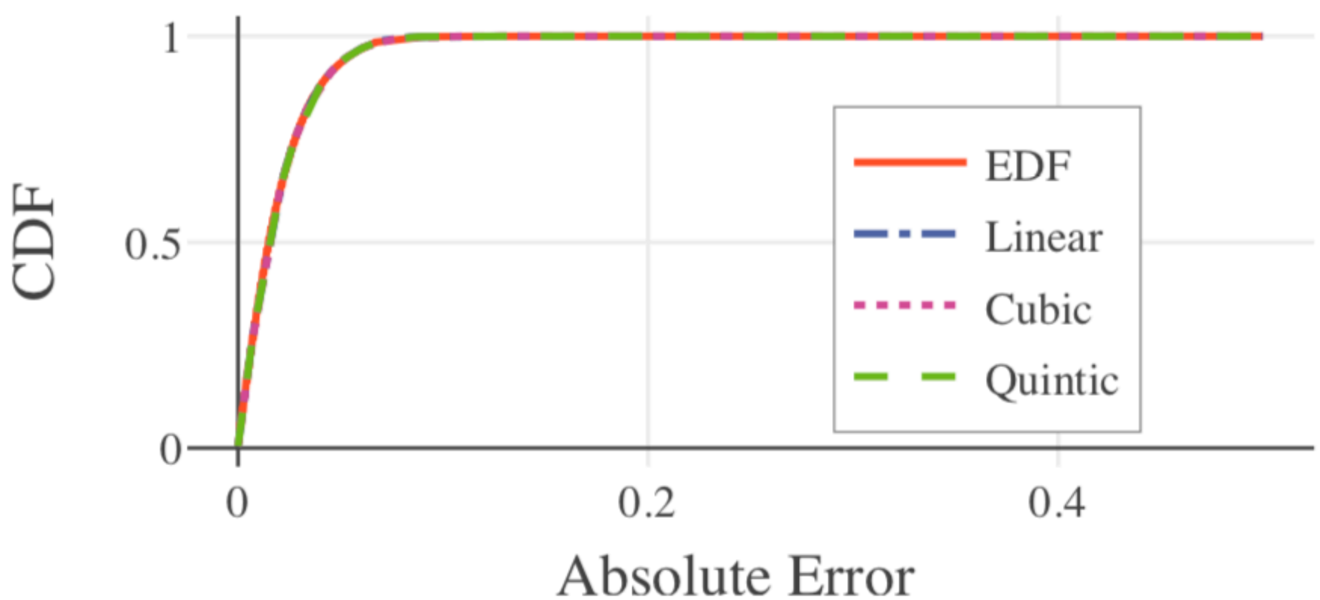




The Errors when Approximating True Distributions with All Techniques

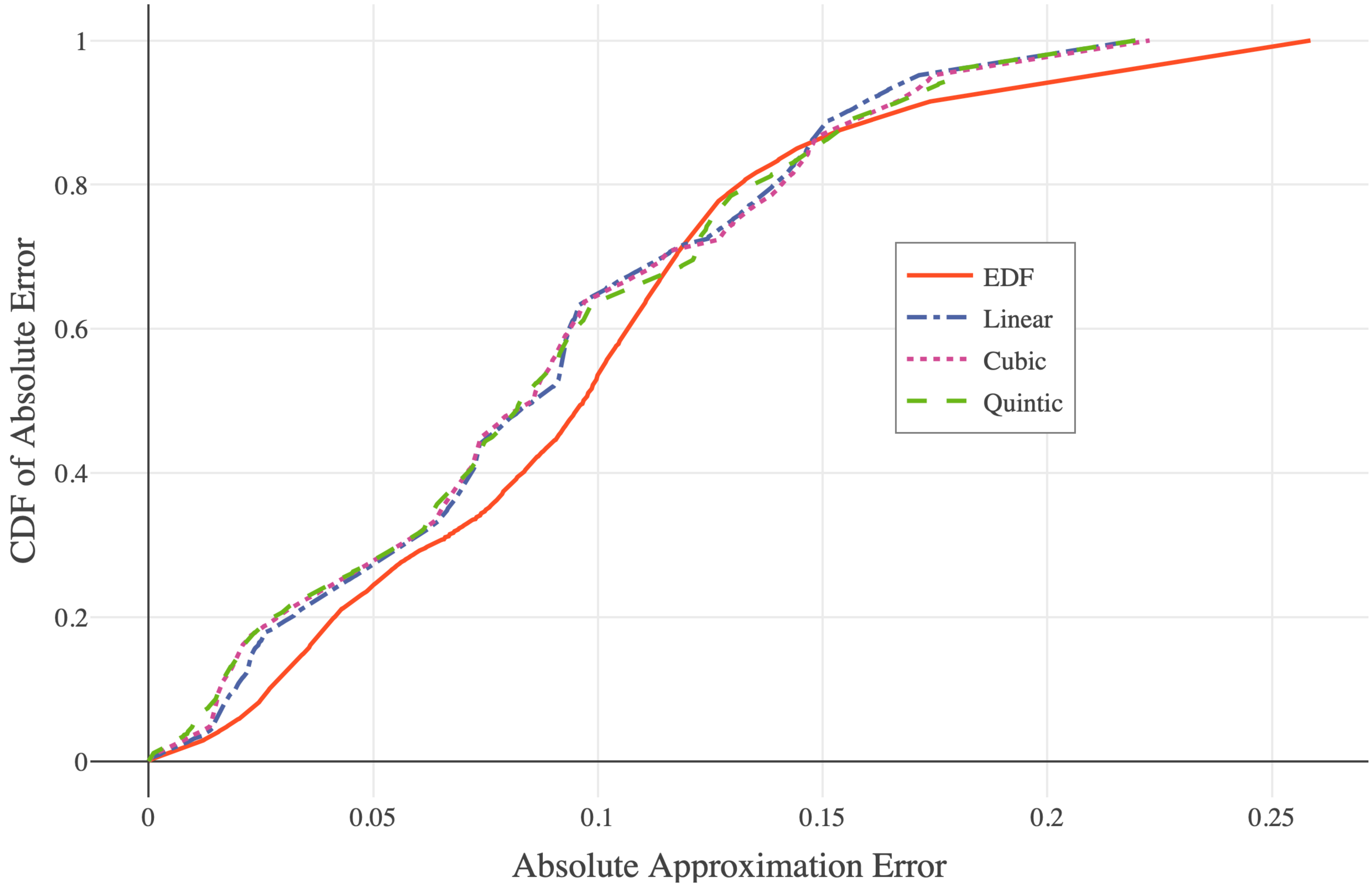






Given 10, 50, and 200 samples respectively, the absolute errors in the approximated distribution look the same for the classic EDF as well as linear, cubic, and quintic interpolants.

The Quintic is Often Better for Tested VarSys Data



Last Goal Before Final Defense

Incorporate additional comparison between regressors and interpolants with analytic function to test theoretical bound.

Produce a mathematical software for the computation of monotone quintic spline interpolants.

Apply this technique as an improved CDF estimator for the VarSys distribution approximation test problems.

Extend empirical VarSys test, write a journal paper, and submit as a mathematical software package to TOMS.