SIAM J. NUMER. ANAL.
Vol. 59, No. 5, pp. 2346–2367


# COMPUTATION OF THE COMPLEX ERROR FUNCTION USING MODIFIED TRAPEZOIDAL RULES*

MOHAMMAD AL AZAH† AND SIMON N. CHANDLER-WILDE‡

**Abstract.** In this paper we propose a method for computing the Faddeeva function $w(z) := \mathrm{e}^{-z^2}\mathrm{erfc}(-\mathrm{i}\,z)$ via truncated modified trapezoidal rule approximations to integrals on the real line. Our starting point is the method due to Matta and Reichel [*Math. Comp.*, 25 (1971), pp. 339–344] and Hunter and Regan [*Math. Comp.*, 26 (1972), pp. 339–541]. Addressing shortcomings flagged by Weideman [*SIAM. J. Numer. Anal.*, 31 (1994), pp. 1497–1518], we construct approximations which we prove are exponentially convergent as a function of $N+1$, the number of quadrature points, obtaining error bounds which show that accuracies of $2\times 10^{-15}$ in the computation of $w(z)$ throughout the complex plane are achieved with $N=11$; this is confirmed by computations. These approximations, moreover, provably achieve small relative errors throughout the upper complex half-plane where $w(z)$ is nonzero. Numerical tests suggest that this new method is competitive, in accuracy and computation times, with existing methods for computing $w(z)$ for complex $z$.

**Key words.** trapezoidal rule, complementary error function, Faddeeva function

**AMS subject classifications.** 65D20, 65D30, 30D10, 30E20

**DOI.** 10.1137/20M1373037

## 1. Introduction.

The complementary error function is defined by [18, (7.2.2)]

$$\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty \mathrm{e}^{-t^2}\,dt \tag{1}$$

for all $z = x + \mathrm{i}y$ $(x, y \in \mathbb{R})$. This paper is concerned with approximating $\mathrm{erfc}(z)$ through approximating an integral representation for the related Faddeeva function, defined by [18, (7.2.3)]

$$w(z) := \mathrm{e}^{-z^2}\mathrm{erfc}(-\mathrm{i}\,z). \tag{2}$$

It is well known [3, (7.1.4)] that

$$w(z) = \frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-t^2}}{z-t}\,dt = \frac{\mathrm{i}z}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-t^2}}{z^2-t^2}\,dt, \quad \mathrm{Im}(z) > 0, \tag{3}$$

and this is our starting point. It is sufficient to devise methods to compute $w(z)$ for $z$ in the first quadrant since values in the other quadrants can be obtained using the symmetries [19, (3.1) and (3.2)]

$$w(-z) = 2\mathrm{e}^{-z^2} - w(z) \quad \text{and} \quad w(\overline{z}) = \overline{w(-z)}. \tag{4}$$

It follows from (2)–(3) that

$$\mathrm{erfc}(z) = \frac{z\,\mathrm{e}^{-z^2}}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-t^2}}{z^2+t^2}\,dt, \quad x = \mathrm{Re}(z) > 0. \tag{5}$$

*Received by the editors October 12, 2020; accepted for publication (in revised form) June 23, 2021; published electronically September 9, 2021.
https://doi.org/10.1137/20M1373037

†School of Social and Basic Sciences, Al Hussein Technical University (HTU), Amman, Jordan (mohammad.azah@htu.edu.jo).

‡Department of Mathematics and Statistics, University of Reading, Whiteknights, P.O. Box 220, Reading RG6 6AX, UK (S.N.Chandler-Wilde@reading.ac.uk).

Starting from this integral representation, Chiarella and Reichel [7] and Matta and Reichel [15] showed, by the contour integration arguments that we recall in section 2.1, that, for $x > 0$,

$$\operatorname{erfc}(z) = \frac{hz\mathrm{e}^{-z^2}}{\pi} \sum_{k=-\infty}^{\infty} \frac{\mathrm{e}^{-k^2h^2}}{z^2 + k^2h^2} + \frac{2\,H(\pi/h - x)}{1 - \mathrm{e}^{2\pi z/h}} + E(h). \tag{6}$$

Here the first term is the trapezoidal rule approximation to (5) using a step size $h > 0$, the second is a modification that arises from Cauchy's residue theorem (expressed using the standard Heaviside step function $H$, with $H(0) = 1/2$), and $E(h)$ is a small error term. Hunter and Regan [12] (correcting the argument in [7, 15]) show, for $x > 0$ with $x \neq \pi/h$, that

$$|E(h)| \leq \frac{2|z\mathrm{e}^{-z^2}|\,\mathrm{e}^{-\pi^2/h^2}}{\pi^{1/2}(1 - \mathrm{e}^{-2\pi^2/h^2})|x^2 - \pi^2/h^2|}. \tag{7}$$

Thus, for every fixed $z = x + \mathrm{i}y$ with $x > 0$, the modified trapezoidal rule approximation obtained by neglecting $E(h)$ in (6) is very rapidly convergent indeed as $h \to 0$.

As the bound (7) suggests, neglecting the error term $E(h)$ in (6) gives a very accurate approximation also for $x = \operatorname{Re}(z) = 0$, except that the approximation is undefined if $z = \mathrm{i}kh$ for some $k \in \mathbb{Z}$, and there are stability issues in evaluation if $z$ is close to one of these points. It is suggested in [12] to solve this issue by switching to the composite midpoint rule where needed. Precisely, Hunter and Regan [12] propose using the formula (6) (neglecting the error term $E(h)$) if $1/4 \leq \varphi(y/h) \leq 3/4$, where $y = \operatorname{Im}(z)$ and

$$\varphi(t) := t - \lfloor t \rfloor \in [0, 1) \tag{8}$$

is the fractional part of $t$. Otherwise they suggest using the midpoint rule–based formula

$$\operatorname{erfc}(z) = \frac{hz\mathrm{e}^{-z^2}}{\pi} \sum_{k=-\infty}^{\infty} \frac{\mathrm{e}^{-(k+1/2)^2h^2}}{z^2 + (k+1/2)^2h^2} + \frac{2\,H(\pi/h - x)}{1 + \mathrm{e}^{2\pi z/h}} + E'(h), \tag{9}$$

neglecting the corresponding error term $E'(h)$ which they show satisfies the same bound (7) as $E(h)$.

These proposals from [12] are our starting point. In a practical implementation the sums in (6) and (9) must be truncated, say to $-N \leq k \leq N$. The contributions of this paper are to

(i) convert a modified version of the proposals of Hunter and Regan [12] into a fully specified algorithm, making clear how the choice of $h > 0$ should be related to $N$ for optimal accuracy;

(ii) provide error estimates for the approximations we propose for $w(z)$, proving that the maximum absolute error (and the maximum relative error in the upper half-plane) decrease exponentially with $N$, reducing by a factor $\mathrm{e}^{\pi} \approx 23.1$ for each additional quadrature point;

(iii) demonstrate that the claimed exponential convergence in absolute and relative errors is achieved numerically;

(iv) present numerical results that suggest that the simple approximation formulae we propose are competitive in accuracy and computation time with existing methods for computing $w(z)$.

In carrying out (i) we are addressing earlier criticisms of the algorithms in [7, 15, 12] made by Weideman [25], who observes that the formula (6) with the summation reduced to $-N \le k \le N$

> *is very accurate, provided for given $z$ and $N$ the optimal step size $h$ is selected. It is not easy, however, to determine this optimal $h$ a priori.*

Our recommendations address this issue, detailing which approximation formula and what step size $h$ should be used for each $N$ and $z$. (It turns out that, to obtain high accuracy, the approximation formula chosen must depend on both $z$ and $N$, but $h$ only needs to depend on $N$.)

The bounds we obtain in carrying out (ii) prove that the absolute error in our approximation for $w(z)$ tends to zero exponentially with $N$, uniformly in the complex plane. This is a substantial improvement on the existing bound (7), which blows up when $x = \pi/h$, and does not capture the additional truncation errors due to replacing infinite sums by finite sums in the approximations (6) and (9).

Concretely, our proposed approximation to $w(z)$, for $z = x + \mathrm{i}y$, with $x, y \ge 0$, is

$$w_N(z) := \begin{cases} w_N^{\mathrm{M}}(z) & \text{if } y \ge \max(x, \pi/h), \\ w_N^{\mathrm{MT}}(z) & \text{if } y < x \quad \text{and} \quad 1/4 \le \varphi(x/h) \le 3/4, \\ w_N^{\mathrm{MM}}(z) & \text{otherwise,} \end{cases} \tag{10}$$

where $\varphi$ is defined by (8), $N \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$,

$$h := \sqrt{\pi/(N+1)}, \tag{11}$$

$$w_N^{\mathrm{M}}(z) := \frac{2\mathrm{i}\,h\,z}{\pi} \sum_{k=0}^{N} \frac{\mathrm{e}^{-t_k^2}}{z^2 - t_k^2}, \tag{12}$$

$$w_N^{\mathrm{MM}}(z) := \frac{2\,\mathrm{e}^{-z^2}}{1 + \mathrm{e}^{-2\mathrm{i}\pi z/h}} + w_N^{\mathrm{M}}(z), \tag{13}$$

$$w_N^{\mathrm{MT}}(z) := \frac{2\,\mathrm{e}^{-z^2}}{1 - \mathrm{e}^{-2\mathrm{i}\pi z/h}} + \frac{\mathrm{i}\,h}{\pi z} + \frac{2\mathrm{i}\,h\,z}{\pi} \sum_{k=1}^{N} \frac{\mathrm{e}^{-\tau_k^2}}{z^2 - \tau_k^2}, \tag{14}$$

$$t_k := (k + 1/2)h, \quad \text{and} \quad \tau_k := kh. \tag{15}$$

We extend the approximation to the full complex plane by using the symmetries (4)—precisely by defining

$$w_N(z) := \overline{w_N(-\overline{z})} \text{ if } y \ge 0 \text{ and } x < 0, \quad w_N(z) := 2\mathrm{e}^{-z^2} - w_N(-z) \text{ if } y < 0. \tag{16}$$

We supply in Table SM1 of the supplementary materials the MATLAB code implementing the approximation $w_N(z)$ that we use for the computations in section 4.[1]

The main error estimate that we prove, using standard complex analysis arguments including a Phragmén–Lindelöf principle, is the following.

THEOREM 1.1. *Suppose $w_N(z)$ is given by* (10) *and* (16). *Then, for $N \in \mathbb{N}_0$ and $z = x + \mathrm{i}y$,*

$$|w(z) - w_N(z)| \le C_1\,\mathrm{e}^{-\pi N} \quad \textit{for all } x, y \in \mathbb{R}, \quad \textit{and}$$
$$\frac{|w(z) - w_N(z)|}{|w(z)|} \le C_2\sqrt{N+1}\,\mathrm{e}^{-\pi N} \quad \textit{if } x \in \mathbb{R} \textit{ and } y \ge 0,$$

[1]The codes in these supplementary materials are also available at github; see https://github.com/sms03snc/Faddeeva.

*where* $C_1 \approx 0.67$ *is given by* (75) *and* $C_2 \approx 3.97$ *by* (76).

We remark that approximation of $w(z)$, for $z \in \mathbb{C}$, provides an effective route to the computation of other special functions, including Fresnel integrals (e.g., [5]), Dawson's integral [18, equation (7.5.1)], and the Voigt functions [18, equation (7.19.3)]. Indeed, we have previously used, in the restricted case $\arg(z) = \pi/4$, an approximation resembling $w_N^{\mathrm{MM}}(z)$ when approximating Fresnel integrals [5], proving results in the spirit of Theorem 1.1.

Let us summarize the rest of the paper. In the largest section, section 2, we derive the above formulae and error bounds. In section 3 we review the existing, alternative approximate methods for computing $\mathrm{erfc}(z)$ and $w(z)$ for complex $z$. For none of these has an error bound been proved, similar to Theorem 1.1. In section 4 we carry out numerical experiments that confirm the accuracy of $w_N(z)$, showing that its absolute error is $< 2 \times 10^{-15}$ throughout the complex plane with $N = 11$, and that the same bound holds for the relative error in the upper half-plane. We also show that our new approximation is competitive in accuracy and appears to be competitive in computing times with the methods that we survey in section 3, specifically those of [25, 29, 27, 2].

We note that this paper is based, in significant part, on Chapter 3 of the first author's thesis [4].

**2. The proposed approximation and its error bounds.** In this section we derive the approximation given by (10) based on modified trapezoidal rules. We also derive the error bounds of Theorem 1.1 that demonstrate that the absolute and relative errors in $w_N(z)$ both decrease exponentially as $N$ increases.

**2.1. The contour integral argument and its history.** Given any $f \in C(\mathbb{R})$ that decays sufficiently rapidly at infinity, let

$$I[f] := \int_{-\infty}^{\infty} f(t)\, dt,$$

and, for $h > 0$ and $\alpha \in [0, 1)$, define the *generalized trapezoidal rule approximation* to $I[f]$ by

$$I_{h,\alpha}[f] := h \sum_{k \in \mathrm{Z}} f((k - \alpha)h). \tag{17}$$

We note that $I_{h,\alpha}[f] = I_{h,0}[f_\alpha]$, where $f_\alpha(t) := f(t - \alpha h)$ for $t \in \mathbb{R}$, and that $I_{h,0}[f]$ is the trapezoidal rule approximation to $I[f]$ and $I_{h,1/2}[f]$ its composite midpoint rule approximation.

The approximation (17) for $I[f]$ converges exponentially when the integrand is analytic in a strip surrounding the real axis and has sufficient decay at $\pm\infty$. The derivation of this result, using contour integration and Cauchy's residue theorem, dates back, for a particular case, at least to Turing [24], and has been analyzed in more general cases by Goodwin [11], Davis [9], McNamee [16], Schwartz [21], and Stenger [22]. For a detailed history and discussion see Trefethen and Weideman [23].

The rate of exponential convergence depends on the width of the strip of analyticity around the real axis, and the accuracy of $I_{h,\alpha}[f]$ deteriorates when $f$ has singularities close to the real line. But, in the case when these singularities are poles, the contour integral method for establishing the exponential convergence of $I_{h,\alpha}[f]$, that we will recall in Proposition 2.1 below, leads naturally to corrections for modifying the trapezoidal rule and recovering rapid convergence; these corrections are

expressed in terms of residues of $f$ at these poles. This appears to have been observed explicitly first by Chiarella and Reichel [7], in the context of evaluating (5) (see also Matta and Reichel [15], Hunter and Regan [12], and Mori [17]), and has been developed into a general theory by Bialecki [6] (see also La Porte [14]).

It is convenient to recall in a proposition the standard arguments (see [7, 15, 12] and cf. [23, pp. 402–403]) that are made to prove exponential convergence, since we will use these arguments below. These use, for given $h > 0$ and $\alpha \in [0, 1)$, the function $g(\zeta)$ defined by

$$g(\zeta) := \mathrm{i} \cot\left(\pi\left(\alpha + \zeta/h\right)\right). \tag{18}$$

This is a meromorphic function with simple poles at $\zeta = (k - \alpha)h$, $k \in \mathbb{Z}$, and the properties that, for $\zeta \in \mathbb{C}$ with $\eta = \mathrm{Im}(\zeta)$,

$$|1 - g(\zeta)| \le \frac{2e^{-2\pi\eta/h}}{1 - e^{-2\pi\eta/h}} \quad \text{if } \eta > 0, \quad |1 + g(\zeta)| \le \frac{2e^{2\pi\eta/h}}{1 - e^{2\pi\eta/h}} \quad \text{if } \eta < 0. \tag{19}$$

We also use, for $H \in \mathbb{R}$, the notation $\Gamma_H$ for the path $\{t + \mathrm{i}H : t \in \mathbb{R}\}$ traversed in the direction of increasing $t$. It is enough for our purposes to suppose that the poles of $f$ are simple. For the case of poles of arbitrary order see [6, Theorem 2.2].

PROPOSITION 2.1. *Suppose that, for some $H > 0$, $f$ is analytic in the strip $S_H := \{\zeta \in \mathbb{C} : |\mathrm{Im}(\zeta)| < H\}$, except for a finite number of simple poles at $\zeta_1, \ldots, \zeta_m \in S_H$, with $\eta_k := \mathrm{Im}(\zeta_k) \neq 0$, for $k = 1, \ldots, m$. Suppose also that $f$ is continuous in $\overline{S_H} \setminus \{\zeta_1, \ldots, \zeta_m\}$ and that, for some $r > 1$, $f(\zeta) = O(|\zeta|^{-r})$ as $|\mathrm{Re}(\zeta)| \to \infty$, uniformly in $S_H$. Then*

$$I[f] - I_{h,\alpha}[f] = \frac{1}{2}\left(\int_{\Gamma_H} f(\zeta)(1 - g(\zeta))\, d\zeta + \int_{\Gamma_{-H}} f(\zeta)(1 + g(\zeta))\, d\zeta\right) + C_{h,\alpha,H}[f],$$

*where*

$$C_{h,\alpha,H}[f] := \pi\mathrm{i} \sum_{k=1}^{m} \left(\mathrm{sign}(\eta_k) - g(\zeta_k)\right) R_k,$$

*and $R_k := \mathrm{Res}(f, \zeta_k) = \lim_{\zeta \to \zeta_k}(\zeta - \zeta_k) f(\zeta)$ denotes the residue of $f$ at $\zeta_k$.*

*Proof.* Let $A_k := \left(k - \alpha + \frac{1}{2}\right) h$ for $k \in \mathbb{Z}$. Let $C_{\widetilde{H},i,j}$ denote the positively oriented rectangular contour with corners at $A_i \pm \mathrm{i}\widetilde{H}$ and $A_j \pm \mathrm{i}\widetilde{H}$, choosing $\widetilde{H} \in (0, H)$ and the integers $i < 0$ and $j > 0$ so that $C_{\widetilde{H},i,j}$ encloses the poles $\zeta_1, \ldots, \zeta_m$ (see Figure 1). Noting that $\mathrm{Res}(g, (k-\alpha)h) = \mathrm{i}h/\pi$, for $k \in \mathbb{Z}$, we apply Cauchy's residue theorem to

$$\int_{C_{\widetilde{H},i,j}} f(\zeta) g(\zeta)\, d\zeta.$$

We take the limit $\widetilde{H} \to H$, and then the limit as $i \to -\infty$ and $j \to +\infty$ which leads (see [23, pp. 402–403] for more detail) to

$$\int_{\Gamma_{-H}} f(\zeta) g(\zeta)\, d\zeta - \int_{\Gamma_H} f(\zeta) g(\zeta)\, d\zeta = -2I_{h,\alpha}[f] + 2\pi\mathrm{i} \sum_{k=1}^{m} g(\zeta_k) R_k.$$

Making similar applications of Cauchy's residue theorem, we obtain also that

$$\int_{\Gamma_H} f(\zeta)\, d\zeta + \int_{\Gamma_{-H}} f(\zeta)\, d\zeta = 2I[f] - 2\pi\mathrm{i} \sum_{k=1}^{m} \mathrm{sign}(\eta_k) R_k,$$

and the result follows. ∎

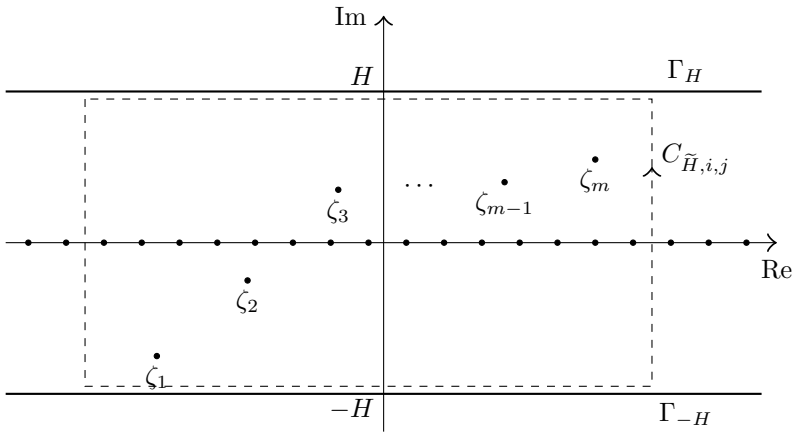


FIG. 1. *The contour $C_{\widetilde{H},i,j}$ used in the proof of Proposition* 2.1. *The dots on the real line are the poles of $g(\zeta)$ at $(k-\alpha)h$ for $k \in \mathbb{Z}$.*

Thanks to the bounds (19) it follows from the above proposition that

$$|I[f] - I^*_{h,\alpha,H}[f]| \le \frac{e^{-2\pi H/h}}{1 - e^{-2\pi H/h}} \int_{\Gamma_H} (|f(\zeta)| + |f(-\zeta)|)\ |d\zeta|, \tag{20}$$

where

$$I^*_{h,\alpha,H}[f] := I_{h,\alpha}[f] + C_{h,\alpha,H}[f] \tag{21}$$

is what we will call the *modified generalized trapezoidal rule*. If $f$ is analytic in $S_H$, so that $I^*_{h,\alpha,H}[f] = I_{h,\alpha}[f]$, this bound reduces to [23, Theorem 5.1] and proves that $I_{h,\alpha}[f]$ is exponentially convergent as $h \to 0$. In the more general case that $f$ has simple pole singularities in $S_H$, the bound (20) proves the same rate of exponential convergence for $I^*_{h,\alpha,H}[f]$, the trapezoidal rule having been modified to take into account these poles of $f$.

Our application of the above proposition and bound will be to the integrals given by (5) and (3). In these cases (cf. Goodwin [11]) we have additionally that

$$f(\zeta) = \mathrm{e}^{-\zeta^2} F(\zeta), \tag{22}$$

where $F$ is even and $F(\zeta) = O(1)$ as $|\mathrm{Re}(\zeta)| \to \infty$, uniformly in $S_H$. This satisfies the conditions of the above proposition, and, since $|\exp(-\zeta^2)| = \mathrm{e}^{H^2 - t^2}$ for $\zeta = t + \mathrm{i}H$, and $\int_{-\infty}^{\infty} \exp(-t^2)\, dt = \sqrt{\pi}$, the bound (20) implies in this case that

$$\left|I[f] - I^*_{h,\alpha,H}[f]\right| \le \frac{2\mathrm{e}^{H^2 - 2\pi H/h}}{1 - \mathrm{e}^{-2\pi H/h}} \int_{-\infty}^{\infty} \mathrm{e}^{-t^2} |F(t + \mathrm{i}H)|\, dt \le \frac{2\sqrt{\pi} M \mathrm{e}^{H^2 - 2\pi H/h}}{1 - \mathrm{e}^{-2\pi H/h}}, \tag{23}$$

where

$$M := \sup_{t \in \mathbb{R}} |F(t + \mathrm{i}H)|. \tag{24}$$

An important observation, particularly when applying the above bound in cases where $F$ is meromorphic in the whole complex plane, is that the exponent $H^2 -$

$2\pi H/h$ is minimized by the choice $H = \pi/h$, in which case $H^2 - 2\pi H/h = -\pi^2/h^2$. The representations (6) and (9), with the bound (7) (which bounds both $|E(h)|$ and $|E'(h)|$), follow from applying (23) to (5). Precisely, we apply (23) with $H = \pi/h$ and $\alpha = 0$ and $\alpha = 1/2$, in the respective cases (6) and (9), noting that (5) can be written as $\operatorname{erfc}(z) = I[f]$ with $f$ given by (22) and $F(t) := z\exp(-z^2)/(\pi(z^2 + t^2))$. To obtain the bound (7) from (23) and (24) we note further that, for $z = x + \mathrm{i}y$ and $\zeta = t + \mathrm{i}\pi/h$, $|z^2 + \zeta^2| = |z + \mathrm{i}\zeta|\,|z - \mathrm{i}\zeta| \geq |x - \pi/h||x + \pi/h| = |x^2 - \pi^2/h^2|$. (We remark that the blow-up when $x = \pi/h$ in the bound (7) is an artifact of the method of argument that the contour $\Gamma_H$ with $H = \pi/h$ passes through a pole of the function $F$ when $x = \pi/h$.)

**2.2. Trapezoidal rule error estimates.** In this subsection we apply the above methods and bounds to obtain uniform absolute and relative error estimates as a replacement for the bound (7). Except where explicitly indicated otherwise, all the results we prove hold for all $h > 0$ and all $\alpha \in [0, 1)$. In the estimates we obtain in Proposition 2.2 we avoid the equivalent of the blow-up seen in (7) when $x = \pi/h$ by making two separate applications of the contour integral argument, with $H = \pi/h \pm \epsilon$. It turns out that the $\epsilon$ correction to the optimal choice $H = \pi/h$ incurs only an $O(1)$ factor correction to the error bound as long as $\epsilon$ is fixed independently of $h$. At the same time, if we take a fixed $\epsilon$ and switch as needed between $H = \pi/h + \epsilon$ and $H = \pi/h - \epsilon$, we can ensure that $\Gamma_H$ remains at least an $O(1)$ distance from the poles of the integrand as we vary $z$.

For $z = x + \mathrm{i}y$ with $y > 0$ we write (3) as $w(z) = I[f_z]$ where

$$f_z(t) := e^{-t^2} F_z(t) \quad \text{and} \quad F_z(t) := \frac{\mathrm{i}\,z}{\pi(z^2 - t^2)}. \tag{25}$$

The even function $f_z$ is meromorphic with simple poles at $t = \pm z$ and residues

$$R_1 := \operatorname{Res}(f_z, z) = \frac{-\mathrm{i}\,\mathrm{e}^{-z^2}}{2\pi} \quad \text{and} \quad R_2 := \operatorname{Res}(f_z, -z) = -R_1. \tag{26}$$

Thus, using the notation of Proposition 2.1, we have

$$C_{h,\alpha,H}[f_z] = \begin{cases} 2\mathrm{e}^{-z^2}/(1 - \mathrm{e}^{-2\mathrm{i}\pi(\alpha + z/h)}) & \text{if } H > y, \\ 0 & \text{if } H < y, \end{cases} \tag{27}$$

and the trapezoidal rule approximation to $w(z) = I[f_z]$ is

$$I_{h,\alpha}[f_z] = h\sum_{k\in\mathbb{Z}} \frac{\mathrm{i}z\,\mathrm{e}^{-(k-\alpha)^2h^2}}{\pi(z^2 - (k-\alpha)^2h^2)} = \begin{cases} \dfrac{h\mathrm{i}z}{\pi}\left[\dfrac{1}{z^2} + 2\displaystyle\sum_{k=1}^{\infty} \frac{\mathrm{e}^{-\tau_k^2}}{z^2 - \tau_k^2}\right], & \alpha = 0, \\ \dfrac{2h\mathrm{i}z}{\pi}\displaystyle\sum_{k=0}^{\infty} \frac{\mathrm{e}^{-t_k^2}}{z^2 - t_k^2}, & \alpha = \frac{1}{2}, \end{cases} \tag{28}$$

where $t_k$ and $\tau_k$ are as defined in (15). It is useful also to introduce the notation

$$C_{h,\alpha}[f_z] := \lim_{H\to\infty} C_{h,\alpha,H}[f_z] = \frac{2\mathrm{e}^{-z^2}}{1 - \mathrm{e}^{-2\mathrm{i}\pi(\alpha + z/h)}} \tag{29}$$

and

$$I^*_{h,\alpha}[f_z] := \lim_{H\to\infty} I^*_{h,\alpha,H}[f_z] = I_{h,\alpha}[f_z] + C_{h,\alpha}[f_z], \tag{30}$$

where $I^*_{h,\alpha,H}[f_z]$ is the modified trapezoidal rule defined by (21). Note that

$$C_{h,\alpha}[f_z] = C_{h,\alpha,H}[f_z] \text{ and } I^*_{h,\alpha}[f_z] = I^*_{h,\alpha,H}[f_z] \quad \text{for } y < H \tag{31}$$

and that $I^*_{h,\alpha,H}[f_z] = I_{h,\alpha}[f_z]$ for $y > H$. We note also that, as a function of $z$, $I^*_{h,\alpha}[f_z]$ is entire: $I_{h,\alpha}[f_z]$ and $C_{h,\alpha}[f_z]$ are both meromorphic with simple poles at $z = \pm(k-\alpha)h$, $k \in \mathbb{Z}$, and it is easy to check that the pole contributions cancel in the sum (30): the singularities in $I^*_{h,\alpha}[f_z]$ are removable.

The following proposition (cf. [17, section 2]) bounds $|w(z) - I^*_{h,\alpha,H}[f_z]|$ for $H = \pi/h$. It also bounds the relative error $|w(z) - I^*_{h,\alpha,H}[f_z]|/|w(z)|$ using the lower bound [5, Theorem 6]

$$|w(z)| \geq \frac{1}{1+\sqrt{\pi}|z|}, \quad \mathrm{Im}(z) \geq 0, \tag{32}$$

this being sharp for small and large $z$ as $w(0) = 1$ and $w(z) \sim \mathrm{i}/(\sqrt{\pi}z)$ as $z \to \infty$ [10, (2.6)].

PROPOSITION 2.2. *Suppose that $z = x + \mathrm{i}y$ with $0 \leq x \leq y$. Then*

$$\left|w(z) - I^*_{h,\alpha}[f_z]\right| \leq 2\sqrt{\frac{\mathrm{e}}{\pi}}\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}} \tag{33}$$

*and*

$$\frac{\left|w(z) - I^*_{h,\alpha}[f_z]\right|}{|w(z)|} \leq \frac{4\sqrt{2\pi\mathrm{e}}}{h}\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}} \tag{34}$$

*if $0 \leq y \leq \pi/h$, while*

$$|w(z) - I_{h,\alpha}[f_z]| \leq 4\sqrt{\frac{\mathrm{e}}{\pi}}\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}} \tag{35}$$

*and*

$$\frac{|w(z) - I_{h,\alpha}[f_z]|}{|w(z)|} \leq \frac{4\sqrt{2\pi\mathrm{e}}(1+\sqrt{\pi})}{h}\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}} \tag{36}$$

*if $y \geq \pi/h$ and $h < \sqrt{2}\,\pi$.*

*Proof.* For $H > 0$ and $y > 0$ with $H \neq y$ we have the bound (23) with

$$M = \sup_{t\in\mathbb{R}} |F_z(t+\mathrm{i}H)| = \frac{|z|}{\pi} \sup_{t\in\mathbb{R}} \frac{1}{|z^2 - (t+\mathrm{i}H)^2|}. \tag{37}$$

Since $|z^2 - (t+\mathrm{i}H)^2| = |z - t - \mathrm{i}H|\,|z + t + \mathrm{i}H| \geq |y - H|\,|y + H|$, it follows that $M \leq |z|/(\pi|y^2 - H^2|)$, so that

$$\left|w(z) - I^*_{h,\alpha,H}[f_z]\right| \leq \frac{2|z|\,\mathrm{e}^{H^2-2\pi H/h}}{\sqrt{\pi}\left(1-\mathrm{e}^{-2\pi H/h}\right)|y^2-H^2|} \leq \frac{2\sqrt{2}\,y\,\mathrm{e}^{H^2-2\pi H/h}}{\sqrt{\pi}\left(1-\mathrm{e}^{-2\pi H/h}\right)|y^2-H^2|} \tag{38}$$

since $0 \leq x \leq y$ so that $|z| \leq \sqrt{2}y$. Further, using (32) and $0 \leq x \leq y$,

$$\frac{\left|w(z) - I^*_{h,\alpha,H}[f_z]\right|}{|w(z)|} \leq \frac{2\sqrt{2}(1+\sqrt{2\pi}y)y\,\mathrm{e}^{H^2-2\pi H/h}}{\sqrt{\pi}\left(1-\mathrm{e}^{-2\pi H/h}\right)|y^2-H^2|}. \tag{39}$$

Now suppose that $0 < y \le \pi/h$, and take $H = \pi/h + \varepsilon$ for some $\varepsilon > 0$. Then, by (31), $I^*_{h,\alpha}[f_z] = I^*_{h,\alpha,H}[f_z]$, and since $y/(H^2 - y^2)$ and $(1 + \sqrt{2\pi}y)y/(H^2 - y^2)$ are increasing as functions of $y$ on $[0, H)$, it follows from (38) and (39) with $H = \pi/h + \varepsilon$ that

$$\left|w(z) - I^*_{h,\alpha}[f_z]\right| \le \frac{2\sqrt{2\pi}\,\mathrm{e}^{-\pi^2/h^2+\varepsilon^2}}{\left(1 - \mathrm{e}^{-2\pi^2/h^2-2\pi\varepsilon/h}\right)\varepsilon(2\pi + \varepsilon h)} \tag{40}$$

and

$$\frac{\left|w(z) - I^*_{h,\alpha}[f_z]\right|}{|w(z)|} \le \frac{2\sqrt{2\pi}\,(h + \sqrt{2\pi}\pi)\,\mathrm{e}^{-\pi^2/h^2+\varepsilon^2}}{h\left(1 - \mathrm{e}^{-2\pi^2/h^2-2\pi\varepsilon/h}\right)\varepsilon(2\pi + \varepsilon h)}. \tag{41}$$

Choosing $\varepsilon = 1/\sqrt{2}$ to minimize $\exp(\varepsilon^2)/\varepsilon$, we obtain, for $0 < y < \pi/h$, the bound (33), and also the bound (34) on noting that $(h + \sqrt{2\pi}\,\pi)/(2\pi + h/\sqrt{2}) \le \sqrt{2}$; these bounds hold also for $y = 0$ and $y = \pi/h$ since the left-hand sides of the bounds depend continuously on $y$ on $[0, \pi/h]$ (recall that $I^*_{h,\alpha}[f_z]$ is an entire function of $z$ and that $w(z)$ is bounded below on $y \ge 0$ by (32)).

Now suppose that $y > \pi/h$, and take $H = \pi/h - \varepsilon$ for some $\varepsilon \in (0, \pi/h)$. Then $I^*_{h,\alpha,H}[f_z] = I_{h,\alpha}[f_z]$, and since $y/(y^2 - H^2)$ and $(1 + \sqrt{2\pi}y)y/(y^2 - H^2)$ are decreasing as functions of $y$ on $(H, \infty]$, it follows from (38) and (39) with $H = \pi/h - \varepsilon$ that

$$|w(z) - I_{h,\alpha}[f_z]| \le \frac{2\sqrt{2\pi}\,\mathrm{e}^{-\pi^2/h^2+\varepsilon^2}}{\left(1 - \mathrm{e}^{-2\pi^2/h^2+2\pi\varepsilon/h}\right)\varepsilon(2\pi - \varepsilon h)} \tag{42}$$

and

$$\frac{|w(z) - I_{h,\alpha}[f_z]|}{|w(z)|} \le \frac{2\sqrt{2\pi}\,(h + \sqrt{2\pi}\pi)\,\mathrm{e}^{-\pi^2/h^2+\varepsilon^2}}{h\left(1 - \mathrm{e}^{-2\pi^2/h^2+2\pi\varepsilon/h}\right)\varepsilon(2\pi - \varepsilon h)}. \tag{43}$$

If $\pi/h > 1/\sqrt{2}$, we can again choose $\varepsilon = 1/\sqrt{2}$, obtaining the bounds (35) and (36) for $y > \pi/h$; these bounds hold also for $y = \pi/h$ since the left-hand sides of the bounds depend continuously on $y$ on $[\pi/h, \infty)$. ∎

It follows immediately from the definition (29) that, for $x \in \mathbb{R}$, $y > 0$,

$$|C_{h,\alpha}[f_z]| \le \frac{2\,\mathrm{e}^{-2\pi y/h}}{1 - \mathrm{e}^{-2\pi y/h}}\,\mathrm{e}^{y^2-x^2}. \tag{44}$$

Since $|I^*_{h,\alpha}[f_z]| \le |I_{h,\alpha}[f_z]| + |C_{h,\alpha}[f_z]|$, the following corollary follows from the above proposition, (44), and (32). The purpose of this corollary is to provide bounds on errors on part of the boundary of the infinite sector $\Omega$ (as defined in Lemma 2.4 with $a = 4$). This will lead to absolute and relative error bounds for $I^*_{h,\alpha}[f_z]$ as an approximation on the whole of $\overline{\Omega}$ via the Phragmén–Lindelöf principle of Lemma 2.4; see Proposition 2.6 below.

COROLLARY 2.3. *If $z = x + \mathrm{i}y$ with $x = y \ge 0$ and $h < \sqrt{2}\,\pi$, then*

$$\left|w(z) - I^*_{h,\alpha}[f_z]\right| \le c_a\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1 - \mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}} \tag{45}$$

*and*

$$\frac{\left|w(z) - I^*_{h,\alpha}[f_z]\right|}{|w(z)|} \le \frac{c_r}{h}\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1 - \mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}}, \tag{46}$$

*where*

$$c_a := \frac{2(2\mathrm{e}+\sqrt{\pi})}{\sqrt{\mathrm{e}\pi}} \approx 4.934 \quad \textit{and} \quad c_r := \frac{2\sqrt{2\pi}\,(1+\sqrt{\pi})(2\mathrm{e}+\sqrt{\pi})}{\sqrt{\mathrm{e}}} \approx 60.77. \tag{47}$$

*Proof.* For $0 \le x = y \le \pi/h$ these bounds follow immediately from the sharper bounds (33) and (34). Suppose now that $x = y \ge \pi/h$ and $h < \sqrt{2}\,\pi$. Then it follows from (44) that

$$|C_{h,\alpha}[f_z]| \le \frac{2\,\mathrm{e}^{-2\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}} \le \frac{2\,\mathrm{e}^{-\pi^2/h^2}}{\sqrt{\mathrm{e}}\left(1-\mathrm{e}^{-2\pi^2/h^2}\right)}.$$

Further, since $(1+\sqrt{2\pi}y)/(\mathrm{e}^{2\pi y/h}-1)$ is decreasing as a function of $y$ on $[\pi/h,\infty)$, it follows from (44) and (32) that

$$\frac{|C_{h,\alpha}[f_z]|}{|w(z)|} \le \frac{2\,(h+\sqrt{2\pi}\pi)\mathrm{e}^{-2\pi^2/h^2}}{h\left(1-\mathrm{e}^{-2\pi^2/h^2}\right)} \le \frac{2\sqrt{2}\,\pi(1+\sqrt{\pi})\,\mathrm{e}^{-\pi^2/h^2}}{h\sqrt{\mathrm{e}}\left(1-\mathrm{e}^{-2\pi^2/h^2}\right)}.$$

Since $|I^*_{h,\alpha}[f_z]| \le |I_{h,\alpha}[f_z]| + |C_{h,\alpha}[f_z]|$, the required bounds for $x = y \ge \pi/h$ follow from the above bounds and (35) and (36) in Proposition 2.2. ∎

Proposition 2.2 tells us that $I^*_{h,\alpha,H}[f_z]$, with $H = \pi/h$, is an approximation to $w(z)$ with controllable absolute and relative errors for $\pi/4 \le \arg(z) \le \pi/2$. To complement this result we will show in Proposition 2.6 below that $I^*_{h,\alpha}[f_z]$ is an approximation to $w(z)$ with controllable absolute and relative errors for $0 \le \arg(z) \le \pi/4$, so that $I^*_{h,\alpha,\pi/h}[f_z]$ and $I^*_{h,\alpha}[f_z]$ together provide trapezoidal rule–based approximations to $w(z)$ across the whole first quadrant (and, via the symmetries (4), across the whole complex plane). We will prove Proposition 2.6 via the following Phragmén–Lindelöf principle applied with $a = 4$ to the left-hand sides of (42) and (43).

LEMMA 2.4 (see [8, Chapter VI, Cor. 4.2]). *Let $a \ge 1/2$, and put*

$$\Omega := \left\{ z = r\mathrm{e}^{\mathrm{i}\theta} : r > 0 \textit{ and } 0 < \theta < \pi/a \right\}.$$

*Suppose that $f$ is analytic on $\Omega$ and continuous in $\overline{\Omega}$ and that there is a constant $P$ such that $|f(z)| \le P$ for all $z \in \partial\Omega$. If there are positive constants $Q$ and $b < a$ such that $|f(z)| \le Q\,\exp(|z|^b)$ for all $z \in \Omega$, then $|f(z)| \le P$ for all $z \in \overline{\Omega}$.*

The main step in proving Proposition 2.6 via this lemma is to show that the left-hand sides of (45) and (46) are bounded on $\partial\Omega$ when $a = 4$. We have bounded these left-hand sides already on $\{r\mathrm{e}^{\mathrm{i}\pi/4} : r \ge 0\}$ in Corollary 2.3. It remains to bound them on the positive real axis, which we do in the next proposition.

PROPOSITION 2.5. *If $x \ge 0$, then*

$$\left|w(x) - I^*_{h,\alpha}[f_x]\right| \le \frac{2h\mathrm{e}^{-\pi^2/h^2}}{\pi^{3/2}\left(1-\mathrm{e}^{-2\pi^2/h^2}\right)} \tag{48}$$

*and*

$$\frac{\left|w(x) - I^*_{h,\alpha}[f_x]\right|}{|w(x)|} \le \left[8 + \frac{10h}{\pi^{3/2}}\right] \frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}}. \tag{49}$$

*Proof.* Arguing as in the proof of Proposition 2.2, for $x \geq 0$ and $0 < y < \pi/h$ we have the bound (23) with $H = \pi/h$. Thus, and since (recall (31)) $I^*_{h,\alpha,H}[f_z] = I^*_{h,\alpha}[f_z]$ for $y < H$, it follows that

$$\left|I[f] - I^*_{h,\alpha}[f_z]\right| \leq \frac{2\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}} \int_{-\infty}^{\infty} \mathrm{e}^{-t^2} |F_z(t+\mathrm{i}H)|\,dt \leq \frac{2\sqrt{\pi}M\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}}, \tag{50}$$

where $M$ is given by (37). Since $M$ and the left- and right-hand sides of the first of these inequalities depend continuously on $y$ on $[0, \pi/h)$, the above inequalities hold also for $y = 0$. Since, also, for $x \geq 0$ and $t \in \mathbb{R}$ we have that

$$\begin{aligned} |x^2 - (t+\mathrm{i}\pi/h)^2| &= |x^2 - (|t| + \mathrm{i}\pi/h)^2| \\ &= |x - |t| - \mathrm{i}\pi/h|\,|x + |t| + \mathrm{i}\pi/h| \geq \frac{\pi}{h}|x + \mathrm{i}\pi/h| \geq \frac{\pi}{h}|x|, \end{aligned} \tag{51}$$

it follows that $M \leq h/\pi^2$ for $x \geq 0$, and (48) follows from (50) with $z = x \geq 0$.

From (32) and (50), with $z = x \geq 0$ and $H = \pi/h$, it follows that

$$\frac{\left|I[f] - I^*_{h,\alpha}[f_x]\right|}{|w(x)|} \leq \frac{2(1+\sqrt{\pi}\,x)\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2}} \int_{-\infty}^{\infty} G_x(t)\,dt \tag{52}$$

for $x \geq 0$, where

$$G_x(t) := \mathrm{e}^{-t^2}|F_x(t+\mathrm{i}\pi/h)| = \frac{x\mathrm{e}^{-t^2}}{\pi\,|x^2 - (t+\mathrm{i}\pi/h)^2|}.$$

If $x \geq 0$ and $-x/2 \leq t \leq x/2$, then

$$|x^2 - (t+\mathrm{i}\pi/h)^2| = |x - t - \mathrm{i}\pi/h|\,|x + t + \mathrm{i}\pi/h| \geq |x/2 - \mathrm{i}\pi/h|\,|x/2 + \mathrm{i}\pi/h| = \frac{x^2 + 4\pi^2/h^2}{4},$$

so that

$$\int_{-x/2}^{x/2} G_x(t)\,dt \leq \frac{4x}{\pi(x^2+4\pi^2/h^2)} \int_{-\infty}^{\infty} \mathrm{e}^{-t^2}\,dt = \frac{4x}{\sqrt{\pi}(x^2+4\pi^2/h^2)}.$$

On the other hand, using (51), we see that $G_x(t) \leq h\mathrm{e}^{-t^2}/\pi^2$ for all $x \geq 0$, $t \in \mathbb{R}$, so that

$$\int_{\mathbb{R}\setminus[-x/2,x/2]} G_x(t)\,dt \leq \frac{2h}{\pi^2} \int_{x/2}^{\infty} \mathrm{e}^{-t^2}\,dt < \frac{2h}{\pi^2 x}\mathrm{e}^{-x^2/4} \leq \frac{2h}{\pi^2 x}$$

for $x > 0$ since, for $a > 0$,

$$\int_a^{\infty} \mathrm{e}^{-t^2}\,dt = \frac{\mathrm{e}^{-a^2}}{2a} - \frac{1}{2}\int_a^{\infty} \frac{\mathrm{e}^{-t^2}}{t^2}dt < \frac{\mathrm{e}^{-a^2}}{2a}. \tag{53}$$

Moreover,

$$\int_{\mathbb{R}\setminus[-x/2,x/2]} G_x(t)\,dt \leq \frac{2h}{\pi^2} \int_0^{\infty} \mathrm{e}^{-t^2}\,dt = \frac{h}{\pi^{3/2}}.$$

Thus, and since $\min(a,b) \leq 2ab/(a+b)$ if $a \geq 0$, $b \geq 0$, and $a + b > 0$, it follows that

$$\int_{\mathbb{R}\setminus[-x/2,x/2]} G_x(t)\,dt \leq \frac{h}{\pi^2}\,\min(\sqrt{\pi}, 2x^{-1}) \leq \frac{4h}{\pi^{3/2}(2+\sqrt{\pi}\,x)}$$

so that

$$\int_{-\infty}^{\infty} G_x(t)\,dt \le \frac{4}{\sqrt{\pi}}\left(\frac{x}{x^2+4\pi^2/h^2}+\frac{h}{\pi(2+\sqrt{\pi}\,x)}\right)$$

and (49) follows from (52) on noting that $x^2/(x^2+4\pi^2/h^2)\le 1$ and $x/(x^2+4\pi^2/h^2)\le h/(4\pi)$. ∎

In our final proposition of this subsection (cf. [4, Proposition 3.3.4]) we combine Corollary 2.3, Proposition 2.5, and Lemma 2.4 to bound approximations to $w(z)$ in $0\le\arg(z)\le\pi/4$, complementing the bounds in Proposition 2.2 for $\pi/4\le\arg(z)\le\pi/2$.

PROPOSITION 2.6. *Suppose that $h<\sqrt{2}\,\pi$ and $z=x+\mathrm{i}y$ with $0\le y\le x$. Then the bounds* (45) *and* (46) *hold with $c_a$ and $c_r$ given by* (47).

*Proof.* We will prove this proposition by applying Lemma 2.4 with $a=4$, so that $\Omega=\{z=x+\mathrm{i}y : 0<y<x\}$, to the functions $E_h(z):=w(z)-I^*_{h,\alpha}[f_z]$ and $e_h(z):=E_h(z)/w(z)=(w(z)-I^*_{h,\alpha}[f_z])/w(z)$.

We have remarked already that $w(z)$ and $I^*_{h,\alpha}[f_z]$ are entire as a function of $z$, so that $E_h$ is entire and, noting (32), $e_h$ is analytic in $\mathrm{Im}(z)>0$ and continuous in $\mathrm{Im}(z)\ge 0$. In particular, $E_h$ and $e_h$ are continuous in $\overline{\Omega}$ and analytic in $\Omega$. Further, if $h<\sqrt{2}\,\pi$, it follows from Corollary 2.3 and Proposition 2.5, on noting that the bounds in Proposition 2.5 are smaller than those in Corollary 2.3, that $E_h$ and $e_h$ satisfy the bounds claimed in the proposition when $z\in\partial\Omega$, i.e., for $z$ on $\{x+\mathrm{i}x : x\ge 0\}$ and on the positive real axis. Thus the proposition follows by Lemma 2.4 if we can show that $E_h(z)$ and $e_h(z)$ do not grow too rapidly as $z\to\infty$ in $\Omega$.

But, if $h<\sqrt{2}\,\pi$, it follows from (31) and (38) applied with $H=3$ that, for some constant $C>0$ independent of $z$, $|E_h(z)|\le C|z|$ if $z\in\Omega$ with $y\le 2$. Similarly, since $I^*_{h,\alpha,H}[f_z]=I_{h,\alpha}[f_z]$ if $y>H$ and $I^*_{h,\alpha}[f_z]=I_{h,\alpha}[f_z]+C_{h,\alpha}[f_z]$, it follows from (38) applied with $H=1$ and (44) that, for some constant $\widetilde{C}>0$ independent of $z$, $|E_h(z)|\le\widetilde{C}|z|$ if $z\in\Omega$ with $y\ge 2$. Thus $|E_h(z)|\le C^*|z|$ for $z\in\Omega$, where $C^*:=\max(C,\widetilde{C})$, so that also, applying (32), $|e_h(z)|\le C^*|z|(1+\sqrt{\pi}\,|z|)$ for $z\in\Omega$. Thus the proposition follows by applying Lemma 2.4. ∎

The following corollary summarizes and simplifies, at the cost of a little sharpness, the results of Propositions 2.2 and 2.6 and of this subsection.

COROLLARY 2.7. *Suppose that $z=x+\mathrm{i}y$ with $x\ge 0$, $y\ge 0$, and $h<\sqrt{2}\,\pi$. Then the bounds* (45) *and* (46) *hold with $c_a$ and $c_r$ given by* (47) *if $y\le\max(x,\pi/h)$. The same bounds hold as bounds on $|w(z)-I_{h,\alpha}[f_z]|$ and $|w(z)-I_{h,\alpha}[f_z]|/|w(z)|$, respectively, with the same values of $c_a$ and $c_r$, if $y\ge\max(x,\pi/h)$.*

*Proof.* The first claim of the corollary follows from Proposition 2.6 and (33) and (34), and the second follows from (35) and (36). ∎

**2.3. Truncating the infinite series.** Propositions 2.2 and 2.6 together provide accurate trapezoidal rule–based approximations to $w(z)$ in the first quadrant of the complex plane that can be extended to the whole complex plane using the symmetries (4). But in implementation the infinite series in these approximations must be truncated. We estimate the additional error this introduces in this subsection.

At this point, since we wish to use the fact that $f_z$ is even to reduce computation, we restrict our attention to the cases $\alpha=0$ and $\alpha=1/2$, in which cases the trapezoidal

rule approximation reduces to (28). In these cases we approximate $I_{h,\alpha}[f_z]$ by

$$I^N_{h,\alpha}[f_z] := \begin{cases} hf_z(0) + 2h\displaystyle\sum_{k=1}^{N} f_z(\tau_k), & \alpha = 0, \\ 2h\displaystyle\sum_{k=0}^{N} f_z(t_k), & \alpha = 1/2, \end{cases} \tag{54}$$

with $f_z$ given by (25) and $\tau_k$ and $t_k$ defined as in (15). We will call the error in approximating $I_{h,\alpha}[f_z]$ by $I^N_{h,\alpha}[f_z]$ the truncation error, given by

$$T^N_{h,\alpha}[f_z] := 2h \sum_{k=N+1}^{\infty} f_z(s_k), \tag{55}$$

where $s_k := (k+\alpha)h$. This is also the error in approximating $I^*_{h,\alpha}[f_z]$ by $I^{*,N}_{h,\alpha}[f_z]$, where $I^*_{h,\alpha}[f_z]$ is defined as in (30) and

$$I^{*,N}_{h,\alpha}[f_z] := I^N_{h,\alpha}[f_z] + C_{h,\alpha}[f_z]. \tag{56}$$

The following result [4, Proposition 3.3.7] bounds $T^N_{h,\alpha}[f_z]$ for $\pi/4 \le \arg(z) \le \pi/2$. We use in this proposition the estimate, obtained since $\exp(-t^2)$ is decreasing on $(0,\infty)$ and noting (53), that

$$2h \sum_{k=M}^{\infty} \mathrm{e}^{-s_k^2} \le 2h\mathrm{e}^{-s_M^2} + 2\int_{s_M}^{\infty} \mathrm{e}^{-t^2}\,dt \le \frac{2hs_M + 1}{s_M}\,\mathrm{e}^{-s_M^2}, \quad M \in \mathbb{N}. \tag{57}$$

PROPOSITION 2.8. *Suppose $\alpha = 0$ or $1/2$ and $z = x + \mathrm{i}y$ with $y \ge x \ge 0$. Then, for $N \in \mathbb{N}_0$,*

$$|T^N_{h,\alpha}[f_z]| \le \frac{(1+2h\,\tau_{N+1})}{\pi\tau^2_{N+1}}\,\mathrm{e}^{-\tau^2_{N+1}} \quad \textit{and} \tag{58}$$

$$\frac{|T^N_{h,\alpha}[f_z]|}{|w(z)|} \le \frac{(1+2h\,\tau_{N+1})(1+2\sqrt{\pi}\,\tau_{N+1})}{\pi\tau^2_{N+1}}\,\mathrm{e}^{-\tau^2_{N+1}}. \tag{59}$$

*Proof.* For $z = x + \mathrm{i}y$ with $y \ge x \ge 0$,

$$|z^2 - s_k^2|^2 = y^4 + s_k^4 + x^4 + 2x^2y^2 + 2s_k^2(y^2 - x^2) \ge y^4 + s_k^4 \ge y^4 + \tau_k^4.$$

Thus, and recalling (25) and using (57) with $M = N+1$,

$$|T^N_{h,\alpha}[f_z]| \le \frac{2\sqrt{2}h\,y}{\pi} \sum_{k=N+1}^{\infty} \frac{\mathrm{e}^{-\tau_k^2}}{\sqrt{y^4+\tau_k^4}} \le \frac{\sqrt{2}y\,(1+2h\,\tau_{N+1})}{\pi\tau_{N+1}\sqrt{y^4+\tau^4_{N+1}}}\,\mathrm{e}^{-\tau^2_{N+1}}.$$

Moreover,

$$y\Big/\sqrt{y^4+\tau^4_{N+1}} \le \frac{1}{\sqrt{2}\,\tau_{N+1}} \quad \text{and} \quad y^2\Big/\sqrt{y^4+\tau^4_{N+1}} \le 1$$

so that (58) follows and also

$$y\,|T^N_{h,\alpha}[f_z]| \le \frac{\sqrt{2}(1+2h\,\tau_{N+1})}{\pi\tau_{N+1}}\,\mathrm{e}^{-\tau^2_{N+1}}. \tag{60}$$

Since, by (32), $|w(z)|^{-1} \le 1 + \sqrt{2\pi}y$ for $0 \le x \le y$, (59) follows from (58) and (60). ∎

The following result [4, Propositions 3.3.5, 3.3.6] bounds $T^N_{h,\alpha}[f_z]$ for $0 \le \arg(z) \le \pi/4$ so that Propositions 2.8 and 2.9 together bound the absolute and relative truncation errors in the first quadrant. The case $0 \le \arg(z) \le \pi/4$ is more subtle because $T^N_{h,\alpha}[f_z]$ is unbounded, it has simple poles at $z = s_k$ for $k \ge N+1$, and our bound requires that the distance of $z$ from this set of poles is $\ge h/4$. Despite this restriction, we can construct accurate approximations covering the whole region $0 \le \arg(z) \le \pi/4$ because $s_k = \tau_k = kh$ for $\alpha = 0$ while $s_k = t_k = \tau_k + h/2$ for $\alpha = 1/2$, so that $\{z : |z - s_k| \ge h/4$ for either $s_k = \tau_k$ or $t_k$, for $k \ge N+1\}$ includes the whole of $\{z : 0 \le \arg(z) \le \pi/4\}$.

PROPOSITION 2.9. *Suppose $\alpha = 0$ or $1/2$ and $z = x + \mathrm{i}y$ with $0 \le y \le x$ and $|z - s_k| \ge h/4$ for $k \ge N+1$. Then, for $N \in \mathbb{N}_0$,*

$$|T^N_{h,\alpha}[f_z]| \le c(h,N,\alpha) := \frac{2\sqrt{2}\,(1+2hs_{N+1})\,(h+4s_{N+1})}{\pi h\, s^2_{N+1}}\,\mathrm{e}^{-s^2_{N+1}} \quad \textit{and} \tag{61}$$

$$\frac{|T^N_{h,\alpha}[f_z]|}{|w(z)|} \le (1+\sqrt{2\pi}s_{N+1})\,c(h,N,\alpha). \tag{62}$$

*Proof.* From (25) and (55), for $0 \le y \le x$,

$$|T^N_{h,\alpha}[f_z]| \le \frac{2h|z|}{\pi}\sum_{k=N+1}^{\infty}\frac{\mathrm{e}^{-s_k^2}}{|z^2 - s_k^2|} \le \frac{2\sqrt{2}hx}{\pi(x+s_{N+1})}\sum_{k=N+1}^{\infty}\frac{\mathrm{e}^{-s_k^2}}{|z-s_k|}. \tag{63}$$

Thus, and noting (32), the bounds (61) and (62) hold if $x = 0$.

Choose $\theta$ with $0 < \theta < 1$. Recalling $z = x + \mathrm{i}y$, in the case that $x > 0$ let $M$ be the smallest integer $\ge N+1$ such that $s_M > \theta x$, so that, if $M > N+1$, it holds that $s_k \le \theta x$ and $|z - s_k| \ge (1-\theta)x$ for $k < M$. If $M > N+1$, it follows, using the bound (57), that

$$2hx\sum_{k=N+1}^{M-1}\frac{\mathrm{e}^{-s_k^2}}{|z-s_k|} \le \frac{2h}{1-\theta}\sum_{k=N+1}^{\infty}\mathrm{e}^{-s_k^2} \le \frac{2hs_{N+1}+1}{(1-\theta)s_{N+1}}\,\mathrm{e}^{-s^2_{N+1}},$$

while, for $M \ge N+1$, assuming $|z - s_k| \ge h/4$ for $k \ge N+1$ and again using (57),

$$2hx\sum_{k=M}^{\infty}\frac{\mathrm{e}^{-s_k^2}}{|z-s_k|} \le 8x\sum_{k=M}^{\infty}\mathrm{e}^{-s_k^2} \le \frac{4x(2hs_M+1)}{hs_M}\mathrm{e}^{-s_M^2} \le \frac{4(2hs_M+1)}{h\theta}\,\mathrm{e}^{-s_M^2}.$$

Thus, and since $(2ht+1)\exp(-t^2)$ is decreasing as a function of $t$ on $[h,\infty)$ and $s_M \ge s_{N+1} \ge h$,

$$\begin{aligned} 2hx\sum_{k=N+1}^{\infty}\frac{\mathrm{e}^{-s_k^2}}{|z-s_k|} &\le (2hs_{N+1}+1)\left[\frac{4}{h\theta} + \frac{1}{(1-\theta)s_{N+1}}\right]\mathrm{e}^{-s^2_{N+1}} \\ &= \frac{2(2hs_{N+1}+1)(h+4s_{N+1})}{hs_{N+1}}\,\mathrm{e}^{-s^2_{N+1}}, \end{aligned} \tag{64}$$

on choosing $\theta = 4s_{N+1}/(h+4s_{N+1})$ so as to equalize the terms in the square brackets. From (63) and (64), on noting that $x + s_{N+1} \ge s_{N+1}$ and $x/(x+s_{N+1}) \le 1$, we see that, for $x > 0$, (61) holds and also $x|T^N_{h,\alpha}[f_z]| \le s_{N+1}c(h,N)$. From these bounds and that $|w(z)|^{-1} \le 1 + \sqrt{2\pi}x$ for $0 \le y \le x$ by (32), the bound (62) follows. ∎

The following corollary summarizes and simplifies, at the cost of a little sharpness, the results of Propositions 2.8 and 2.9 and this subsection.

COROLLARY 2.10. *Suppose* $\alpha = 0$ *or* $1/2$, $z = x + \mathrm{i}y$ *with* $x \geq 0$, $y \geq 0$, *and either* $y \geq x$ *or* $|z - s_k| \geq h/4$ *for* $k \geq N+1$. *Then, for* $N \in \mathbb{N}_0$,

$$|T^N_{h,\alpha}[f_z]| \leq c(h,N,0) = \frac{2\sqrt{2}\,(1+2h\tau_{N+1})\,(h+4\tau_{N+1})}{\pi h\,\tau^2_{N+1}}\,\mathrm{e}^{-\tau^2_{N+1}}, \quad \textit{and also} \tag{65}$$

$$\frac{|T^N_{h,\alpha}[f_z]|}{|w(z)|} \leq (1+\sqrt{2\pi}\,\tau_{N+1})\,c(h,N,0), \quad \textit{provided} \quad h \geq 1/(N+1). \tag{66}$$

*Proof.* Proposition 2.8 implies the above bounds hold when $y \geq x$. If $y \leq x$, the above bounds are immediate from Proposition 2.9 in the case $\alpha = 0$. They hold also when $\alpha = 1/2$ as (i) $c(h,N,1/2) \leq c(h,N,0)$, since $s^{-m}\mathrm{e}^{-s^2}$ decreases as $s$ increases on $(0,\infty)$ for $m = 0,1,2$; (ii) $t_{N+1}c(h,N,1/2) \leq \tau_{N+1}c(h,N,0)$ if $\tau_{N+1} \geq 1$ (i.e., $h \geq 1/(N+1)$), since also $s\mathrm{e}^{-s^2}$ decreases as $s$ increases on $[1,\infty)$. ∎

**2.4. Proof of the main theorem.** In this subsection we bring together the bounds on the error in the trapezoidal rule approximation (Corollary 2.7) and on the truncation error (Corollary 2.10) to bound the errors in the truncated trapezoidal rule approximation (54) and its modification (56). Clearly,

$$|w(z) - I^N_{h,\alpha}[f_z]| \leq |w(z) - I_{h,\alpha}[f_z]| + |T^N_{h,\alpha}[f_z]| \quad \text{and} \tag{67}$$

$$|w(z) - I^{*,N}_{h,\alpha}[f_z]| \leq |w(z) - I^*_{h,\alpha}[f_z]| + |T^N_{h,\alpha}[f_z]|. \tag{68}$$

Applying Corollaries 2.7 and 2.10 it follows that, for $\alpha = 0$ and $1/2$ and $z = x + \mathrm{i}y$ with $x \geq 0$, $0 \leq y \leq \max(x, \pi/h)$,

$$|w(z) - I^{*,N}_{h,\alpha}[f_z]| \leq c_a\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}} + c(h,N,0), \quad \text{where} \tag{69}$$

$$c(h,N,0) := \frac{2\sqrt{2}\,(1+2h\tau_{N+1})\,(h+4\tau_{N+1})}{\pi h\,\tau^2_{N+1}}\,\mathrm{e}^{-\tau^2_{N+1}},$$

provided $y \geq x$ or $|z - s_k| \geq h/4$ for $k \geq N+1$. Similarly, applying Corollaries 2.7 and 2.10,

$$|w(z) - I^N_{h,\alpha}[f_z]| \leq c_a\,\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}} + c(h,N,0) \tag{70}$$

for $\alpha = 0$ and $1/2$ if $z = x + \mathrm{i}y$ with $x \geq 0$ and $y \geq \max(x, \pi/h)$.

We choose the step size $h$, as a function of $N$, to approximately balance the contributions from the trapezoidal rule error and the truncation error in the above error bounds. Precisely, we choose $h$ so that the exponents of $\mathrm{e}^{-\pi^2/h^2}$ and $e^{-\tau^2_{N+1}}$ are equal; i.e., we define $h := \sqrt{\pi/(N+1)}$ as in (11). With this choice of $h$ we have $\pi/h = \tau_{N+1} = \sqrt{(N+1)\pi}$,

$$c(h,N,0) = \frac{2\sqrt{2}\,(1+2\pi)\,(5+4N)}{\pi^2(N+1)}\,\mathrm{e}^{-(N+1)\pi} \leq \frac{10\sqrt{2}\,(1+2\pi)}{\pi^2}\,\mathrm{e}^{-(N+1)\pi}, \tag{71}$$

and

$$\frac{\mathrm{e}^{-\pi^2/h^2}}{1-\mathrm{e}^{-2\pi^2/h^2+\sqrt{2}\pi/h}} = \frac{\exp(-(N+1)\pi)}{1-\exp\left(-2(N+1)\pi+\sqrt{2(N+1)\pi}\right)} \leq c^*\,\mathrm{e}^{-(N+1)\pi}, \tag{72}$$

where

$$c^* := \left(1 - \exp\left(-2\pi + \sqrt{2\pi}\right)\right)^{-1} \approx 1.0234. \tag{73}$$

This leads to our main theorem.

THEOREM 2.11. *Suppose $\alpha = 0$ or $1/2$, $N \in \mathbb{N}_0$, $h$ is defined by (11), and $z = x + \mathrm{i}y$. If $\pi/h \geq y \geq x \geq 0$ or $0 \leq y \leq x$ and $|z - s_k| \geq h/4$ for $k \geq N+1$, then*

$$|w(z) - I^{*,N}_{h,\alpha}[f_z]| \leq C_1\, \mathrm{e}^{-\pi N} \quad \textit{and} \quad \frac{|w(z) - I^{*,N}_{h,\alpha}[f_z]|}{|w(z)|} \leq C_2 \sqrt{N+1}\, \mathrm{e}^{-\pi N}, \tag{74}$$

*where*

$$C_1 := \frac{2(2\mathrm{e} + \sqrt{\pi})}{\mathrm{e}^{\pi}\sqrt{\mathrm{e}\pi}\left(1 - \exp\left(-2\pi + \sqrt{2\pi}\right)\right)} + \frac{10\sqrt{2}\,(1+2\pi)}{\mathrm{e}^{\pi}\pi^2} \approx 0.6692 \quad \textit{and} \tag{75}$$

$$C_2 := \frac{2\sqrt{2}\,(1+\sqrt{\pi})(2\mathrm{e} + \sqrt{\pi})}{\mathrm{e}^{\pi}\sqrt{\mathrm{e}}\left(1 - \exp\left(-2\pi + \sqrt{2\pi}\right)\right)} + \frac{10(1+2\pi)(2\pi + \sqrt{2})}{\mathrm{e}^{\pi}\pi^2} \approx 3.971. \tag{76}$$

*Further, if $x \geq 0$ and $y \geq \max(x, \pi/h)$, then*

$$|w(z) - I^{N}_{h,\alpha}[f_z]| \leq C_1\, \mathrm{e}^{-\pi N} \quad \textit{and} \quad \frac{|w(z) - I^{N}_{h,\alpha}[f_z]|}{|w(z)|} \leq C_2 \sqrt{N+1}\, \mathrm{e}^{-\pi N}. \tag{77}$$

*Proof.* The bounds on $|w(z) - I^{*,N}_{h,\alpha}[f_z]|$ and $|w(z) - I^{N}_{h,\alpha}[f_z]|$ follow from (69), (70), (71), (72), and the definitions, (47) and (73), of $c_a$ and $c^*$. With $h$ defined by (11), it follows that $h \geq 1/(N+1)$ and

$$1 + \sqrt{2\pi}\,\tau_{N+1} = 1 + \pi\sqrt{2(N+1)} \leq \frac{\sqrt{2} + 2\pi}{\sqrt{2}}\sqrt{N+1}.$$

Thus Corollaries 2.7 and 2.10, together with (67), (68), (71), and (72), imply, with $h$ defined by (11), that $|w(z) - I^{*,N}_{h,\alpha}[f_z]|/|w(z)|$ and $|w(z) - I^{N}_{h,\alpha}[f_z]|/|w(z)|$ are both bounded above by

$$\left(\frac{c_r c^*}{h} + \frac{10(1+2\pi)(2\pi+\sqrt{2})}{\pi^2}\sqrt{N+1}\right)\mathrm{e}^{-(N+1)\pi} = C_2\sqrt{N+1}\,\mathrm{e}^{-\pi N},$$

under their respective constraints on $x$ and $y$. ∎

The above theorem justifies approximating $w(z)$ by $I^N_{h,\alpha}[f_z]$, with $\alpha = 0$ or $1/2$, if $h$ is given by (11), $x \geq 0$, and $y \geq \max(x, \pi/h)$; we choose, arbitrarily, the midpoint rule–based approximation $w^{\mathrm{M}}_N(z) := I^N_{h,1/2}[f_z]$ given explicitly by (12). If $y < \max(x, \pi/h)$, the above theorem suggests approximating by $I^{*,N}_{h,\alpha}[f_z]$, with $\alpha = 0$ or $1/2$. For $0 \leq x \leq y < \pi/h$ we choose the modified midpoint rule–based approximation $w^{\mathrm{MM}}_N(z) := I^{*,N}_{h,1/2}[f_z]$ given explicitly by (13). This choice ensures that the distance of $z$ from the set of quadrature points $\{t_0, \ldots, t_N\}$ is $\geq h/(2\sqrt{2})$, so that the size of the largest term in the sum (12) does not exceed $1/(\pi\sqrt{2})$ and there is no loss of precision through cancellation between the two terms in the sum (13). For $0 \leq y < x$ we approximate either by $w^{\mathrm{MM}}_N(z)$ or by the modified trapezoidal-based approximation $w^{\mathrm{MT}}_N(z) := I^{*,N}_{h,0}[f_z]$, written explicitly in (14). Which of these we use

is determined by the rule (10). This rule ensures that $|z - s_k| \geq h/4$ for $k \in \mathbb{N}_0$, where $s_k = t_k$ when $\alpha = 1/2$, and $s_k = \tau_k$ when $\alpha = 0$, so that Theorem 2.11 applies and, in the use of both (13) and (14), we avoid loss of precision through cancellation between nearly equal terms.

*Proof of Theorem* 1.1. Noting the discussion in the above paragraph, we see that the bounds in Theorem 1.1 follow, for $x \geq 0$, $y \geq 0$, immediately from those in Theorem 2.11. That the absolute error bound holds in the whole complex plane, and that the bound on the relative error holds in $\{z : y \geq 0\}$, follows from the bounds in the first quadrant and the symmetry relations and definitions (4) and (16). ∎

**3. Survey of existing methods.** There are a number of other schemes for the computation of $w(z)$ for complex $z$, and we briefly summarize the best of these, making connections with (10). Most use variations on polynomial or rational approximation, with different schemes in different regions of the first quadrant (leading, through (4), to approximation in the whole complex plane). Indeed, our own approximation (10) uses three formulae (12)–(14), with (12) rational and (13)–(14) rational with meromorphic corrections in terms of exponential functions.

Gautschi [10] advocated, for larger $z$, the rational approximation

$$w(z) \approx \frac{\mathrm{i}/\sqrt{\pi}}{z-}\,\frac{1/2}{z-}\,\frac{2/2}{z-}\,\frac{3/2}{z-}\cdots\frac{(n-1)/2}{z}, \tag{78}$$

the $n$th convergent of the beautiful Laplace continued fraction representation for $w(z)$ (specifically suggesting $n = 9$). Gautschi notes that (i) by construction the $n$th convergent is asymptotically accurate, with error $O(|z|^{-2n-1})$ as $|z| \to \infty$, uniformly in the first quadrant; (ii) the $n$th convergent converges to $w(z)$ as $n \to \infty$ if and only if $\mathrm{Im}(z) > 0$; (iii) remarkably, for $\mathrm{Im}(z) > 0$, the $n$th convergent coincides with the approximation obtained by approximating (3) by an $n$-point Gauss–Hermite rule. For smaller $z$ Gautschi [10] proposed (rational) approximations that are truncated Taylor expansions with the coefficients approximated by convergents of continued fractions.

This methodology, carefully tuned, is the basis of TOMS Algorithm 680 (Poppe and Wijers [19]) which achieves a relative error of $10^{-14}$ over nearly all the complex plane using, in the first quadrant, (i) Maclaurin polynomials of degree $\leq 55$ for the odd function $\mathrm{erf}(-\mathrm{i}z)$ (substituted into (2)) in an ellipse around the origin; (ii) the convergents (78) with $n \leq 18$ outside a larger ellipse; (iii) the more expensive mix of Taylor expansion and continued fraction calculation proposed by Gautschi [10] in between. This algorithm has been used as a benchmark by several later authors.

Weideman [25, 26] proposed (the derivation starts from (3)) the single rational approximation

$$w(z) \approx \frac{1}{\sqrt{\pi}(L - \mathrm{i}z)} + \frac{2}{(L - \mathrm{i}z)^2}\sum_{n=0}^{N-1} a_{n+1}\left(\frac{L + \mathrm{i}z}{L - \mathrm{i}z}\right)^n \quad \text{for } \mathrm{Im}(z) \geq 0, \tag{79}$$

where the size of $N$ controls the accuracy of the approximation, $L := 2^{-1/4}N^{1/2}$, and the $a_n$ are discrete Fourier coefficients that can be precomputed by the FFT. He argues, based on operation counts, that, for intermediate values of $|z|$, the work required to compute $w(z)$ to $10^{-14}$ relative accuracy is much smaller for (79) than for the Poppe and Wijers algorithm [19].

Zaghloul and Ali proposed a method (see TOMS Algorithm 916 [29] and the

refinements in [27], and cf. [20] and [3, (7.1.29)]) starting from

$$\operatorname{erf}(z) = \operatorname{erf}(x) + \frac{2\mathrm{e}^{-x^2}}{\sqrt{\pi}} \int_0^y \mathrm{e}^{t^2} \sin(2xt)\, dt + \frac{2\mathrm{i}\,\mathrm{e}^{-x^2}}{\sqrt{\pi}} \int_0^y \mathrm{e}^{t^2} \cos(2xt)\, dt \tag{80}$$

for $z = x + \mathrm{i}y$. They approximate

$$w(z) \approx u(x,y) + \mathrm{i}v(x,y), \quad x, y \geq 0, \tag{81}$$

where

$$u(x,y) := \mathrm{e}^{-x^2}\operatorname{erfcx}(y)\cos(2xy) + \frac{2a\sin^2(xy)}{\pi y}\,\mathrm{e}^{-x^2} + \frac{ay}{\pi}\left(-2\cos(2xy)\,S_1 + S_2 + S_3\right),$$

$$v(x,y) := -\mathrm{e}^{-x^2}\operatorname{erfcx}(y)\sin(2xy) + \frac{a\sin(2xy)}{\pi y}\,\mathrm{e}^{-x^2} + \frac{a}{\pi}\left(2y\sin(2xy)\,S_1 - S_4 + S_5\right),$$

$\operatorname{erfcx}(y) := \mathrm{e}^{y^2}\operatorname{erf}(y)$, and $S_j$, $j = 1, \ldots, 5$, are the following summations reminiscent of the trapezoidal rule approximations (6):

$$\begin{aligned} S_1 &:= \sum_{k=1}^{\infty} \left(\frac{1}{a^2k^2 + y^2}\right) \mathrm{e}^{-(a^2k^2 + x^2)}, \quad S_2 := \sum_{k=1}^{\infty} \left(\frac{1}{a^2k^2 + y^2}\right) \mathrm{e}^{-(ak+x)^2}, \\ S_3 &:= \sum_{k=1}^{\infty} \left(\frac{1}{a^2k^2 + y^2}\right) \mathrm{e}^{-(ak-x)^2}, \quad S_4 := \sum_{k=1}^{\infty} \left(\frac{ak}{a^2k^2 + y^2}\right) \mathrm{e}^{-(ak+x)^2}, \\ S_5 &:= \sum_{k=1}^{\infty} \left(\frac{ak}{a^2k^2 + y^2}\right) \mathrm{e}^{-(ak-x)^2}. \end{aligned} \tag{82}$$

The authors have supplied us with their MATLAB implementation `Faddeyeva_v2(z,M)` [27], where the parameter $M$ is the number of accurate significant figures required, and the code enforces $4 \leq M \leq 13$. In this code the choice $a = 1/2$ is made and the sums in (82) are truncated, the number of terms retained depending on $M$. Zaghloul and Ali [29] (see also [27]) present numerical evidence that the approximation (81), with $a = 1/2$ and appropriate truncation of the infinite sums (82), is more accurate and faster than TOMS Algorithm 680 [19]. This algorithm and code have been used by Zaghloul [28] to benchmark a more efficient, low accuracy ($< 4 \times 10^{-5}$ maximum relative error in both real and imaginary parts) approximation to $w(z)$ for $z$ in the first quadrant.

Abrarov, Quine, and Jagpal [2] (also see [1]) proposed recently another method for computing $w(z)$ using modified rational approximations, namely

$$w(z) \approx \begin{cases} \psi_1(z) \text{ if } z \in D_1, \\ \psi_2(z) \text{ if } z \in D_2, \\ \psi_3(z) \text{ if } z \in D_3, \end{cases} \tag{83}$$

where $D_1 := \{z = x + \mathrm{i}y : |z| < 8 \text{ and } y > 0.05x\}$, $D_2 := \{z = x + \mathrm{i}y : |z| < 8 \text{ and } y \leq 0.05x\}$, $D_3 := \{z : |z| \geq 8\}$,

$$\psi_1(z) := \sum_{m=1}^{M} \frac{A_m + B_m(z + \mathrm{i}\alpha/2)}{C_m^2 - (z + \mathrm{i}\alpha/2)^2}, \quad \psi_2(z) := \mathrm{e}^{-z^2} + z\sum_{m=1}^{M+2} \frac{\alpha_m - \beta_m z^2}{\gamma_m - \theta_m z^2 + z^4}, \tag{84}$$

the coefficients $A_m$, $B_m$, $C_m$, $\alpha_m$, $\beta_m$, $\gamma_m$, and $\theta_m$ are specified in [2], and $\psi_3(z)$ is approximately (78) with $n = 10$ (see [2, equation (9)]). Abrarov, Quine, and Jagpal [2] present numerical evidence to show that (83) achieves an accuracy of $10^{-13}$ using $\alpha = 2.75$ and $M = 23$ in (84).

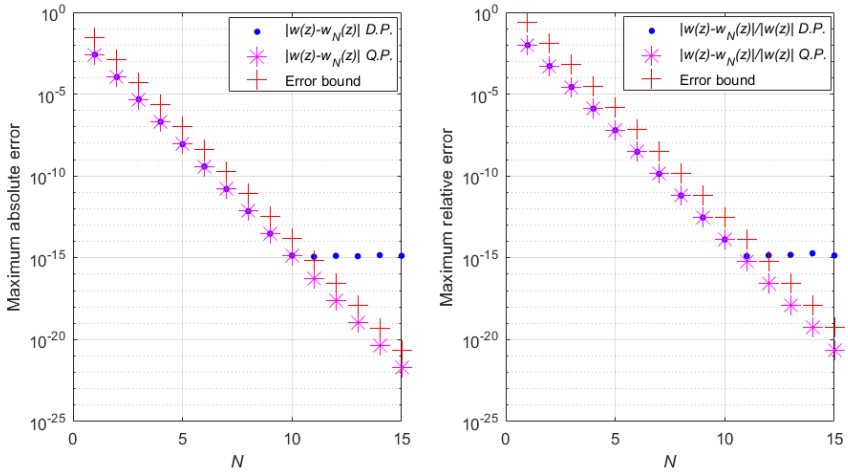


FIG. 2. *Maximum absolute and relative errors in the approximation* (10) *and the error bounds of Theorem* 1.1*, plotted against* $N$*, with* $w_N(z)$ *calculated in standard double precision (D.P.) and in quadruple precision (Q.P.). The maxima are taken over a large number of points* $z$ *in the first quadrant, ranging from exponentially small to exponentially large (see text for details).*

**4. Numerical results.** In this section we show calculations that illustrate and support Theorem 1.1, and that compare the accuracy and efficiency of our approximation $w_N(z)$ given by (10) to those of the approximations (79), (81), and (83). We omit comparison with the method of [28] because of its limited accuracy, and we omit comparison with the older algorithm of [19] because of evidence, discussed in section 3, that the approximation (81) is superior.

In Figure 2 we plot estimates of the maximum values in the first quadrant of the absolute and relative errors in our approximation (10) to $w(z)$. We show results for two implementations. The first of these uses the MATLAB code `wTrap(z,N)`, provided in Table SM1 of the supplementary material to this paper, which computes our approximation $w_N(z)$ using standard double precision arithmetic in MATLAB. The second implementation uses the MATLAB code `wTrap_Q(z,N)` in Table SM4 of the supplementary material, which computes $w_N(z)$ in quadruple precision arithmetic using the Multiprecision Computing Toolbox ADVANPIX (http://www.advanpix.com). In each case the maximum values we plot are discrete maxima taken over the $1,602,801$ points $z = 10^p e^{i\theta}$, with $p = -6(0.0006)6$ and $\theta = 0(\pi/1600)\pi/2$, a superset of the $40,401$ test values in Weideman [25, 26]. Whichever approximation for $w_N(z)$ is used, we use as the exact value for $w(z)$ the approximation $w_{20}(z)$, given by (10) and computed by a call to `wTrap_Q(z,N)` with $N = 20$. This approximation is predicted by Theorem 1.1 to have absolute and relative errors of $< 3.5 \times 10^{-28}$ and $< 9.4 \times 10^{-27}$, respectively, throughout the first quadrant. We expect these error bounds to be achieved working in quadruple precision arithmetic where the machine epsilon is $2^{-112} \approx 1.9 \times 10^{-34}$.

We observe in Figure 2 the rate of exponential convergence predicted by Theorem 1.1. When evaluated in standard double precision arithmetic the approximation $w_N(z)$ achieves, with $N = 11$ over this set of discrete points in the first quadrant, maximum absolute and relative errors which are $< 2 \times 10^{-15}$.

In Table 1 we compare the accuracy and efficiency of our approximation and

TABLE 1
*Maximum absolute and relative errors for the MATLAB codes implementing the approximations* (10), (79), (83), *and* (81). *The computing times are mean and s.d. of* 25 *executions.*

| Algorithm | Maximum abs. error | Maximum rel. error | Computing time (seconds) |
|---|---|---|---|
| `wTrap(z,11)` | $1.19 \times 10^{-15}$ | $1.31 \times 10^{-15}$ | 4.29 ($\pm$0.08) |
| `cef(z,40)` [25] | $1.33 \times 10^{-15}$ | $1.33 \times 10^{-15}$ | 4.20 ($\pm$0.02) |
| `fadsamp(z)` [2] | $3.78 \times 10^{-14}$ | $3.78 \times 10^{-14}$ | 5.74 ($\pm$0.04) |
| `Faddeyeva_v2(z,13)` [27] | $4.07 \times 10^{-15}$ | $1.71 \times 10^{-13}$ | 11.00 ($\pm$0.11) |

our (double precision) MATLAB code `wTrap(z,N)` with (double precision) MATLAB implementations of the approximations (79), (81), and (83). Results are shown in Table 1 for the following:

1. our approximation $w_N(z)$ with $N = 11$ implemented by the call `wTrap(z,11)` to the MATLAB code provided in Table SM1 of the supplementary material;
2. Weideman's approximation (79) with $N = 40$ (this choice of $N$ ensures high accuracy throughout the whole first quadrant; see [25, 26, Figure 8], [5, Figure 2]), implemented by the call `cef(z,40)` to the MATLAB code in [25, Table 1];
3. the approximation of Abrarov, Quine, and Jagpal, implemented by the call `fadsamp(z)` to the MATLAB function in [2, Appendix], which uses the method (83) with $\alpha = 2.75$ and $M = 23$ in the formulae for $\psi_1(z)$ and $\psi_2(z)$;
4. the approximation (81) of Zaghloul and Ali [29], implemented by the call `Faddeyeva_v2(z,M)` with $M = 13$ (the maximum value permitted by the code) to the MATLAB code described in [27]: here $M$ is the number of accurate significant figures required.

For these approximations Table 1 shows estimated maximum absolute and relative errors in the first quadrant, and computation times (mean and standard deviation of 25 executions, each measured by MATLAB `timeit`) running MATLAB version 9.3.0.713579 (R2017b) on a laptop with a single Intel64 Family 6 Model 78 2.40 GHz processor (and with `maxNumCompThreads` set to its default value of 2). The estimated maximum errors are discrete maxima over the same $1,602,801$ points as above, and we again use as the exact value the approximation $w_{20}(z)$ given by (10), implemented in quadruple precision by `wTrap_Q(z,20)`. The computation times are for the case where `z` is a matrix containing the $16,008,001$ points $z = x+\mathrm{i}y$, with $x = 0(0.0025)10$, $y = 0(0.0025)10$.

As measured by maximum absolute and relative errors over this large discrete set of $z$ values covering the first quadrant, our approximation $w_{11}(z)$ implemented as `wTrap(z,11)` is marginally the most accurate, though Weideman's approximation is essentially as accurate and all four methods achieve $< 4 \times 10^{-14}$ maximum absolute error and $< 2 \times 10^{-13}$ maximum relative error.

In these specific calculations our MATLAB code is, on average over 25 realizations, marginally slower than that of Weideman, but faster than the other two. But little should be read into these timings beyond, possibly, that the first three of the methods are about equally fast, and the fourth is a little slower. In particular we note the following:

(i) Repeating these timings, even on exactly the same machine with exactly the same version of MATLAB, will give slightly different results (see the additional timings results table in the supplementary materials).

(ii) These comparisons are made with the specific MATLAB implementations published by the various proposers of the algorithms. It may be that these timings could be reduced for any or all of these methods by better MATLAB implementation strategies.

(iii) Our timing comparisons in Table 1 are for computing $w(z)$ on a uniformly distributed grid of points across the square $\{z = x + \mathrm{i}y : 0 \le x \le 10, 0 \le y \le 10\}$. A different choice of points would lead to different results, at least for our approximation $w_N(z)$ and for the approximation of Abrarov, Quine, and Jagpal, given by (83), since both these approximations use somewhat different formulae in different regions of the first quadrant.

We leave to future publications more detailed comparisons of operation counts, timings, and accuracy of the above methods for computing $w(z)$ for complex $z$, along the lines of [25, 26] or [29, 27]. Such publications might also study alternative, potentially more efficient, implementations, for example using SIMD-aware C++ codes. We remark that C++ implementations of the method of Zaghloul and Ali [29] and of the continued fraction approximation (78) advocated by Gautschi [10] (which is used in the `fadsamp` code of Abrarov, Quine, and Jagpal [2]) are the basis of the widely used (though unpublished) Faddeeva Package of Johnson [13].

**Acknowledgment.** We would like to acknowledge the detailed and helpful comments of the two anonymous referees, whose input has led to significant improvements.

## REFERENCES

[1] S. M. Abrarov and B. M. Quine, *Sampling by incomplete cosine expansion of the sinc function: Application to the Voigt/complex error function*, Appl. Math. Comput., 258 (2015), pp. 425–435.

[2] S. M. Abrarov, B. M. Quine, and R. K. Jagpal, *A sampling-based approximation of the complex error function and its implementation without poles*, Appl. Numer. Math., 129 (2018), pp. 181–191.

[3] M. Abromowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1968.

[4] M. Al Azah, *Fast Evaluation of Special Functions by the Modified Trapezium Rule*, Ph.D. thesis, University of Reading, Reading, UK, 2017.

[5] M. Alazah, S. N. Chandler-Wilde, and S. La Porte, *Computing Fresnel integrals via modified trapezium rules*, Numer. Math., 128 (2014), pp. 635–661.

[6] B. Bialecki, *A modified sinc quadrature rule for functions with poles near the arc of integration*, BIT, 29 (1989), pp. 464–476.

[7] C. Chiarella and A. Reichel, *On the evaluation of integrals related to the error function*, Math. Comp., 22 (1968), pp. 137–143.

[8] J. B. Conway, *Functions of One Complex Variable*, 2nd ed., Grad. Texts in Math. 11, Springer, New York, 1978.

[9] P. J. Davis, *On the numerical integration of periodic analytic functions*, in Proceedings of a Symposium on Numerical Approximations, R. E. Langer, ed., University of Wisconsin Press, Madison, WI, 1959, pp. 45–59.

[10] W. Gautschi, *Efficient computation of the complex error function*, SIAM J. Numer. Anal., 7 (1970), pp. 187–198, https://doi.org/10.1137/0707012.

[11] E. Goodwin, *The evaluation of integrals of the form* $\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-x^2}dx$, Math. Proc. Cambridge Philos. Soc., 45 (1949), pp. 241–245.

[12] D. Hunter and T. Regan, *A note on the evaluation of the complementary error function*, Math. Comp., 26 (1972), pp. 539–541.

[13] S. G. Johnson, *Faddeeva Package*, 2012, http://ab-initio.mit.edu/wiki/index.php/Faddeeva_Package; downloaded 8 June 2021.

[14] S. La Porte, *Modified Trapezium Rule Methods for the Efficient Evaluation of Green's Functions in Acoustics*, Ph.D. thesis, Brunel University, London, 2007.

[15] F. Matta and A. Reichel, *Uniform computation of the error function and other related functions*, Math. Comp., 25 (1971), pp. 339–344.

[16] J. McNamee, *Error-bounds for the evaluation of integrals by the Euler-Maclaurin formula and by Gauss-type formulae*, Math. Comp., 18 (1964), pp. 368–381.
[17] M. Mori, *A method for evaluation of the error function of real and complex variable with high relative accuracy*, Publ. Res. Inst. Math. Sci., 19 (1983), pp. 1081–1094.
[18] *NIST Digital Library of Mathematical Functions*, Release 1.0.28 of 2020-09-15, F. W. J. Olver et al., eds., http://dlmf.nist.gov/.
[19] G. Poppe and C. Wijers, *More efficient computation of the complex error function*, ACM Trans. Math. Software, 16 (1990), pp. 38–46.
[20] H. E. Salzer, *Formulas for calculating the error function of a complex variable*, Math. Tables Aids Comp., 5 (1951), pp. 67–70.
[21] C. Schwartz, *Numerical integration of analytic functions*, J. Comput. Phys., 4 (1969), pp. 19–29.
[22] F. Stenger, *Integration formulae based on the trapezoidal formula*, IMA J. Appl. Math., 12 (1973), pp. 103–114.
[23] L. N. Trefethen and J. A. C. Weideman, *The exponentially convergent trapezoidal rule*, SIAM Rev., 56 (2014), pp. 385–458, https://doi.org/10.1137/130932132.
[24] A. M. Turing, *A method for the calculation of the zeta-function*, Proc. London Math. Soc. (2), 48 (1943), pp. 180–197.
[25] J. A. C. Weideman, *Computation of the complex error function*, SIAM J. Numer. Anal., 31 (1994), pp. 1497–1518, https://doi.org/10.1137/0731077.
[26] J. A. C. Weideman, *Erratum: Computation of the complex error function*, SIAM J. Numer. Anal., 32 (1995), pp. 330–331, https://doi.org/10.1137/0732014.
[27] M. R. Zaghloul, *Remark on "Algorithm* 916: *Computing the Faddeyeva and Voigt functions": Efficiency improvements and Fortran translation*, ACM Trans. Math. Software, 42 (2016), 26.
[28] M. R. Zaghloul, *Algorithm* 985: *simple, efficient, and relatively accurate approximation for the evaluation of the Faddeyeva function*, ACM Trans. Math. Software, 44 (2017), 22.
[29] M. R. Zaghloul and A. N. Ali, *Algorithm* 916: *Computing the Faddeyeva and Voigt functions*, ACM Trans. Math. Software, 38 (2011), 15.