Computing the Tracy–Widom Distribution for Arbitrary $\beta > 0$

Thomas TROGDON and Yiting ZHANG

Department of Applied Mathematics, University of Washington, Seattle, Washington, USA
E-mail: *trogdon@uw.edu, yitinz91@uw.edu*

Abstract. We compute the Tracy–Widom distribution describing the asymptotic distribution of the largest eigenvalue of a large random matrix by solving a boundary-value problem posed by Bloemendal in his Ph.D. Thesis (2011). The distribution is computed in two ways. The first method is a second-order finite-difference method and the second is a highly accurate Fourier spectral method. Since β is simply a parameter in the boundary-value problem, any $\beta > 0$ can be used, in principle. The limiting distribution of the nth largest eigenvalue can also be computed. Our methods are available in the JULIA package `TracyWidomBeta.jl`.

Key words: numerical differential equation; Tracy–Widom distribution; Fourier transformation

2020 Mathematics Subject Classification: 65M06; 60B20; 60H25

1   Introduction

Tracy and Widom [23, 24, 25] introduced the Tracy–Widom distribution that gives the limiting distribution of the rescaled largest eigenvalue of a random matrix taken from an appropriate symmetry class. More precisely, the largest eigenvalue $\lambda_{\max}$ satisfies the following fundamental limit

$$\lim_{n \to \infty} \mathbb{P}\big(n^{1/6}(\lambda_{\max}(A_n) - 2\sqrt{n}) \leq x\big) = F_\beta(x) \qquad \text{for all } x \in \mathbb{R},$$

where F_β is the Tracy–Widom distribution and $\beta = 1$, 2, 4 if $A_n \sim$ Gaussian orthogonal ensemble, Gaussian unitary ensemble, Gaussian symplectic ensemble, respectively.

For an $n \times n$ Gaussian ensemble A_n with ordered eigenvalues $\{\lambda_j\}_{j=1}^n$, the joint probability density function (jpdf) for its eigenvalues is given by

$$\rho(\lambda_1, \lambda_2, \ldots, \lambda_n) = Z_{n,\beta}^{-1} \prod_i \mathrm{e}^{-\beta \lambda_i^2/4} \prod_{i<j} |\lambda_j - \lambda_i|^\beta, \qquad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n, \tag{1.1}$$

where $Z_{n,\beta}$ is the partition function. For $\beta = 1, 2, 4$, (1.1) is solvable: all correlation functions in finite dimensions can be explicitly calculated using Hermite polynomials. This provides definitive local limit theorems and establishes clear limits for the random points [15]. Importantly, (1.1) also describes a one-dimensional Coulomb gas at inverse temperature β for any $\beta > 0$. If $\beta \neq 1, 2, 4, 6$, there are no known explicit formulae which appear amenable to asymptotic analysis (see, for example, [10, 13, 20] for $\beta = 6$). One, in general, must resort to numerical computations, see [3] for $\beta = 1, 2, 4$.

This paper is a contribution to the Special Issue on Evolution Equations, Exactly Solvable Models and Random Matrices in honor of Alexander Its' 70th birthday. The full collection is available at https://www.emis.de/journals/SIGMA/Its.html

Dumitriu and Edelman [7] established that a family of symmetric tridiagonal (Jacobi) matrix models have (1.1) as their eigenvalue density for any $\beta > 0$. More specifically, the matrix given by

$$H_n^\beta = \frac{1}{\sqrt{\beta}} \begin{bmatrix} g & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & g & \chi_{(n-2)\beta} & & \\ & \chi_{(n-2)\beta} & g & \ddots & \\ & & \ddots & \ddots & \chi_\beta \\ & & & \chi_\beta & g \end{bmatrix}, \tag{1.2}$$

has (1.1) as the jpdf of its eigenvalues. Here the entries are independent random variables, up to symmetry, $g \sim N(0,2)$ and $\chi_k \sim \text{Chi}(k)$. We call H_n^β the β-Hermite ensemble.

Sutton and Edelman [8, 21] then presented an argument[1] describing how the spectrum of the rescaled operator

$$\tilde{H}_n^\beta = n^{1/6}\big(2\sqrt{n}I - H_n^\beta\big),$$

where I is the $n \times n$ identity matrix, should be described by the spectrum of the stochastic Airy operator

$$\mathcal{H}_\beta = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x + \frac{2}{\sqrt{\beta}}b_x',$$

as $n \to \infty$, where b_x' is standard Gaussian white noise.

As a result, the following eigenvalue problem was considered in [19]

$$\mathcal{H}_\beta f = \Lambda f \qquad \text{on } L^2(\mathbb{R}_+)$$

with a Dirichlet boundary condition $f(0) = 0$. Ramírez, Rider, and Virág [19] proved the following theorem.

Theorem 1.1. *With probability one, for each $k \geq 0$, the set of eigenvalues of $\mathcal{H}_\beta$ has a well-defined $(k+1)$st lowest element Λ_k. Moreover, let $\lambda_1 \geq \lambda_2 \geq \cdots$ denote the eigenvalues of H_n^β. Then the vector*

$$\big(n^{1/6}(2\sqrt{n} - \lambda_l)\big)_{l=1,\ldots,k}$$

converges in distribution to $(\Lambda_0, \Lambda_1, \ldots, \Lambda_{k-1})$ as $n \to \infty$.

Ramírez, Rider, and Virág showed that the distribution of $-\Lambda_0$ is a consistent definition of Tracy–Widom($\cdot$) for general $\beta > 0$. Based on [19, 26], Bloemendal and Viraǵ [1, 2] considered a generalized eigenvalue problem

$$\mathcal{H}_\beta f = \Lambda f \qquad \text{on } L^2(\mathbb{R}_+)$$

with boundary condition $f'(0) = \omega f(0)$, where $\omega \in \mathbb{R}$ represents a scaling parameter. Let $\mathcal{H}_{\beta,\omega}$ denote $\mathcal{H}_\beta$ together with this boundary condition. According to [19], the distribution $F_{\beta,\omega}$ of $-\Lambda_0$ in the case $\omega = +\infty$ thus coincides with Tracy–Widom for general $\beta > 0$.

Bloemendal and Virág [2] further showed that $F = F_{\beta,\omega}$ can be characterized using the solution of a boundary value problem, as we now discuss in the following section. We also point out that Bloemendal [1] provided MATHEMATICA code to approximate F. The goal of the current work is to expand and improve upon this scheme.

This paper is laid out as follows. In Section 2, we outline our two algorithms to compute the Tracy–Widom distribution. In Section 3, we validate and compare our methods. In Section 4, we present a number of additional numerical results. We also include two appendices to discuss some nuances in the numerical algorithms (Appendix A) and to discuss the large β limit (Appendix B). Code to produce all figures in this paper can be found here [27].

---

[1] A version of this argument was first presented by Edelman at the SIAM Conference on Applied Linear Algebra held at the College of William & Mary in 2003.

2 Algorithm description

The Tracy–Widom distribution function F_β can be characterized as follows [1]. Consider

$$\frac{\partial F}{\partial x} + \frac{2}{\beta}\frac{\partial^2 F}{\partial \omega^2} + \left(x - \omega^2\right)\frac{\partial F}{\partial \omega} = 0 \qquad \text{for } (x,\omega) \in \mathbb{R}^2, \tag{2.1}$$

with boundary conditions given by

$$F(x,\omega) \to 1 \qquad \text{as } x,\omega \to \infty \text{ together,}$$
$$F(x,\omega) \to 0 \qquad \text{as } \omega \to -\infty \text{ with} x \text{ bounded above.}$$

Then [2, Theorem 1.7]

$$F_\beta(x) = \lim_{\omega \to \infty} F(x,\omega). \tag{2.2}$$

Using $\omega = -\cot\theta$, we rewrite (2.1) as

$$\frac{\partial H}{\partial x} + \left(\frac{2}{\beta}\sin^4\theta\right)\frac{\partial^2 H}{\partial \theta^2} + \left[\left(x + \frac{2}{\beta}\sin 2\theta\right)\sin^2\theta - \cos^2\theta\right]\frac{\partial H}{\partial \theta} = 0, \tag{2.3}$$

with boundary condition

$$H(x,0) = 0.$$

To address the boundary condition as $x \to \infty$ and $\theta \to \pi$, we truncate the domain at a finite value, denoted as $x = x_0 > 0$. We then employ the Gaussian asymptotics established by Bloemendal [1, Theorem 4.1.1] to obtain the following approximate asymptotic initial condition [1, p. 103]:

$$H(x_0,\theta) = \begin{cases} \Phi\left(\dfrac{x_0 - \cot^2\theta}{\sqrt{(4/\beta)\cot\theta}}\right), & 0 \le \theta \le \pi/2, \\ 1, & \pi/2 \le \theta. \end{cases} \tag{2.4}$$

Here Φ denotes the standard normal distribution function. Note that $H(x_0,\theta)$ is continuous in both $x_0 > 0$ and $\theta \ge 0$. One then approximates the undeformed Tracy–Widom distribution $F_\beta(x)$ at $\theta = \pi$ by $\omega = -\cot\theta$ and (2.2), i.e., $F_\beta(x) \approx H(x,\pi)$.

2.1 Finite-difference discretization

One way to solve this boundary-value problem is by discretizing it using finite differences, see, for example, [12]. For $0 \le n \le N$, $0 \le m \le M$, define

$$x_n = x_0 + n\Delta x, \qquad \theta_m = mh.$$

Given that $\beta > 0$, one integrates equation (2.3) backward in "time" with respect to the time-like variable x to guarantee its well-posedness: $\Delta x < 0$. Denote the approximation of $H(x,\theta_m)$ by $H_m(x)$. We then replace partial derivatives of H with respect to θ by centered differences, with grid spacing h. The method of lines formulation reads

$$\frac{\partial \boldsymbol{H}_M}{\partial x} = -\left(\frac{2\sin^4\theta}{\beta h^2}\right)\check{T}\boldsymbol{H}_M + \frac{\left[\left(x + \frac{2}{\beta}\sin 2\theta\right)\sin^2\theta - \cos^2\theta\right]}{2h}\check{U}\boldsymbol{H}_M, \tag{2.5}$$

where $\boldsymbol{H}_M(x) = [H_1(x), H_2(x), \dots, H_M(x)]^\mathsf{T}$, $Mh = \theta_M$, and

$$
\check{T} = \begin{bmatrix}
-2 & 1 & & & & \\
1 & -2 & 1 & & & \\
& 1 & -2 & 1 & & \\
& & \ddots & \ddots & \ddots & \\
& & & 1 & -2 & 1 \\
& & & & 1 & -2
\end{bmatrix}, \qquad
\check{U} = \begin{bmatrix}
0 & -1 & & & & \\
1 & 0 & -1 & & & \\
& 1 & 0 & -1 & & \\
& & \ddots & \ddots & \ddots & \\
& & & 1 & 0 & -1 \\
& & & & -1 & 4 & -3
\end{bmatrix}. \tag{2.6}
$$

$\boldsymbol{H}_M$ excludes H_0 since $H_0(x) = 0$. The coefficients in the last row of $\check{U}$ come from the parameters of the two-step backward difference formula (BDF2). Note that there is no need to replace the last row of $\check{T}$ using a backward difference formula since $\theta_M = \pi$ and, mathematically, there is no need to set an additional boundary condition since (2.5) has vanishing diffusivity for $\theta = \pi$.

Let

$$
T(\beta, \boldsymbol{\theta}_M, h) := -\left(\frac{2\sin^4 \boldsymbol{\theta}_M}{\beta h^2}\right)\check{T} = \begin{bmatrix}
\dfrac{-2\sin^4 \theta_1}{\beta h^2} & & \\
& \ddots & \\
& & \dfrac{-2\sin^4 \theta_M}{\beta h^2}
\end{bmatrix}\check{T}, \tag{2.7}
$$

and

$$
\begin{aligned}
U(\beta, x, \boldsymbol{\theta}_M, h) &:= \frac{\left[\left(x + \frac{2}{\beta}\sin 2\boldsymbol{\theta}_M\right)\sin^2 \boldsymbol{\theta}_M - \cos^2 \boldsymbol{\theta}_M\right]}{2h}\check{U} \\
&= \begin{bmatrix}
\dfrac{\left(x + \frac{2}{\beta}\sin 2\theta_1\right)\sin^2 \theta_1 - \cos^2 \theta_1}{2h} & & \\
& \ddots & \\
& & \dfrac{\left(x + \frac{2}{\beta}\sin 2\theta_M\right)\sin^2 \theta_M - \cos^2 \theta_M}{2h}
\end{bmatrix}\check{U},
\end{aligned} \tag{2.8}
$$

where $\boldsymbol{\theta}_M = [\theta_1, \theta_2, \dots, \theta_M]^\mathsf{T}$. Then

$$
\frac{\partial \boldsymbol{H}_M}{\partial x} = [T(\beta, \boldsymbol{\theta}_M, h) + U(\beta, x, \boldsymbol{\theta}_M, h)]\,\boldsymbol{H}_M. \tag{2.9}
$$

Noting that x is the time-like variable, we apply the trapezoidal rule with time step $\Delta x < 0$ to (2.9) yielding

$$
\begin{aligned}
\frac{\boldsymbol{H}_M^{n+1} - \boldsymbol{H}_M^n}{\Delta x} &= \frac{1}{2}\big[T(\beta, \boldsymbol{\theta}_M, h)\boldsymbol{H}_M^n + U(\beta, x_n, \boldsymbol{\theta}_M, h)\boldsymbol{H}_M^n\big] \\
&\quad + \frac{1}{2}\big[T(\beta, \boldsymbol{\theta}_M, h)\boldsymbol{H}_M^{n+1} + U(\beta, x_{n+1}, \boldsymbol{\theta}_M, h)\boldsymbol{H}_M^{n+1}\big],
\end{aligned}
$$

where $\boldsymbol{H}_M^n \approx \boldsymbol{H}_M(x_n)$. Upon rearranging, this gives

$$
\begin{aligned}
\left[I - \frac{\Delta x}{2}T(\beta, \boldsymbol{\theta}_M, h) - \frac{\Delta x}{2}U(\beta, x_{n+1}, \boldsymbol{\theta}_M, h)\right]\boldsymbol{H}_M^{n+1} \\
= \left[I + \frac{\Delta x}{2}T(\beta, \boldsymbol{\theta}_M, h) + \frac{\Delta x}{2}U(\beta, x_n, \boldsymbol{\theta}_M, h)\right]\boldsymbol{H}_M^n.
\end{aligned} \tag{2.10}
$$

Finally, we obtain $F_\beta(x) \approx \tilde{F}_\beta(x) := H_M(x) \approx H_M^n$, $n = 0, 1, \dots, N$. Below, we also explore other time integration methods, in addition to the trapezoidal method. We use the trapezoidal method as our default due to its A-stability, but it can be outperformed by BDF methods in this context.

2.2   Spectral discretization

To obtain better accuracy, we apply a Fourier spectral method. As before, for $0 \leq n \leq N$, define $x_n = x_0 + n\Delta x$. Suppose $H(x, \theta) = \int_0^\theta \rho(x, \theta') \, d\theta'$, then we rewrite (2.3) as

$$\frac{\partial H}{\partial x} + \left(\frac{2}{\beta} \sin^4 \theta \right) \frac{\partial \rho}{\partial \theta} + \left[\left(x + \frac{2}{\beta} \sin 2\theta \right) \sin^2 \theta - \cos^2 \theta \right] \rho = 0.$$

Upon taking the derivative with respect to θ on both sides, we arrive at a partial differential equation for $\rho(x, \theta)$,

$$
\begin{aligned}
&\frac{\partial \rho}{\partial x} + \left(\frac{8}{\beta} \sin^3 \theta \cos \theta \right) \frac{\partial \rho}{\partial \theta} + \left(\frac{2}{\beta} \sin^4 \theta \right) \frac{\partial^2 \rho}{\partial \theta^2} \\
&\quad + \left[(2x + 2) \sin \theta \cos \theta + \frac{2}{\beta} \left(2 \sin^2 \theta \cos 2\theta + 2 \sin \theta \cos \theta \sin 2\theta \right) \right] \rho \\
&\quad + \left[\left(x + \frac{2}{\beta} \sin 2\theta \right) \sin^2 \theta - \cos^2 \theta \right] \frac{\partial \rho}{\partial \theta} = 0.
\end{aligned}
\tag{2.11}
$$

Now, suppose

$$\rho(x, \theta) \approx \sum_{m=-M}^{M} a_m(x) e^{2im\theta/l}, \qquad \theta \in [0, l\pi), \quad \theta_M = l\pi. \tag{2.12}$$

Substituting (2.12) in (2.11) gives a system of ordinary differential equations for $a_m(x)$, $m = -M, \ldots, M$, after truncation

$$\frac{d\boldsymbol{a}_M(x)}{dx} = (A + xB)\boldsymbol{a}_M(x), \qquad \boldsymbol{a}_M(x) = \begin{bmatrix} a_{-M}(x) \\ \vdots \\ a_M(x) \end{bmatrix}, \tag{2.13}$$

where

$$
\begin{aligned}
A = &-\frac{2}{\beta} \left(\frac{3}{8} I - \frac{1}{4} S_{-l} - \frac{1}{4} S_{+l} + \frac{1}{16} S_{+2l} + \frac{1}{16} S_{-2l} \right) D_2 \\
&+ \left[\frac{1}{2} I + \frac{1}{4} S_{-l} + \frac{1}{4} S_{+l} - \frac{2}{\beta} \left(\frac{1}{2i} S_{-l} - \frac{1}{2i} S_{+l} \right) \left(\frac{1}{2} I - \frac{1}{4} S_{+l} - \frac{1}{4} S_{-l} \right) \right] D_1 \\
&- \left[\frac{8}{\beta} \left(\frac{1}{-8i} S_{+3l/2} + \frac{1}{8i} S_{-3l/2} + \frac{3}{8i} S_{-l/2} - \frac{3}{8i} S_{+l/2} \right) \left(\frac{1}{2} S_{-l/2} + \frac{1}{2} S_{+l/2} \right) \right] D_1 \\
&- \frac{4}{\beta} \left(\frac{1}{2} I - \frac{1}{4} S_{+l} - \frac{1}{4} S_{-l} \right) \left(\frac{1}{2} S_{+l} + \frac{1}{2} S_{-l} \right) \\
&- \frac{4}{\beta} \left(\frac{1}{2i} S_{-l/2} - \frac{1}{2i} S_{+l/2} \right) \left(\frac{1}{2} S_{+l/2} + \frac{1}{2} S_{-l/2} \right) \left(\frac{1}{2i} S_{-l} - \frac{1}{2i} S_{+l} \right) \\
&- 2 \left(\frac{1}{2i} S_{-l/2} - \frac{1}{2i} S_{+l/2} \right) \left(\frac{1}{2} S_{+l/2} + \frac{1}{2} S_{-l/2} \right),
\end{aligned}
\tag{2.14}
$$

and

$$B = \left(-\frac{1}{2} I + \frac{1}{4} S_{-l} + \frac{1}{4} S_{+l} \right) D_1 - 2 \left(\frac{1}{2i} S_{-l/2} - \frac{1}{2i} S_{+l/2} \right) \left(\frac{1}{2} S_{+l/2} + \frac{1}{2} S_{-l/2} \right). \tag{2.15}$$

Here $S_{\pm k}$ represent the (Fourier modes) shift matrices, $S_{\pm k} = S_{\pm 1}^k$, where

$$
S_{+1} = \begin{bmatrix} 0 & 1 & & & & 0 \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & 0 & 1 \\ 1 & & & & & 0 \end{bmatrix}, \qquad S_{-1} = \begin{bmatrix} 0 & & & & & 1 \\ 1 & 0 & & & & \\ 0 & 1 & 0 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & 0 & 0 \\ 0 & & & & 1 & 0 \end{bmatrix}.
$$

Also D_1 and D_2 represent the first and second-order differentiation matrices in the Fourier space, i.e., $D_2 = D_1^2$, where

$$
D_1 = \begin{bmatrix} -2\mathrm{i}M/l & & & & \\ & -2\mathrm{i}(M-1)/l & & & \\ & & \ddots & & \\ & & & 2\mathrm{i}(M-1)/l & \\ & & & & 2\mathrm{i}M/l \end{bmatrix}.
$$

The Fourier coefficients of the initial condition are obtained via

$$
a_m(x_0) = \frac{1}{l\pi} \int_0^{l\pi} \rho(x_0, \theta) \mathrm{e}^{-2\mathrm{i}m\theta/l} \, \mathrm{d}\theta. \tag{2.16}
$$

Instead of applying the second-order accurate trapezoidal rule to integrate the system of ODEs for $\rho(x, \theta)$, we suggest the use of a five-step backward differentiation formula method (BDF5), see [12, Section 8.4]. To use BDF5, we need four more starting conditions, i.e., $\boldsymbol{a}_M(x_0 - i\Delta x)$, $i = 1, 2, 3, 4$, which can be obtained by (2.4) and (2.16). We then proceed with BDF5 to solve for $\boldsymbol{a}_M^n \approx \boldsymbol{a}_M(x_n)$

$$
[137I - 60\Delta x \, (A + x_n B)] \, \boldsymbol{a}_M^n = 300\boldsymbol{a}_M^{n-1} - 300\boldsymbol{a}_M^{n-2} + 200\boldsymbol{a}_M^{n-3} - 75\boldsymbol{a}_M^{n-4} + 12\boldsymbol{a}_M^{n-5}. \tag{2.17}
$$

Other time integration methods can be used, and this will be discussed further below.

Finally, the value of $H(x, \theta)$ can be recovered from

$$
H(x, \theta) \approx \int_0^\theta \sum_{m=-M}^M a_m(x) \mathrm{e}^{2\mathrm{i}m\theta'/l} \, \mathrm{d}\theta' = \sum_{m=-M}^M a_m(x) \int_0^\theta \mathrm{e}^{2\mathrm{i}m\theta'/l} \, \mathrm{d}\theta'.
$$

The Tracy–Widom distribution can then be approximated by setting $\theta = \pi$,

$$
F_\beta(x) \approx \tilde{F}_\beta(x) = \sum_{m=-M}^M a_m(x) \int_0^\pi \mathrm{e}^{2\mathrm{i}m\theta'/l} \, \mathrm{d}\theta' = \sum_{m=-M}^M a_m(x) \frac{l}{2\mathrm{i}m} \big(\mathrm{e}^{2\mathrm{i}m\pi/l} - 1 \big),
$$

from which we find

$$
F_\beta(x_n) \approx \tilde{F}_\beta(x_n) \approx \sum_{m=-M}^M a_m^n \frac{l}{2\mathrm{i}m} \big(\mathrm{e}^{2\mathrm{i}m\pi/l} - 1 \big), \qquad n = 0, 1, \ldots, N.
$$

3 Algorithm validation and comparison

At this point, we have approximated the values of $F_\beta(x)$ on equally-spaced grid points x_0, x_1, $\ldots, x_N$. To obtain a continuous function, we perform Fourier interpolation. Since Fourier interpolation has high accuracy when the function being interpolated is periodic, we define

$$
\varphi(x) = \frac{\mathrm{erf}(x) + 1}{2}, \qquad x \in \mathbb{R},
$$

where erf denotes the error function [18]. Then consider

$$G_\beta(x) := \tilde{F}_\beta(x) - \varphi(x),$$

which is nearly a periodic function. Note that other functions can also be used instead of the error function $\mathrm{erf}(x)$ in constructing a periodic function. We perform Fourier interpolation to $G_\beta(x)$ on the grid $x_1, \ldots, x_N$. The resulting interpolant is then evaluated on a shifted and scaled Chebyshev grid on $[x_N, x_0]$. By adding back in $\varphi(x)$ evaluated on the same Chebyshev grid and then interpolating with Chebyshev polynomials, we obtain a useful high-accuracy approximation of $F_\beta(x)$. The number of Chebyshev coefficients is user-decided (we use 10^3 by default).

We point out that using either (2.5) or (2.13), we can also obtain the approximation of $F'_\beta(x)$ on the grid $x_0, x_1, \ldots, x_N$ at nearly no extra cost. We then perform the same procedure of interpolation to get an approximation of $F'_\beta(x)$, without using $\varphi(x)$ since the function being interpolated is already nearly periodic.

Algorithm 1 shows the pseudocode for the Fourier interpolation, where T is the grid points $x_N, x_{N-1}, \ldots, x_0$ with grid spacing Δx, K is an integer that gives the number of Chebyshev coefficients, F is the value of $\tilde{F}_\beta(x_j)$, $j = 1, \ldots, N$, and f is the value of $\tilde{F}'_\beta(x_j)$. We think of F, f as functions on T and need to extend them to all of $[x_N, x_0]$.

Algorithm 1 Fourier_interp(T, K, F, f)

1: Set D=$[\min \mathtt{T}, \max \mathtt{T}]$
2: Set G(T)=F(T)$-\varphi$(T)
3: Perform Fourier interpolation to G(T) and f(T) to obtain the interpolant G(x) and f(x), $x \in$ D
4: Set F̃=G($\bar{x}$)+φ($\bar{x}$) and f̃($\bar{x}$)=f($\bar{x}$), where $\bar{x}$ is a Chebyshev grid on D with K points
5: Perform Chebyshev interpolation to F̃ and f̃ and return the interpolants

---

For the finite-difference discretization using trapezoidal method to compute either the cumulative distribution function (cdf) or the probability density function (pdf) of F_β using $\mathtt{TW}(\beta)$, as provided in `TracyWidomBeta.jl`, the following default values for the parameters are used:

$$x_0 = \left\lfloor \frac{13}{\sqrt{\beta}} \right\rfloor, \quad x_N = -10, \quad \Delta x = -10^{-3}, \quad \theta_M = \pi, \quad K = 10^3, \quad M = 10^3, \tag{3.1}$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

Notation. Throughout we use $\mathtt{TW}(\beta; \mathtt{params})$ to refer to our implementation with different choices of parameters. For example,

$$\mathtt{TW}(\beta; \mathtt{method="finite"}, \mathtt{step="trapz"}, \mathtt{pdf=true}),$$

refers to using the finite-difference discretization, time-stepped with the trapezoidal method, outputting an approximation of F'_β with the default parameters (3.1).

The default values of x_0 and x_N are chosen to be $\lfloor 13/\sqrt{\beta} \rfloor$ and -10 respectively so that $F_\beta(x_0) \approx 1$ and $F_\beta(x_N) \approx 0$ for $\beta \geq 1$. Though not optimal, larger values of x_0 and smaller values of x_N can also be used. See Section 3.3.3 for a discussion on the selection of the default value for x_0. For $0 < \beta < 1$, a larger domain for x should be used. The values of Δx and M are chosen so that $M = \lfloor -1/\Delta x \rfloor$. In this way, the local truncation error of the trapezoidal method is of the optimal order. Algorithm 2 shows the pseudocode for $\mathtt{TW}(\beta; \mathtt{pdf=true})$. Step 6 may be replaced by solving an alternate discretization of (2.9).

Similarly, for the spectral discretization using BDF5 method to compute either the cdf or the pdf using

$$\mathtt{TW}(\beta; \mathtt{method="spectral"}, \mathtt{step="bdf5"}) \quad \text{or}$$
$$\mathtt{TW}(\beta; \mathtt{method="spectral"}, \mathtt{step="bdf5"}, \mathtt{pdf=true}),$$

Algorithm 2 TW(β; `pdf=true`)

1: Set the values of x_0, x_N, Δx, θ_{M}, K, and M as in (3.1)
2: Set x=$x_0 : \Delta x : x_N$
3: Set up the initial condition as in (2.4)
4: Set up the matrices as in (2.6), (2.7), and (2.8)
5: **for** m=1,..., `length`(x)-1 **do**
6:     Solve (2.10) and denote the result by H
7:     Use (2.5) and denote the result by h
8: **end for**
9: Take the values of H and h for θ=π and denote the results by F and f
10: Perform Algorithm 1 with F, f, and T=x
11: Return the interpolant for $\tilde{\mathrm{f}}$

---

the following values for the parameters are used:

$$x_0 = \left\lfloor \frac{13}{\sqrt{\beta}} \right\rfloor, \quad x_N = -10, \quad \Delta x = -10^{-3}, \quad \theta_M = 20\pi, \quad K = 10^3, \quad M = 8 \times 10^3. \quad (3.2)$$

One thing to note here is that θ_M is set to be 20π, which is much larger than π as used for finite-difference discretization. This setting is necessary since a periodic boundary condition (2.12) is imposed on $\rho(x,\theta)$, or equivalently, on $H(x,\theta)$. If the length of the domain, θ_M, is not large enough, due to periodicity, the approximation to $H(x,\theta)$ will propagate to the end of the domain and reappear at the lower boundary $\theta = 0$. In other words, θ_M depends on the speed of propagation of the approximation to $H(x,\theta)$, and it turns out that setting $\theta_M = 20\pi$ is sufficient. M is set to be 8×10^3 regardless of the value of Δx to make sure we have large enough number of Fourier modes to represent the initial condition. See Section 3.4 for details on which value of M to use for each method in terms of both accuracy and computation time. Algorithm 3 shows the pseudocode for TW(β; `method="spectral"`, `step="bdf5"`, `pdf=true`). As with the finite-difference method, step 11 can be replaced with solving an alternate discretization of (2.13).

Algorithm 3 TW(β; `method="spectral"`, `step="bdf5"`, `pdf=true`)

1: Set the values of x_0, x_N, Δx, θ_{M}, K, and M as in (3.2)
2: Set x=$x_0 : \Delta x : x_N$
3: Set D=$0 : h : \theta_{\mathrm{M}} - h$ with h=$\theta_{\mathrm{M}}/$M
4: Set up the initial conditions using (2.4) and (2.16).
5: Construct matrices A and B as in (2.14) and (2.15)
6: Set N=`-floor`(M/2) $: 1 :$ `floor`((M$-$1)/2)
7: **for** m $\in$ N **do**
8:     Compute the vector `integ`=$\frac{l}{2im}\left(e^{2im\pi/l}-1\right)$ with l=θ_{M}/π
9: **end for**
10: **for** k=$2 :$ `length`(x) **do**
11:     Solve (2.17) to find $\mathbf{a}_{\mathrm{M}}^n$
12:     Use (2.13) and denote the result by $\mathbf{b}_{\mathrm{M}}^n$
13: **end for**
14: Compute F=$\langle \mathbf{a}_{\mathrm{M}}^n, \mathtt{integ} \rangle$ and f=$\langle \mathbf{b}_{\mathrm{M}}^n, \mathtt{integ} \rangle$
15: Perform Algorithm 1 with F, f, and T=x
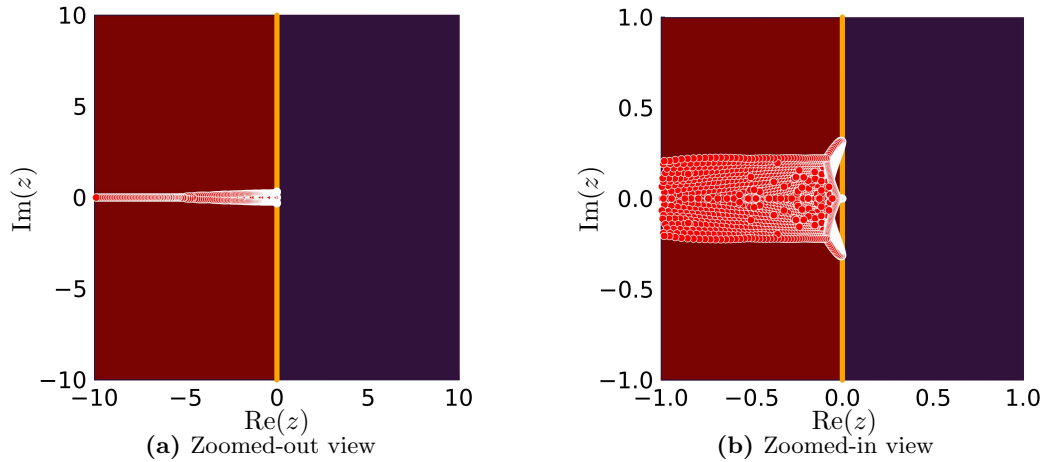16: Return the interpolant for $\tilde{\mathrm{f}}$

Figure 1. The absolute stability region of trapezoidal method applied to the finite-difference discretization (the maroon-shaded region along with the orange boundary curve) and the eigenvalues of $\Delta x(T+U)$ (the red dots with white boundary) for $\beta = 2$, $x = x_0, x_0 - 1, \ldots, x_N$, $\Delta x = -10^{-3}$, and $M = 10^3$.
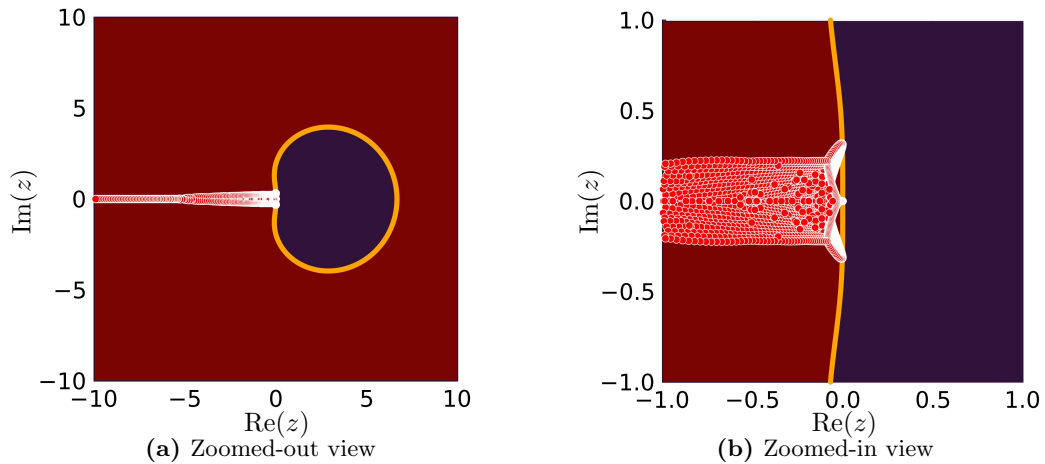

Figure 2. The absolute stability region of BDF3 applied to the finite-difference discretization (the maroon-shaded region along with the orange boundary curve) and the eigenvalues of $\Delta x(T+U)$ (the red dots with white boundary) for $\beta = 2$, $x = x_0, x_0 - 1, \ldots, x_N$, $\Delta x = -10^{-3}$, and $M = 10^3$.

Remark 3.1. To optimize the use of the spectral method, one should consider θ from the interval $[k_1(x)\pi, k_2(x)\pi]$, where k_1, k_2 are x-dependent integers that adapt to where the solution is non-zero, see Figure 10.

Note that when β is very large, numerical instabilities can occur. This instability arises primarily because the initial condition closely resembles a step function. To get an accurate result in this case, first of all, one needs to use finer grid for x and larger value for M, i.e., refinement in both time and space. Also, one needs to use more Chebyshev coefficients. With current values for the parameters, `TW(`β`)` and `TW(`β`; method="spectral", step="bdf5")` exhibits stability roughly for $1 \leq \beta \leq 30$.

3.1 Eigenvalues of $\Delta x[T(\beta, \theta_M, h) + U(\beta, x, \theta_M, h)]$ for finite-difference discretization

The default time-stepping routine for the finite-difference discretization is the trapezoidal method but BDF3, BDF4, BDF5, and BDF6 can also be used on (2.5) to solve for $\boldsymbol{H}_M$. It turns out that with values for the parameters as in (3.1), convergence occurs for all these methods.

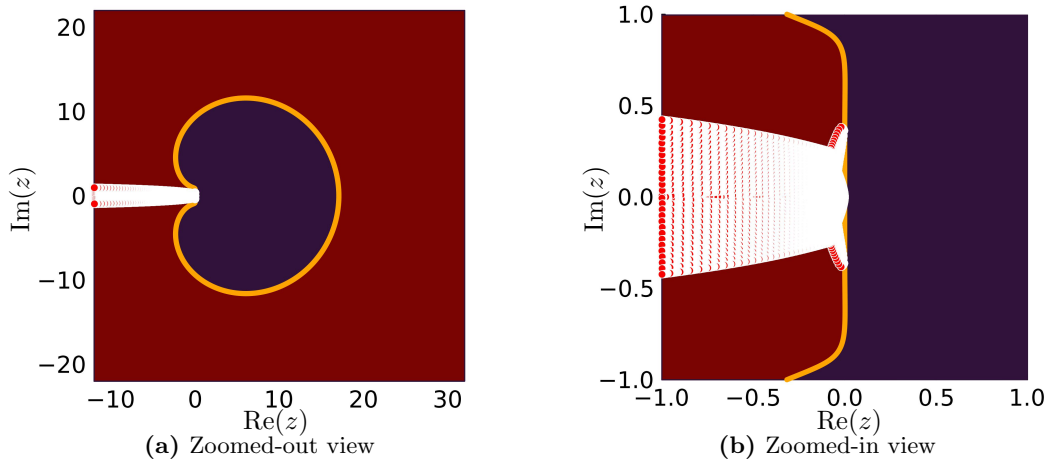**(a)** Zoomed-out view

(b) Zoomed-in view

Figure 3. The absolute stability region of BDF5 applied to the spectral discretization (the maroon-shaded region along with the orange boundary curve) and the eigenvalues of $\Delta x(A + xB)$ (the red dots with white boundary) for $\beta = 2$, $x = x_0, x_0 - 1, \ldots, x_N$, $\Delta x = -10^{-3}$, and $M = 8 \times 10^3$.
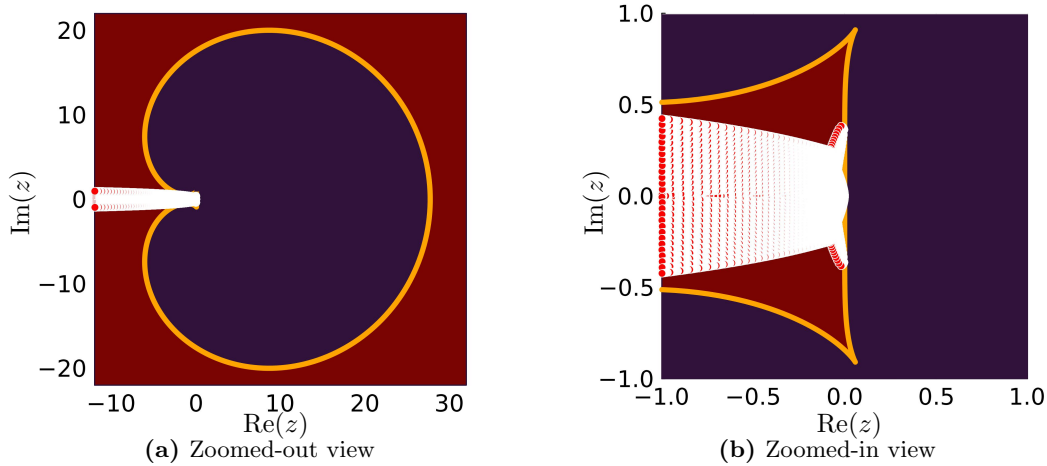

(a) Zoomed-out view

(b) Zoomed-in view

Figure 4. The absolute stability region of BDF6 applied to the spectral discretization (the maroon-shaded region along with the orange boundary curve) and the eigenvalues of $\Delta x(A + xB)$ (the red dots with white boundary) for $\beta = 2$, $x = x_0, x_0 - 1, \ldots, x_N$, $\Delta x = -10^{-3}$, and $M = 8 \times 10^3$.

Since with the finite-difference discretization, the trapezoidal method and BDF3 are usually used, Figures 1 and 2 show the absolute stability regions of trapezoidal method and BDF3 along with eigenvalues of $\Delta x(T + U)$ for $\beta = 2$, $x = x_0, x_0 - 1, \ldots, x_N$, $\Delta x = -10^{-3}$, and $M = 10^3$. The eigenvalues in the left-half plane are firmly within the region of absolute stability in each case.

3.2   Eigenvalues of $\Delta x(A + xB)$ for the spectral discretization

Figures 3 and 4 show the absolute stability regions of BDF5 and BDF6 along with eigenvalues of $\Delta x(A + xB)$ for $\beta = 2$, $x = x_0, x_0 - 1, \ldots, x_N$, $\Delta x = -10^{-3}$, and $M = 8 \times 10^3$. It is clear that, with the values of the parameters given as in (3.2), the eigenvalues in the left-half plane are firmly within the region of absolute stability for these methods. We choose BDF5 over BDF6 because with the same values of these parameters, the performance is roughly the same yet BDF5 has a larger absolute stability region. Selecting Δx is more nuanced than one might think, see Appendix A for more details.

3.3 Error analysis

We now compare the accuracy of these two algorithms when computing the cdf for $\beta = 1, 2, 4$. For reference solutions, we implement the Fredholm determinant representations for $\beta = 1, 2, 4$ [3, 4, 9] by porting the code in the JULIA package `RandomMatrices` to MATHEMATICA and implementing it in high-precision arithmetic. This ensures that our reference solutions are accurate to beyond the $\approx 10^{-16}$ machine precision for standard double precision. Observe that for $F_4(x)$, x should be divided by a factor of $2^{1/6}$. This difference in variance convention has also been underscored in [16, p. 47], [14, 17], and [1, Remark 5.1.4].

3.3.1 Error across the domain
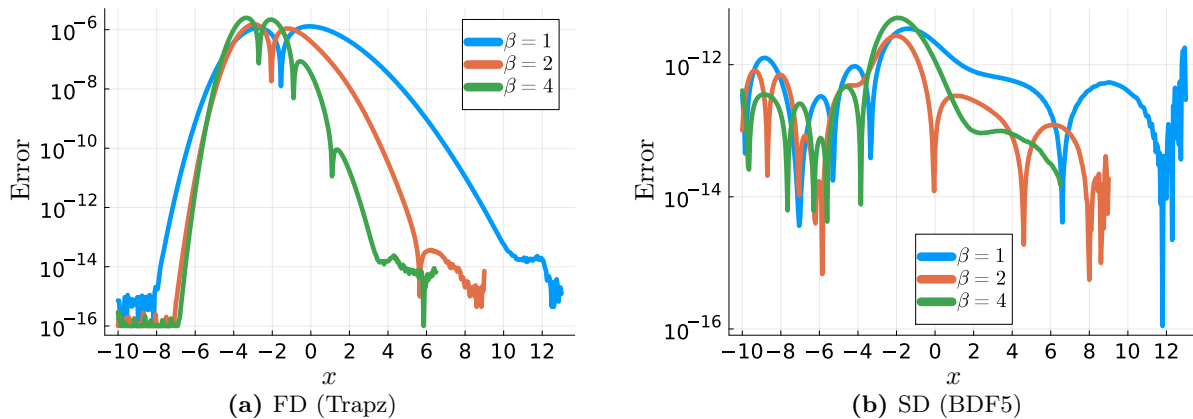

(a) FD (Trapz)   **(b)** SD (BDF5)

Figure 5. Absolute errors from the two algorithms are presented over the domain $x \in [-10, 13]$ for $\beta = 1, 2, 4$ with $x_0 = \lfloor 13/\sqrt{\beta} \rfloor$ and $\Delta x = -10^{-3}$. For the finite-difference discretization, $M = 10^3$ is used, while for the spectral discretization, $M = 8 \times 10^3$ is employed.

With values of the parameters given as in (3.1) and (3.2), Figure 5 shows the absolute errors of the two algorithms across the domain. Tables 1 and 2 show the absolute errors of some selected x-values with same values for the parameters.

For finite-difference discretization, from either Figure 5 (a) or Table 1, the error is roughly on the order of 10^{-7} around the peak of the distribution, and it improves when x approaches either end of the domain. This is due to the fact that exact values, 0 and 1, for the initial condition are imposed on the endpoints of the domain, and the Dirichlet boundary condition ensures that the solution tends to zero.

For spectral discretization, from either Figure 5 (b) or Table 2, the error is roughly on the order of 10^{-13} throughout the entire domain. Regardless of the order of error, the reason for this discrepancy from finite-difference discretization near the edges of the domain is that an error is introduced for the initial condition when we represent $\rho(x, \theta)$ in terms of a truncated Fourier series.

3.3.2 Order of error

For $\beta = 2$ evaluated at $x = -2$, Figure 6 shows the order of error plots of finite-difference discretization and spectral discretization. We choose $x = -2$ since it is near the peak of the distribution. With $\Delta x = -0.004, -0.002, -0.001$ for finite-difference discretization and $\Delta x = -0.2, -0.1, -0.05$ for spectral discretization, Figure 7 shows the change of error as the value of $|\Delta x|$ decreases.

β \ x	1	2	4
-8	$1.832(-15)$	$5.024(-17)$	$1.950(-17)$
-6	$7.278(-10)$	$9.602(-12)$	$4.766(-16)$
-4	$3.715(-7)$	$3.600(-7)$	$1.807(-7)$
-2	$6.553(-7)$	$1.027(-7)$	$2.020(-6)$
0	$1.295(-6)$	$4.014(-7)$	$3.319(-8)$
2	$4.116(-7)$	$1.137(-8)$	$6.185(-12)$
4	$3.352(-8)$	$3.286(-11)$	$1.565(-14)$
6	$9.575(-10)$	$3.186(-14)$	$5.995(-15)$

Table 1. Absolute errors of finite-difference discretization using trapezoidal method with values of the parameters as in (3.1).

β \ x	1	2	4
-8	$5.912(-13)$	$6.870(-13)$	$3.509(-13)$
-6	$3.278(-13)$	$3.470(-14)$	$2.021(-13)$
-4	$9.057(-13)$	$4.580(-13)$	$4.525(-13)$
-2	$2.663(-12)$	$2.740(-12)$	$4.809(-12)$
0	$2.162(-12)$	$1.090(-13)$	$1.135(-12)$
2	$7.924(-13)$	$3.514(-13)$	$9.293(-14)$
4	$5.719(-13)$	$1.423(-13)$	$7.527(-14)$
6	$1.819(-13)$	$7.094(-14)$	$6.439(-15)$

Table 2. Absolute errors of spectral discretization using BDF5 with values of the parameters as in (3.2).

For both the finite-difference discretization and the spectral discretization, from Figure 6, the error has the expected order with respect to $|\Delta x|$. However, if we decrease the value of $|\Delta x|$ further, each BDF method has a corresponding range for Δx that causes instability. Moreover, once the value of Δx exits this range, the corresponding BDF method becomes accurate again. See Appendix A for more details.

Remark 3.2. In addition to the trapezoidal and BDF methods, one could consider A-stable and L-stable diagonally implicit Runge–Kutta methods as alternative options for time-stepping. Furthermore, one could even consider adaptive time-stepping. While these alternative methods can effectively address the stability concerns arising from the spectrum of $\Delta x(A + xB)$ and $\Delta x(T + U)$, they do tend to increase the computation time. These methods likely require more than one linear solve at each time step (possibly one matrix factorization and multiple uses of it). And our target time step is $|\Delta x| = 10^{-3}$ and, with this time step, BDF5 is stable with the default parameters, requiring only one (sparse) linear solve per time step. Then the comparison of methods becomes a complicated competition between more expensive, higher-order methods that allow a larger time step and less expensive, but still fairly high-order, methods requiring a smaller time step. It is possible that savings can be achieved here, but we did not see it in our experiments.

3.3.3   Error with respect to x_0

Figure 8 shows how the maximum error over $x = -8, -7, \ldots, x_0$ changes with respect to the value of x_0. For $\beta = 4$ with spectral discretization using BDF5, the error for $x_0 = 12, 13$ is roughly $\mathcal{O}(10^{-11})$, which can be improved to $\mathcal{O}(10^{-12})$ using more Fourier modes. Except for
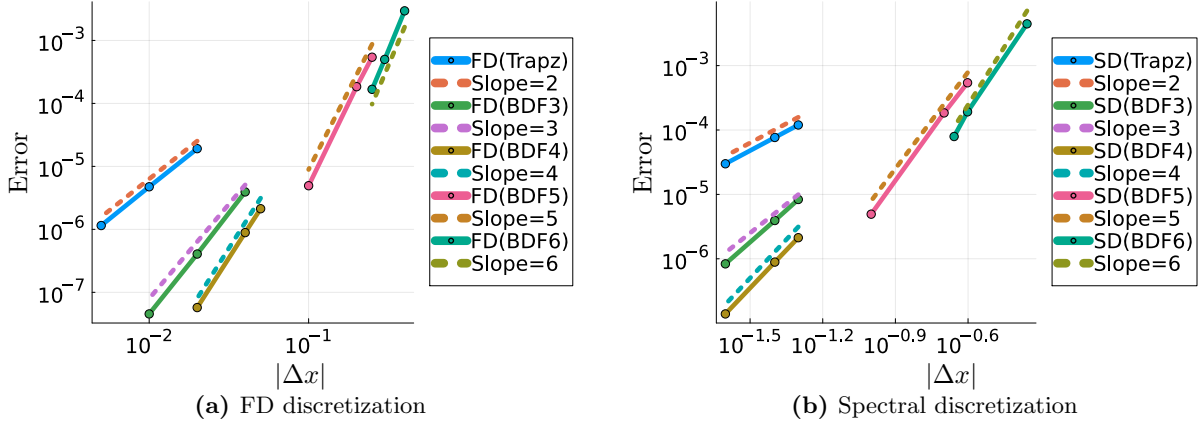
(a) FD discretization   **(b)** Spectral discretization

Figure 6. The order of accuracy for the finite-difference (FD) discretization and the spectral discretization. (a) The finite-difference discretization. The errors for trapezoidal (Trapz) method are computed by treating $\texttt{TW}\big(\beta\texttt{=2}; \ x_0 = \lfloor 13/\sqrt{\beta} \rfloor, \ \texttt{step="trapz"}, \ \texttt{interp=false}, \ \Delta x\texttt{=} - 10^{-3}, \ M\texttt{=}10^3\big)$ as the reference and compared with $\Delta x = -0.005, -0.01, -0.02$ at $x = -2$. For BDF methods, the same reference is used with $\texttt{step}$ changed accordingly. For BDF3, $\Delta x = -0.01, -0.02, -0.04$ are used. For BDF4, $\Delta x = -0.02, -0.04, -0.05$ are used. For BDF5, $\Delta x = -0.1, -0.2, -0.25$ are used. For BDF6, $\Delta x = -0.25, -0.3, -0.4$ are used. (b) The spectral discretization. The errors for trapezoidal (Trapz) method are computed by treating $\texttt{TW}\big(\beta\texttt{=2}; \ x_0 = \lfloor 13/\sqrt{\beta} \rfloor, \ \texttt{method="spectral"}, \ \texttt{step="trapz"}, \ \texttt{interp=false}, \ \Delta x\texttt{=} - 10^{-3}, \ M\texttt{=8000}\big)$ as the reference and compared with $\Delta x = -0.025, -0.04, -0.05$ at $x = -2$. For BDF methods, the same reference is used with $\texttt{step}$ changed accordingly. For BDF3 and BDF4, $\Delta x = -0.025, -0.04, -0.05$ are used. For BDF5, $\Delta x = -0.1, -0.2, -0.25$ are used. For BDF6, $\Delta x = -0.22, -0.25, -0.44$ are used.
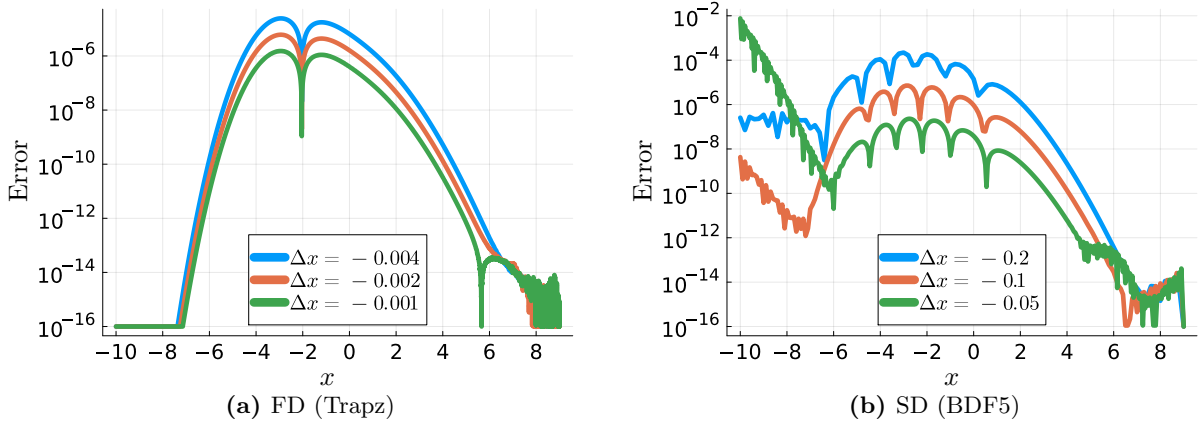

(a) FD (Trapz)   **(b)** SD (BDF5)

Figure 7. Change of error of finite-difference discretization using trapezoidal method with $\Delta x = -0.004, -0.002, -0.001$ and spectral discretization using BDF5 with $\Delta x = -0.2, -0.1, -0.05$. Values of the other parameters are used as in (3.1) and (3.2). When applying BDF5 to the spectral discretization, errors are more pronounced for smaller time steps at the left edge due to the onset of instability discussed in the appendix.

these two cases, for both methods, as the value of β increases, the value of $\tilde{x}_0$, the minimum value of x_0 that can be used without affecting the accuracy, decreases. This is exactly what we expect since the larger value of β is, the more concentrated the distribution becomes. Based on [5, 6], one expects $\tilde{x}_0 \approx C/\beta^{\frac{2}{3}}$ for some constant C. This suggests a way to choose the optimal value of x_0 in terms of accuracy and computation time. For our algorithms, we choose $C = 13$ since for $\beta = 1$, $F_\beta(13) \approx 1$. In practice, instead of using $\beta^{\frac{2}{3}}$ for the denominator, we are more
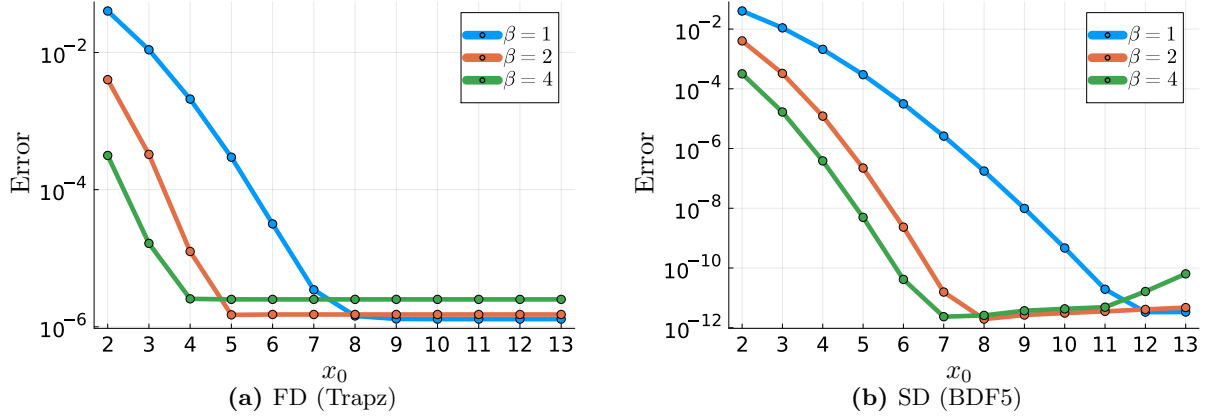
(a) FD (Trapz)   **(b)** SD (BDF5)

Figure 8. Change of error of finite-difference discretization using trapezoidal method and spectral discretization using BDF5 with respect to the value of x_0. $\beta = 1, 2, 4$ and $x_0 = 13, 12, \ldots, 2$ are used. Values of the other parameters are the same as in (3.1) and (3.2). For each value of β and x_0, the error is taken to be the maximum over $x = -8, -7, \ldots, x_0$.

Discretization	Method	Time (default)	Error (default)	Time $\left(10^{-6}\right)$
Finite difference	Trapz	12.684s	$1.027(-7)$	0.203s
	BDF3	10.731s	$5.461(-8)$	0.138s
	BDF4	10.458s	$5.457(-8)$	0.142s
	BDF5	10.947s	$5.538(-8)$	0.141s
	BDF6	11.067s	$5.539(-8)$	0.140s
Spectral	Trapz	222.135s	$4.790(-8)$	10.906s
	BDF3	205.888s	$4.298(-11)$	1.949s
	BDF4	209.175s	$2.346(-12)$	1.388s
	BDF5	201.819s	$2.663(-12)$	1.008s
	BDF6	204.295s	$2.677(-12)$	204.295s

Table 3. For each discretization in the first column and time-stepping method in the second column, computation time to get the interpolated cdf for the Tracy–Widom distribution and the corresponding error at $x = -2$ for $\beta = 2$ are recorded in the next two columns. The computation times in the third column are generated using the default parameters as in (3.1) and (3.2), with the corresponding errors in the fourth column. If we aim for an error of $\mathcal{O}\left(10^{-6}\right)$, the last column shows the corresponding minimum observed computation times for each discretization and time-stepping method to achieve this error.

conservative and use $\beta^{\frac{1}{2}}$ since, based on Figure 8 (b), when $\beta = 4$, using $x_0 = 13/\beta^{\frac{2}{3}} \approx 5.16$ brings in an error of $\mathcal{O}\left(10^{-9}\right)$.

3.4 Computation time

The values in Table 3 are obtained by running on a computer with processor: 11th Gen Intel(R) Core(TM) i7-11800H 2.30GHz with 16.0GB of RAM.

Table 3 shows the computation time with default values of the parameters as in (3.1) and (3.2) for $\beta = 2$ along with the corresponding error at $x = -2$. The last column provides the computation time if we aim for an error of $\mathcal{O}\left(10^{-6}\right)$ at $x = -2$. For finite-difference discretization, to have the error of $\mathcal{O}\left(10^{-6}\right)$ at $x = -2$, $\Delta x = -0.01$ can be used instead of $\Delta x = -0.001$. For spectral discretization, to have the error of $\mathcal{O}\left(10^{-6}\right)$ at $x = -2$, $M = 4 \times 10^3$ can be used instead of $M = 8 \times 10^3$ with $\Delta x = -0.01$ for trapezoidal method, $\Delta x = -0.05$ for BDF3, $\Delta x = -0.07$ for BDF4, and $\Delta x = -0.1$ for BDF5. It turns out that the error using BDF6 will jump from

$\mathcal{O}(10^{-5})$ to $\mathcal{O}(10^{-12})$. From Table 3, it takes about 204.295s using BDF6 to have the error of $\mathcal{O}(10^{-12})$ with parameters as in (3.2). To have the error of $\mathcal{O}(10^{-5})$, it takes about 0.639s with $M = 5 \times 10^3$ and $\Delta x = -0.2$. Work to improve the speed of the spectral discretization is ongoing.

Using the same machine, Bloemendal's code (see [1, Section 6]), which uses MATHEMATICA's `NDSolve`, takes approximately 3 seconds to compute the approximation to the function H. We point out that MATHEMATICA 13.2, generates warnings from `NDSolve` regarding the errors in the approximate solution. The accuracy of this approach, while simple, appears to be limited to a maximum of 7 digits. Yet, the fact that this is indeed so simple, and works, is an important reminder of the conceptual simplicity of this representation of the Tracy–Widom distribution function.

Our methods expand upon this, providing both the cdfs and pdfs as output and producing high-accuracy interpolants while leaving the trade-off between accuracy and speed to the user's discretion. And since we consider the general-β Tracy–Widom distribution functions as important nonlinear special functions, developing methods, with the highest possible accuracy, is critical. See Section 5 for some thoughts on further improvements.

4 Additional numerical results

In this section, we present additional plots to demonstrate the power and flexibility of the code.

4.1 Comparison with large random matrices

We verify numerically that the pdf generated by our algorithm agrees with the model presented by Dumitriu and Edelman [7]. Recall that H_n^β in (1.2) has (1.1) as the jpdf for its eigenvalues, and the distribution of its largest eigenvalue, after rescaling, converges to F_β for any $\beta > 0$ as $n \to \infty$.

Histograms in Figure 9 are the normalized histograms for $n^{1/6}\big(\lambda_{\max}\big(H_n^\beta\big) - 2\sqrt{n}\big)$, where $\lambda_{\max}\big(H_n^\beta\big)$ denotes the largest eigenvalue of the β-Hermite ensemble H_n^β.

4.2 Evolution of the density $\rho(x,\theta) := \partial H(x,\theta)/\partial\theta$

Figure 10 shows the waterfall plots of the approximation of $\rho(x,\theta) := \partial H(x,\theta)/\partial\theta$ for $\theta \in [0, 10\pi]$ and $x = 0, -2, \ldots, -10$ using finite-difference discretization with trapezoidal method on (2.11) with $\theta_M = 10\pi$ and $M = 10^4$. Values of the other parameters are the same as in (3.1). As the value of x decreases, the density of the solution propagates to the right.

Figure 11 shows the contour plots of the approximation of $H(x,\theta)$ using finite-difference discretization with trapezoidal method with values of the parameters given as in (3.1). For each contour plot, the initial condition $H(x_0,\theta)$ is found along the top of the plot and the approximate Tracy–Widom distribution, $F_\beta(x) \approx H(x,\theta_M)$, is obtained on the right-hand side of the plot.

4.3 Density of the Tracy–Widom distribution

Figure 12 (a) shows the plot of the approximation of F_β' for $\beta = 1$ to 10 and $x \in [-6, 4]$ using finite-difference discretization with trapezoidal method with values of the other parameters as in (3.1). As we can see from the plot, as the value of β increases, F_β' becomes more concentrated, and its peak moves leftwards (See Appendix B for a discussion of the exact limiting behavior as $\beta \to \infty$).

Figure 12 (b) is a two-dimensional version of Figure 12 (a), which is also generated using the finite-difference discretization with trapezoidal method. It provides a closer view of F_β' for $\beta = 1$ to 4 with step size $= 0.2$. The red curves from right to left correspond to $\beta = 1, 2, 4$ respectively. The black curves show F_β' for the other values of β.
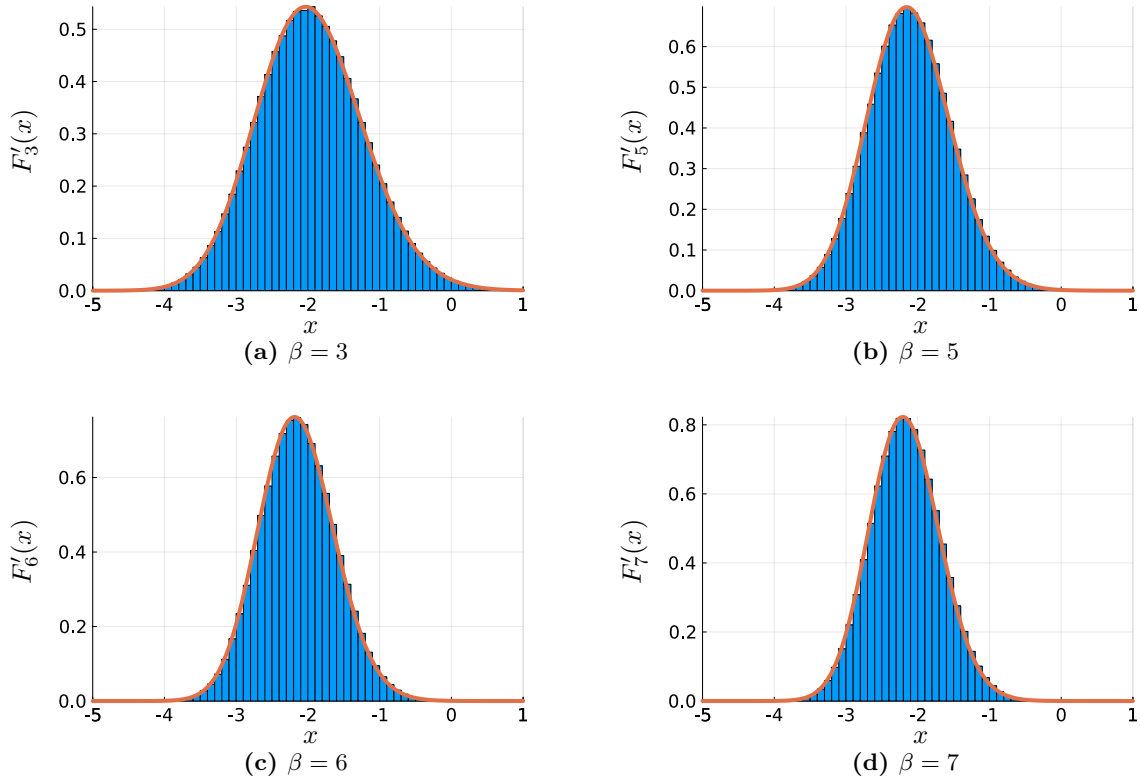
(a) $\beta = 3$

(b) $\beta = 5$

(c) $\beta = 6$

(d) $\beta = 7$

Figure 9. The histograms for $n^{1/6}\big(\lambda_{\max}\big(H_n^\beta\big) - 2\sqrt{n}\big)$ using 10^6 samples with $n = 10^4$. The density $F'_\beta(x)$ is generated by `TW`$(\beta; \texttt{pdf=true})$ for $\beta = 3, 5, 6, 7$. The small positive bias in each histogram can be improved by using larger values of n.
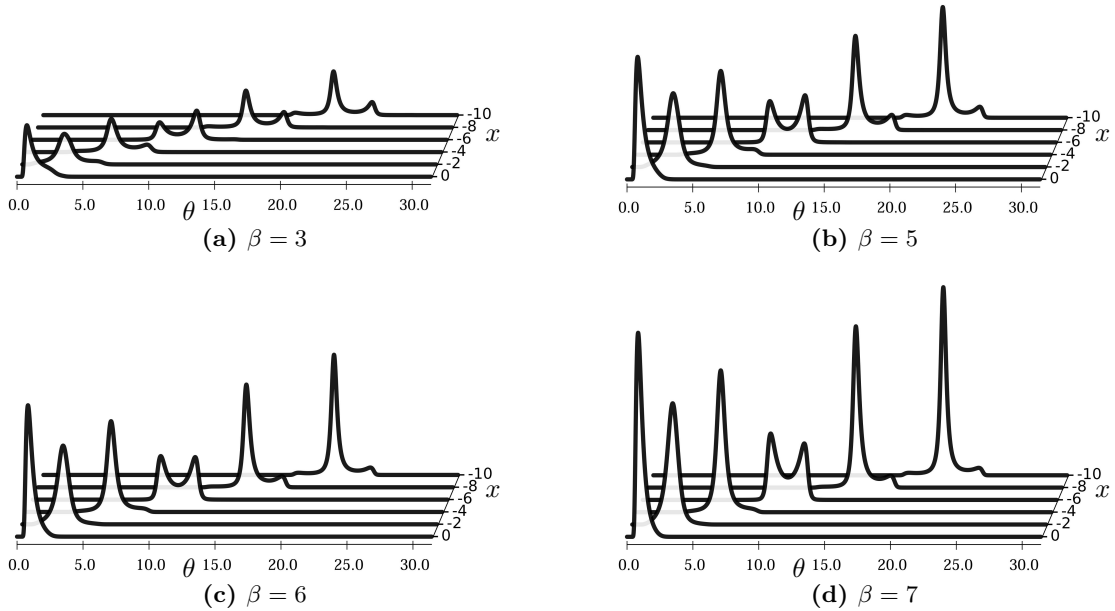

(a) $\beta = 3$

(b) $\beta = 5$

(c) $\beta = 6$

(d) $\beta = 7$

Figure 10. The evolution of $\rho(x, \theta) := \frac{\partial H}{\partial \theta}$ as x decreases from $x = 0$ to -10 with step size $= -2$ for $\beta = 3, 5, 6, 7$. Finite-difference discretization with trapezoidal method is used on (2.11) with $\Delta x = -10^{-3}$ and $M = 10^4$ for $\theta \in [0, 10\pi]$.

(a) $\beta = 3$

(b) $\beta = 5$

(c) $\beta = 6$

(d) $\beta = 7$
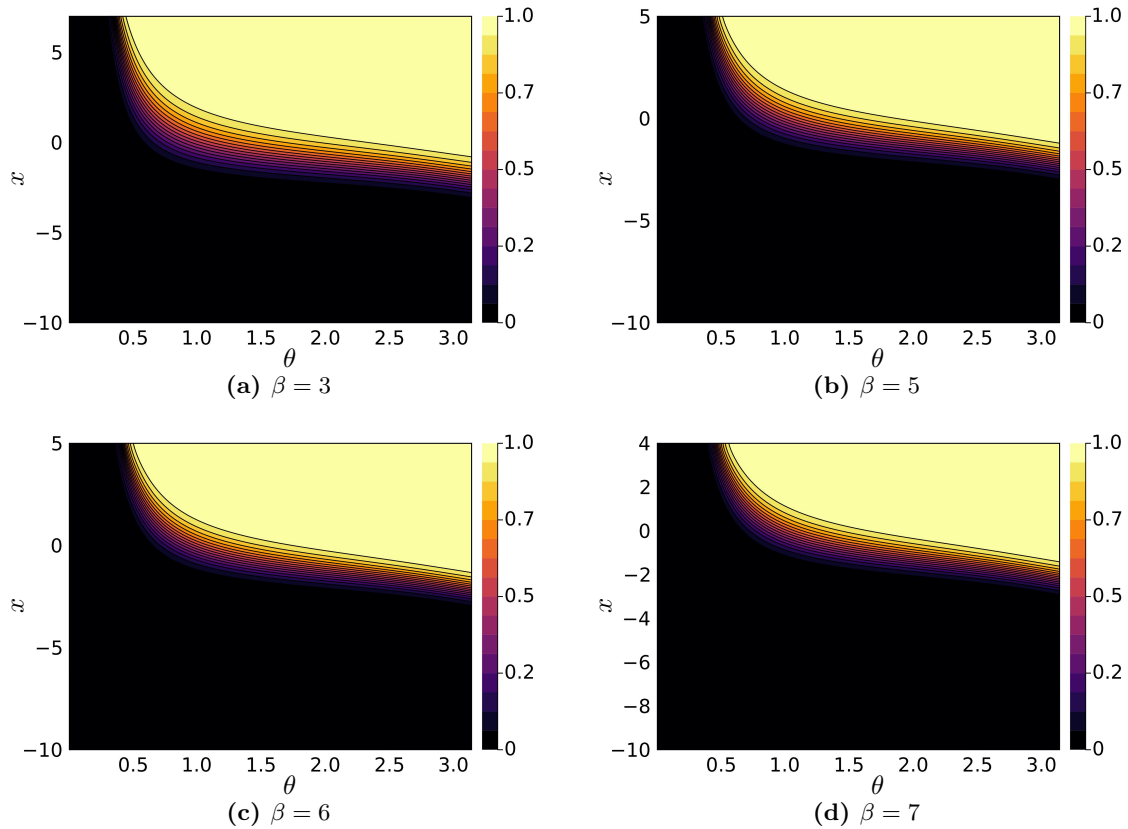
Figure 11. The evolution of $H(x,\theta)$ as x decreases from $\lfloor 13/\sqrt{\beta} \rfloor$ to -10 for $\beta = 3, 5, 6, 7$. Finite-difference discretization with trapezoidal method is used with $\Delta x = -10^{-3}$ and $M = 10^3$.
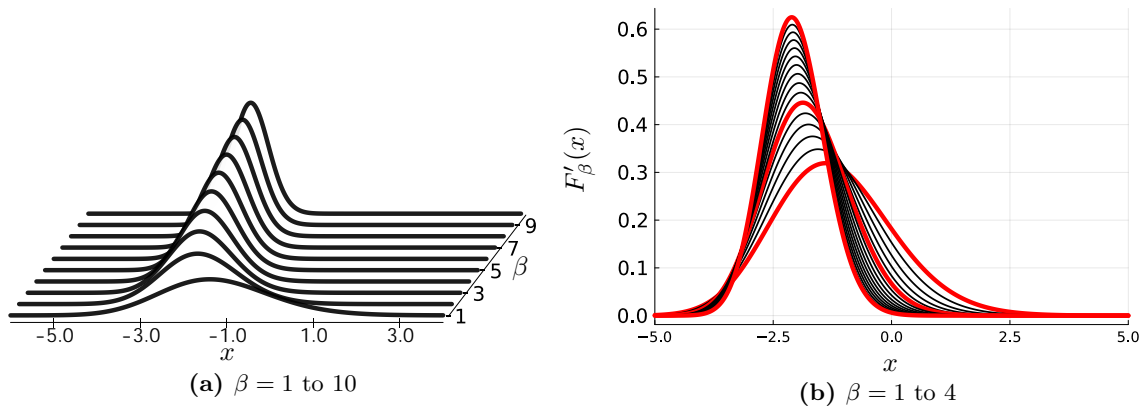

(a) $\beta = 1$ to 10

(b) $\beta = 1$ to 4

Figure 12. Approximation of the density of the Tracy–Widom distribution for different values of β using the finite-difference discretization with trapezoidal method with values for the other parameters as in (3.1).

4.4 Limiting densities of other eigenvalues

By [1, Theorem 2.4.3], one finds the limiting density of the kth largest eigenvalue, after rescaling, of the β-Hermite ensemble H_n^β at $\theta = k\pi$.

Using finite-difference discretization with trapezoidal method with values for the parameters as in (3.1), Figure 13 shows the limiting densities of the largest three eigenvalues of H_n^β, namely, $F'(\cdot, \pi)$, $F'(\cdot, 2\pi)$, and $F'(\cdot, 3\pi)$ from right to left respectively. They can also be interpreted as the densities of $-\Lambda_0$, $-\Lambda_1$, and $-\Lambda_2$ of the stochastic Airy operator $\mathcal{H}_\beta$. As the value of k increases, the limiting distribution becomes more concentrated.
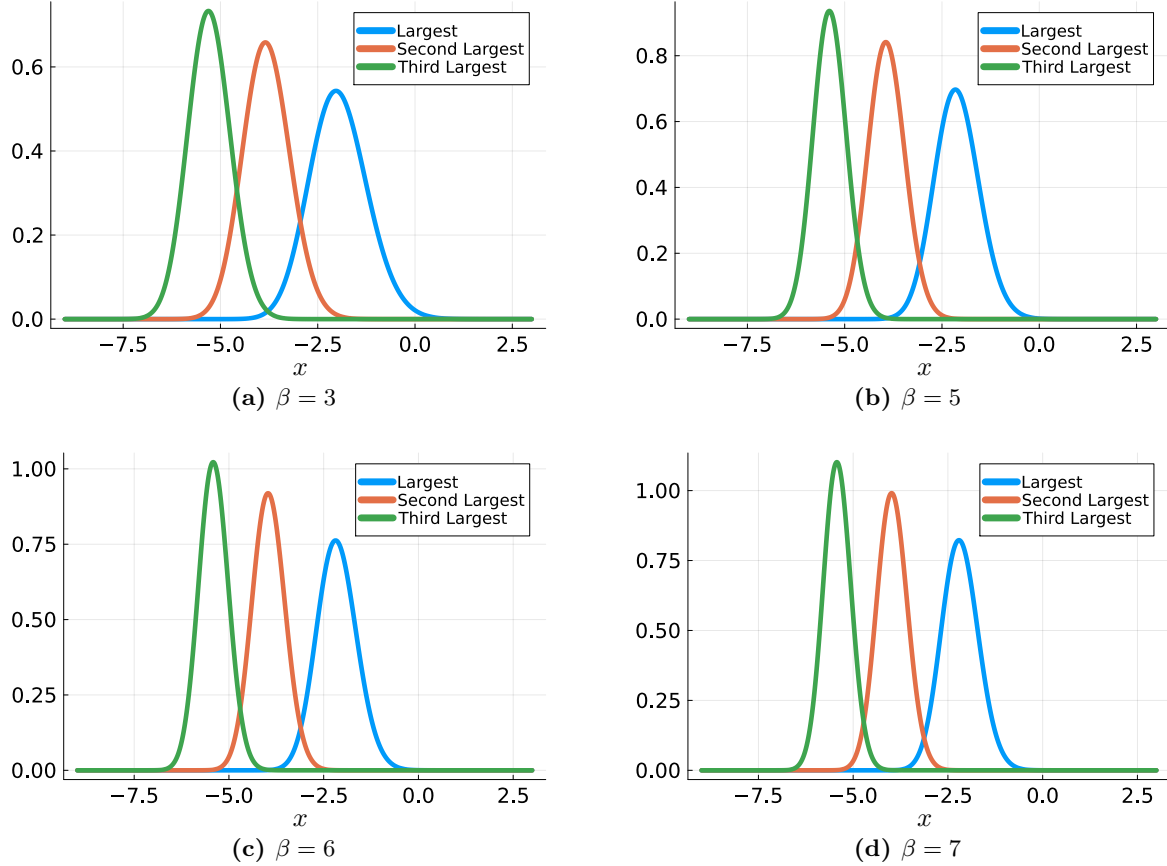
Figure 13. Densities of $-\Lambda_0$, $-\Lambda_1$, and $-\Lambda_2$ of $\mathcal{H}_\beta$ for $\beta = 3, 5, 6, 7$. Finite-difference discretization with trapezoidal method is used with values of the parameters as in (3.1).

5   Outlook and open questions

The error analysis in this paper was empirical, as a proof of convergence is elusive. It is likely that eigenvalue perturbation theory could be used to help show that the spectrum of the x-dependent families of matrices we consider remains close to, or inside of, the stability regions for the time-steppers we have chosen — at least for sufficiently small time steps. A challenge here is to obtain quantitative bounds, making "sufficiently small" precise, and giving useful conclusions. Furthermore, even if the spectrum is understood, it is currently not known how to estimate the eigenvector condition numbers and pseudospectra of these families of matrices.

We also believe it to be possible to achieve better accuracy by adapting the global spectral method of Olver and Townsend [22] and future work will be in this direction. A high-precision implementation of this idea could help corroborate conjectures about the tail behavior of the Tracy–Widom distribution.

A   The range of Δx that causes instability

When applying BDF methods to the spectral discretization, we find that each BDF method has a corresponding range for Δx that causes instability. Moreover, once the value of Δx exits this range, convergence appears to occur at the expected rate. As a result, when an error plot is generated with respect to the value of $|\Delta x|$, we will observe a non-monotonic pattern of
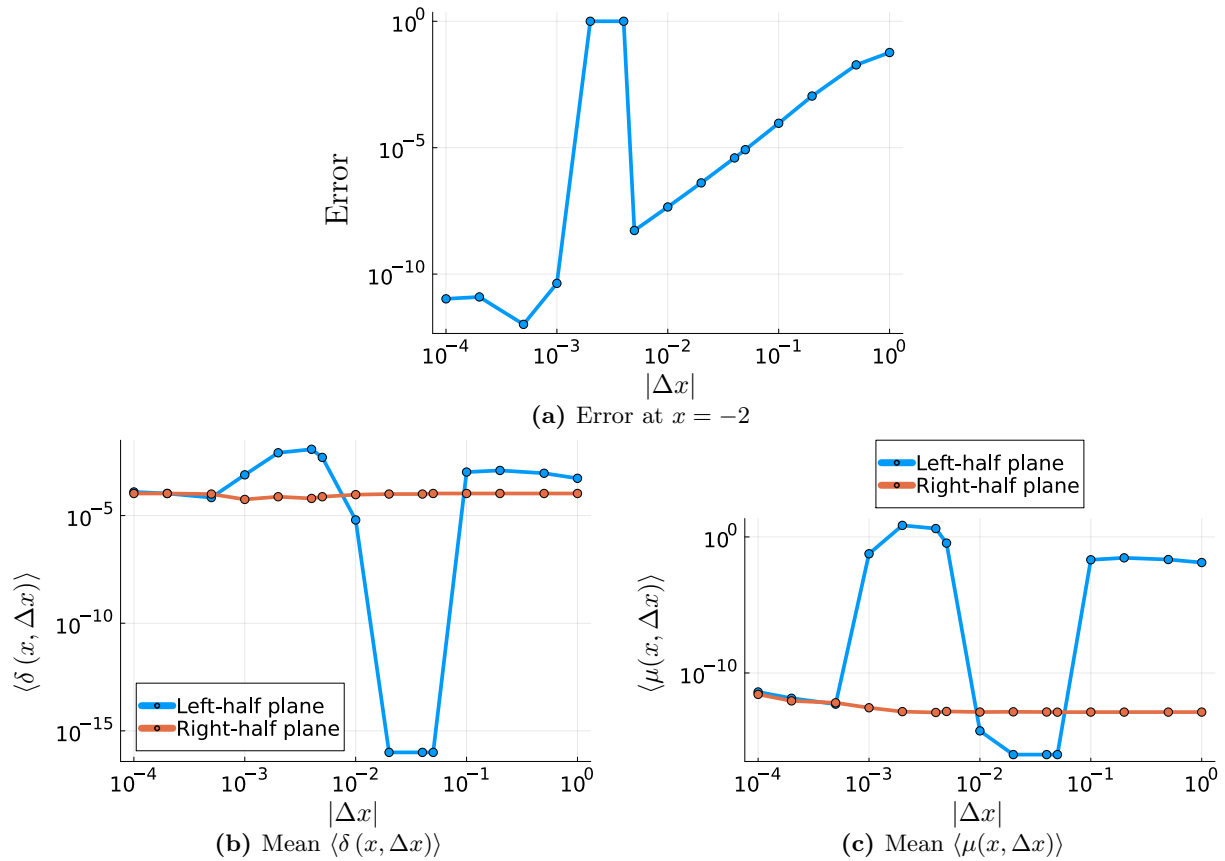
(a) Error at $x = -2$

(b) Mean $\langle \delta(x, \Delta x) \rangle$

(c) Mean $\langle \mu(x, \Delta x) \rangle$

Figure 14. The non-monotonic convergence for BDF3 along with the mean fraction $\langle \delta(x, \Delta x) \rangle$ of the unstable eigenvalues and the mean value $\langle \mu(x, \Delta x) \rangle$. (a) `TW`($\beta$=2; $x_0 = \lfloor 13/\sqrt{\beta} \rfloor$, `method="spectral"`, `step="bdf3"`, `interp=false`, Δx, `M`=8000) is implemented with $\Delta x = -1$, $-0.5, -0.2, -0.1, -0.05, -0.04, -0.02, -0.01, -0.005, -0.004, -0.002, -0.001, -0.0005, -0.0002, -0.0001$. Measured errors are capped at one for readability. (b) The blue (red) dots correspond to the mean fraction $\langle \delta_l(x, \Delta x) \rangle$($\langle \delta_r(x, \Delta x) \rangle$) as defined in (A.3) ((A.4)). (c) The blue (red) dots correspond to the mean value $\langle \mu_l(x, \Delta x) \rangle$($\langle \mu_r(x, \Delta x) \rangle$) as defined in (A.1)((A.2)).

(a) Zoomed-out view
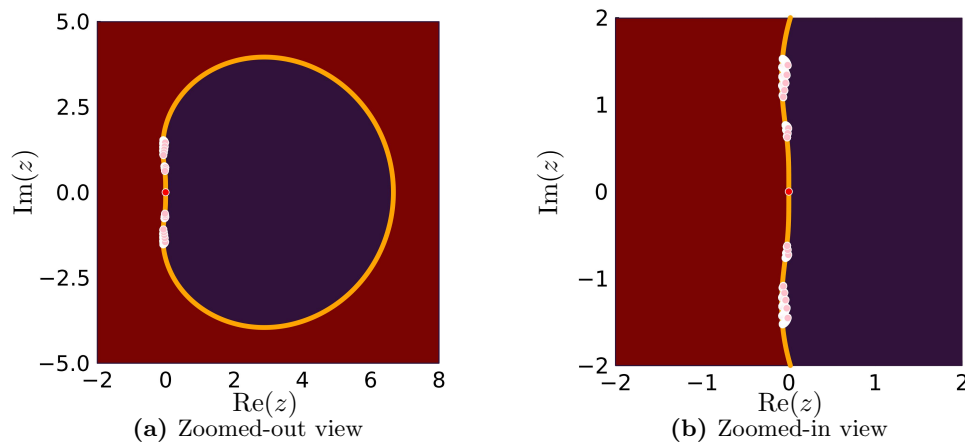

(b) Zoomed-in view

Figure 15. The absolute stability region of BDF3 applying on spectral discretization (the maroon-shaded region along with the orange boundary curve) and the eigenvalues z of $\Delta x(A + xB)$ (the red dots with white boundary correspond to $\text{Re}(z) > 0$, and the pink dots with white boundary correspond to $\text{Re}(z) < 0$) for $\beta = 2$, $x = x_N, -9, \ldots, x_0$, $\Delta x = -0.004, -0.002$, and $M = 8 \times 10^3$.
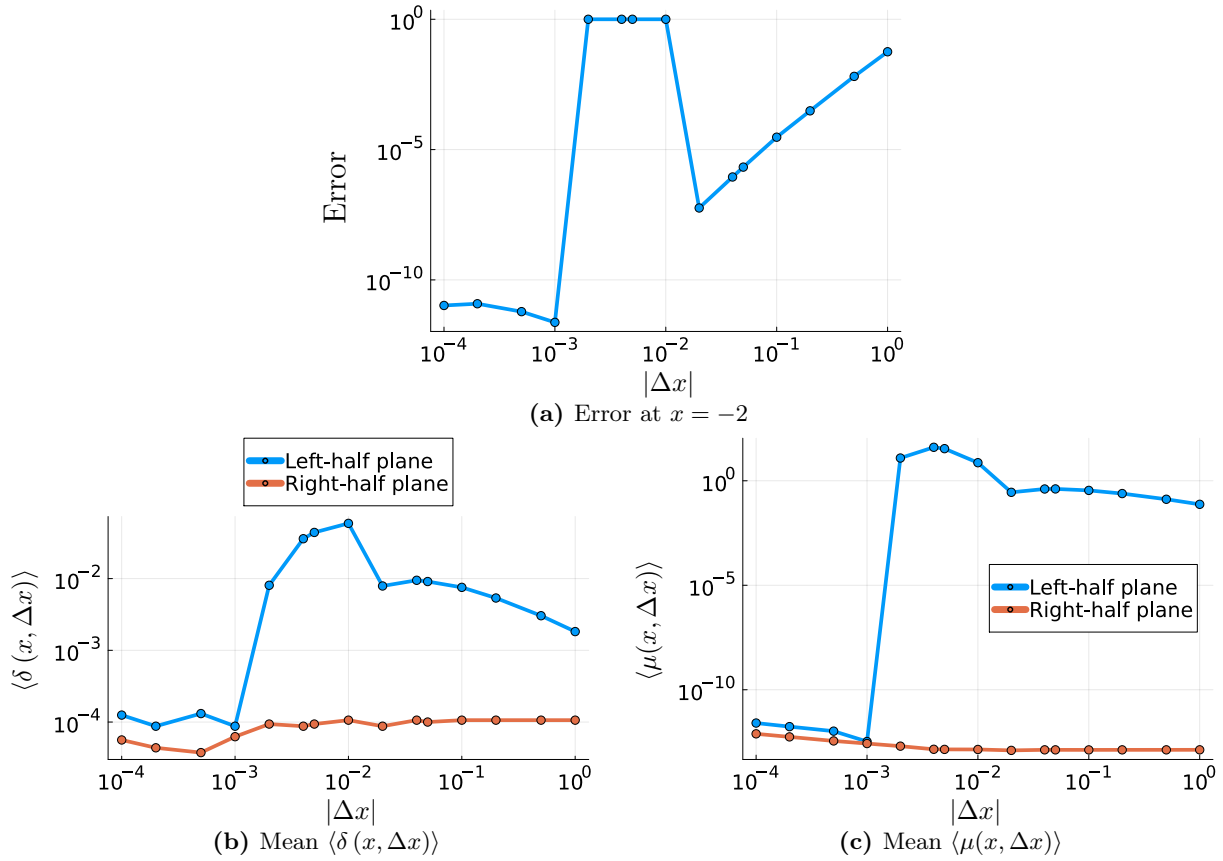
(a) Error at $x = -2$

(b) Mean $\langle \delta \left(x, \Delta x \right) \rangle$

(c) Mean $\langle \mu(x, \Delta x) \rangle$

Figure 16. The non-monotonic convergence for BDF4 along with the mean fraction $\langle \delta \left(x, \Delta x \right) \rangle$ of the unstable eigenvalues and the mean value $\langle \mu(x, \Delta x) \rangle$. The detailed computing process for each plot is the same as for BDF3, see the caption of Figure 14.
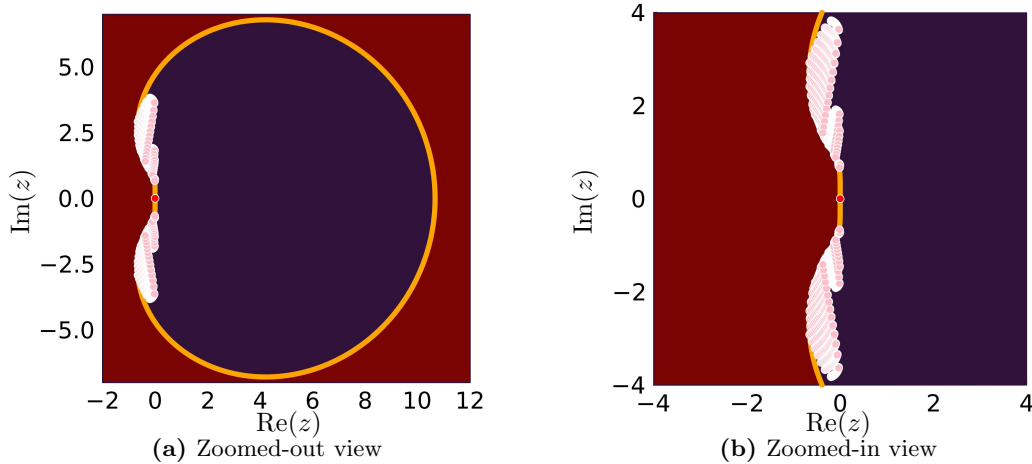

(a) Zoomed-out view

(b) Zoomed-in view

Figure 17. The absolute stability region of BDF4 (the maroon-shaded region along with the orange boundary curve) and the eigenvalues z of $\Delta x(A + xB)$ (the red dots with white boundary correspond to $\mathrm{Re}(z) > 0$, and the pink dots with white boundary correspond to $\mathrm{Re}(z) < 0$ for $\beta = 2$, $x = x_N, -9, \ldots, x_0$, $\Delta x = -0.01, -0.005, -0.004, -0.002$, and $M = 8 \times 10^3$.

convergence. Applying a r-step linear method with the form [12, Section 7.3]

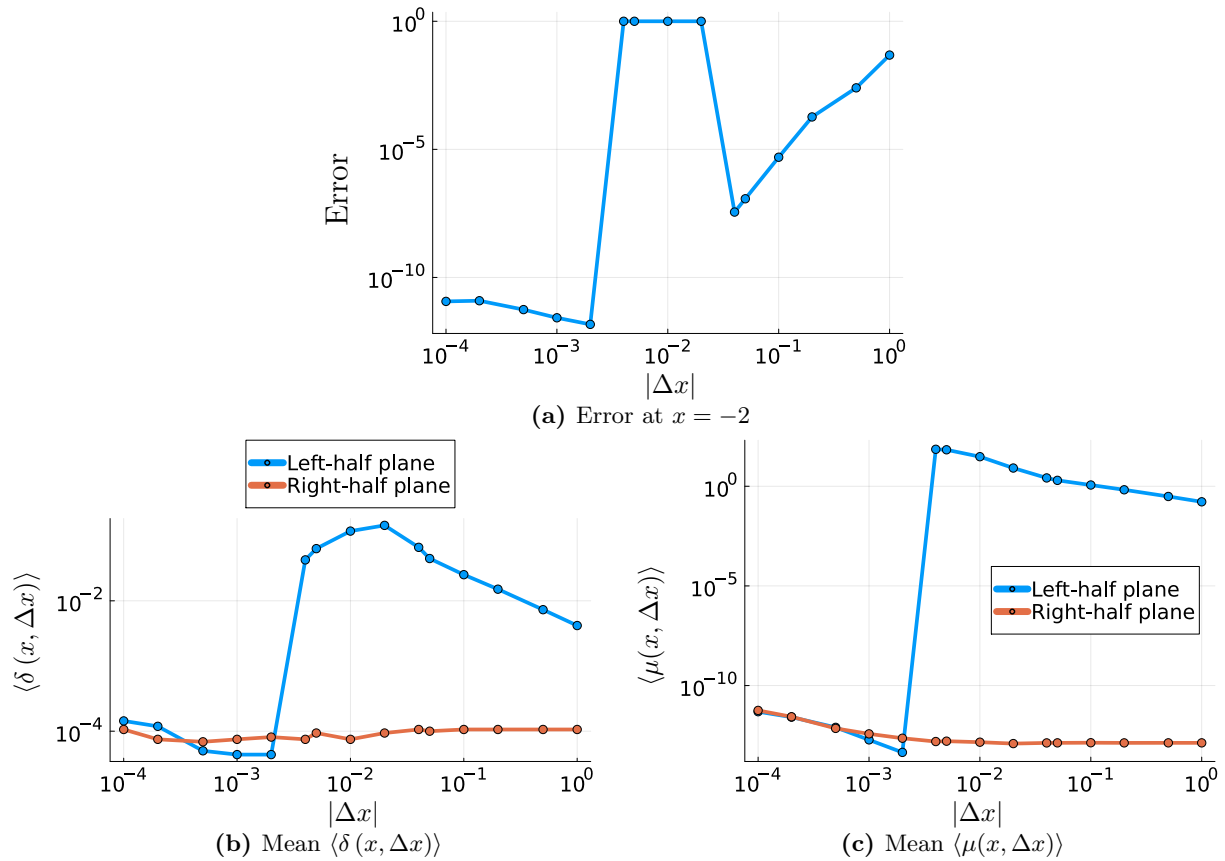$$\sum_{j=0}^{r} \alpha_j U^{n+j} = k \sum_{j=0}^{r} \beta_j f\left(U^{n+j}, t_{n+j} \right)$$

(a) Error at $x = -2$

(b) Mean $\langle \delta \left(x, \Delta x \right) \rangle$

(c) Mean $\langle \mu(x, \Delta x) \rangle$

Figure 18. The non-monotonic convergence for BDF5 along with the mean fraction $\langle \delta \left(x, \Delta x \right) \rangle$ of the unstable eigenvalues and the mean value $\langle \mu(x, \Delta x) \rangle$. The detailed computing process for each plot is the same as for BDF3, see the caption of Figure 14.
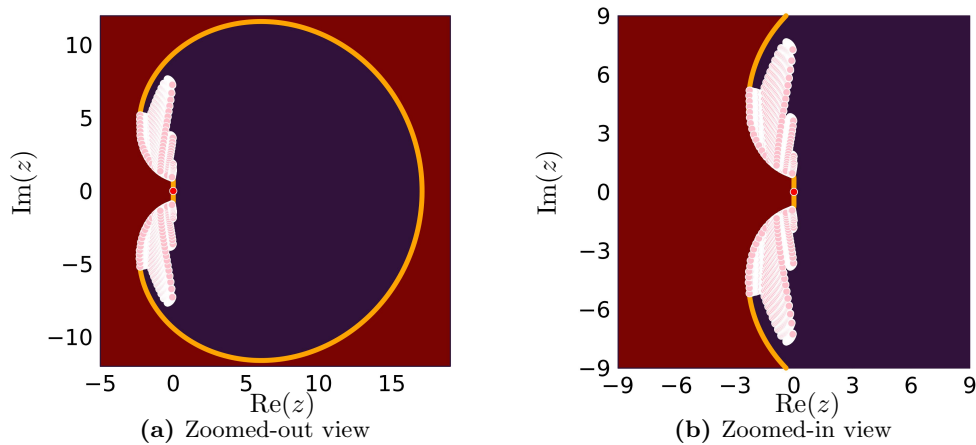

(a) Zoomed-out view

(b) Zoomed-in view

Figure 19. The absolute stability region of BDF5 (the maroon-shaded region along with the orange boundary curve) and the eigenvalues z of $\Delta x(A + xB)$ (the red dots with white boundary correspond to $\mathrm{Re}(z) > 0$, and the pink dots with white boundary correspond to $\mathrm{Re}(z) < 0$ for $\beta = 2$, $x = x_N, -9, \ldots, x_0$, $\Delta x = -0.02, -0.01, -0.005, -0.004$, and $M = 8 \times 10^3$.

to $u' = \lambda u$ gives

$$\sum_{j=0}^{r} \alpha_j U^{n+j} = k \sum_{j=0}^{r} \beta_j \lambda U^{n+j} \implies \sum_{j=0}^{r} (\alpha_j - z\beta_j) U^{n+j} = 0,$$

(a) Error at $x = -2$

(b) Mean $\langle \delta(x, \Delta x) \rangle$

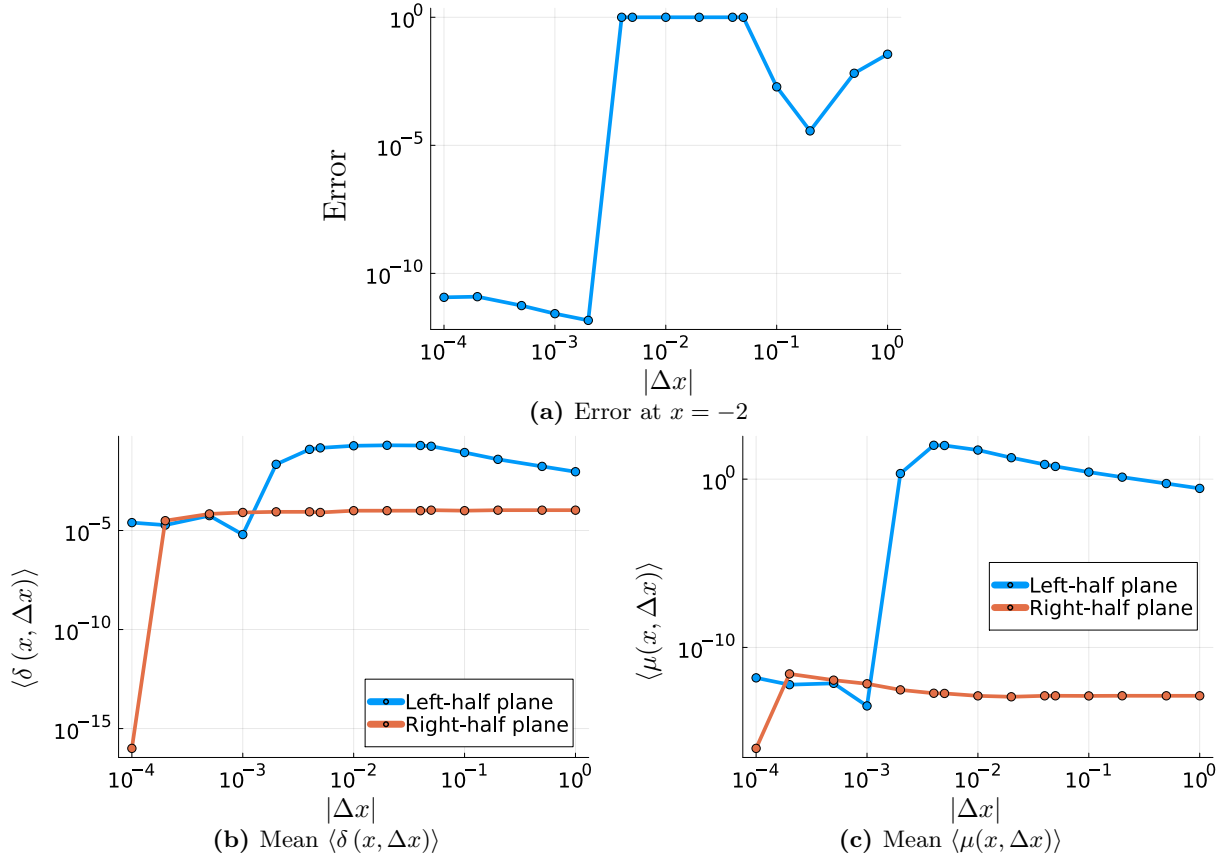**(c)** Mean $\langle \mu(x, \Delta x) \rangle$

Figure 20. The non-monotonic convergence for BDF6 along with the mean fraction $\langle \delta(x, \Delta x) \rangle$ of the unstable eigenvalues and the mean value $\langle \mu(x, \Delta x) \rangle$. The detailed computing process for each plot is the same as for BDF3, see the caption of Figure 14.
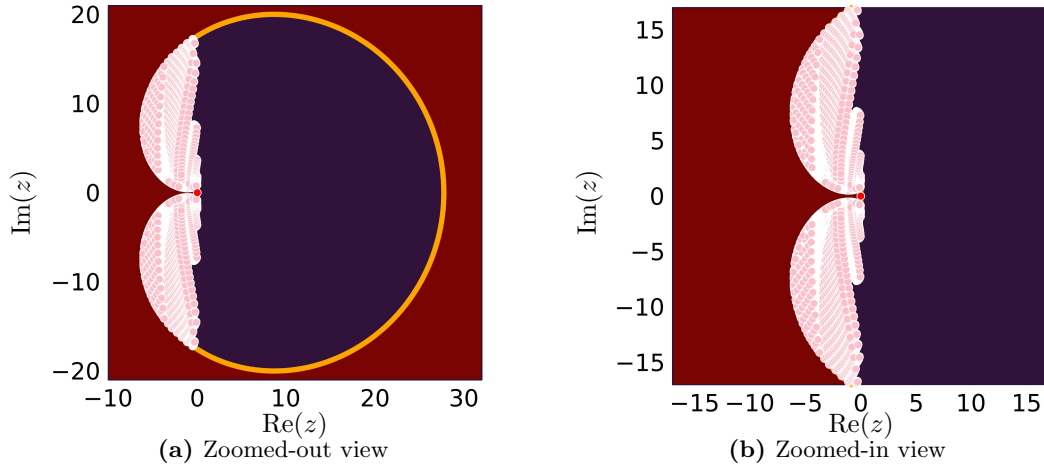

(a) Zoomed-out view

(b) Zoomed-in view

Figure 21. The absolute stability region of BDF6 (the maroon-shaded region along with the orange boundary curve) and the eigenvalues z of $\Delta x (A + x B)$ (the red dots with white boundary correspond to $\mathrm{Re}(z) > 0$, and the pink dots with white boundary correspond to $\mathrm{Re}(z) < 0$) for $\beta = 2$, $x = x_N, -9, \ldots, x_0$, $\Delta x = -0.1, -0.05, -0.04, -0.02, -0.01, -0.005, -0.004$, and $M = 8 \times 10^3$.

where $z \equiv k\lambda$. We call $\sum_{j=0}^{r}(\alpha_j - z\beta_j)\xi^j$ the stability polynomial of this method and denote it by $\pi(\xi; z)$ [12, Section 7.3]. The absolute stability region of a general r-step method consists of the values of z such that the roots $\xi_i(z)$, $i = 1, 2, \ldots, r$ of $\pi(\xi; z)$ satisfy the following two

conditions:

(1) $|\xi_i| \leq 1$ for $i = 1, 2, \ldots, r$,

(2) If ξ_i is a repeated root, then $|\xi_i| < 1$.

See, for example, [12, Definition 6.2] for the definition. To analyze the roots of $\pi(\xi; z)$ for each BDF method as a function of x and Δx, suppose the eigenvalues of $(A + xB)$ are given by $\lambda_1, \lambda_2, \ldots, \lambda_{2M+1}$ and define

$$\mu_l(x, \Delta x) := \max_j \left| \frac{\max_i |\xi_i(\Delta x \lambda_j)| - 1}{\Delta x} \right| \qquad \text{for } \text{Re}(\Delta x \lambda_j) < 0,$$

$$\mu_r(x, \Delta x) := \max_j \left| \frac{\max_i |\xi_i(\Delta x \lambda_j)| - 1}{\Delta x} \right| \qquad \text{for } \text{Re}(\Delta x \lambda_j) > 0.$$

Similarly, define

$$\delta_l(x, \Delta x) := \frac{1}{2M+1} \#\big\{ j \colon \max_i |\xi_i(\Delta x \lambda_j)| > 1, \text{Re}(\Delta x \lambda_j) < 0 \big\},$$

$$\delta_r(x, \Delta x) := \frac{1}{2M+1} \#\big\{ j \colon \max_i |\xi_i(\Delta x \lambda_j)| > 1, \text{Re}(\Delta x \lambda_j) > 0 \big\},$$

where $A + xB$ is the coefficient matrix of the spectral discretization and $\#$ gives the cardinality of the set. The function $\mu_l(x, \Delta x)$ $(\mu_r(x, \Delta x))$ helps capture the maximum growth rate instabilities in the numerical method caused by the eigenvalues of $\Delta x(A + xB)$ in the left(right)-half plane and $\delta_l(x, \Delta x)$ $(\delta_r(x, \Delta x))$ gives the fraction of eigenvalues of $\Delta x(A + xB)$ in the left(right)-half plane that give rise to a positive growth rate.

We then compute the mean value $\langle \mu_l(x, \Delta x) \rangle$, $\langle \mu_r(x, \Delta x) \rangle$, $\langle \delta_l(x, \Delta x) \rangle$, and $\langle \delta_r(x, \Delta x) \rangle$ respectively over $x = -10, -9, \ldots, \lfloor 13/\sqrt{\beta} \rfloor$:

$$\langle \mu_l(x, \Delta x) \rangle := \frac{1}{10 + \lfloor 13/\sqrt{\beta} \rfloor} \sum_{x=-10}^{\lfloor 13/\sqrt{\beta} \rfloor} \mu_l(x, \Delta x), \tag{A.1}$$

$$\langle \mu_r(x, \Delta x) \rangle := \frac{1}{10 + \lfloor 13/\sqrt{\beta} \rfloor} \sum_{x=-10}^{\lfloor 13/\sqrt{\beta} \rfloor} \mu_r(x, \Delta x), \tag{A.2}$$

$$\langle \delta_l(x, \Delta x) \rangle := \frac{1}{10 + \lfloor 13/\sqrt{\beta} \rfloor} \sum_{x=-10}^{\lfloor 13/\sqrt{\beta} \rfloor} \delta_l(x, \Delta x), \tag{A.3}$$

$$\langle \delta_r(x, \Delta x) \rangle := \frac{1}{10 + \lfloor 13/\sqrt{\beta} \rfloor} \sum_{x=-10}^{\lfloor 13/\sqrt{\beta} \rfloor} \delta_r(x, \Delta x). \tag{A.4}$$

Figures 14–21 show the non-monotonic convergence for each BDF method along with the mean fraction $\langle r(x, \Delta x) \rangle$ of the unstable eigenvalues of $\Delta x(A + xB)$, the value of $\langle \mu(x, \Delta x) \rangle$, and the eigenvalues of $\Delta x(A + xB)$ within the unstable range for Δx. On one hand, it is not the unstable eigenvalues in the right-half plane causing the non-monotonic convergence. On the other hand, the range for Δx that causes instability coincides with the range for Δx whose values of $\langle \mu(x, \Delta x) \rangle$ in the left-half plane are larger than $\exp(1)$ and have the largest magnitudes. Moreover, it also coincides with the range for Δx that contains the largest mean fraction $\langle r(x, \Delta x) \rangle$ of the unstable eigenvalues from the left-half plane. Note that for BDF6, though the error at $x = -0.1$ does not blow up, it still suggests that $x = -0.1$ causes instability. For BDF3, the range for Δx that causes instability is approximately $[-0.004, -0.002]$. For BDF4, the range is approximately $[-0.01, -0.002]$. For BDF5, the range is approximately $[-0.02, -0.004]$. For BDF6, the range is approximately $[-0.1, -0.004]$.

B The case when $\beta = \infty$

We consider the limiting behavior of F_β as $\beta \to \infty$. If we let $\beta = \infty$, the original boundary value problem becomes

$$\frac{\partial F}{\partial x} + \left(x - \omega^2\right)\frac{\partial F}{\partial \omega} = 0 \qquad \text{for } (x, \omega) \in \mathbb{R}^2, \tag{B.1}$$

with the following boundary conditions:

$$F(x, \omega) \to 1 \qquad \text{as } x, \omega \to \infty \text{ together,} \tag{B.2}$$

$$F(x, \omega) \to 0 \qquad \text{as } \omega \to -\infty \text{ with } x \text{ bounded above.} \tag{B.3}$$

Using the method of characteristics [11], rewrite (B.1) as

$$\frac{\mathrm{d}F}{\mathrm{d}x} = 0,$$

where

$$\frac{\mathrm{d}F}{\mathrm{d}x} = \frac{\partial F}{\partial x} + \frac{\mathrm{d}\omega}{\mathrm{d}x}\frac{\partial F}{\partial \omega}$$

along curves for which

$$\frac{\mathrm{d}\omega}{\mathrm{d}x} = x - \omega^2. \tag{B.4}$$

Substituting $\omega = u'/u$ in (B.4) yields the Airy equation

$$u'' - xu = 0,$$

whose solution is given by

$$u = c_1 \operatorname{Ai}(x) + c_2 \operatorname{Bi}(x),$$

a linear combination of $\operatorname{Ai}(x)$, the Airy function of the first kind, and $\operatorname{Bi}(x)$, the Airy function of the second kind [18, Chapter 9]. Thus, the solution of (B.4) is given by

$$\omega = \frac{u'}{u} = \frac{c_1 \operatorname{Ai}'(x) + c_2 \operatorname{Bi}'(x)}{c_1 \operatorname{Ai}(x) + c_2 \operatorname{Bi}(x)}, \tag{B.5}$$

which are the characteristic curves of $F(x, \omega)$.

Claim B.1. *There exists a unique characteristic curve for $F(x, \omega)$ along which $\omega \to -\infty$ as $x \to \infty$.*

Proof. For existence, consider the characteristic curves

$$\tilde{\omega}(x) = \frac{\operatorname{Ai}'(x)}{\operatorname{Ai}(x)}, \tag{B.6}$$

which is obtained from (B.5) by setting $c_2 = 0$ and $c_1 \neq 0$. By [18, Section 9.7 (ii)],

$$\lim_{x \to \infty} \tilde{\omega}(x) = \lim_{x \to \infty} \frac{\operatorname{Ai}'(x)}{\operatorname{Ai}(x)} = -\infty. \tag{B.7}$$

Therefore, among the characteristic curves represented by (B.6), there exists one along which $\omega \to -\infty$ as $x \to \infty$.

For uniqueness, note that for fixed values of c_1 and $c_2 \neq 0$, as a well-defined function of x, $\omega \to \infty$ as $x \to \infty$. Thus, combined with (B.7), we conclude that only one such characteristic curve exists. ∎
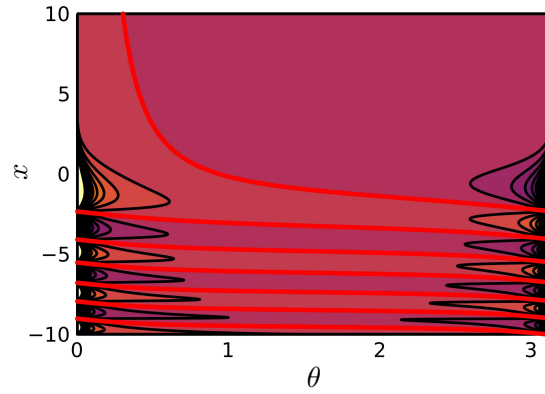
Figure 22. Contour plot of (B.8) with respect to x and θ for $x \in [-10, 10]$ and $\theta \in [0, \pi]$. The red curves correspond to the level zero. The values of x where the red curves cross $\theta = \pi$ coincide with the zeros of $\mathrm{Ai}(x)$.

After the change of variable $\omega = -\cot \theta$, we obtain

$$\cot \theta = -\frac{c_1 \, \mathrm{Ai}'(x) + c_2 \, \mathrm{Bi}'(x)}{c_1 \, \mathrm{Ai}(x) + c_2 \, \mathrm{Bi}(x)}, \tag{B.8}$$

and the boundary conditions become

$$F(x, \theta) \to 1 \qquad \text{as } x \to \infty \text{ and } \theta \to \pi,$$
$$F(x, \theta) \to 0 \qquad \text{as } \theta \to 0 \text{ with } x \text{ bounded above.}$$

Figure 22 shows the contour plot of (B.8) with respect to x and θ for $x \in [-10, 10]$ and $\theta \in [0, \pi]$. The red curves correspond to the level zero, which occurs when $c_1 \neq 0$ and $c_2 = 0$. We can see that there is one and only one red curve along which $\theta \to 0$ as $x \to \infty$. Denote the first real zero (closest to $x = 0$) of $\mathrm{Ai}(x)$ by $x_1 \approx -2.33811$. It can be verified from (B.8) by setting $c_2 = 0$ and $c_1 \neq 0$ that $x \to x_1^+$ as $\theta \to \pi^-$. Set $F(x, \omega)$ in the following way:

$$F(x, \omega) = \begin{cases} 0, & \omega < \dfrac{\mathrm{Ai}'(x)}{\mathrm{Ai}(x)}, \\[2mm] 1, & \omega > \dfrac{\mathrm{Ai}'(x)}{\mathrm{Ai}(x)}. \end{cases} \tag{B.9}$$

Then (B.9) is the solution of (B.1) satisfying both boundary conditions (B.2) and (B.3).

Acknowledgements

References

[1] Bloemendal A., Finite rank perturbations of random matrices and their continuum limits, Ph.D. Thesis, University of Toronto, Canada, 2011.

[2] Bloemendal A., Virág B., Limits of spiked random matrices I, *Probab. Theory Related Fields* **156** (2013), 795–825, arXiv:1011.1877.

[3] Bornemann F., On the numerical evaluation of distributions in random matrix theory: a review, *Markov Process. Related Fields* **16** (2010), 803–866, arXiv:0904.1581.

[4] Bornemann F., On the numerical evaluation of Fredholm determinants, *Math. Comp.* **79** (2010), 871–915, arXiv:0804.2543.

[5] Borot G., Nadal C., Right tail asymptotic expansion of Tracy–Widom beta laws, *Random Matrices Theory Appl.* **1** (2012), 1250006, 23 pages.

[6] Dumaz L., Virág B., The right tail exponent of the Tracy–Widom β distribution, *Ann. Inst. H. Poincaré Probab. Statist.* **49** (2013), 915–933, arXiv:1102.4818.

[7] Dumitriu I., Edelman A., Matrix models for beta ensembles, *J. Math. Phys.* **43** (2002), 5830–5847, arXiv:math-ph/0206043.

[8] Edelman A., Sutton B.D., From random matrices to stochastic operators, *J. Stat. Phys.* **127** (2007), 1121–1165, arXiv:math-ph/0607038.

[9] Ferrari P.L., Spohn H., A determinantal formula for the GOE Tracy–Widom distribution, *J. Phys. A* **38** (2005), L557–L561, arXiv:math-ph/0505012.

[10] Grava T., Its A., Kapaev A., Mezzadri F., On the Tracy–Widom$_\beta$ distribution for $\beta = 6$, *SIGMA* **12** (2016), 105, 26 pages, arXiv:1607.01351.

[11] Kevorkian J., Partial differential equations. Analytical solution techniques, Wadsworth & Brooks/Cole Math. Ser., Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1990.

[12] LeVeque R.J., Finite difference methods for ordinary and partial differential equations. Steady-state and time-dependent problems, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2007.

[13] Li Y., On the open question of the Tracy–Widom distribution of β-ensemble with $\beta = 6$, arXiv:1812.00522.

[14] Mays A., Ponsaing A., Schehr G., Tracy–Widom distributions for the Gaussian orthogonal and symplectic ensembles revisited: a skew-orthogonal polynomials approach, *J. Stat. Phys.* **182** (2021), 28, 55 pages, arXiv:2007.14597.

[15] Mehta M.L., Random matrices, *Pure Appl. Math. (Amsterdam)*, Vol. 142, 3rd ed., Elsevier/Academic Press, Amsterdam, 2004.

[16] Nadal C., Matrices aléatoires et leurs applications à la physique statistique et quantique, Ph.D. Thesis, Université Paris Sud - Paris XI, 2011, available at https://theses.hal.science/tel-00633266.

[17] Nadal C., Majumdar S.N., A simple derivation of the Tracy–Widom distribution of the maximal eigenvalue of a Gaussian unitary random matrix, *J. Stat. Mech. Theory Exp.* **2011** (2011), P04001, 29 pages, arXiv:1102.0738.

[18] Olver F.W.J., Olde Daalhuis A.B., Lozier D.W., Schneider B.I., Boisvert R.F., Clark C.W., Miller B.R., Saunders B.V., Cohl H.S., McClain M.A., NIST digital library of mathematical functions, Release 1.1.8 of 2022-12-15, aviable at http://dlmf.nist.gov/.

[19] Ramírez J.A., Rider B., Virág B., Beta ensembles, stochastic Airy spectrum, and a diffusion, *J. Amer. Math. Soc.* **24** (2011), 919–944, arXiv:math.PR/0607331.

[20] Rumanov I., Painlevé representation of Tracy–Widom$_\beta$ distribution for $\beta = 6$, *Comm. Math. Phys.* **342** (2016), 843–868, arXiv:1408.3779.

[21] Sutton B.D., The stochastic operator approach to random matrix theory, Ph.D. Thesis, Massachusetts Institute of Technology, 2005.

[22] Townsend A., Olver S., The automatic solution of partial differential equations using a global spectral method, *J. Comput. Phys.* **299** (2015), 106–123, arXiv:1409.2789.

[23] Tracy C.A., Widom H., Level-spacing distributions and the Airy kernel, *Phys. Lett. B* **305** (1993), 115–118, arXiv:hep-th/9210074.

[24] Tracy C.A., Widom H., Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* **159** (1994), 151–174, arXiv:hep-th/9211141.

[25] Tracy C.A., Widom H., On orthogonal and symplectic matrix ensembles, *Comm. Math. Phys.* **177** (1996), 727–754, arXiv:solv-int/9509007.

[26] Valkó B., Virág B., Continuum limits of random matrices and the Brownian carousel, *Invent. Math.* **177** (2009), 463–508, arXiv:0712.2000.

[27] Zhang Y., Trogdon T., TracyWidomBeta, 2023, available at https://github.com/Yiting687691/TracyWidomBeta.jl.