

In this paper, we present an adaptive delaminating Levin method for evaluating bivariate oscillatory integrals over rectangular domains. Whereas previous analyses of Levin methods impose non-resonance conditions that exclude stationary and resonance points, we rigorously establish the existence of a slowly-varying, approximate solution to the Levin PDE across all frequency regimes, even when the non-resonance condition is violated. This allows us to derive error estimates for the numerical solution of the Levin PDE via the Chebyshev spectral collocation method, and for the evaluation of the corresponding oscillatory integrals, showing that high accuracy can be achieved regardless of whether or not stationary and resonance points are present. We show that, by incorporating adaptive subdivision in both two and one dimensions, as well as delaminating Chebyshev spectral collocation, our algorithm is effective in both the presence and absence of stationary and resonance points, which we demonstrate with a number of numerical experiments.

An adaptive delaminating Levin method in two dimensions

Shukai Chen^{†◊*}, Kirill Serkh^{‡†◊},
James Bremer^{‡◊}, Murdock Aubry[†]
May 27, 2025

◊ This author's work was supported in part by the NSERC Discovery Grants RGPIN-2020-06022 and DGEER-2020-00356.

◊ This author's work was supported in part by the NSERC Discovery Grant RGPIN-2021-02613.

† Dept. of Computer Science, University of Toronto, Toronto, ON M5S 2E4

‡ Dept. of Mathematics, University of Toronto, Toronto, ON M5S 2E4

* Corresponding author. Email: shukai.chen@mail.utoronto.ca

Keywords: *numerical integration, multivariate oscillatory integrals, Levin method, Chebyshev polynomials, adaptive integration*

1 Introduction

The subject of this paper is the efficient evaluation of oscillatory integrals of the form

$$I[f, g] = \int_{-1}^1 \int_{-1}^1 f(x, y) \exp(ig(x, y)) \, dx \, dy, \quad (1)$$

where $f: [-1, 1]^2 \rightarrow \mathbb{C}$ and $g: [-1, 1]^2 \rightarrow \mathbb{R}$ are smooth, slowly-varying functions, and ∇g can be of large magnitude. We refer to f as the amplitude function, and g as the phase function.

When evaluating the integral $I[f, g]$ using the most straightforward quadrature methods, adaptive Gauss-Legendre quadrature for example, the incurred cost grows at best linearly and at worst quadratically with the magnitude of ∇g . When the magnitude of ∇g is large, this exorbitant cost renders such methods impractical, motivating the development of many approaches for the numerical integration of highly oscillatory functions, including asymptotic expansions, numerical steepest descent, Filon methods, and Levin methods. For a review of these methods, we refer to the monograph [1] and references therein. This paper focuses on Levin methods, which date back to David Levin's seminal work [2]. These methods reformulate the problem of evaluating an oscillatory integral as a problem of solving a differential equation with slowly-varying coefficients, whose solution can be computed to a fixed desired accuracy in time independent of the magnitude of ∇g .

The classical Levin method, originally introduced in [2], works by solving the second-order partial differential equation (PDE)

$$p_{xy} + ig_y p_x + ig_x p_y + (ig_{xy} - g_x g_y) p = f \text{ in } (-1, 1)^2 \quad (2)$$

by collocation with the monomial basis, so that the slowly-varying solution p satisfies

$$\frac{\partial^2}{\partial x \partial y} (p(x, y) \exp(ig(x, y))) = f(x, y) \exp(ig(x, y)), \quad (3)$$

for all $(x, y) \in (-1, 1)^2$. The value of the integral $I[f, g]$ is then readily obtained as

$$I[f, g] = -p(-1, 1) \exp(ig(-1, 1)) - p(1, -1) \exp(ig(1, -1)) \\ + p(1, 1) \exp(ig(1, 1)) + p(-1, -1) \exp(ig(-1, -1)). \quad (4)$$

In order to guarantee the existence of a slowly-varying solution to (2), which can then be approximated by a basis of slowly-varying functions, Levin's method imposes a non-resonance condition requiring that both g_x and g_y are nonvanishing and that their maximum absolute values over the domain are large. This non-resonance condition excludes stationary points of g , where both g_x and g_y vanish, as well as resonance points, which are boundary points where ∇g aligns with the normal vector.

A generalization of the classical Levin method, applicable to non-polytopal domains, was later proposed by Olver in [3]. Instead of the PDE (2), Olver's method considers the vector-valued first-order PDE,

$$\mathcal{L}[\mathbf{p}] = \nabla \cdot \mathbf{p} + i \nabla g \cdot \mathbf{p} = f \text{ in } \Omega, \quad (5)$$

which we refer to as the Levin PDE, where $\Omega \subset \mathbb{R}^n$ is a connected, open, and bounded set with piecewise smooth boundary Γ . If $\mathbf{p} = (p_1, \dots, p_n)^T: \bar{\Omega} \rightarrow \mathbb{C}^n$ is a solution to (5), then

$$\nabla \cdot (\mathbf{p}(\mathbf{x}) \exp(ig(\mathbf{x}))) = f(\mathbf{x}) \exp(ig(\mathbf{x})), \quad (6)$$

for all $\mathbf{x} \in \Omega$, and, by the divergence theorem,

$$\int_{\Omega} f(\mathbf{x}) \exp(ig(\mathbf{x})) \, dA = \int_{\Gamma} \mathbf{p}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \exp(ig(\mathbf{x})) \, d\ell. \quad (7)$$

For $\Omega = (-1, 1)^2$, the value of $I[f, g]$ is given by

$$\begin{aligned} I[f, g] = & - \int_{-1}^1 p_2(x, -1) \exp(ig(x, -1)) \, dx + \int_{-1}^1 p_1(1, y) \exp(ig(1, y)) \, dy \\ & + \int_{-1}^1 p_2(x, 1) \exp(ig(x, 1)) \, dx - \int_{-1}^1 p_1(-1, y) \exp(ig(-1, y)) \, dy. \end{aligned} \quad (8)$$

In order for Olver's method to be applicable, it must be possible to both accurately solve (5) over the domain, and to efficiently evaluate the boundary integral in (7). To ensure the solvability of (5) using collocation, Olver's method eliminates the null space of the differential operator \mathcal{L} appearing in the left-hand side of the PDE, which consists of vector fields of the form

$$\mathbf{p}(\mathbf{x}) = \mathbf{q}(\mathbf{x}) \exp(-ig(\mathbf{x})), \quad (9)$$

with $\nabla \cdot \mathbf{q}(\mathbf{x}) = 0$, by imposing a non-resonance condition analogous to the condition found in the classical Levin method. Olver's non-resonance condition requires that ∇g is both of large magnitude and is nowhere orthogonal to the boundary of the domain. This non-resonance condition again excludes both stationary points and resonance points, noting that, in Olver's formulation, resonance points are precisely stationary points of the phase function in the right-hand side of (7).

When the stationary points have locations which are known beforehand, certain methods, such as the modified Levin-GMRES method [4] and the augmented Levin method [5], can be used to efficiently evaluate the corresponding integrals in the univariate case. Besides that these methods have yet to be generalized to higher dimensions, they are not applicable when the number and locations of the stationary points of the phase function are unknown a priori. These challenges have led to the long-standing belief that developing efficient multivariate Levin methods is particularly difficult when stationary and resonance points are present, and that their absence is essential for the success of such methods.

It has been observed however that, in the univariate case, Levin methods are much more effective in the presence of stationary points than was previously believed. For univariate oscillatory integrals of the form

$$\int_{-1}^1 f(x) \exp(ig(x)) \, dx, \quad (10)$$

a Chebyshev spectral collocation method was used by Li et al. in [6, 7] to discretize the Levin ordinary differential equation (ODE)

$$p' + ig'p = f \text{ in } (-1, 1), \quad (11)$$

whose solution leads to the formula

$$\begin{aligned} \int_{-1}^1 f(x) \exp(ig(x)) dx &= \int_{-1}^1 \frac{d}{dx} (p(x) \exp(ig(x))) dx \\ &= p(1) \exp(ig(1)) - p(-1) \exp(ig(-1)). \end{aligned} \quad (12)$$

Li et al. present numerical evidence showing that, when the resulting linear system is solved via a truncated singular value decomposition, high accuracy is obtained even in the presence of stationary points. In fact, a similar Levin collocation method was previously shown to be effective in Moylan's thesis [8], in which it was combined with an adaptive subdivision strategy. Recently, Chen et al. [9] proved the existence of a slowly-varying, approximate solution to the ODE (11) in the presence of stationary points, and demonstrated that the univariate Levin method, which discretizes (11) via the Chebyshev spectral collocation method and solves the resulting linear system using a truncated singular value decomposition, can be applied adaptively to accurately and efficiently evaluate the integral (10). We claim that, as a result, it is not actually necessary to exclude the presence of resonance points when computing $I[f, g]$, since the univariate adaptive Levin method can be used to evaluate the boundary integrals in (7) efficiently.

In the presence of stationary points, ill-conditioned linear systems are encountered when solving the PDE (5) using collocation methods, since the null space of the differential operator is fully resolved by any collocation grid dense enough to resolve f and ∇g . Ashton [10] observes that, with appropriate Cauchy data, the collocation system can be transformed into a well-conditioned problem, from which a solution \mathbf{p} to (5) with $\|\nabla \cdot \mathbf{p}\|_{L^2(\Omega)}^2 = \mathcal{O}(\|\nabla g\|_{L^\infty(\Omega)}^{-1/2})$ can be recovered. In particular, Ashton studies the PDE

$$\mathbf{c} \cdot \nabla p + i(\mathbf{c} \cdot \nabla g)p = f \text{ in } \Omega, \quad (13)$$

where Ω is an n -simplex in \mathbb{R}^n and g has a stationary point at a vertex of Ω . This PDE can be derived from (5) by considering solutions of the form $\mathbf{p} = \mathbf{c}p$, where \mathbf{c} is a nonzero constant and $p: \bar{\Omega} \rightarrow \mathbb{C}$. Under a transversality condition on \mathbf{c} and g , Ashton constructs initial data that determines a solution p satisfying $\|\mathbf{c} \cdot \nabla p\|_{L^2(\Omega)}^2 = \mathcal{O}(\|\nabla g\|_{L^\infty(\Omega)}^{-1/2})$, and uses the initial data to obtain a well-conditioned collocation system. We note that their solution may have higher order derivatives that grow with $\|\nabla g\|_{L^\infty(\Omega)}$.

Instead of constructing an appropriate initial condition for (13), Li et al. [11] address the problem of ill-conditioned linear systems using truncated singular value decompositions. Specifically, they consider the PDE

$$p_y + ig_y p = f \text{ in } (-1, 1)^2, \quad (14)$$

which is a special case of (13) with $\mathbf{c} = (0, 1)^T$ and $\Omega = (-1, 1)^2$, and present numerical evidence showing that, when (14) is discretized using the Chebyshev spectral collocation method and the resulting linear system is solved using a truncated singular value decomposition, the integral $I[f, g]$ can be computed much more accurately and efficiently than

the classical Levin method. The delaminating quadrature method of Li et al. exploits the block diagonal structure of the resulting linear system, where each block corresponds to a univariate fiber of the domain, in order to reduce computational complexity by solving each block independently. Furthermore, it is shown in [11] that their method can converge rapidly in the presence of stationary and resonance points, provided that the domain is adaptively refined around them.

In this paper, we prove that the Levin PDE admits a well-behaved, approximate solution, both in the case where ∇g is nonvanishing, and in the case where ∇g is of small magnitude and g can have stationary points. We further prove that, when the Levin PDE is discretized using the Chebyshev spectral collocation method and the resulting linear system is solved numerically using a truncated singular value decomposition, high accuracy can be achieved regardless of whether or not g has stationary and resonance points. Finally, we demonstrate through a set of numerical experiments that an adaptive delaminating Levin method, which solves the Levin PDE using collocation and computes the boundary integrals in (8) using the univariate adaptive Levin method presented in [9], is highly effective for evaluating integrals of the form (1). We note in passing the great generality of integrals expressible in the form (1), as most oscillators of interest admit phase function representations that can be efficiently constructed using the algorithms of [12, 13, 14].

The remainder of this paper is structured as follows. Section 2 reviews the necessary mathematical and numerical preliminaries. In Section 3, we prove the existence of a slowly-varying, approximate solution to the Levin PDE. In Section 4, we derive error estimates for the numerical solution of the Levin PDE via the Chebyshev spectral collocation method, and for the computed value of the integral. We give a detailed description of the adaptive delaminating Levin method in Section 5, and present the results of numerical experiments demonstrating the properties of the algorithm in Section 6. Finally, we provide a few comments regarding this work and future directions for research in Section 7.

2 Preliminaries

2.1 Notation and conventions

An n -tuple of the form $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ is called a multi-index of order $|\boldsymbol{\alpha}|$, where

$$|\boldsymbol{\alpha}| = \alpha_1 + \dots + \alpha_n. \quad (15)$$

Given a multi-index $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, we define

$$\mathbf{x}^\boldsymbol{\alpha} = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}, \quad (16)$$

and

$$D^\boldsymbol{\alpha} u(\mathbf{x}) = \frac{\partial^{|\boldsymbol{\alpha}|} u(\mathbf{x})}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}. \quad (17)$$

For $u: U \subset \mathbb{R}^n \rightarrow \mathbb{C}$, the notation

$$\frac{\partial u}{\partial x}, \quad \frac{\partial^2 u}{\partial x \partial y} \quad (18)$$

is used interchangeably with u_x, u_{xy} to denote partial derivatives. For $\mathbf{u}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, we employ the notation $D\mathbf{u}$ for the Jacobian matrix of \mathbf{u} .

If X is a topological space, we denote by $C(X)$ the space of continuous complex-valued functions on X . For a locally compact Hausdorff space X , we denote the space of complex Radon measures on X by $M(X)$, and use $|\mu|$ for the total variation of $\mu \in M(X)$.

For an open subset U of \mathbb{R}^n , we denote by $C^k(U)$ the space of k -times continuously differentiable functions, and by $C^k(\bar{U})$ the subspace of functions u in $C^k(U)$ such that $D^\alpha u$ is uniformly continuous on bounded subsets of U , and hence extends continuously to \bar{U} , for all multi-indices α with $|\alpha| \leq k$. We denote by $C_b^k(U)$ the subspace of functions u in $C^k(U)$ such that $D^\alpha u$ is bounded for all multi-indices α with $|\alpha| \leq k$, and by $C_c^k(U)$ the subspace of functions in $C^k(U)$ with compact support. We use the notations $C^\infty(U) = \bigcap_{k=0}^\infty C^k(U)$ and $C^\infty(\bar{U}) = \bigcap_{k=0}^\infty C^k(\bar{U})$ for the spaces of infinitely differentiable functions. For a Lebesgue measurable subset U of \mathbb{R}^2 , we denote by $L^\infty(U)$ the Banach space of measurable functions which are essentially bounded on U .

We let $\mathcal{S}(\mathbb{R}^n)$ denote the space of infinitely differentiable functions φ such that

$$\|\varphi\|_{\alpha, \beta} = \sup_{\mathbf{x} \in \mathbb{R}^n} |\mathbf{x}^\alpha D^\beta \varphi(\mathbf{x})| < \infty, \quad (19)$$

for all multi-indices α, β . Its topological dual, which is the space of tempered distributions, is denoted by $\mathcal{S}'(\mathbb{R}^n)$. We write $\langle \mathcal{T}, \varphi \rangle$ for the action of the tempered distribution \mathcal{T} on the test function $\varphi \in \mathcal{S}(\mathbb{R}^n)$. The support of $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^n)$ is defined as the complement of the union of all open subsets U of \mathbb{R}^n such that $\langle \mathcal{T}, \varphi \rangle = 0$ whenever $\varphi \in \mathcal{S}(\mathbb{R}^n)$ is supported in U . We use $\mathcal{E}'(\mathbb{R}^n)$ to denote the subspace of $\mathcal{S}'(\mathbb{R}^n)$ consisting of compactly supported tempered distributions. We say that a tempered distribution $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^n)$ is of order at most m if, for each compact set $K \subset \mathbb{R}^n$, there exists a constant M_K such that

$$|\langle \mathcal{T}, \varphi \rangle| \leq M_K \sup_{|\alpha| \leq m} \sup_{\mathbf{x} \in K} |D^\alpha \varphi(\mathbf{x})| \quad (20)$$

for every function $\varphi \in \mathcal{S}(\mathbb{R}^n)$ supported in K . The order of a tempered distribution \mathcal{T} is the least nonnegative integer m with the property above, noting that all tempered distributions are of finite order.

We use the convention

$$\widehat{\varphi}(\xi_1, \xi_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y) \exp(-i(\xi_1 x + \xi_2 y)) \, dx \, dy \quad (21)$$

for the Fourier transform of $\varphi \in \mathcal{S}(\mathbb{R}^2)$, and, accordingly, we have

$$\varphi(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widehat{\varphi}(\xi_1, \xi_2) \exp(i(\xi_1 x + \xi_2 y)) \, d\xi_1 \, d\xi_2. \quad (22)$$

The Fourier transform is extended to a topological automorphism in $\mathcal{S}'(\mathbb{R}^2)$ via the formula $\langle \widehat{\mathcal{T}}, \varphi \rangle = \langle \mathcal{T}, \widehat{\varphi} \rangle$. We use the operators

$$\mathcal{F}_y[\varphi](x, \xi_2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(x, y) \exp(-i\xi_2 y) \, dy, \quad (23)$$

and

$$\mathcal{F}_y^{-1}[\widehat{\varphi}](\xi_1, y) = \int_{-\infty}^{\infty} \widehat{\varphi}(\xi_1, \xi_2) \exp(i\xi_2 y) \, d\xi_2 \quad (24)$$

to denote the partial Fourier transform with respect to y and its inverse. If the Fourier transform of φ is supported in $[-c, c]^2$, we say that φ has bandlimit c . Note that we do not require c to be the smallest positive real number with this property.

We use bold capital letters for matrices, and bold lowercase letters for vectors. For $1 \leq p \leq \infty$, we denote by $\|\mathbf{x}\|_p$ the p -norm of a vector $\mathbf{x} \in \mathbb{C}^n$, and by $\|\mathbf{A}\|_p$ the matrix norm of a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ induced by the vector p -norms.

The notation $x \lesssim y$ is used to indicate that there is some constant $C > 0$, not depending on either x or y , such that $x \leq Cy$. For $\mathbf{x} \in \mathbb{C}^n$, the notations $\|\mathbf{x}\|_p \lesssim y$ and $\|\mathbf{x}\|_p \gtrsim y$ are used to indicate that there is some nonvanishing polynomial $C(n)$ such that $\|\mathbf{x}\|_p \leq C(n)y$ and $C(n)\|\mathbf{x}\|_p \geq y$, respectively. For $\mathbf{A} \in \mathbb{C}^{m \times n}$, the notations $\|\mathbf{A}\|_p \lesssim y$ and $\|\mathbf{A}\|_p \gtrsim y$ are used to indicate that there is some nonvanishing polynomial $C(m, n)$ such that $\|\mathbf{A}\|_p \leq C(m, n)y$ and $C(m, n)\|\mathbf{A}\|_p \geq y$, respectively.

2.2 Chebyshev interpolation

We denote by T_n the Chebyshev polynomial of degree n . For $k > 1$, the extrema of T_{k-1} on the interval $[-1, 1]$ are termed the Chebyshev nodes, which form the k -point Chebyshev grid, and are denoted by

$$-1 = t_{1,k} < t_{2,k} < \dots < t_{k,k} = 1, \quad (25)$$

where

$$t_{j,k} = \cos\left(\frac{k-j}{k-1}\pi\right), \quad j = 1, \dots, k. \quad (26)$$

The $k \times k$ tensor product Chebyshev grid on $[-1, 1]^2$ is given by

$$\left\{ \mathbf{x}_{i,k}^{(j)} \right\}_{i,j=1}^k, \quad \text{where } \mathbf{x}_{i,k}^{(j)} = (t_{i,k}, t_{j,k}), \quad (27)$$

so that i indexes nodes in the x -direction and j in the y -direction. We will find it useful to represent the Chebyshev grid as a single vector of length k^2 , for which we use the notation

$$\mathbf{x}_{i+k(j-1),k^2} = \mathbf{x}_{i,k}^{(j)}, \quad i, j = 1, \dots, k. \quad (28)$$

For $f \in C^1([-1, 1]^2)$, we denote by $\mathbf{P}_k[f]$ the unique Chebyshev expansion of the form

$$\sum_{m,n=0}^{k-1} a_{mn} T_m(x) T_n(y), \quad a_{mn} \in \mathbb{C}, \quad (29)$$

which agrees with f at the nodes (27). It is well-known that $\mathbf{P}_k[f]$ converges to f uniformly as $k \rightarrow \infty$. For bivariate polynomials, we use the term ‘‘degree’’ to mean the maximum of the degrees in x and the degrees in y , whichever is larger, of all terms in the polynomial, rather than the total degree of the polynomial. Evidently, if f is a bivariate polynomial of degree less than k , then

$$\mathbf{P}_k[f](x, y) = f(x, y), \quad (30)$$

and

$$\mathbf{P}_k \left[\frac{\partial f}{\partial x} \right] (x, y) = \frac{\partial f}{\partial x} (x, y) = \frac{\partial}{\partial x} \mathbf{P}_k[f](x, y). \quad (31)$$

We denote by \mathbf{D}_k the $k \times k$ Chebyshev differentiation matrix which takes the vector

$$(f(t_{1,k}) \quad f(t_{2,k}) \quad \cdots \quad f(t_{k,k}))^T \quad (32)$$

of the values of a univariate polynomial $f(x)$ of degree less than k at the Chebyshev nodes to the vector

$$(f'(t_{1,k}) \quad f'(t_{2,k}) \quad \cdots \quad f'(t_{k,k}))^T \quad (33)$$

of the values of its derivative at the same nodes. Since the nodes (27) are indexed in the usual lexicographical order on \mathbb{R}^2 , the matrix

$$(\mathbf{D}_x)_k = \mathbf{I}_k \otimes \mathbf{D}_k \quad (34)$$

computes the values of the partial derivative of a bivariate polynomial $f(x, y)$ of degree less than k with respect to x on the $k \times k$ tensor product Chebyshev grid, where the symbol \otimes denotes the Kronecker product of matrices and \mathbf{I}_k is the $k \times k$ identity matrix. We denote by $\mathbf{P}_{(k,\ell)}$ the $k^2 \times \ell^2$ interpolation matrix which takes the vector

$$(f(\mathbf{x}_{1,\ell^2}) \quad f(\mathbf{x}_{2,\ell^2}) \quad \cdots \quad f(\mathbf{x}_{\ell^2,\ell^2}))^T \quad (35)$$

of the values of a bivariate polynomial $f(x, y)$ of degree less than ℓ on the $\ell \times \ell$ tensor product Chebyshev grid to the vector

$$(f(\mathbf{x}_{1,k^2}) \quad f(\mathbf{x}_{2,k^2}) \quad \cdots \quad f(\mathbf{x}_{k^2,k^2}))^T \quad (36)$$

of the values of the polynomial on the $k \times k$ tensor product Chebyshev grid.

It is well known from classic Sturm-Liouville theory (see, for instance, Section 4 of [15]) that, if $f \in C^\infty([-1, 1]^2)$, then it admits a uniformly convergent Chebyshev series

$$f(x, y) = \sum_{m,n=0}^{\infty} b_{mn} T_m(x) T_n(y), \quad (37)$$

such that $|b_{mn}|$ decays super-algebraically as $m, n \rightarrow \infty$. It follows from the aliasing formula for Chebyshev coefficients that

$$\|\mathbf{P}_k[f] - f\|_{L^\infty([-1,1]^2)} \leq 2 \sum_{\max\{m,n\} \geq k} |b_{mn}|, \quad (38)$$

and

$$\left\| \frac{\partial}{\partial x} \mathbf{P}_k[f] - \frac{\partial f}{\partial x} \right\|_{L^\infty([-1,1]^2)} \leq 2 \sum_{\max\{m,n\} \geq k} m^2 |b_{mn}|. \quad (39)$$

Lemma 2.1. *Suppose that f is a bivariate polynomial of degree less than k , and that*

$$\mathbf{f} = (f(\mathbf{x}_{1,k^2}) \quad f(\mathbf{x}_{2,k^2}) \quad \cdots \quad f(\mathbf{x}_{k^2,k^2}))^T. \quad (40)$$

Then,

$$\|f\|_{L^\infty([-1,1]^2)} \leq \Lambda_k^2 \|\mathbf{f}\|_\infty, \quad (41)$$

where Λ_k is the Lebesgue constant for the k -point Chebyshev grid on $[-1, 1]$.

Proof. Let $(x, y) \in [-1, 1]^2$ with

$$|f(x, y)| = \left| \sum_{i=1}^k \sum_{j=1}^k f(t_{i,k}, t_{j,k}) \ell_i(x) \ell_j(y) \right| = \|f\|_{L^\infty([-1,1]^2)}, \quad (42)$$

where ℓ_1, \dots, ℓ_k are the Lagrange polynomials associated with the k -point Chebyshev grid. The conclusion of the lemma follows immediately from this and the definition

$$\Lambda_k = \max_{-1 \leq x \leq 1} \sum_{i=1}^k |\ell_i(x)|. \quad (43)$$

■

2.3 Tempered distributions

In this subsection, we collect some standard results from the theory of distributions that we will use in the analysis of the Levin PDE.

Definition 2.1. For $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, the convolution $\mathcal{T} * \varphi$ is defined as $\mathcal{T} * \varphi(\mathbf{x}) = \langle \mathcal{T}, \tau_{\mathbf{x}} \tilde{\varphi} \rangle$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\tau_{\mathbf{x}} \varphi(\mathbf{y}) = \varphi(\mathbf{y} - \mathbf{x})$ denotes translation by \mathbf{x} and $\tilde{\varphi}(\mathbf{x}) = \varphi(-\mathbf{x})$ denotes reflection through the origin.

The following lemma is reproduced from Proposition 9.10 of [16].

Lemma 2.2. *If $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then $\mathcal{T} * \varphi$ is a slowly increasing C^∞ function, and satisfies*

$$D^\alpha(\mathcal{T} * \varphi) = (D^\alpha \mathcal{T}) * \varphi = \mathcal{T} * (D^\alpha \varphi) \quad (44)$$

for all multi-indices α . Furthermore,

$$\text{supp}(\mathcal{T} * \varphi) \subset \text{supp} \mathcal{T} + \text{supp} \varphi. \quad (45)$$

The following provides the standard definition for the tensor product of two tempered distributions.

Definition 2.2. For $\mathcal{S} \in \mathcal{S}'(\mathbb{R}^m)$ and $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^n)$, the tensor product $\mathcal{S} \otimes \mathcal{T}$ is a tempered distribution on $\mathbb{R}^m \times \mathbb{R}^n$ defined via the formula

$$\langle \mathcal{S}(\mathbf{x}) \otimes \mathcal{T}(\mathbf{y}), \varphi(\mathbf{x}, \mathbf{y}) \rangle = \langle \mathcal{S}(\mathbf{x}), \langle \mathcal{T}(\mathbf{y}), \varphi(\mathbf{x}, \mathbf{y}) \rangle \rangle \quad (46)$$

for all $\varphi(\mathbf{x}, \mathbf{y}) \in \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$.

The notation we have used has the obvious meaning that, for each fixed $\mathbf{x} \in \mathbb{R}^m$, $\langle \mathcal{T}(\mathbf{y}), \varphi(\mathbf{x}, \mathbf{y}) \rangle$ is the action of the distribution \mathcal{T} on the function $\varphi(\mathbf{x}, \cdot) \in \mathcal{S}(\mathbb{R}^n)$. A proof of the fact that $\langle \mathcal{T}(\mathbf{y}), \varphi(\mathbf{x}, \mathbf{y}) \rangle$ defines a function in $\mathcal{S}(\mathbb{R}^m)$ can be found in Chapter 7 of [17].

A proof of the following lemma, which states that the tensor product commutes, can be found in the same place.

Lemma 2.3. *The tensor product is commutative, that is, $\mathcal{S}(\mathbf{x}) \otimes \mathcal{T}(\mathbf{y}) = \mathcal{T}(\mathbf{y}) \otimes \mathcal{S}(\mathbf{x})$.*

In Theorem 2.6, we prove the structure theorem for tempered distributions that are compactly supported in a bounded convex set with nonempty interior.

A key ingredient in our proof is the continuity of the following extension operator E , which is constructed in Chapter VI of [18].

Lemma 2.4. *Let $U \subset \mathbb{R}^n$ be an open, bounded, convex set. Then there exists a linear extension operator $E: C_b(U) \rightarrow C_b(\mathbb{R}^n)$ such that E maps $C_b^m(U)$ continuously into $C_b^m(\mathbb{R}^n)$ for all nonnegative integers m , and $E[f]$, $f \in C_b(U)$, is supported in a fixed compact set that depends only on the domain U .*

Note that, if $\mathcal{T} \in \mathcal{E}'(\mathbb{R}^n)$ is of order at most m , then formula (20), together with the fact that $C_c^\infty(\mathbb{R}^n)$ is dense in $C^m(\mathbb{R}^n)$, means that \mathcal{T} can be continuously extended to $C^m(\mathbb{R}^n)$. The following lemma can be proven by constructing a sequence of bells $(\rho_n)_{n=1}^\infty$ in $C_c^\infty(\mathbb{R}^n)$ such that $\{\mathbf{x} : \rho_n(\mathbf{x}) = 1\}^\circ \supset \text{supp } \mathcal{T}$ and $\bigcap_{n=1}^\infty \text{supp } \rho_n = \text{supp } \mathcal{T}$, and showing that $T(\varphi) = T(\rho_n \varphi) \rightarrow 0$.

Lemma 2.5. *Let $\mathcal{T} \in \mathcal{E}'(\mathbb{R}^n)$ be of order at most m , and $\varphi \in C^m(\mathbb{R}^n)$ with $D^\alpha \varphi = 0$ on $\text{supp } \mathcal{T}$ for all $|\alpha| \leq m$. Then $\langle \mathcal{T}, \varphi \rangle = 0$.*

The following theorem provides a structure theorem for tempered distributions that are compactly supported in a bounded convex set with nonempty interior. The proof closely follows Appendix 2 of [17].

Theorem 2.6. *Let $U \subset \mathbb{R}^n$ be an open, bounded, convex set, and suppose that $\mathcal{T} \in \mathcal{E}'(\mathbb{R}^n)$ is of order at most m with $\text{supp } \mathcal{T} \subset \bar{U}$. Then there exist complex Radon measures $\mu_\alpha \in M(\bar{U})$, corresponding to the set of multi-indices α with $|\alpha| \leq m$, such that*

$$\langle \mathcal{T}, \varphi \rangle = \sum_{|\alpha| \leq m} \int_{\bar{U}} D^\alpha \varphi(\mathbf{x}) d\mu_\alpha(\mathbf{x}) \quad (47)$$

for all $\varphi \in C^m(\mathbb{R}^n)$.

Proof. Using the extension operator E in Lemma 2.4, for each $\varphi \in C^m(\mathbb{R}^n)$, we see that there exists a function $\tilde{\varphi} = E[\varphi|_U]$ compactly supported in a set $K \supset \bar{U}$ such that $D^\alpha \tilde{\varphi} = D^\alpha \varphi$ on $\text{supp } \mathcal{T}$ for all $|\alpha| \leq m$, and

$$\sup_{|\alpha| \leq m} \sup_{\mathbf{x} \in K} |D^\alpha \tilde{\varphi}(\mathbf{x})| \leq A_m \sup_{|\alpha| \leq m} \sup_{\mathbf{x} \in \bar{U}} |D^\alpha \varphi(\mathbf{x})| \quad (48)$$

for some constant A_m that depends on the order m .

Combining Lemma 2.5 with the estimate (48), we obtain the bound

$$\begin{aligned} |\langle \mathcal{T}, \varphi \rangle| &= |\langle \mathcal{T}, \tilde{\varphi} \rangle| \leq M_K \sup_{|\alpha| \leq m} \sup_{\mathbf{x} \in K} |D^\alpha \tilde{\varphi}(\mathbf{x})| \\ &\leq A_m M_K \sup_{|\alpha| \leq m} \sup_{\mathbf{x} \in \bar{U}} |D^\alpha \varphi(\mathbf{x})|. \end{aligned} \quad (49)$$

Let $N = |\{\alpha: |\alpha| \leq m\}|$, and S be the union of N disjoint copies of \bar{U} endowed with the disjoint union topology. The space $C(S)$ is the Banach space of all continuous complex-valued functions $\psi = (\psi_1, \dots, \psi_N)$ on S , where $\psi_k \in C(\bar{U})$ for all $1 \leq k \leq N$. Let $Y \subset C(S)$ be the subspace consisting of all m -jets $(D^\alpha \varphi)_{|\alpha| \leq m}$ for $\varphi \in C^m(\mathbb{R}^n)$.

From (49), it is clear that \mathcal{T} induces a bounded linear operator on Y . By the Hahn-Banach theorem, \mathcal{T} has an extension to a bounded linear operator on $C(S)$. Since S is compact, it then follows from the Riesz representation theorem that there exists a complex Radon measure $\mu \in M(S)$ such that, for all $\varphi \in C^m(\mathbb{R}^n)$,

$$\langle \mathcal{T}, \varphi \rangle = \sum_{|\alpha| \leq m} \int_{\bar{U}} D^\alpha \varphi(\mathbf{x}) d\mu_\alpha(\mathbf{x}), \quad (50)$$

where $\mu_\alpha \in M(\bar{U})$ are the restrictions of μ to the disjoint copies of \bar{U} . ■

2.4 Bandlimited approximation of bivariate functions

In this subsection, we provide an explicit construction of certain bandlimited approximations of a function f defined on a closed subset of $[-1, 1]^2$, under the mild regularity condition that f extends smoothly to a neighborhood of its domain.

Theorem 2.7. *Let $\Omega \subset [-1, 1]^2$ be closed. Suppose that $f: \Omega \rightarrow \mathbb{C}$ admits an infinitely differentiable extension to an open neighborhood of Ω . Then, for each positive integer m and each real number $c > 1$, there exist a constant $C(m, f)$, depending on m and f , but not c , and a function $f_b \in \mathcal{S}(\mathbb{R}^2)$ such that*

1. \widehat{f}_b is supported in $[-c-4, c+4]^2$,
2. $\|f_b - f\|_{L^\infty(\Omega)} \leq \frac{C(m, f)}{c^{2m}}$,
3. $\|f_b\|_{L^\infty(\mathbb{R}^2)} \leq 2\|f\|_{L^\infty(\Omega)} + \frac{C(m, f)}{c^{2m}}$,
4. $\left\| D^\alpha \widehat{f}_b \right\|_{L^\infty(\mathbb{R}^2)} \leq 10\|f\|_{L^\infty(\Omega)}$ for all $|\alpha| \leq 2$.

Proof. Let $\tilde{f}: U \rightarrow \mathbb{C}$ be an infinitely differentiable extension of f to an open neighborhood U of Ω . Let $M = \|f\|_{L^\infty(\Omega)}$ and choose an open neighborhood V of Ω such that $V \subset U \cap [-2, 2]^2$ and \tilde{f} is bounded in magnitude by $2M$ on V .

By the C^∞ Urysohn's lemma (see Chapter 8 of [16]), there exists an infinitely differentiable function $g_1: \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $0 \leq g_1(x, y) \leq 1$ for all $(x, y) \in \mathbb{R}^2$, $g_1(x, y) = 1$ on Ω and $\text{supp } g_1 \subset V$. Then $f_1(x, y) = f(x, y)g_1(x, y)$ is an element of

$\mathcal{S}(\mathbb{R}^2)$, and so is its Fourier transform \widehat{f}_1 . Consequently, there exists a constant k_1 (depending on m and f) such that

$$\left| \widehat{f}_1(\xi_1, \xi_2) \right| \leq \frac{k_1}{(1 + \xi_1^2 + \xi_2^2)^{m+1}}, \quad (51)$$

for all $(\xi_1, \xi_2) \in \mathbb{R}^2$. Since g_1 is bounded in magnitude by 1, we have

$$|f_1(x, y)| \leq 2M, \quad (52)$$

for all $(x, y) \in \mathbb{R}^2$. Since $\text{supp } f_1 \subset V \subset [-2, 2]^2$, it follows that, for all multi-indices α ,

$$\begin{aligned} \left| D^\alpha \widehat{f}_1(\xi_1, \xi_2) \right| &= \left| \frac{(-i)^{|\alpha|}}{4\pi^2} \int_V \mathbf{x}^\alpha f_1(x, y) \exp(-i(\xi_1 x + \xi_2 y)) \, dx \, dy \right| \\ &\leq 2^{|\alpha|} \frac{8}{\pi^2} M \end{aligned} \quad (53)$$

holds for all $(\xi_1, \xi_2) \in \mathbb{R}^2$, where $\mathbf{x} = (x, y) \in \mathbb{R}^2$.

Using the infinitely differentiable ramp function

$$H(x) = \begin{cases} 0, & x \leq -1, \\ \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x}{\sqrt{1-x^2}} \right) \right), & x \in (-1, 1), \\ 1, & x \geq 1, \end{cases} \quad (54)$$

suggested in [19], one can construct an infinitely differentiable window function $g_2: \mathbb{R} \rightarrow \mathbb{R}$ such that $0 \leq g_2(x) \leq 1$, $|g_2'(x)| \leq 1$ and $|g_2''(x)| \leq 1$ for all $x \in \mathbb{R}$, $g_2(x) = 1$ for all $|x| \leq c$ and $g_2(x) = 0$ for all $|x| \geq c + 4$. We then define f_b by its Fourier transform

$$\widehat{f}_b(\xi_1, \xi_2) = \widehat{f}_1(\xi_1, \xi_2) g_2(\|(\xi_1, \xi_2)\|), \quad (55)$$

so that the first condition listed above is satisfied. From (51) and the definition of g_2 , it is clear that, for all $(x, y) \in \mathbb{R}^2$,

$$\begin{aligned} &|f_1(x, y) - f_b(x, y)| \\ &= \left| \int_{\mathbb{R}^2} \widehat{f}_1(\xi_1, \xi_2) (1 - g_2(\|(\xi_1, \xi_2)\|)) \exp(i(\xi_1 x + \xi_2 y)) \, d\xi_1 \, d\xi_2 \right| \\ &\leq \int_0^{2\pi} \int_c^\infty \frac{k_1 r}{(1 + r^2)^{m+1}} \, dr \, d\theta \\ &= \frac{k_1 \pi}{m(1 + c^2)^m}. \end{aligned} \quad (56)$$

This inequality leads directly to the second condition listed above, and we obtain the third condition by combining it with (52). Using (53) and the properties of g_2 , it is straightforward to verify that the remaining conditions listed above are satisfied by repeated differentiation of (55). \blacksquare

Definition 2.3. Let $\Omega \subset [-1, 1]^2$ be closed. Suppose that $f: \Omega \rightarrow \mathbb{C}$ admits an infinitely differentiable extension to an open neighborhood of Ω . For $0 < \varepsilon < 1$, we say that a function $f_b \in \mathcal{S}(\mathbb{R}^2)$ is an ε -bandlimited approximation of f if f_b is a bandlimited function with the following properties:

1. $\|f_b - f\|_{L^\infty(\Omega)} \leq \varepsilon \|f\|_{L^\infty(\Omega)}$,
2. $\|f_b\|_{L^\infty(\mathbb{R}^2)} \leq 3 \|f\|_{L^\infty(\Omega)}$,
3. $\|D^\alpha \widehat{f_b}\|_{L^\infty(\mathbb{R}^2)} \leq 10 \|f\|_{L^\infty(\Omega)}$ for all $|\alpha| \leq 2$.

Furthermore, we denote by $c_f(\varepsilon)$ the infimum of the set of positive real numbers c such that there exists an ε -bandlimited approximation of f with bandlimit c . Note that Theorem 2.7 shows that this set is nonempty, hence $c_f(\varepsilon) < \infty$.

Throughout this paper, we will assume that $c_f(\varepsilon)$ is positive because it is zero only in the trivial case where f is the zero function.

2.5 Legendre expansions of bandlimited functions

Here we characterize the super-exponential decay of the coefficients of the Legendre expansion of a bandlimited function in terms of its bandlimit. Throughout this subsection, P_n denotes the Legendre polynomial of degree n , and j_ν denotes the spherical Bessel function of order ν .

Lemma 2.8. *Let $\varphi \in \mathcal{S}'(\mathbb{R}^2)$. If the Fourier transform of φ is a tempered distribution of order at most p supported in $[-c, c]^2$, then φ coincides with an entire function of the form*

$$\varphi(x, y) = \sum_{|\alpha| \leq p} \mathbf{x}^\alpha \int_{[-c, c]^2} \exp(i(\xi_1 x + \xi_2 y)) d\mu_\alpha(\xi_1, \xi_2), \quad (57)$$

where $\mu_\alpha \in M([-c, c]^2)$ for all multi-indices α with $|\alpha| \leq p$.

Proof. In view of the formula

$$\varphi(x, y) = \langle \widehat{\varphi}(\xi_1, \xi_2), \exp(i(\xi_1 x + \xi_2 y)) \rangle, \quad (58)$$

the conclusion of the lemma is a direct consequence of Theorem 2.6. ■

Theorem 2.9. *Let $\varphi \in \mathcal{S}'(\mathbb{R}^2)$ such that the Fourier transform of φ is a tempered distribution of order at most p supported in $[-c, c]^2$ for $c \geq 1$. Then φ admits a uniformly convergent Legendre expansion*

$$\varphi(x, y) = \sum_{m, n=0}^{\infty} a_{mn} P_m(x) P_n(y), \quad (59)$$

and there exists a constant $C(\varphi, p)$, depending only on φ and p , but not m and n , such that

$$|a_{mn}| \leq C(\varphi, p) \frac{(c/2)^{m+n+p}}{\Gamma(m-p+1)\Gamma(n-p+1)} \quad (60)$$

for all $\min\{m, n\} \geq p$, and

$$|a_{mn}| \leq C(\varphi, p) \frac{(c/2)^{p'+p}}{\Gamma(p'-p+1)} \quad (61)$$

for all $p' = \max\{m, n\} \geq p$.

Proof. By Lemma 2.8,

$$\begin{aligned} & \left| \int_{-1}^1 \int_{-1}^1 \varphi(x, y) P_m(x) P_n(y) dx dy \right| \\ & \leq \max_{|\alpha| \leq p} |\mu_\alpha| ([-c, c]^2) \sum_{0 \leq k+\ell \leq p} \max_{\xi \in [-c, c]} |I_{km}(\xi)| \max_{\xi \in [-c, c]} |I_{\ell n}(\xi)|, \end{aligned} \quad (62)$$

where

$$I_{k\ell}(\xi) = \int_{-1}^1 x^k P_\ell(x) \exp(i\xi x) dx. \quad (63)$$

For $k \leq \ell$, we have the expansion

$$x^k P_\ell(x) = \sum_{\nu=\ell-k}^{\ell+k} b_{k\ell\nu} P_\nu(x), \quad (64)$$

where

$$b_{k\ell\nu} = \frac{2\nu+1}{2} \int_{-1}^1 x^k P_\ell(x) P_\nu(x) dx. \quad (65)$$

Using the three-term recurrence relation for Legendre polynomials, it can be shown that $|b_{k\ell\nu}| \leq 1$. It then follows from Formula 10.54.2 and 10.14.4 in [20] that

$$\begin{aligned} |I_{k\ell}(\xi)| & \leq \sum_{\nu=\ell-k}^{\ell+k} \left| \int_{-1}^1 P_\nu(x) \exp(i\xi x) dx \right| \\ & = 2 \sum_{\nu=\ell-k}^{\ell+k} |j_\nu(|\xi|)| \\ & \leq 2 \sum_{\nu=\ell-k}^{\ell+k} \frac{|\xi/2|^\nu}{\Gamma(\nu+1)} \\ & \leq (4k+2) \frac{(c/2)^{\ell+k}}{\Gamma(\ell-k+1)} \end{aligned} \quad (66)$$

holds for all $|\xi| \leq c$. Finally, inserting the bound

$$|I_{k\ell}(\xi)| \leq 2 \quad (67)$$

and (66) into (62) yields the conclusion of the theorem. \blacksquare

Remark 2.1. Using the asymptotic expansion

$$j_\nu(\xi) \sim \frac{\sqrt{e/2}}{2\nu+1} \left(\frac{e\xi}{2\nu+1} \right)^\nu, \quad \text{as } \nu \rightarrow \infty, \quad (68)$$

which is an application of Formula 10.19.1 in [20], in place of the inequality used in (66), we obtain the approximation

$$\begin{aligned} |I_{k\ell}(\xi)| &\leq 2 \sum_{\nu=\ell-k}^{\ell+k} |j_\nu(|\xi|)| \\ &\sim 2 \sum_{\nu=\ell-k}^{\ell+k} \frac{\sqrt{e/2}}{2\nu+1} \left(\frac{e|\xi|}{2\nu+1} \right)^\nu \\ &\leq \sqrt{2e} \sum_{\nu=\ell-k}^{\ell+k} \frac{1}{2\nu+1} \left(\frac{ec}{2\nu+1} \right)^\nu, \end{aligned} \quad (69)$$

as $\ell \rightarrow \infty$, which holds for all $|\xi| \leq c$. Applying the bounds (67) and (69) to (62), together with the condition

$$\max\{m, n\} > \frac{e}{2}c + p - \frac{1}{2}, \quad (70)$$

we see that

$$\frac{ec}{2\nu+1} < 1 \quad (71)$$

holds for all ν involved, so we expect the coefficients in the Legendre expansion (59) of φ to decay super-exponentially as soon as (70) is satisfied.

2.6 Truncated singular value decompositions

If \mathbf{A} is a complex-valued $m \times n$ matrix with $m \geq n$, then any decomposition of the form

$$\mathbf{A} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m) \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_n & \\ 0 & 0 & \cdots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 0 & \end{pmatrix} (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n)^*, \quad (72)$$

where $\sigma_1, \dots, \sigma_n \geq 0$, $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ is an orthonormal basis of \mathbb{C}^m , and $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthonormal basis of \mathbb{C}^n , is known as a singular value decomposition of \mathbf{A} . The scalars $\sigma_1, \dots, \sigma_n$ are uniquely determined up to ordering, and they are known as the singular values of \mathbf{A} . It is conventional to arrange them in descending order, and we will assume that this is the case with all singular value decompositions that we consider.

A k -truncated singular value decomposition of \mathbf{A} is an approximation of the form

$$\mathbf{A} \approx \mathbf{A}_k = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_k) \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_k \end{pmatrix} (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_k)^*, \quad (73)$$

where (72) is a singular value decomposition of \mathbf{A} with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ and $1 \leq k \leq n$. Typically, a desired precision $\varepsilon > 0$ is specified and k is taken to be the least integer between 1 and $n - 1$ such that $\sigma_{k+1} < \varepsilon$, if such an integer exists, or $k = n$ otherwise. In this case, we say that (73) is a singular value decomposition which has been truncated at precision ε . We call the vector $\mathbf{x} = \mathbf{A}_k^\dagger \mathbf{y}$ given by

$$\mathbf{x} = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_k) \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_k} \end{pmatrix} (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_k)^* \mathbf{y} \quad (74)$$

the solution of the linear system $\mathbf{A}\mathbf{x} = \mathbf{y}$ obtained from the truncated singular value decomposition (73).

The following theorem implies that, when a linear system admits an approximate solution with a modest norm, and it is solved numerically using a truncated singular value decomposition, the computed solution will have both a small residual and a modest norm. We refer the reader to [21] for a proof of the theorem.

Theorem 2.10. *Suppose that $\mathbf{A} \in \mathbb{C}^{m \times n}$, where $m \geq n$, and let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ be the singular values of \mathbf{A} . Let $\mathbf{b} \in \mathbb{C}^m$, and suppose that $\mathbf{x} \in \mathbb{C}^n$ satisfies*

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \quad (75)$$

Let $\varepsilon > 0$, and suppose further that

$$\widehat{\mathbf{x}}_k = (\mathbf{A} + \mathbf{E})_k^\dagger (\mathbf{b} + \mathbf{e}), \quad (76)$$

where $(\mathbf{A} + \mathbf{E})_k^\dagger$ is the pseudoinverse of the k -truncated singular value decomposition of $\mathbf{A} + \mathbf{E}$, so that

$$\widehat{\sigma}_k \geq \varepsilon \geq \widehat{\sigma}_{k+1}, \quad (77)$$

where $\widehat{\sigma}_k$ and $\widehat{\sigma}_{k+1}$ are the k th and $(k + 1)$ th largest singular values of $\mathbf{A} + \mathbf{E}$, defining $\widehat{\sigma}_{n+1} = 0$, and where $\mathbf{E} \in \mathbb{C}^{m \times n}$ and $\mathbf{e} \in \mathbb{C}^m$, with $\|\mathbf{E}\|_2 < \varepsilon/2$. Then

$$\|\widehat{\mathbf{x}}_k\|_2 \leq \frac{1}{\widehat{\sigma}_k} (2\varepsilon \|\mathbf{x}\|_2 + \|\mathbf{e}\|_2) + \|\mathbf{x}\|_2, \quad (78)$$

and

$$\|\mathbf{A}\widehat{\mathbf{x}}_k - \mathbf{b}\|_2 \leq 5\varepsilon \|\mathbf{x}\|_2 + \frac{3}{2} \|\mathbf{e}\|_2. \quad (79)$$

3 Analysis of the Levin PDE

In this section, we prove that, under mild conditions on g_x , the Levin PDE

$$\nabla \cdot \mathbf{p} + i \nabla g \cdot \mathbf{p} = f \text{ in } (-1, 1)^2 \quad (80)$$

admits a well-behaved, approximate solution of the form

$$\mathbf{p}(x, y) = \begin{pmatrix} p_b(x, y) \\ 0 \end{pmatrix}, \quad (81)$$

where $p_b: [-1, 1]^2 \rightarrow \mathbb{C}$ is an infinitely differentiable function.

The main results of this section are Theorem 3.3 and Theorem 3.4. Theorem 3.3 shows that, when g_x is nonzero over the domain and the ratio of its maximum absolute value to its minimum absolute value is small, the Levin PDE admits a well-behaved, approximate solution of the form (81), whose complexity is independent of the magnitude of g_x . Theorem 3.4 shows that, when g_x is of small magnitude over the domain, the Levin PDE also admits a well-behaved, approximate solution of the form (81), whose complexity decreases with the magnitude of g_x .

We begin our analysis with the following technical lemma, which applies when the domain is an arbitrary closed subset of $[-1, 1]^2$ and g_x is equal to a nonzero constant W .

Lemma 3.1. *Let $\Omega \subset [-1, 1]^2$ be closed. Suppose that $f: \Omega \rightarrow \mathbb{C}$ admits an infinitely differentiable extension to an open neighborhood of Ω and $W \neq 0$. Suppose further that $0 < \varepsilon < 1$, and let $W_0 = c_f(\varepsilon)$. Then there exist a right-continuous function $C: (0, \infty) \rightarrow \mathbb{R}$ and, for each $\delta > 0$, an entire function $p_{b,\delta} \in \mathcal{S}'(\mathbb{R}^2)$ such that*

1. $\widehat{p}_{b,\delta}$ is a tempered distribution of order 1 supported in $[-W_0 - \delta, W_0 + \delta]^2$,
2. $\left| \frac{\partial p_{b,\delta}}{\partial x}(x, y) + iW p_{b,\delta}(x, y) - f(x, y) \right| \leq \varepsilon \|f\|_{L^\infty(\Omega)}$ for all $(x, y) \in \Omega$,
3. $\|p_{b,\delta}\|_{L^\infty(\Omega)} \leq C(W_0 + \delta) \min \left\{ 1, \frac{1}{|W|} \right\} \|f\|_{L^\infty(\Omega)}$,
4. $\left\| \frac{\partial p_{b,\delta}}{\partial x} \right\|_{L^\infty(\Omega)} \leq C(W_0 + \delta) \min \left\{ 1, \frac{1}{|W|} \right\} \|f\|_{L^\infty(\Omega)}$.

Proof. We let $f_{b,\delta} \in \mathcal{S}(\mathbb{R}^2)$ be an ε -bandlimited approximation of f with bandlimit $W_0 + \delta$, and define the function $p_{b,\delta}$ via the formula

$$p_{b,\delta}(x, y) = \int_{-\infty}^{\infty} \left[\text{p.v.} \int_{-\infty}^{\infty} \frac{\widehat{f}_{b,\delta}(\xi_1, \xi_2)}{i(\xi_1 + W)} \exp(i(\xi_1 x + \xi_2 y)) d\xi_1 \right] d\xi_2. \quad (82)$$

Let \mathcal{S} denote the tempered distribution induced by the constant function 1, and \mathcal{T}_W denote the tempered distribution defined via the formula

$$\langle \mathcal{T}_W, \varphi \rangle = \text{p.v.} \int_{-\infty}^{\infty} \frac{\varphi(\xi)}{i(\xi + W)} d\xi. \quad (83)$$

Since $\widehat{f}_{b,\delta}$ is compactly supported, $p_{b,\delta}$ is exactly the inverse Fourier transform of the product of $\mathcal{T}_W \otimes \mathcal{S}$ and the infinitely differentiable function $\widehat{f}_{b,\delta}$. Since \mathcal{T}_W is of order 1 and \mathcal{S} is of order 0, and $f_{b,\delta}$ has bandlimit $W_0 + \delta$, it is clear that $\widehat{p}_{b,\delta}$ is a tempered distribution of order 1 which is supported in $[-W_0 - \delta, W_0 + \delta]^2$, so the first condition listed above is satisfied. In particular, the Schwartz-Paley-Wiener theorem asserts that $p_{b,\delta}$ is an entire function.

From Lemma (2.3), we arrive at the alternative formula

$$\begin{aligned} p_{b,\delta}(x, y) &= \left\langle \mathcal{T}_W(\xi_1), \left\langle \mathcal{S}(\xi_2), \widehat{f}_{b,\delta}(\xi_1, \xi_2) \exp(i(\xi_1 x + \xi_2 y)) \right\rangle \right\rangle \\ &= \left\langle \mathcal{T}_W(\xi_1), \mathcal{F}_y^{-1} \left[\widehat{f}_{b,\delta} \right] (\xi_1, y) \exp(i\xi_1 x) \right\rangle \\ &= \text{p.v.} \int_{-\infty}^{\infty} \frac{\mathcal{F}_y^{-1} \left[\widehat{f}_{b,\delta} \right] (\xi_1, y)}{i(\xi_1 + W)} \exp(i\xi_1 x) d\xi_1. \end{aligned} \quad (84)$$

Then we see that, for all $(x, y) \in \Omega$,

$$\begin{aligned} &\frac{\partial p_{b,\delta}}{\partial x}(x, y) + iW p_{b,\delta}(x, y) \\ &= \text{p.v.} \int_{-\infty}^{\infty} \frac{i(\xi_1 + W) \mathcal{F}_y^{-1} \left[\widehat{f}_{b,\delta} \right] (\xi_1, y)}{i(\xi_1 + W)} \exp(i\xi_1 x) d\xi_1 \\ &= f_{b,\delta}(x, y), \end{aligned} \quad (85)$$

so the second condition listed above follows from this and property (1) in Definition 2.3.

To establish the remaining conditions listed above, we will make use of the following bounds on the partial inverse Fourier transform of $\widehat{f}_{b,\delta}$ and its derivative. From property (3) in Definition 2.3, we have

$$\left| \mathcal{F}_y^{-1} \left[\widehat{f}_{b,\delta} \right] (\xi_1, y) \right| \leq 20(W_0 + \delta) \|f\|_{L^\infty(\Omega)}, \quad (86)$$

and

$$\left| \frac{\partial}{\partial \xi_1} \mathcal{F}_y^{-1} \left[\widehat{f}_{b,\delta} \right] (\xi_1, y) \right| \leq 20(W_0 + \delta) \|f\|_{L^\infty(\Omega)}. \quad (87)$$

Then an analogous proof to that of Lemma 4 in [9] shows that

$$\|p_{b,\delta}\|_{L^\infty(\Omega)} \leq C(W_0 + \delta) \min \left\{ 1, \frac{1}{|W|} \right\} \|f\|_{L^\infty(\Omega)}, \quad (88)$$

and

$$\left\| \frac{\partial p_{b,\delta}}{\partial x} \right\|_{L^\infty(\Omega)} \leq C(W_0 + \delta) \min \left\{ 1, \frac{1}{|W|} \right\} \|f\|_{L^\infty(\Omega)}, \quad (89)$$

where C is a right-continuous function. ■

The construction of $p_{b,\delta}$ outlined above fails in the case of $\delta = 0$, since the definition of $c_f(\varepsilon)$ does not guarantee the existence of an ε -bandlimited approximation of f with bandlimit $c_f(\varepsilon)$. Nonetheless, the topological structure of $\mathcal{S}'(\mathbb{R}^2)$ ensures the existence of a well-behaved, approximate solution p_b to the Levin PDE with bandlimit $c_f(\varepsilon)$. In the following corollary, we present the necessary topological arguments to extract p_b .

Corollary 3.2. *Under the assumptions of Lemma 3.1, there exist an entire function $p_b \in \mathcal{S}'(\mathbb{R}^2)$ and a constant $C(W_0)$ such that p_b satisfies the properties of $p_{b,\delta}$ in Lemma 3.1 for $\delta = 0$, with the exception that \widehat{p}_b is of order at most 1, rather than of order exactly 1.*

Proof. For each $n \in \mathbb{N}$, we invoke Lemma 3.1 to obtain an ε -bandlimited approximation $f_{b,1/n} \in \mathcal{S}(\mathbb{R}^2)$ of f with bandlimit $W_0 + \frac{1}{n}$ and an entire function $p_{b,1/n} \in \mathcal{S}'(\mathbb{R}^2)$ such that

1. $\widehat{p}_{b,1/n}$ is a tempered distribution of order 1 supported in $[-W_0 - \frac{1}{n}, W_0 + \frac{1}{n}]^2$,
2. $\left| \frac{\partial p_{b,1/n}}{\partial x}(x, y) + iW p_{b,1/n}(x, y) - f(x, y) \right| \leq \varepsilon \|f\|_{L^\infty(\Omega)}$ for all $(x, y) \in \Omega$,
3. $\|p_{b,1/n}\|_{L^\infty(\Omega)} \leq C \left(W_0 + \frac{1}{n} \right) \min \left\{ 1, \frac{1}{|W|} \right\} \|f\|_{L^\infty(\Omega)}$,
4. $\left\| \frac{\partial p_{b,1/n}}{\partial x} \right\|_{L^\infty(\Omega)} \leq C \left(W_0 + \frac{1}{n} \right) \min \left\{ 1, \frac{1}{|W|} \right\} \|f\|_{L^\infty(\Omega)}$.

Recall that the weak dual topology on $\mathcal{S}'(\mathbb{R}^2)$, which is the topology of pointwise convergence in $\mathcal{S}(\mathbb{R}^2)$, is defined by the family of seminorms $\{q_\varphi : \varphi \in \mathcal{S}(\mathbb{R}^2)\}$, where $q_\varphi(\mathcal{T}) = |\langle \mathcal{T}, \varphi \rangle|$ for $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^2)$.

To see that the sequence $(\widehat{p}_{b,1/n})_{n=1}^\infty$ is weakly bounded, we observe that

$$\begin{aligned} q_\varphi(\widehat{p}_{b,1/n}) &= \left| \int_{-W_0 - \frac{1}{n}}^{W_0 + \frac{1}{n}} \langle \mathcal{T}_W(\xi_1), \widehat{f}_{b,1/n}(\xi_1, \xi_2) \varphi(\xi_1, \xi_2) \rangle d\xi_2 \right| \\ &\leq 4 \left(W_0 + \frac{1}{n} \right) \left(\left\| \widehat{f}_{b,1/n} \varphi \right\|_{\mathbf{e}_1, \mathbf{0}} + \left\| \widehat{f}_{b,1/n} \varphi \right\|_{\mathbf{0}, \mathbf{e}_1} \right) \\ &\leq 120 (W_0 + 1)^2 \|f\|_{L^\infty(\Omega)} \sup_{|\beta| \leq 1} \|\varphi\|_{\mathbf{0}, \beta}, \end{aligned} \tag{90}$$

where $\mathbf{e}_1 = (1, 0)$, and where the first inequality follows from the fact that

$$\left| \left\langle \text{p.v.} \frac{1}{x}, \varphi \right\rangle \right| \leq 2 \left(\|\varphi\|_{1,0} + \|\varphi\|_{0,1} \right) \tag{91}$$

and second inequality follows from property (3) in Definition 2.3. Since $\mathcal{S}(\mathbb{R}^2)$ is barreled, it follows that, since $(\widehat{p}_{b,1/n})_{n=1}^\infty$ is weakly bounded, it is also relatively compact in the weak dual topology (see Theorem 33.2 of [22]), and hence contains a convergent subnet $(\widehat{p}_{b,1/F(\nu)})_{\nu \in \mathcal{I}}$ for some monotone final function $F: \mathcal{I} \rightarrow \mathbb{N}$ and directed system \mathcal{I} . We denote by \widehat{p}_b the limit of the subnet, and define p_b as the inverse Fourier transform of \widehat{p}_b .

From (90) and the pointwise convergence of $(\widehat{p}_{b,1/F(\nu)})_{\nu \in \mathcal{I}}$, we observe that

$$|\langle \widehat{p}_b, \varphi \rangle| \leq 120 (W_0 + 1)^2 \|f\|_{L^\infty(\Omega)} \sup_{|\beta| \leq 1} \|\varphi\|_{\mathbf{0}, \beta} \tag{92}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^2)$, which implies that \widehat{p}_b is of order at most 1. For each test function φ whose support is disjoint from $[-W_0, W_0]^2$, there exists some $n \in \mathbb{N}$ such that $\widehat{p}_{b,1/n}$ and φ have disjoint support, so we can find some $\nu' \in \mathcal{I}$ such that $q_\varphi(\widehat{p}_{b,1/F(\nu)}) = 0$ for all $\nu \succ \nu'$. This implies that the limit $q_\varphi(\widehat{p}_b)$ must be zero, and thus the support of \widehat{p}_b is

contained in $[-W_0, W_0]^2$. It is then a direct consequence of the Schwartz-Paley-Wiener theorem that p_b is an entire function.

Owing to the formula

$$p_{b,1/F(\nu)}(x, y) = \langle \widehat{p}_{b,1/F(\nu)}, \exp(i(\xi_1 x + \xi_2 y)) \rangle, \quad (93)$$

we conclude that $(D^\alpha p_{b,1/F(\nu)})_{\nu \in \mathcal{I}}$ converges pointwise to $D^\alpha p_b$ for all multi-indices α . By combining the pointwise convergence of $D^\alpha p_{b,1/F(\nu)}$ with the right-continuity of C , the remaining conditions follow immediately. \blacksquare

We now consider the case where the domain is $[-1, 1]^2$, and g_x is nonconstant with no zeros over the domain. We suppose that $f: [-1, 1]^2 \rightarrow \mathbb{C}$ and $g: [-1, 1]^2 \rightarrow \mathbb{R}$ admit infinitely differentiable extensions to an open neighborhood of $[-1, 1]^2$, and that the extension of g_x is nonzero in the neighborhood. We let

$$G_0 = \min_{(x,y) \in [-1,1]^2} \left| \frac{\partial g}{\partial x}(x, y) \right|, \quad G_1 = \max_{(x,y) \in [-1,1]^2} \left| \frac{\partial g}{\partial x}(x, y) \right|, \quad (94)$$

and

$$W = \max_{y \in [-1,1]} \left| \int_{-1}^1 \frac{\partial g}{\partial x}(x, y) dx \right|. \quad (95)$$

Furthermore, we define $\mathbf{u}: [-1, 1]^2 \rightarrow \mathbb{R}^2$ via the formula

$$\mathbf{u}(x, y) = \left(\frac{1}{W} \int_{-1}^x \frac{\partial g}{\partial x}(t, y) dt, y \right). \quad (96)$$

Since g_x is nonvanishing on $[-1, 1]^2$, \mathbf{u} is injective and has invertible Jacobian on $[-1, 1]^2$. By the generalized inverse function theorem (see Chapter 1 of [23]), \mathbf{u} extends to a diffeomorphism in an open neighborhood of $[-1, 1]^2$. We denote the image of $(-1, 1)^2$ under \mathbf{u} by $\Omega_{\mathbf{u}}$, noting that $\bar{\Omega}_{\mathbf{u}} \subset [-1, 1]^2$, and define $h: \bar{\Omega}_{\mathbf{u}} \rightarrow \mathbb{R}^2$ via the formula

$$h(x', y') = f(\mathbf{u}^{-1}(x', y')) \det(D\mathbf{u}^{-1})(x', y'). \quad (97)$$

Notice that both \mathbf{u} and h are independent of the magnitude of g_x , in the sense that they remain unchanged under any rescaling of g_x by a nonzero constant.

We are now ready to prove the first main result of this section.

Theorem 3.3. *Suppose that $0 < \varepsilon < 1$, and let $W_0 = c_h(\varepsilon)$. Then there exists an infinitely differentiable function $p_b: [-1, 1]^2 \rightarrow \mathbb{C}$ and a constant $C(W_0)$, depending only on W_0 , such that*

1. *the Fourier transform of $p_b(\mathbf{u}^{-1}(x', y'))$ is a tempered distribution of order at most 1 supported in $[-W_0, W_0]^2$,*
2. $\left| \frac{\partial p_b}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p_b(x, y) - f(x, y) \right| \leq \varepsilon \frac{G_1}{G_0} \|f\|_{L^\infty([-1,1]^2)}$ for all $(x, y) \in [-1, 1]^2$,
3. $\|p_b\|_{L^\infty([-1,1]^2)} \leq C(W_0) \frac{W}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1,1]^2)}$,

$$4. \left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1,1]^2)} \leq C(W_0) \frac{G_1}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1,1]^2)}.$$

Proof. Using the change of variables $(x', y') = \mathbf{u}(x, y)$, the PDE

$$\frac{\partial p}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p(x, y) = f(x, y) \text{ in } (-1, 1)^2 \quad (98)$$

can be reduced to the simpler form

$$\frac{\partial p}{\partial x'}(x', y') + iW p(x', y') = h(x', y') \text{ in } \Omega_{\mathbf{u}}. \quad (99)$$

Since h admits an infinitely differentiable extension to an open neighborhood of $\bar{\Omega}_{\mathbf{u}}$, an application of Corollary 3.2 to (99) shows that there exist an entire function p_1 and a constant $C(W_0)$ such that

1. \hat{p}_1 is a tempered distribution of order at most 1 supported in $[-W_0, W_0]^2$,
2. $\left| \frac{\partial p_1}{\partial x'}(x', y') + iW p_1(x', y') - h(x', y') \right| \leq \varepsilon \|h\|_{L^\infty(\bar{\Omega}_{\mathbf{u}})}$ for all $(x', y') \in \bar{\Omega}_{\mathbf{u}}$,
3. $\|p_1\|_{L^\infty(\bar{\Omega}_{\mathbf{u}})} \leq C(W_0) \min \left\{ 1, \frac{1}{W} \right\} \|h\|_{L^\infty(\bar{\Omega}_{\mathbf{u}})}$,
4. $\left\| \frac{\partial p_1}{\partial x'} \right\|_{L^\infty(\bar{\Omega}_{\mathbf{u}})} \leq C(W_0) \min \left\{ 1, \frac{1}{W} \right\} \|h\|_{L^\infty(\bar{\Omega}_{\mathbf{u}})}$.

We define p_b as the composition of p_1 and \mathbf{u} . It is clear that the first condition on p_b is satisfied. From (96) and (97), we see that

$$\|h\|_{L^\infty(\bar{\Omega}_{\mathbf{u}})} \leq \frac{W}{G_0} \|f\|_{L^\infty([-1,1]^2)}, \quad (100)$$

and therefore

$$\|p_b\|_{L^\infty([-1,1]^2)} \leq C(W_0) \frac{W}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1,1]^2)}, \quad (101)$$

and

$$\left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1,1]^2)} \leq C(W_0) \frac{G_1}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1,1]^2)}. \quad (102)$$

Moreover, we have

$$\left| \frac{\partial p_b}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p_b(x, y) - f(x, y) \right| \leq \varepsilon \frac{G_1}{G_0} \|f\|_{L^\infty([-1,1]^2)} \quad (103)$$

for all $(x, y) \in [-1, 1]^2$. ■

We note that Theorem 3.3 does not assert that the function p_b is bandlimited. However, we conclude that its smoothness is characterized independently of the magnitude of g_x . This follows from the observation that p_1 has a bandlimit independent of the magnitude

of g_x and that the function \mathbf{u} is also independent of the magnitude of g_x . It is clear that p_b can be approximated to a fixed relative accuracy via a Legendre expansion at a cost that does not depend on the magnitude of g_x , but which necessarily depends on the complexity of \mathbf{u} and h .

We close this section with the second main result of this section, which proves the existence of a well-behaved, approximate solution to the Levin PDE if g_x is of small magnitude.

Theorem 3.4. *Suppose that $f: [-1, 1]^2 \rightarrow \mathbb{C}$ and $g: [-1, 1]^2 \rightarrow \mathbb{R}$ admit infinitely differentiable extensions to an open neighborhood of $[-1, 1]^2$, and that $G_1 < 1/2$. Then, for $0 < \varepsilon < 1$ and integer n defined via the formula*

$$n = \left\lfloor \frac{\log(\varepsilon)}{\log(2G_1)} \right\rfloor, \quad (104)$$

there exists an entire function $p_b \in \mathcal{S}'(\mathbb{R}^2)$ such that

1. \widehat{p}_b is a tempered distribution of order at most 1 supported in

$$[-c_f(\varepsilon) - nc_{g_x}(\varepsilon), c_f(\varepsilon) + nc_{g_x}(\varepsilon)]^2,$$

2. $\left| \frac{\partial p_b}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p_b(x, y) - f(x, y) \right| \leq 3\varepsilon \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1, 1]^2)}$ for all $(x, y) \in [-1, 1]^2$,

3. $\|p_b\|_{L^\infty([-1, 1]^2)} \leq \frac{2}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)}$,

4. $\left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \leq 4 \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1, 1]^2)}$.

Proof. We will first construct an entire function $p_{b,\delta} \in \mathcal{S}'(\mathbb{R}^2)$ for $\delta > 0$ such that $p_{b,\delta}$ satisfies conditions (2)-(4) listed above, and $\widehat{p}_{b,\delta}$ is a tempered distribution of order at most 1 supported in

$$[-c_f(\varepsilon) - nc_{g_x}(\varepsilon) - (n+1)\delta, c_f(\varepsilon) + nc_{g_x}(\varepsilon) + (n+1)\delta]^2. \quad (105)$$

We let $f_{b,\delta}$ and $(g_x)_{b,\delta}$ be ε -bandlimited approximations of f and g_x with bandlimit $c_f(\varepsilon) + \delta$ and $c_{g_x}(\varepsilon) + \delta$, respectively. We then define $A_\delta: L^\infty([-1, 1]^2) \rightarrow L^\infty([-1, 1]^2)$ via the formula

$$A_\delta[\varphi](x, y) = -i \int_0^x \left(\frac{\partial g}{\partial x} \right)_{b,\delta}(t, y) \varphi(t, y) dt, \quad (106)$$

and define

$$h_\delta(x, y) = \int_0^x f_{b,\delta}(t, y) dt. \quad (107)$$

The function $p_{b,\delta}$ is then defined as

$$p_{b,\delta}(x, y) = \sum_{k=0}^n A_\delta^k[h_\delta](x, y), \quad (108)$$

where A_δ^k denotes the repeated application of the operator A_δ and n is defined by (104). By the same derivation presented in the proof of Theorem 5 in [9], we see that

$$\begin{aligned} & \left| \frac{\partial p_{b,\delta}}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p_{b,\delta}(x, y) - f(x, y) \right| \\ & \leq 3\varepsilon \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1,1]^2)} \end{aligned} \quad (109)$$

for all $(x, y) \in [-1, 1]^2$, and

$$\|p_{b,\delta}\|_{L^\infty([-1,1]^2)} \leq \frac{2}{1 - 2G_1} \|f\|_{L^\infty([-1,1]^2)}, \quad (110)$$

and

$$\left\| \frac{\partial p_{b,\delta}}{\partial x} \right\|_{L^\infty([-1,1]^2)} \leq 4 \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1,1]^2)}. \quad (111)$$

It remains to verify the bandlimit of $p_{b,\delta}$ and the order of its Fourier transform. We first observe that

$$\widehat{h}_\delta = \widehat{f}_{b,\delta}(\mathcal{T}_0 \otimes \mathcal{S}) - \delta(\xi_1) \otimes \left(\left\langle \mathcal{T}_0(\xi), \widehat{f}_{b,\delta}(\xi, \xi_2) \right\rangle \mathcal{S}(\xi_2) \right), \quad (112)$$

where \mathcal{T}_0 is defined by (83), and \mathcal{S} denotes integration. Since $f_{b,\delta}$ has bandlimit $c_f(\varepsilon) + \delta$, it is clear that h_δ has the same bandlimit. Analogously, we have

$$\widehat{A_\delta[\varphi]} = -i \left(\widehat{I_\delta[\varphi]}(\mathcal{T}_0 \otimes \mathcal{S}) - \delta(\xi_1) \otimes \left(\left\langle \mathcal{T}_0(\xi), \widehat{I_\delta[\varphi]}(\xi, \xi_2) \right\rangle \mathcal{S}(\xi_2) \right) \right), \quad (113)$$

where $I_\delta[\varphi]$ denotes the integrand of $A_\delta[\varphi]$. Suppose that φ is a slowly increasing C^∞ function with bandlimit c . Accordingly, as a consequence of Lemma 2.2 and the fact that $(g_x)_{b,\delta} \in \mathcal{S}(\mathbb{R}^2)$ has bandlimit $c_{g_x}(\varepsilon) + \delta$,

$$\widehat{I_\delta[\varphi]}(\mathbf{x}) = \widehat{\varphi} * \widehat{(g_x)_{b,\delta}}(\mathbf{x}) \quad (114)$$

defines a C^∞ function supported in $[-c - c_{g_x}(\varepsilon) - \delta, c + c_{g_x}(\varepsilon) + \delta]^2$. Hence it is clear that $A_\delta[\varphi]$ has bandlimit $c + c_{g_x}(\varepsilon) + \delta$. It then follows by induction that $A_\delta^k[h_\delta]$ has bandlimit $c_f(\varepsilon) + kc_{g_x}(\varepsilon) + (k+1)\delta$. Moreover, $A_\delta^k[h_\delta]^\wedge$ is of order at most 1 since \mathcal{T}_0 is of order 1. We conclude that $p_{b,\delta}$ has bandlimit $c_f(\varepsilon) + nc_{g_x}(\varepsilon) + (n+1)\delta$, and the order of its Fourier transform is at most 1.

Recall that the weak dual topology on $\mathcal{S}'(\mathbb{R}^2)$ is defined by the family of seminorms $\{q_\varphi: \varphi \in \mathcal{S}(\mathbb{R}^2)\}$, where $q_\varphi(\mathcal{T}) = |\langle \mathcal{T}, \varphi \rangle|$ for $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^2)$. To see that the sequence $(\widehat{h}_{1/m})_{m=1}^\infty$ is weakly bounded, we first observe that

$$\begin{aligned} & \left\langle \delta(\xi_1) \otimes \left(\left\langle \mathcal{T}_0(\xi), \widehat{f}_{b,1/m}(\xi, \xi_2) \right\rangle \mathcal{S}(\xi_2) \right), \varphi(\xi_1, \xi_2) \right\rangle \\ & = \left\langle \widehat{f}_{b,1/m}(\xi, \xi_2)(\mathcal{T}_0(\xi) \otimes \mathcal{S}(\xi_2)), \varphi(0, \xi_2) \right\rangle, \end{aligned} \quad (115)$$

and then apply the same reasoning involved in (90) to (112) to arrive at the bound

$$q_\varphi(\widehat{h}_{1/m}) \leq 240 (c_f(\varepsilon) + 1)^2 \|f\|_{L^\infty([-1,1]^2)} \sup_{|\beta| \leq 1} \|\varphi\|_{\mathbf{0},\beta}. \quad (116)$$

Suppose now that $(\widehat{\psi}_m)_{m=1}^\infty$ is a sequence in $\mathcal{E}'(\mathbb{R}^2)$ with bandlimit c satisfying

$$q_\varphi(\widehat{\psi}_m) \leq C \sup_{|\beta| \leq 1} \|\varphi\|_{\mathbf{0},\beta} \quad (117)$$

for some constant C independent of m and φ . From (44), (117) and property (3) of $((g_x)_{b,1/m})_{m=1}^\infty$ in Definition 2.3, we see that $(I_{1/m}[\psi_m]^\wedge)_{m=1}^\infty$ and their first derivatives are all uniformly bounded, and are supported in $[-c - c_{g_x}(\varepsilon) - 1, c + c_{g_x}(\varepsilon) + 1]^2$. It follows from the same arguments again that $(A_{1/m}[\psi_m]^\wedge)_{m=1}^\infty$ is weakly bounded and satisfies a bound of the form (117).

By induction, we see that $(A_{1/m}^k[h_{1/m}]^\wedge)_{m=1}^\infty$ is weakly bounded for all positive k . Consequently, $(\widehat{p}_{b,1/m})_{m=1}^\infty$ is weakly bounded due to subadditivity of seminorms. The existence of an entire function p_b satisfying the conditions listed above can be established by the same arguments used in the proof of Corollary 3.2. \blacksquare

4 Numerical aspects of the Levin method

In this section, we derive error estimates for the Levin method that discretizes the PDE

$$p_x + ig_x p = f \text{ in } (-1, 1)^2 \quad (118)$$

using collocation on a tensor product Chebyshev grid and then solves the resulting linear system numerically using a truncated singular value decomposition.

In this section, we use the symbol \widehat{x} to denote the computed approximation to x , rather than the Fourier transform of x .

Throughout this section, we will assume that $f: [-1, 1]^2 \rightarrow \mathbb{C}$ and $g: [-1, 1]^2 \rightarrow \mathbb{R}$ admit infinitely differentiable extensions to an open neighborhood of $[-1, 1]^2$, and that $0 < \varepsilon < 1/2$. As in the preceding section, we let

$$G_0 = \min_{(x,y) \in [-1,1]^2} \left| \frac{\partial g}{\partial x}(x, y) \right|, \quad G_1 = \max_{(x,y) \in [-1,1]^2} \left| \frac{\partial g}{\partial x}(x, y) \right|, \quad (119)$$

and

$$W = \max_{y \in [-1,1]} \left| \int_{-1}^1 \frac{\partial g}{\partial x}(x, y) dx \right|. \quad (120)$$

4.1 Case I: $G_0 > 0$

We let $W_0 = c_h(\varepsilon)$, where h is defined via the formula (97) and $c_h(\varepsilon)$ is defined in Definition 2.3. According to Theorem 3.3, there exist an infinitely differentiable function $p_b: [-1, 1]^2 \rightarrow \mathbb{C}$ and a constant $C(W_0)$, which we assume to satisfy $C(W_0) \geq 1$ in this subsection, such that

$$\left| \frac{\partial p_b}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p_b(x, y) - f(x, y) \right| \leq \varepsilon \frac{G_1}{G_0} \|f\|_{L^\infty([-1,1]^2)} \quad (121)$$

for all $(x, y) \in [-1, 1]^2$, and both

$$\|p_b\|_{L^\infty([-1, 1]^2)} \leq C(W_0) \frac{W}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1, 1]^2)} \quad (122)$$

and

$$\left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \leq C(W_0) \frac{G_1}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1, 1]^2)}. \quad (123)$$

From (38), (39), and the discussion following Theorem 3.3, we observe that an integer k can be chosen independently of the magnitude of g_x such that

$$\|\mathbf{P}_k[p_b] - p_b\|_{L^\infty([-1, 1]^2)} \leq \varepsilon \|p_b\|_{L^\infty([-1, 1]^2)}, \quad (124)$$

and

$$\left\| \frac{\partial}{\partial x} \mathbf{P}_k[p_b] - \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \leq \varepsilon \left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1, 1]^2)}. \quad (125)$$

Moreover, from the uniform convergence of Chebyshev interpolants, the value of k can be chosen such that

$$\left\| \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] - \frac{\partial g}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \leq \varepsilon \left\| \frac{\partial g}{\partial x} \right\|_{L^\infty([-1, 1]^2)}, \quad (126)$$

and

$$\|\mathbf{P}_{2k-1}[f] - f\|_{L^\infty([-1, 1]^2)} \leq \varepsilon \|f\|_{L^\infty([-1, 1]^2)} \quad (127)$$

are also satisfied. With this choice of k and the assumption that $0 < \varepsilon < 1/2$, we have

$$\begin{aligned} & \left\| \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] \mathbf{P}_k[p_b] - \frac{\partial g}{\partial x} p_b \right\|_{L^\infty([-1, 1]^2)} \\ & \leq \left\| \mathbf{P}_k[p_b] \left(\mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] - \frac{\partial g}{\partial x} \right) + \frac{\partial g}{\partial x} (\mathbf{P}_k[p_b] - p_b) \right\|_{L^\infty([-1, 1]^2)} \\ & \leq 3\varepsilon \left\| \frac{\partial g}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \|p_b\|_{L^\infty([-1, 1]^2)}. \end{aligned} \quad (128)$$

From (122), (123), (125), (128), and the fact that

$$\max\{W, 1\} \min \left\{ 1, \frac{1}{W} \right\} = 1, \quad (129)$$

we see that

$$\begin{aligned} & \left\| \frac{\partial}{\partial x} \mathbf{P}_k[p_b] + i \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] \mathbf{P}_k[p_b] - \left(\frac{\partial p_b}{\partial x} + i \frac{\partial g}{\partial x} p_b \right) \right\|_{L^\infty([-1, 1]^2)} \\ & \leq 3\varepsilon \left\| \frac{\partial g}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \|p_b\|_{L^\infty([-1, 1]^2)} + \varepsilon \left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \\ & \leq 4\varepsilon C(W_0) \frac{G_1}{G_0} \|f\|_{L^\infty([-1, 1]^2)} \end{aligned} \quad (130)$$

holds regardless of the magnitude of g_x . Combining (121) with (127) and (130) yields

$$\begin{aligned} & \left| \frac{\partial}{\partial x} \mathbf{P}_k[p_b](x, y) + i \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] (x, y) \mathbf{P}_k[p_b](x, y) - \mathbf{P}_{2k-1}[f](x, y) \right| \\ & \leq \varepsilon \left(1 + (1 + 4C(W_0)) \frac{G_1}{G_0} \right) \|f\|_{L^\infty([-1,1]^2)} \\ & \leq 6\varepsilon C(W_0) \frac{G_1}{G_0} \|f\|_{L^\infty([-1,1]^2)}, \end{aligned} \quad (131)$$

for all $(x, y) \in [-1, 1]^2$, since $C(W_0) \geq 1$. If we let

$$\mathbf{p}_b = \begin{pmatrix} p_b(\mathbf{x}_{1,k^2}) \\ p_b(\mathbf{x}_{2,k^2}) \\ \vdots \\ p_b(\mathbf{x}_{k^2,k^2}) \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f(\mathbf{x}_{1,(2k-1)^2}) \\ f(\mathbf{x}_{2,(2k-1)^2}) \\ \vdots \\ f(\mathbf{x}_{(2k-1)^2,(2k-1)^2}) \end{pmatrix}, \quad (132)$$

and define the $(2k-1)^2 \times (2k-1)^2$ diagonal matrix \mathbf{G}_{2k-1} with entries given by

$$(\mathbf{G}_{2k-1})_{i,i} = g_x(\mathbf{x}_{i,(2k-1)^2}), \quad (133)$$

for all $1 \leq i \leq (2k-1)^2$, then it follows from (131) that

$$(\mathbf{P}_{(2k-1,k)}(\mathbf{D}_x)_k + i\mathbf{G}_{2k-1}\mathbf{P}_{(2k-1,k)}) \mathbf{p}_b = \mathbf{f} + \mathbf{r} \quad (134)$$

with

$$\|\mathbf{p}_b\|_2 \lesssim C(W_0) \frac{W}{G_0} \min \left\{ 1, \frac{1}{W} \right\} \|f\|_{L^\infty([-1,1]^2)} \quad (135)$$

and

$$\|\mathbf{r}\|_2 \lesssim \varepsilon C(W_0) \frac{G_1}{G_0} \|f\|_{L^\infty([-1,1]^2)}, \quad (136)$$

where $(\mathbf{D}_x)_k$ and $\mathbf{P}_{(2k-1,k)}$ are the Chebyshev differentiation and interpolation matrices defined in Section 2.2.

We introduce the notation

$$\mathbf{A} = \mathbf{P}_{(2k-1,k)}(\mathbf{D}_x)_k + i\mathbf{G}_{2k-1}\mathbf{P}_{(2k-1,k)} \quad (137)$$

in order to simplify the following discussion. We observe that

$$\|\mathbf{A}\|_2 \lesssim \max\{G_1, 1\}, \quad (138)$$

and, by combining Lemma 2.1 with (126), we see that

$$\|\mathbf{A}\|_2 \geq \frac{1}{\sqrt{2k}} \left(1 + \frac{(1-\varepsilon)G_1}{\Lambda_k^2} \right) \geq \frac{1+G_1}{2\sqrt{2k}\Lambda_k^2} \geq \frac{\max\{G_1, 1\}}{2\sqrt{2k}\Lambda_k^2}, \quad (139)$$

and thus $\|\mathbf{A}\|_2 \gtrsim \max\{G_1, 1\}$. We now solve the linear system

$$\mathbf{A}\mathbf{p}_b = \mathbf{f} \quad (140)$$

via a truncated singular value decomposition which has been truncated at precision $\varepsilon\|\mathbf{A}\|_2$, and obtain a solution

$$\widehat{\mathbf{p}} = (\mathbf{A} + \mathbf{E})_{\nu}^{\dagger}(\mathbf{f} + \mathbf{e}), \quad (141)$$

where $\mathbf{E} \in \mathbb{C}^{(2k-1)^2 \times k^2}$ and $\mathbf{e} \in \mathbb{C}^{(2k-1)^2}$ with $\|\mathbf{E}\|_2 < \varepsilon\|\mathbf{A}\|_2/2$ and $\|\mathbf{e}\|_2 \lesssim \varepsilon\|\mathbf{f}\|_2$, and where $(\mathbf{A} + \mathbf{E})_{\nu}^{\dagger}$ is the pseudoinverse of the ν -truncated singular value decomposition of $\mathbf{A} + \mathbf{E}$, so that

$$\widehat{\sigma}_{\nu} \geq \varepsilon\|\mathbf{A}\|_2 \geq \widehat{\sigma}_{\nu+1}, \quad (142)$$

where $\widehat{\sigma}_{\nu}$ and $\widehat{\sigma}_{\nu+1}$ are the ν th and $(\nu + 1)$ th largest singular values of $\mathbf{A} + \mathbf{E}$, defining $\widehat{\sigma}_{k^2+1} = 0$. By Theorem 2.10, we have that

$$\|\widehat{\mathbf{p}}\|_2 \leq \frac{1}{\varepsilon\|\mathbf{A}\|_2} (2\varepsilon\|\mathbf{A}\|_2\|\mathbf{p}_{\mathbf{b}}\|_2 + \|\mathbf{e} - \mathbf{r}\|_2) + \|\mathbf{p}_{\mathbf{b}}\|_2, \quad (143)$$

and

$$\|\mathbf{A}\widehat{\mathbf{p}} - \mathbf{f}\|_2 \leq 5\varepsilon\|\mathbf{A}\|_2\|\mathbf{p}_{\mathbf{b}}\|_2 + \frac{3}{2}\|\mathbf{e} - \mathbf{r}\|_2 + \|\mathbf{r}\|_2. \quad (144)$$

Combining the above inequalities with (135), (136), (138), (139) and the fact that

$$\max\{G_1, 1\} \min\{W, 1\} \leq 2G_1, \quad (145)$$

we conclude that

$$\|\widehat{\mathbf{p}}\|_2 \lesssim C(W_0) \frac{\min\{G_1, 1\}}{G_0} \|f\|_{L^{\infty}([-1,1]^2)}, \quad (146)$$

and

$$\|\mathbf{A}\widehat{\mathbf{p}} - \mathbf{f}\|_2 \lesssim \varepsilon C(W_0) \frac{G_1}{G_0} \|f\|_{L^{\infty}([-1,1]^2)}. \quad (147)$$

We now let \widehat{p} be the unique bivariate polynomial of degree less than k whose values on the $k \times k$ tensor product Chebyshev grid are given by the entries of $\widehat{\mathbf{p}}$, and let \widehat{r} be the unique bivariate polynomial of degree less than $2k - 1$ whose values on the $(2k - 1) \times (2k - 1)$ tensor product Chebyshev grid equal to the entries of the vector $\widehat{\mathbf{r}} = \mathbf{A}\widehat{\mathbf{p}} - \mathbf{f}$. It then follows that

$$\frac{\partial \widehat{p}}{\partial x}(x, y) + i \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right](x, y) \widehat{p}(x, y) = \mathbf{P}_{2k-1}[f](x, y) + \widehat{r}(x, y) \quad (148)$$

for all $(x, y) \in [-1, 1]^2$, since both sides of (148) are bivariate polynomials of degree less than $2k - 1$ which agree on the $(2k - 1) \times (2k - 1)$ tensor product Chebyshev grid. Finally, we combine (126), (127), (146) and (147) in order to conclude that

$$\begin{aligned} & \left| \frac{\partial \widehat{p}}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) \widehat{p}(x, y) - f(x, y) \right| \\ & \leq \left\| \widehat{\mathbf{r}} + (\mathbf{P}_{2k-1}[f] - f) + i \widehat{p} \left(\frac{\partial g}{\partial x} - \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] \right) \right\|_{L^{\infty}([-1,1]^2)} \\ & \lesssim \varepsilon C(W_0) \frac{G_1}{G_0} \|f\|_{L^{\infty}([-1,1]^2)} \end{aligned} \quad (149)$$

for all $(x, y) \in [-1, 1]^2$.

To complete our analysis for the case in which $G_0 > 0$, we let

$$I = \int_{-1}^1 \int_{-1}^1 f(x, y) \exp(ig(x, y)) \, dx \, dy \quad (150)$$

be the true value of the oscillatory integral we aim to evaluate using the Levin method, and consider the estimate \hat{I} of I given by

$$\hat{I} = \int_{-1}^1 \hat{p}(1, y) \exp(ig(1, y)) \, dy - \int_{-1}^1 \hat{p}(-1, y) \exp(ig(-1, y)) \, dy, \quad (151)$$

where we assume that the univariate integrals are computed exactly. We observe that

$$\hat{I} = \int_{-1}^1 \int_{-1}^1 \frac{\partial}{\partial x} (\hat{p}(x, y) \exp(ig(x, y))) \, dx \, dy, \quad (152)$$

and it follows from this and (149) that

$$\begin{aligned} |\hat{I} - I| &\leq \int_{-1}^1 \int_{-1}^1 \left| \frac{\partial \hat{p}}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) \hat{p}(x, y) - f(x, y) \right| \, dx \, dy \\ &\lesssim \varepsilon C(W_0) \frac{G_1}{G_0} \|f\|_{L^\infty([-1, 1]^2)}. \end{aligned} \quad (153)$$

We note that the ratio G_1/G_0 is small whenever g_x does not vary in magnitude too much over the domain, which is a reasonable assumption for an adaptive integration scheme. We conclude that (153) shows that the absolute error in the estimate is bounded independently of the magnitude of g_x , provided that g_x does not vary too much over the domain.

4.2 Case II: $G_1 \leq 1/4$

We now move on to the case in which $G_1 \leq 1/4$. We invoke Theorem 3.4 to see that there exists a bandlimited function p_b such that

$$\begin{aligned} &\left| \frac{\partial p_b}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) p_b(x, y) - f(x, y) \right| \\ &\leq 3\varepsilon \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1, 1]^2)} \end{aligned} \quad (154)$$

for all $(x, y) \in [-1, 1]^2$, and both

$$\|p_b\|_{L^\infty([-1, 1]^2)} \leq \frac{2}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)} \quad (155)$$

and

$$\left\| \frac{\partial p_b}{\partial x} \right\|_{L^\infty([-1, 1]^2)} \leq 4 \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1, 1]^2)}. \quad (156)$$

From (104) and the assumption that $G_1 \leq 1/4$, we see that the maximum bandlimit of p_b is bounded, which implies that the Chebyshev coefficients of p_b are bounded by a

rapidly decaying function which is independent of G_1 . We can thus choose an integer k independently of G_1 such that (124) through (127) hold. Proceeding as in the previous subsection, we see that

$$\begin{aligned} & \left| \frac{\partial}{\partial x} \mathbf{P}_k[p_b](x, y) + i \mathbf{P}_k \left[\frac{\partial g}{\partial x} \right] (x, y) \mathbf{P}_k[p_b](x, y) - \mathbf{P}_{2k-1}[f](x, y) \right| \\ & \lesssim \varepsilon \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1, 1]^2)}, \end{aligned} \quad (157)$$

for all $(x, y) \in [-1, 1]^2$. If we define \mathbf{p}_b , \mathbf{f} , \mathbf{r} and \mathbf{A} as before, then we have that

$$\|\mathbf{p}_b\|_2 \lesssim \frac{1}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)}, \quad (158)$$

and

$$\|\mathbf{r}\|_2 \lesssim \varepsilon \left(1 + \frac{G_1}{1 - 2G_1} \right) \|f\|_{L^\infty([-1, 1]^2)}. \quad (159)$$

We again solve the linear system

$$\mathbf{A}\mathbf{p}_b = \mathbf{f} \quad (160)$$

via a truncated singular value decomposition which has been truncated at precision $\varepsilon\|\mathbf{A}\|_2$, and, by Theorem 2.10, we see that the obtained solution $\widehat{\mathbf{p}}$ satisfies

$$\|\widehat{\mathbf{p}}\|_2 \lesssim \frac{1}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)}, \quad (161)$$

and

$$\|\mathbf{A}\widehat{\mathbf{p}} - \mathbf{f}\|_2 \lesssim \varepsilon \frac{1}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)}. \quad (162)$$

Defining the polynomials \widehat{p} and \widehat{r} as before, we conclude that

$$\left| \frac{\partial \widehat{p}}{\partial x}(x, y) + i \frac{\partial g}{\partial x}(x, y) \widehat{p}(x, y) - f(x, y) \right| \lesssim \varepsilon \frac{1}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)} \quad (163)$$

for all $(x, y) \in [-1, 1]^2$. It then follows from this that the absolute error in the estimate \widehat{I} of I defined by (151) satisfies

$$|\widehat{I} - I| \lesssim \varepsilon \frac{1}{1 - 2G_1} \|f\|_{L^\infty([-1, 1]^2)}. \quad (164)$$

Since $G_1 \leq 1/4$, we have that $1/2 \leq 1 - 2G_1 \leq 1$, so the constant in (164) is small.

5 Algorithm description

In this section, we describe an adaptive delaminating Levin method for the numerical evaluation of the integral

$$\int_{-1}^1 \int_{-1}^1 f(x, y) \exp(ig(x, y)) \, dx \, dy. \quad (165)$$

We recall that the rigorous analysis presented in the preceding section concerns a spectral collocation method in which the solution to the PDE (118) is represented by a bivariate polynomial of degree less than k , while the PDE is enforced on the $(2k - 1) \times (2k - 1)$ tensor product Chebyshev grid. This is necessary to establish (148) since the product $P_k[g_x]\hat{p}$ is of degree less than $2k - 1$. The algorithm we describe in this section only enforces the PDE on the $k \times k$ tensor product Chebyshev grid. This change enables the efficient solution of the PDE using the delaminating Levin method, where the PDE is solved over the fibers $\{\mathbf{x}_{i,k}^{(j)}\}_{i=1}^k$, independently for each $1 \leq j \leq k$. We justify this modification as follows.

When the delaminating Levin method is used adaptively, the integration domain is typically subdivided until both f and g_x can be represented by polynomials of degree much lower than k . This is often the case for adaptive procedures since adaptive subdivision typically leads to over-discretization of the inputs. If g_x is large and has been over-discretized, the obtained solution \hat{p} will also be a polynomial of degree somewhat lower than k . This is because the discretized PDE does not have a nontrivial null space when g_x is large, and so \hat{p} will closely agree with the slowly-varying solution p_b , which can be represented by a polynomial of low degree. So in the case when g_x is large, the product $g_x\hat{p}$ can often be represented accurately by a polynomial of degree less than k .

On the other hand, when g_x is small enough, the discretized null space is nontrivial, but its elements can be represented by polynomials of low degree. Accordingly, in this case, the product $g_x\hat{p}$ can also often be represented by a polynomial of degree less than k .

Only when g_x is of moderate size, not too large or too small, do we expect \hat{p} to be of degree close to $k - 1$, which means that the product $g_x\hat{p}$ is of degree larger than $k - 1$ when g_x is nontrivial. Consequently, when the delaminating Levin method is employed, only those subrectangles on which g_x falls into a relatively narrow range of moderate values will require additional subdivision. Thus, we do not expect many additional subdivisions to occur.

We first describe the adaptive subdivision procedure before detailing the fixed order delaminating Levin method. The adaptive algorithm takes as input a subdivision tolerance parameter $\varepsilon > 0$, an integer k specifying the order of the Chebyshev spectral collocation method used to discretize the PDE on each subrectangle considered, and an external subroutine which evaluates the functions f and g at a specified collection of points. It maintains an estimated value val of (165) and a list of subrectangles. Initially, the list of subrectangles contains only $[-1, 1]^2$ and the value of the estimate is set to zero. While the list of subrectangles is nonempty, the adaptive algorithm repeats the following steps:

1. Remove a subrectangle $[a, b] \times [c, d]$ from the list of subrectangles.
2. Compute an estimate val_0 of

$$\text{val}_0 = \int_c^d \int_a^b f(x, y) \exp(ig(x, y)) \, dx \, dy \quad (166)$$

using the delaminating Levin method.

3. Compute an estimate val_i of

$$\text{val}_i = \int_{R_i} f(x, y) \exp(ig(x, y)) \, dx \, dy \quad (167)$$

for $1 \leq i \leq 4$, where

$$\begin{aligned} R_1 &= \left[a, \frac{a+b}{2} \right] \times \left[c, \frac{c+d}{2} \right], & R_2 &= \left[\frac{a+b}{2}, b \right] \times \left[c, \frac{c+d}{2} \right], \\ R_3 &= \left[a, \frac{a+b}{2} \right] \times \left[\frac{c+d}{2}, d \right], & R_4 &= \left[\frac{a+b}{2}, b \right] \times \left[\frac{c+d}{2}, d \right], \end{aligned} \quad (168)$$

using the delaminating Levin method.

4. If $\text{val}_0 - (\text{val}_1 + \text{val}_2 + \text{val}_3 + \text{val}_4) < \varepsilon$, then update the estimate val by letting $\text{val} = \text{val} + \text{val}_0$. Otherwise, add the subrectangle R_i to the list of subrectangles for all $1 \leq i \leq 4$.

When the list of subrectangles becomes empty, the adaptive algorithm terminates and returns the estimate val for (165).

We now set forth the delaminating Levin method, which estimates the value of

$$\int_c^d \int_a^b f(x, y) \exp(ig(x, y)) dx dy, \quad (169)$$

where $[a, b] \times [c, d]$ is a bounded rectangle in \mathbb{R}^2 . It takes as input the domain $[a, b] \times [c, d]$, an integer k which controls the order of the Chebyshev spectral collocation method, a truncation tolerance parameter $\varepsilon_0 > 0$, and an external subroutine for evaluating the functions f and g . The delaminating Levin method proceeds as follows:

1. Use the external subroutine provided by the user to evaluate the functions f and g at the nodes of the $k \times k$ tensor product Chebyshev grid translated from $[-1, 1]^2$ to $[a, b] \times [c, d]$, which we also denote by

$$\left\{ \mathbf{x}_{i,k}^{(j)} \right\}_{i,j=1}^k, \quad (170)$$

where

$$\mathbf{x}_{i,k}^{(j)} = \mathbf{x}_{i+k(j-1),k^2} = (t_{i,k}, t_{j,k}) \quad (171)$$

for $1 \leq i, j \leq k$, so that $\{t_{i,k}\}_{i=1}^k$ and $\{t_{j,k}\}_{j=1}^k$ are the Chebyshev nodes translated to $[a, b]$ and $[c, d]$, respectively.

2. For each $1 \leq j \leq k$,

- (a) Compute the approximate values

$$\widetilde{g_x(\mathbf{x}_{1,k}^{(j)})}, \dots, \widetilde{g_x(\mathbf{x}_{k,k}^{(j)})} \quad (172)$$

of the partial derivative of g with respect to x using the formula

$$\begin{pmatrix} \widetilde{g_x(\mathbf{x}_{1,k}^{(j)})} \\ \widetilde{g_x(\mathbf{x}_{2,k}^{(j)})} \\ \vdots \\ \widetilde{g_x(\mathbf{x}_{k,k}^{(j)})} \end{pmatrix} = \mathbf{D}_k \begin{pmatrix} g(\mathbf{x}_{1,k}^{(j)}) \\ g(\mathbf{x}_{2,k}^{(j)}) \\ \vdots \\ g(\mathbf{x}_{k,k}^{(j)}) \end{pmatrix}. \quad (173)$$

- (b) Form the matrix $\mathbf{A}_j = \mathbf{D}_k + i\mathbf{G}_j$, where \mathbf{G}_j is the $k \times k$ diagonal matrix with entries given by

$$(\mathbf{G}_j)_{i,i} = g_x(\widetilde{\mathbf{x}}_{i,k}^{(j)}) \quad (174)$$

for all $1 \leq i \leq k$.

- (c) Construct the singular value decomposition

$$\mathbf{A}_j = \mathbf{U}_j \begin{pmatrix} \sigma_1^{(j)} & & & \\ & \sigma_2^{(j)} & & \\ & & \ddots & \\ & & & \sigma_k^{(j)} \end{pmatrix} \mathbf{V}_j^* \quad (175)$$

of the matrix \mathbf{A}_j .

- (d) Find the largest integer $1 \leq \nu_j \leq k$ such that $\sigma_{\nu_j}^{(j)} \geq \varepsilon_0 \|\mathbf{A}_j\|_2$. If no such integer exists, set the values of

$$p(\widetilde{\mathbf{x}}_{1,k}^{(j)}), \dots, p(\widetilde{\mathbf{x}}_{k,k}^{(j)}) \quad (176)$$

to zero.

- (e) Otherwise, let

$$\begin{pmatrix} p(\widetilde{\mathbf{x}}_{1,k}^{(j)}) \\ p(\widetilde{\mathbf{x}}_{2,k}^{(j)}) \\ \vdots \\ p(\widetilde{\mathbf{x}}_{k,k}^{(j)}) \end{pmatrix} = \frac{b-a}{2} (\mathbf{A}_j)_{\nu_j}^\dagger \begin{pmatrix} f(\mathbf{x}_{1,k}^{(j)}) \\ f(\mathbf{x}_{2,k}^{(j)}) \\ \vdots \\ f(\mathbf{x}_{k,k}^{(j)}) \end{pmatrix}, \quad (177)$$

where $(\mathbf{A}_j)_{\nu_j}^\dagger$ is the pseudoinverse of the ν_j -truncated singular value decomposition of \mathbf{A}_j .

3. Let \widehat{p} be the bivariate polynomial of degree less than k whose values on the $k \times k$ tensor product Chebyshev grid are given by

$$\widehat{\mathbf{p}} = \left(p(\widetilde{\mathbf{x}}_{1,k^2}) \quad p(\widetilde{\mathbf{x}}_{2,k^2}) \quad \cdots \quad p(\widetilde{\mathbf{x}}_{k^2,k^2}) \right)^T. \quad (178)$$

Return the estimate

$$\int_c^d \widehat{p}(b, y) \exp(ig(b, y)) dy - \int_c^d \widehat{p}(a, y) \exp(ig(a, y)) dy \quad (179)$$

computed using the univariate adaptive Levin method presented in [9], with subdivision tolerance parameter $\varepsilon^{\text{1D}} = \beta \max\{\|\widehat{\mathbf{p}}\|_\infty / \|\mathbf{f}\|_\infty, 1\} \varepsilon$ and truncation tolerance parameter $\varepsilon_0^{\text{1D}} = \varepsilon_0$, for some safety factor $0 < \beta < 1$.

A distinctive feature of the delaminating Levin method described above is the use of the univariate adaptive Levin method, as opposed to a univariate Levin method with a fixed order. When resonance points are present, but the PDE (118) admits a smooth solution, a univariate Levin method of a fixed order can fail, which will cause the algorithm to subdivide the subrectangle. As shown in [9], the univariate adaptive Levin method can efficiently evaluate univariate oscillatory integrals with stationary points, requiring at worst a number of subdivisions that scales logarithmically with the magnitude of the derivative of the phase function. By using an adaptive method for the univariate integrals in (179), the subrectangles are only subdivided when the PDE does not admit a smooth solution. This is far more efficient than adaptively subdividing in two dimensions in order to accurately compute the univariate integrals in (179) using a method of fixed order.

We note that the choice to delaminate along the x -direction is arbitrary, and that the discussion of the Levin PDE and its numerical discretization presented above can be trivially adapted if delamination is performed along the y -direction instead. In view of the bound (153), it may at first seem advantageous to select the direction in which the ratio of the maximum to minimum value of the corresponding partial derivative of g is small. Note however that evaluating the smaller of the two partial derivatives can be an ill-conditioned procedure, especially when the size of the other partial derivative is much larger. As a result, we find that delaminating along the direction with the largest partial derivative is most efficient, in that it requires the fewest subdivisions.

Remark 5.1. In our implementation of the algorithm, we used a rank-revealing QR decomposition in lieu of the truncated singular value decomposition to solve the linear systems arising from the discretization of the PDE. This was found to be about twice as fast overall and leads to no apparent loss in accuracy.

Remark 5.2. Both the delaminating Levin method and the univariate Levin method involve solving linear systems of the form

$$(\mathbf{D}_k + \mathbf{G})\mathbf{p} = \mathbf{f}, \tag{180}$$

where \mathbf{D}_k is the $k \times k$ Chebyshev differentiation matrix and \mathbf{G} is a $k \times k$ diagonal matrix. When \mathbf{G} is invertible, (180) can be solved extremely rapidly via the iteration

$$\mathbf{p}_0 = \mathbf{0} \quad \text{and} \quad \mathbf{p}_{n+1} = -\mathbf{G}^{-1}\mathbf{D}_k\mathbf{p}_n + \mathbf{G}^{-1}\mathbf{f}, \tag{181}$$

as long as there is some matrix norm $\|\cdot\|_p$ such that $\|\mathbf{G}^{-1}\mathbf{D}_k\|_p \ll 1$. Since

$$\|\mathbf{G}^{-1}\mathbf{D}_k\|_\infty \leq \|\mathbf{G}^{-1}\|_\infty \|\mathbf{D}_k\|_\infty, \tag{182}$$

this condition is straightforward to check in practice. Furthermore, the iteration can be implemented in $\mathcal{O}(k \log k)$ operations using the discrete Chebyshev transform [24], which results in substantial speedup when k is large. We note that this is often the case when Levin methods are used non-adaptively.

6 Numerical experiments

In this section, we present the results of numerical experiments conducted to illustrate the properties of the adaptive delaminating Levin method. We implemented our algorithm

in Fortran and compiled our code with version 14.2.0 of the GNU Fortran compiler. All experiments were performed on a desktop computer equipped with an Intel Core i9-13900K processor and 64GB of memory. No attempt was made to parallelize our code.

In all of our experiments, the value of the parameter k , which determines the order of the bivariate Chebyshev spectral collocation method, was taken to be 7. The subdivision tolerance parameter ε was set to 10^{-12} , and the truncation tolerance parameter ε_0 was set to $\beta_0\varepsilon/\|\mathbf{f}\|_\infty$ with $\beta_0 = 1/2$. For computing the univariate boundary integrals, we used a 12-point Chebyshev spectral collocation method for the univariate adaptive Levin method, and set the safety factor β to 10^{-1} . To account for task scheduling and other vagaries of modern computing environments, all reported times were obtained by averaging the cost of each calculation over 100 runs.

In order to estimate the error in the results produced by the adaptive delaminating Levin method, we compared the results of the adaptive delaminating Levin method with the results produced by adaptive tensor product Gauss-Legendre quadrature for evaluating integrals of the form

$$\int_{-1}^1 \int_{-1}^1 f(x, y) \, dx \, dy. \quad (183)$$

We used this method to estimate the error in every case, except when explicit formulas for the integrals were available. Our implementation of the adaptive tensor product Gauss-Legendre quadrature is written in Fortran and is quite standard. It maintains a list of subrectangles, which is initialized with the single rectangle $[-1, 1]^2$, and a running tally of the value of the integral. As long as the list of subrectangles is nonempty, the algorithm removes a subrectangle $[a, b] \times [c, d]$ from the list, and compares the value of

$$\int_c^d \int_a^b f(x, y) \, dx \, dy \quad (184)$$

as computed by a 10^2 -point tensor product Gauss-Legendre quadrature rule to the value of the sum

$$\sum_{i=1}^4 \int_{R_i} f(x, y) \, dx \, dy, \quad (185)$$

where the subrectangles R_i are defined in (168), and where each integral is separately estimated with a 10^2 -point tensor product Gauss-Legendre quadrature rule. If the difference is larger than ε , where ε is a tolerance parameter specified by the user, then the subrectangles R_i are inserted into the list of subrectangles. Otherwise, the value of (184) is added to the running tally of the integral (183). In all of our experiments, the tolerance parameter for the adaptive Gaussian quadrature code was taken to be $\varepsilon = 10^{-14}$.

6.1 Integrals involving elementary functions, with explicit formulas

In the experiments described in this subsection, we consider the integrals

$$\begin{aligned} I_1(\lambda) &= \int_0^1 \int_0^1 \cos(x + y) \exp(i\lambda(x + y + x^2 + y^2)) \, dx \, dy \\ &= \frac{i\pi}{8\lambda} \sum_{\nu=0}^1 e^{i\left(2\nu - \frac{(1+\lambda)^2}{2\lambda}\right)} \operatorname{erf}\left(-\frac{(-1)^\nu + \lambda}{\sqrt{2\lambda}(1+i)}, -\frac{(-1)^\nu + 3\lambda}{\sqrt{2\lambda}(1+i)}\right)^2 \end{aligned} \quad (186)$$

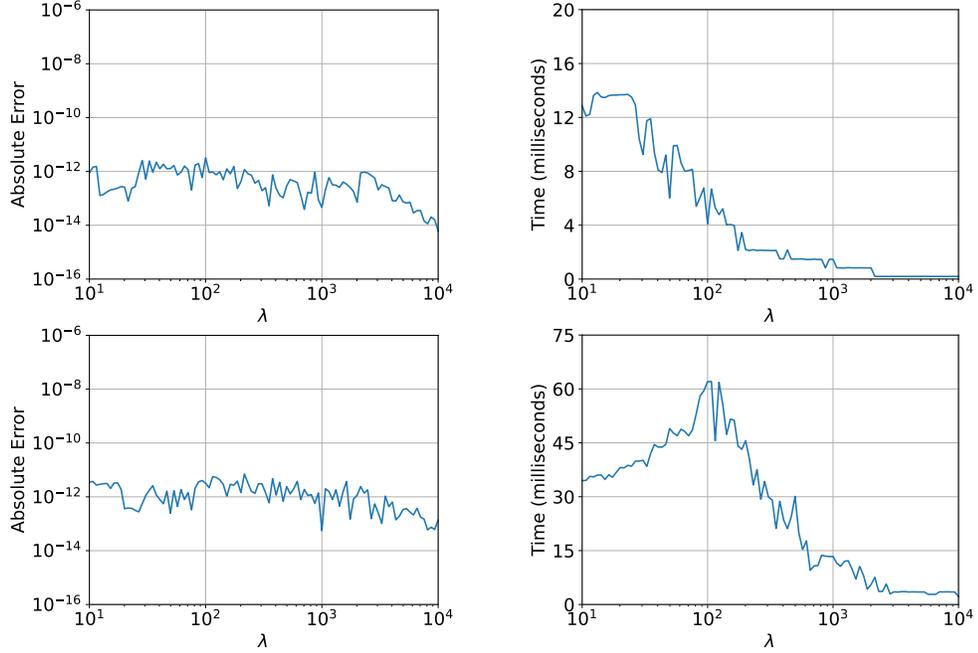


Figure 1: The results of the experiments of Section 6.1. The first row of plots pertains to the integral $I_1(\lambda)$, and the second pertains to $I_2(\lambda)$. In each row, the plot on the left gives the absolute error in the calculation of the integral as a function of λ , and the plot on the right shows the running time in milliseconds as a function of λ .

and

$$\begin{aligned}
 I_2(\lambda) &= \int_0^2 \int_0^2 \frac{\exp(i\lambda(\arctan(x) + \arctan(y)))}{(1+x^2)(1+y^2)} dx dy \\
 &= - \left(\frac{1 - \exp(i\lambda \arctan(2))}{\lambda} \right)^2,
 \end{aligned} \tag{187}$$

where $\text{erf}(z_0, z_1) = \text{erf}(z_1) - \text{erf}(z_0)$ is the generalized error function. The integral $I_1(\lambda)$ can be found in [2, 11], and the integral $I_2(\lambda)$ can be computed via a change of variables.

We sampled $\ell = 100$ equispaced points x_1, \dots, x_ℓ in the interval $[1, 4]$, and used the adaptive delaminating Levin method to evaluate the integrals $I_1(\lambda)$ and $I_2(\lambda)$ for each $\lambda = 10^{x_1}, \dots, 10^{x_\ell}$. Figure 1 gives the time taken to evaluate these integrals and the absolute error in the obtained values.

6.2 Integrals involving the Bessel functions, with explicit formulas

In the experiments described in this subsection, we used the adaptive delaminating Levin method to evaluate the integrals

$$\begin{aligned}
 I_3(\lambda) &= \int_1^2 \int_1^2 xy H_0^{(1)}(\lambda xy) dx dy \\
 &= \frac{-H_0^{(1)}(\lambda) + 2H_0^{(1)}(2\lambda) - H_0^{(1)}(4\lambda)}{\lambda^2},
 \end{aligned} \tag{188}$$

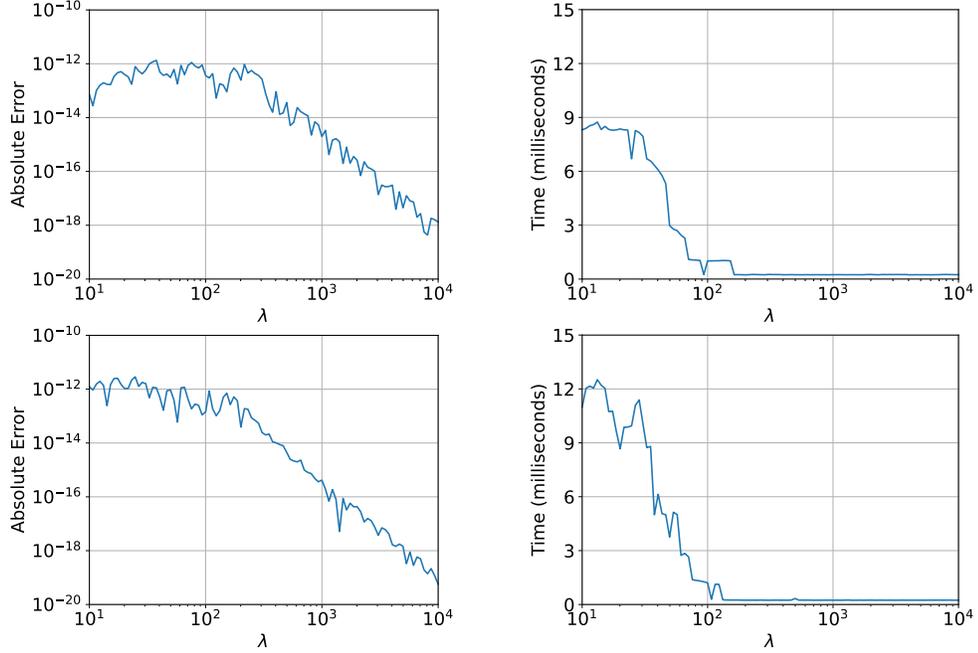


Figure 2: The results of the experiments of Section 6.2. The first row of plots pertains to the integral $I_3(\lambda)$, and the second pertains to $I_4(\lambda)$. In each row, the plot on the left gives the absolute error in the obtained value of the integral as a function of λ , and the plot on the right shows the running time in milliseconds as a function of λ .

and

$$\begin{aligned}
 I_4(\lambda) &= \int_1^2 \int_1^2 x^5 y^3 H_2^{(1)}(\lambda xy) dx dy \\
 &= \frac{H_4^{(1)}(\lambda) - 20H_4^{(1)}(2\lambda) + 64H_4^{(1)}(4\lambda)}{\lambda^2}.
 \end{aligned} \tag{189}$$

Both of the above integrals can be found in [25].

The Hankel functions of the first kind $H_\nu^{(1)}$ are defined via the formula

$$H_\nu^{(1)}(x) = J_\nu(x) + iY_\nu(x), \tag{190}$$

where J_ν and Y_ν are the Bessel function of the first and second kinds of order ν . For each $\nu \in \{0, 2\}$, we constructed a phase function ψ_ν^{bes} for the normal form

$$y''(x) + \left(1 + \frac{\frac{1}{4} - \nu^2}{x^2}\right) y(x) = 0 \text{ in } (0, \infty) \tag{191}$$

of Bessel's differential equation using the algorithm of [13], allowing the Bessel functions of the first and second kinds of order ν to be computed as

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}} \frac{\sin(\psi_\nu^{\text{bes}}(x))}{\sqrt{\frac{d}{dx} \psi_\nu^{\text{bes}}(x)}} \quad \text{and} \quad Y_\nu(x) = -\sqrt{\frac{2}{\pi x}} \frac{\cos(\psi_\nu^{\text{bes}}(x))}{\sqrt{\frac{d}{dx} \psi_\nu^{\text{bes}}(x)}}. \tag{192}$$

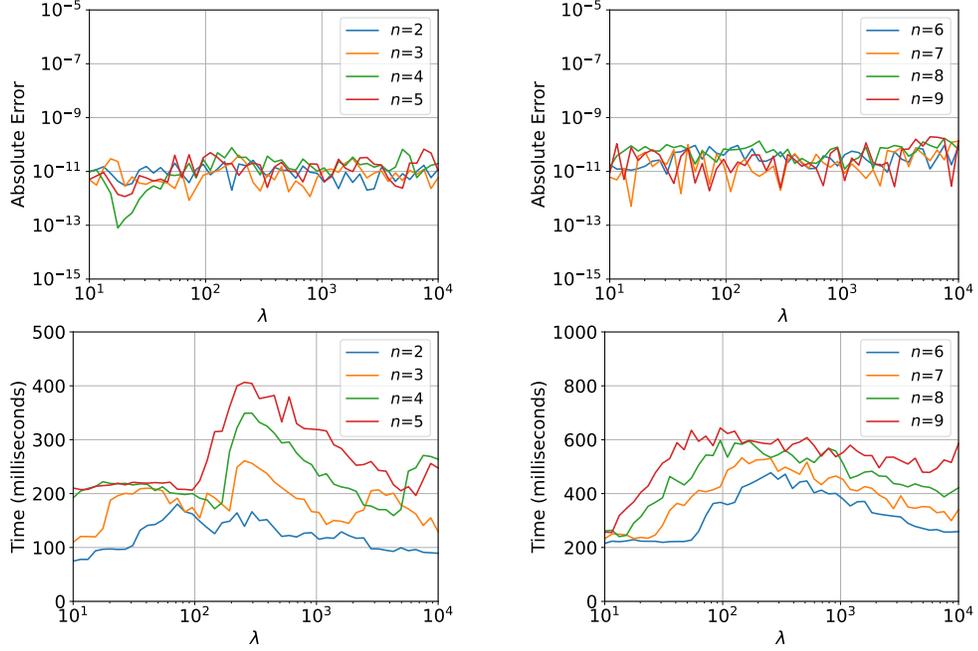


Figure 3: The results of the experiments of Section 6.3. The first row of plots shows the absolute error in the calculated value of $I_5(\lambda, n)$ as a function of λ for the values of n considered. The second row of plots shows the running time of the adaptive delaminating Levin method in milliseconds as a function of λ for $n = 2, 3, 4, 5, 6, 7, 8,$ and 9 .

The representation

$$H_\nu^{(1)}(x) = \sqrt{\frac{2}{\pi x \frac{d}{dx} \psi_\nu^{\text{bes}}(x)}} \exp\left(i\left(\psi_\nu^{\text{bes}}(x) - \frac{\pi}{2}\right)\right) \quad (193)$$

of the Hankel functions of the first kind was then used in conjunction with the adaptive delaminating Levin method to evaluate the above integrals. The construction of the phase functions took 2.79 ms and 1.29 ms for $H_0^{(1)}$ and $H_2^{(1)}$, respectively.

We sampled $\ell = 100$ equispaced points x_1, \dots, x_ℓ in the interval $[1, 4]$, and used the adaptive delaminating Levin method to evaluate the above integrals for each $\lambda = 10^{x_1}, \dots, 10^{x_\ell}$. The results are shown in Figure 2. The reported times account for both the evaluation of the phase functions and the execution of the adaptive delaminating Levin method. Note that the time taken to construct the phase function is not included, as it represents a one-time cost that is independent of λ .

6.3 Behavior in the presence of a stationary point

In the experiments described in this subsection, we consider the integral

$$I_5(\lambda, n) = \int_{-1}^1 \int_{-1}^1 \frac{1}{1+x^2+y^2} \exp(i\lambda(x^n + y^n)) dx dy \quad (194)$$

in order to illustrate the behavior of the adaptive delaminating Levin method in the presence of a stationary point.

We sampled $\ell = 50$ equispaced points x_1, \dots, x_ℓ in the interval $[1, 4]$, and, for each $\lambda = 10^{x_1}, \dots, 10^{x_\ell}$ and $n = 2, 3, 4, 5, 6, 7, 8, 9$, we evaluated $I_5(\lambda, n)$ using the adaptive delaminating Levin method. The results are shown in Figure 3. We observe that the running time of our algorithm increases mildly with n and, for each fixed n , it is largely independent of λ after peaking in the moderate-frequency regime.

This behavior can be explained by the analysis presented in Section 3. Since the phase function has a stationary point at the origin, Theorem 3.4 implies that the delaminating Levin method will yield an accurate result on subrectangles of the form $[-\delta, \delta]^2$, provided that the magnitude of ∇g is sufficiently small on the subrectangle. Theorem 3.3 indicates that the delaminating Levin method will also yield an accurate result on any subrectangle of $[-1, 1]^2$ which is bounded away from the stationary point, provided that the ratio of the maximum to minimum value of the partial derivative of g along the delamination direction is small and that

$$h(x', y') = f(\mathbf{u}^{-1}(x', y')) \det(D\mathbf{u}^{-1})(x', y') \quad (195)$$

can be represented by a well-behaved function with a small bandlimit over the subrectangle. The latter condition is more-or-less equivalent to the requirement that h can be represented by a Chebyshev expansion of small fixed order. Thus, we expect the adaptive delaminating Levin method to subdivide the domain until one or both of Theorems 3.3 and 3.4 apply to the individual subrectangles.

Since $g_x(x, y) = \lambda n x^{n-1}$ and $g_y(x, y) = \lambda n y^{n-1}$, we see that the condition

$$\delta < \left(\frac{C}{n\lambda}\right)^{\frac{1}{n-1}} \quad (196)$$

must be satisfied in order for $|g_x(x, y)|$ and $|g_y(x, y)|$ to be bounded by a constant $C < 1$, so we expect the adaptive delaminating Levin method to subdivide the domain into at least

$$\mathcal{O}\left(\log\left(\frac{1}{\delta}\right)\right) = \mathcal{O}\left(\frac{\log(n\lambda)}{n-1}\right) \quad (197)$$

subrectangles. The algorithm will further subdivide the subrectangles outside this low frequency region until, on each of the resulting subrectangles, the ratio of the maximum to minimum value of the partial derivative of g along the delamination direction is small, and h is accurately represented by a Chebyshev expansion of small fixed order.

Away from the stationary point, the low frequency region is enclosed by a region in which the variation of the partial derivatives of g is bounded, so that, on each subrectangle R , the ratio G_1/G_0 is small, where $G_1 = \max_{(x,y) \in R} |g_v(x, y)|$ and $G_0 = \min_{(x,y) \in R} |g_v(x, y)|$, and where v is the delamination direction. It is not hard to see that the subrectangles must be smallest in the moderate frequency region bordering the low frequency region, and that they become larger away from the stationary point. As a result, the majority of the subrectangles lie on an annulus encircling the low frequency region. From (196), we see that the low frequency region containing the stationary point expands with increasing n , so the area of the annulus—and hence the number of subrectangles contained within—increases with n . It then follows that, when λ is fixed, the total number of subrectangles in the adaptive discretization of the domain must

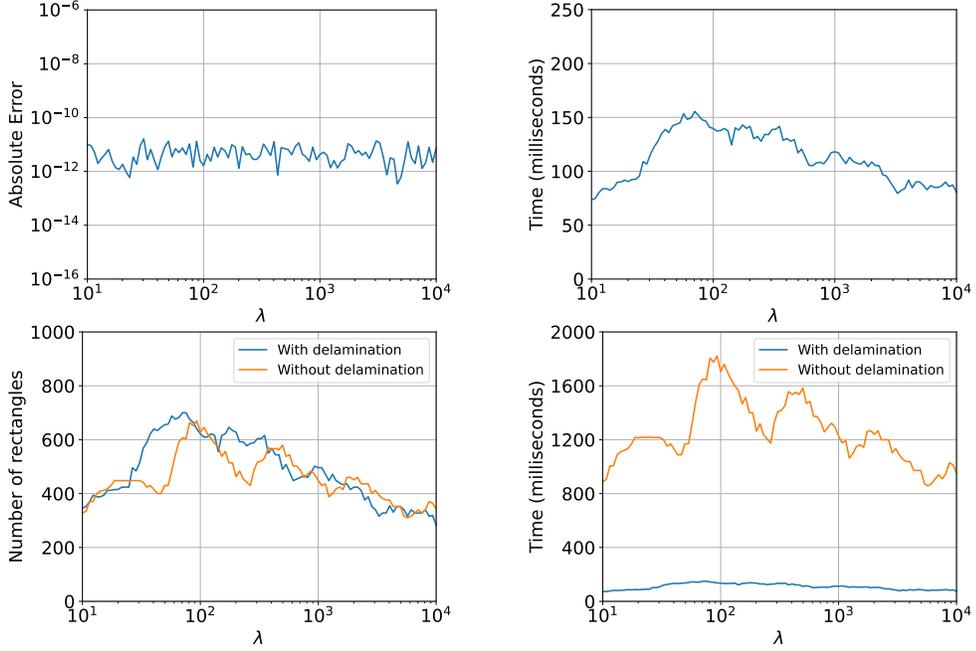


Figure 4: The results of the experiments of Section 6.4. The first row of plots shows the absolute error in the calculated value of the integral $I_6(\lambda)$ and the running time of the adaptive delaminating Levin method in milliseconds as a function of λ . The second row of plots compares both the number of rectangles in the adaptive discretization of the domain and the running time, between the adaptive delaminating Levin method and an adaptive method without delamination.

increase with n , which explains why the running time of our algorithm increases with n . Our analysis is consistent with the results shown in Figure 6, which displays the adaptive discretization of the domain produced by the adaptive delaminating Levin method for $\lambda = 300$ and $n = 2, 3, 4$, and 5 . The subrectangles are colored according to the ratio of the maximum to minimum value of the partial derivative of g along the delamination direction, except for subrectangles where the relative maximum values of the partial derivatives of g are less than $1/4$, which are shown in grey.

6.4 Behavior in the presence of stationary and resonance points

In the experiments described in this subsection, we consider the integral

$$I_6(\lambda) = \int_{-1}^1 \int_{-1}^1 (1 + xy) \exp(i\lambda(x^2 - xy - y^2)) dx dy, \quad (198)$$

which can be found in [26], in order to understand the behavior of the adaptive delaminating Levin method in the presence of both stationary and resonance points. The phase function of $I_6(\lambda)$ is noteworthy due to the stationary point at the origin, and the lines $y = 2x$ and $y = -x/2$ where g_x and g_y vanish, respectively.

In the first set of experiments, we sampled $\ell = 100$ equispaced points x_1, \dots, x_ℓ in the interval $[1, 4]$, and used the adaptive delaminating Levin method to evaluate the integral $I_6(\lambda)$ for each $\lambda = 10^{x_1}, \dots, 10^{x_\ell}$. The first row of Figure 4 presents the results. We note that resonance points arise whenever the lines $y = 2x$ and $y = -x/2$ intersect

the boundaries of a subrectangle. Nevertheless, since the univariate Levin method is used to compute the boundary integrals in (179), the resonance points do not trigger any additional subdivisions of subrectangles, as evidenced by Figure 7.

In a second set of experiments, we evaluated the above integral for each $\lambda = 10^{x_1}, \dots, 10^{x_\ell}$ using an adaptive method which solves a rectangular linear system that enforces the collocation condition on the $(2k - 1) \times (2k - 1)$ tensor product Chebyshev grid. The results are shown in the second row of Figure 4. This approach follows the analysis presented in Section 4 more closely, but the resulting rectangular linear system lacks a block structure and must be solved monolithically. Consequently, it incurs a significantly higher computational cost, especially as the size of the collocation system increases with the value of k and the collocation grid size. We observe that, while the adaptive delaminating Levin method produces slightly more subdivisions in the moderate frequency regime, its overall running time is substantially lower due to the smaller linear systems it solves. Figure 7 shows the adaptive discretization of the domain produced by the adaptive delaminating Levin method and the adaptive method described above, for $\lambda = 10^1, 10^2, 10^3, 10^4$. The subrectangles are colored according to the ratio of the maximum to minimum value of the partial derivative of g along the delamination direction. The subrectangles where the relative maximum values of the partial derivatives of g are less than $1/4$ are colored grey.

6.5 Behavior in the presence of many stationary points

In the experiments described in this subsection, we considered the integral

$$I_7(\lambda, m) = \int_0^1 \int_0^1 \exp\left(i\lambda \left(\sin^2\left(\frac{\pi}{2}mx\right) + \sin^2\left(\frac{\pi}{2}my\right)\right)\right) dx dy, \quad (199)$$

which has $(m + 1)^2$ stationary points in the unit square $[0, 1]^2$.

We sampled $\ell = 50$ equispaced points x_1, \dots, x_ℓ in the interval $[1, 4]$, and for each $\lambda = 10^{x_1}, \dots, 10^{x_\ell}$ and $m = 1, 2, 3, 4, 5, 6, 7, 8$, we evaluated $I_7(\lambda, m)$ using the adaptive delaminating Levin method. The results are shown in Figure 5. We observe that the running time of our algorithm grows linearly with the number of stationary points, and is essentially independent of λ for all values of m considered.

7 Conclusions

We have shown that the Levin PDE admits a slowly-varying, approximate solution across all frequency regimes, regardless of whether stationary and resonance points are present, and that the adaptive delaminating Levin method, when combined with a univariate adaptive Levin method, rapidly and accurately evaluates bivariate oscillatory integrals over rectangular domains. We have also presented numerical experiments demonstrating the effectiveness of the adaptive delaminating Levin method on a large class of bivariate oscillatory integrals, including many with stationary and resonance points.

The numerical method can be readily extended to evaluate bivariate oscillatory integrals over general non-polytopal domains, provided that the integrand is defined on a bounding box containing the domain. With a piecewise smooth parametrization of the boundary, the univariate adaptive Levin method is well-suited for evaluating the boundary

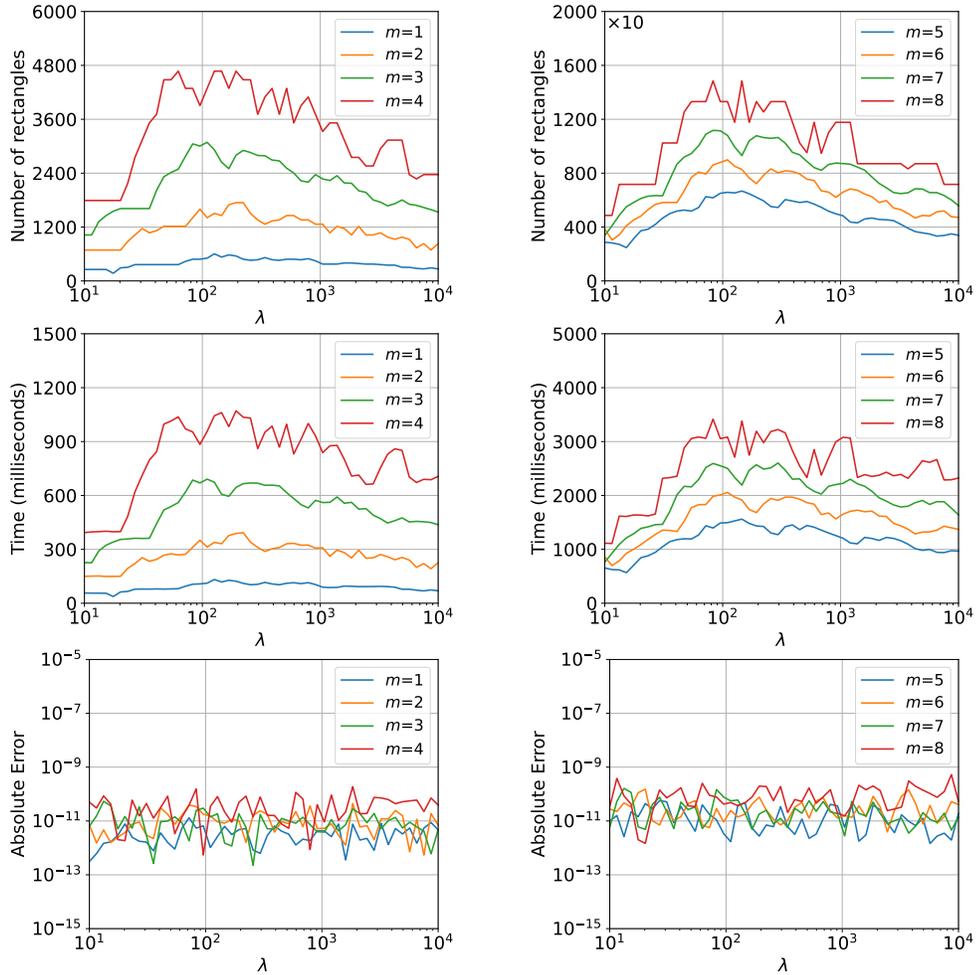


Figure 5: The results of the experiments of Section 6.5. The plots in the first row give the number of rectangles in the adaptively determined discretization of $[0, 1]^2$ used to compute the integral $I_7(\lambda, m)$ as a function of λ for the values of m considered. The plots in the second row show the runtime of the adaptive delaminating Levin method as a function of λ for the values of m considered. The third row of plots gives the absolute error in the calculated value of $I_7(\lambda, m)$ as a function of λ for $m = 1, 2, 3, 4, 5, 6, 7,$ and 8 .

integrals after solving the Levin PDE on the bounding box. When the integrand cannot be evaluated outside the domain, but a smooth mapping exists from a rectangle to the domain, the adaptive delaminating Levin method can be applied on the pullback, with the Jacobian of the mapping appearing in the integrand. As a result, the method can be applied to any domain meshed with curved quadrilateral elements, provided that the mapping functions for each element are sufficiently smooth.

Finally, we note that both the delaminating Levin method and the theoretical analysis presented in this paper are only weakly dependent on the dimensionality of the ambient space, as tensor products are the primary theoretical tool. As a result, our method can be easily generalized to evaluate volumetric oscillatory integrals in dimensions three and higher.

References

- [1] Alfredo Deaño, Daan Huybrechs, and Arieh Iserles. *Computing Highly Oscillatory Integrals*. SIAM, 2018.
- [2] David Levin. Procedures for computing one- and two-dimensional integrals of functions with rapid irregular oscillations. *Math. Comp.*, 38(158): 531–538, 1982.
- [3] Sheehan Olver. On the Quadrature of Multivariate Highly Oscillatory Integrals Over Non-polytope Domains. *Numer. Math.*, 103(4): 643–665, 2006.
- [4] Sheehan Olver. Fast, numerically stable computation of oscillatory integrals with stationary points. *BIT Numer. Math.*, 50(1): 149–171, 2010.
- [5] Yinkun Wang and Shuhuang Xiang. Fast and stable augmented Levin methods for highly oscillatory and singular integrals. *Math. Comp.*, 91(336): 1893–1923, 2022.
- [6] JianBing Li, XueSong Wang, and Tao Wang. A universal solution to one-dimensional oscillatory integrals. *Sci. China Ser. F-Inf. Sci.*, 51(10): 1614–1622, 2008.
- [7] Jianbing Li, Xuesong Wang, Tao Wang, and Shunping Xiao. An improved Levin quadrature method for highly oscillatory integrals. *Appl. Numer. Math.*, 60(8): 833–842, 2010.
- [8] Andrew James Moylan. *Highly Oscillatory Integration, Numerical Wave Optics, and the Gravitational Lensing of Gravitational Waves*. PhD thesis. The Australian National University, 2008.
- [9] Shukui Chen, Kirill Serkh, and James Bremer. On the adaptive Levin method. *Numer. Math.*, 156(6): 1927–1985, 2024.
- [10] Anthony Ashton. Cauchy data for Levin’s method. *IMA J. Numer. Anal.*, 45(1): 87–125, 2025.
- [11] Jianbing Li, Xuesong Wang, Tao Wang, and Chun Shen. Delaminating quadrature method for multi-dimensional highly oscillatory integrals. *Appl. Math. Comput.*, 209(2): 327–338, 2009.
- [12] James Bremer. On the numerical solution of second order ordinary differential equations in the high-frequency regime. *Appl. Comput. Harmon. Anal.*, 44(2): 312–349, 2018.
- [13] James Bremer. Phase function methods for second order linear ordinary differential equations with turning points. *Appl. Comput. Harmon. Anal.*, 65: 137–169, 2023.
- [14] Murdock Aubry and James Bremer. A solver for linear scalar ordinary differential equations whose running time is bounded independent of frequency. *arXiv preprint*, arXiv:2311.08578, 2023.
- [15] Christine Bernardi and Yvon Maday. Properties of Some Weighted Sobolev Spaces and Application to Spectral Approximations. *SIAM J. Numer. Anal.*, 26(4): 769–829, 1989.

- [16] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Wiley, 1999.
- [17] J. Ian Richards and Heekyung Youn. *Theory of Distributions: A Non-Technical Introduction*. Cambridge University Press, 1995.
- [18] Elias M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series Vol. 30. Princeton University Press, 2016.
- [19] John P. Boyd. Approximation of an analytic function on a finite real interval by a bandlimited function and conjectures on properties of prolate spheroidal functions. *Appl. Comput. Harmon. Anal.*, 15(2): 168–176, 2003.
- [20] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.
- [21] Mohan Zhao and Kirill Serkh. On the approximation of singular functions by series of noninteger powers. *IMA J. Numer. Anal.*, 2025.
- [22] François Trèves. *Topological Vector Spaces, Distributions and Kernels*. Dover Publications, 2006.
- [23] Victor Guillemin and Alan Pollack. *Differential Topology*. AMS, 2010.
- [24] John C. Mason and David C. Handscomb. *Chebyshev Polynomials*. CRC, 2003.
- [25] Ruyun Chen and Shuhuang Xiang. Note on the homotopy perturbation method for multivariate vector-value oscillatory integrals. *Appl. Math. Comput.*, 215(1): 78–84, 2009.
- [26] Daan Huybrechs and Stefan Vandewalle. The Construction of cubature rules for multivariate highly oscillatory integrals. *Math. Comp.*, 76(260): 1955–1981, 2007.

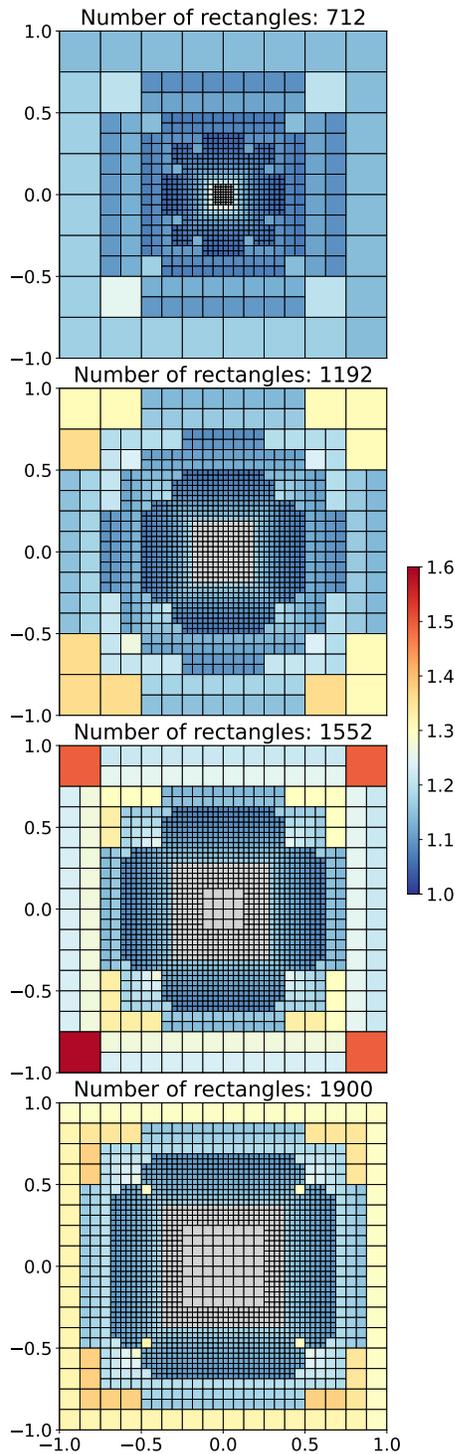


Figure 6: The plots show the adaptive discretization of the domain produced by the adaptive delaminating Levin method for computing the integral $I_5(300, n)$ with $n = 2, 3, 4,$ and 5 . The subrectangles are colored according to the ratio of the maximum to minimum value of the partial derivative of g along the delamination direction. The subrectangles are grey if the relative maximum values of the partial derivatives of g are less than $1/4$.

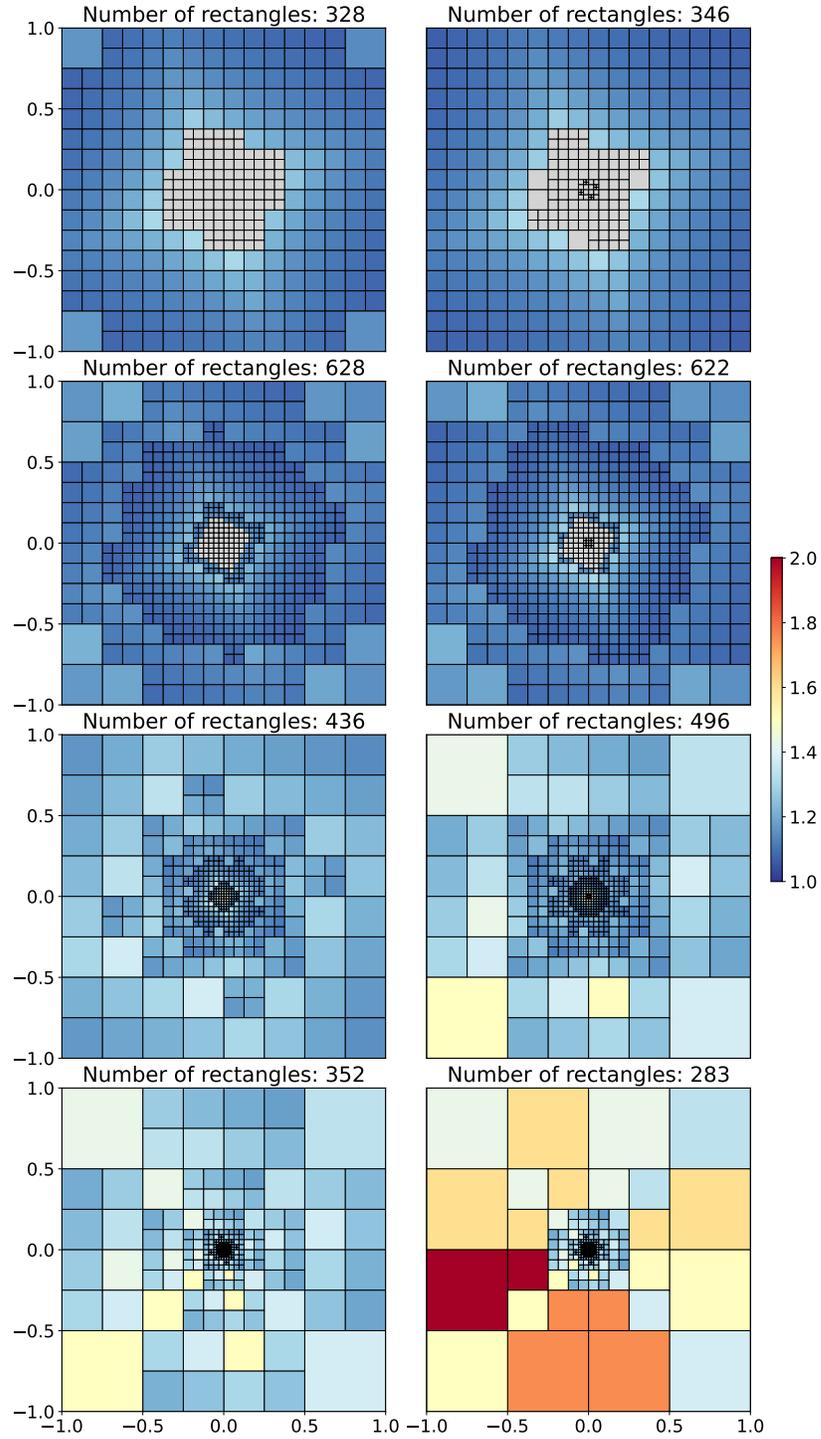


Figure 7: The plots in the column on the left and right show the adaptive discretization of the domain produced by the adaptive method that solves the rectangular systems without delamination, and by the adaptive delaminating Levin method, respectively, for computing the integral $I_6(\lambda)$ with $\lambda = 10^1, 10^2, 10^3, 10^4$. The subrectangles are colored according to the ratio of the maximum value of the partial derivative of g along the delamination direction to the minimum value of the partial derivative of g along the delamination direction. The subrectangles are grey if the relative maximum values of the partial derivatives of g are less than $1/4$.